AIP | Journal of Mathematical Physics

Stochastic Processes and Their Representations in Hilbert Space

F. G. Hall and R. E. Collins

Citation: Journal of Mathematical Physics **12**, 100 (1971); doi: 10.1063/1.1665465
View online: http://dx.doi.org/10.1063/1.1665465
View Table of Contents: http://scitation.aip.org/content/aip/journal/jmp/12/1?ver=pdfcov
Published by the AIP Publishing

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 1 JANUARY 1971

Stochastic Processes and Their Representations in Hilbert Space*

F. G. HALL AND R. E. COLLINS
Department of Physics, University of Houston, Houston, Texas 77004

(Received 8 April 1970)

Beginning with an intuitive consideration of sequences of measurements, we define a time-ordered event space representing the collection of all imaginable outcomes for measurement sequences. We then postulate the *generalized distributive relation* on the event space and examine the class of measurements for which this relation can be experimentally validated. The generalized distributive relation is shown to lead to a σ-additive conditional probability on the event space and to a *predictive* and *retrodictive* formalism for stochastic processes. We then show that this formalism has a predictive and a retrodictive representation in a separable Hilbert space $\mathcal{H}$, which has no counterpart in unitary quantum dynamics.

INTRODUCTION

A recent series of papers[1–6] has developed the idea that much of the formal mathematical structure of physical theory can be deduced directly from the statistical nature of experimental data. The present paper presents that portion of these studies which bears directly on the evolution of irreversible physical processes.

We begin the study of the evolution of a system by insisting that if we are to say we have observed the dynamic behavior of the system, then we must monitor the system by a sequence of time-documented measurements $\{M_0 \to M_1 \to M_2 \to \cdots \to M_L\}$.

With each of the measurements in the sequence, we associate in our mind a collection of *possible outcomes*, the collection being determined, of course, by the properties of the measuring apparatus. We may also associate a collection of possible outcomes with the entire measurement sequence.

We assume that all experimental data is statistical in nature, i.e., each outcome in the collection of possible outcomes is a random event. This assumption leads us to consider probability theory as a mathematical model for the kinematics of a system.

Since our imagination, at least for physical-experimental situations, seems to be conditioned by conventional logic, we will assume that a σ-algebra describes the collection of imaginable outcomes (event space) of a measurement and that the frequency of outcomes can be described by a σ-additive measure of unit norm whose domain is the σ-algebra.

This approach does not differ from conventional approaches except, as we will show, in the definition of the σ-algebra of possible outcomes for measurement sequence and the conditional probability defined on this σ-algebra.

We will show that an equivalence relation must be defined on the σ-algebra for the measurement sequences in order to obtain the predictive and retrodictive random walk formulation for stochastic processes. This equivalence relation, the *generalized distributive relation*, is empirical in nature and is not deducible from the logical structure of the mathematics describing the measurement sequence.

We will then show that the predictive and retrodictive random walk formulations for the dynamics of a physical system have representations in a separable Hilbert space $\mathcal{H}$, which differ considerably from the conventional quantum representation. It appears that the dynamical laws of conventional quantum theory are not the most general representation of the random walk formulation in $\mathcal{H}$.

THE MEASUREMENT

For the sake of clarity and brevity in the following discussions, we will begin by defining the measurement process.

We assume that an experimental situation may be completely described by a countable, functionally independent set of real-valued functions $(f_1, f_2, f_3, \cdots)$, which may be arbitrarily partitioned into two functionally independent sets; one set, a K-tuple $(f'_1, f'_2, \cdots, f'_K)$ describing the results of K simultaneous measurements, and one set $(f''_1, f''_2, f''_3, \cdots)$ describing the environment conditioning the measurement. (This simply states that we must be satisfied to determine a finite number of system properties.)

We suppose that a measurement is always limited to some finite resolution, and thus each of $f'_1, f'_2, \cdots, f'_K$ has a countable range $R_{f_1}, R_{f_2}, \cdots, R_{f_K}$, respectively. Since each of $f'_1, f'_2, \cdots, f'_K$ has a countable range, there exists a *countable collection*

$$\{(p_1, p_2, \cdots, p_K)\}_{p_1 \in R_{f_1}, p_2 \in R_{f_2}, \cdots, p_K \in R_{f_K}}$$

of K-tuples of real numbers (denoted $\{\hat{p}_k\}_{k=1,2,3,\ldots}$) which contains all possible K-tuples of real numbers in the range of $(f'_1, f'_2, \cdots, f'_K)$.

Such assumptions lead us to make the following definitions:

A *measurement* of a system is an operation performed on a system which assigns a *configuration* $\hat{p}_k \in \{\hat{p}_k\}_{k=1,2,3,\ldots}$ to the system.

The *spectrum* of a measurement is the collection of all possible configurations $\{\hat{p}_k\}_{k=1,2,3,\ldots}$. For example, if we are interested in the pressure and volume of a system, then a configuration assigned to the system is a 2-tuple of real numbers (P_i, V_i) in the range of the functions P and V, respectively.

We may now define the *event space* as the collection of all imaginable outcomes for a measurement. Let C_i denote the spectrum of a measurement process M_i. The event space $\{E_i(C_i)\}$ is the σ-algebra[7] of subsets of C_i. The motivations for such a choice for the event space are discussed in several texts[8,9]; arguments against such a choice have been discussed by Jauch.[10] We will assume the σ-algebra to be a valid representation since as we will see there seem to be many physical situations for which the σ-algebra is appropriate and yields results not obtainable by conventional quantum theory.

Here we will refer to the members of $\{E_i(C_i)\}$ as *events* and define the probability for an event as a σ-additive measure P of unit norm on $\{E_i(C_i)\}$. Such a function has the following properties:

(i) If $E \in \{E_i(C_i)\}$, $0 \leq P(E) \leq 1$;

(ii) $P(\phi) = 0$; ϕ is the null event corresponding to the empty set in $\{E_i(C_i)\}$;

(iii) $P(C_i) = 1$; $C_i = \bigcup_{k_i=1}^{\infty} (\hat{p}_{k_i})$ (set union is interpreted as logical or);

(iv) if $\{E_j\}_{j=1,2,3,\ldots}$ is a disjoint sequence of sets in $\{E_i(C_i)\}$, then

$$P\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} P(E_j).$$

There is a much wider agreement on the properties of P than the event space because of obvious physical interpretations. Axioms (i) and (ii) follow from the operational definition of probability. Axiom (iii) simply states that some value in the spectrum must be obtained as a result of M_i, and axiom (iv) is the mathematical statement of the familiar mutually exclusive rule in probability theory.

With this brief introduction we may now consider sequences of measurement operations.

SEQUENCES OF MEASUREMENTS

We wish now to consider the time-documented sequence of measurements $\{M_0 \to M_1 \to M_2 \to \cdots \to M_L\}$. By time documented we mean M_i occurs at t_i and, in case $i < j$, then $t_i < t_j$. Since each M_i has an associated event space $\{E_i(C_i)\}$, the collection of all imaginable outcomes for the ordered sequence $\{M_i\}_{i=1,L}$ is a physically meaningful notion; thus we proceed to define the event space $\{E(C)\}$ for $\{M_i\}_{i=1,L}$. Let C denote the Cartesian product space for the sequence of σ-algebras $\{\{E_0(C_0)\} \to \{E_1(C_1)\} \to \{E_2(C_2)\} \to \cdots \to \{E_L(C_L)\}\}$, i.e.,

$$C = \{E_0(C_0)\} \otimes \{E_1(C_1)\} \otimes \cdots \otimes \{E_L(C_L)\}. \quad (1)$$

$\{E(C)\}$ is the event space for $\{M_i\}_{i=1,L}$ means that E is an event in $\{E(C)\}$ only in case E is a subset of C.

That $\{E(C)\}$ contains the imaginable paths of outcomes for the measurement sequence can be seen from recognizing that $\{E(C)\}$ contains the collection $\{S_n\}$ of *simple paths* $\{(\hat{p}_{k_0} \to \hat{p}_{k_1} \to \cdots \to \hat{p}_{k_L})\}$ (which are read as "$\hat{p}_{k_0}$ occurred, then $\hat{p}_{k_1}$ occurred, then, $\cdots$, then $\hat{p}_{k_L}$ occurred"), the *compound paths* such as

$$\left\{\left(\bigcup_{k_0=1}^{l_0} \hat{p}_{k_0} \to \bigcup_{k_1=1}^{l_1} \hat{p}_{k_1} \to \cdots \to \bigcup_{k_L=1}^{l_L} \hat{p}_{k_L}\right)\right\},$$

and the unions and intersections of the compound paths, for example,

$$\left(\hat{p}_{k_0} \to \bigcup_{k_1=1}^{l_1} \hat{p}_{k_1} \to \cdots \to \hat{p}_{k_L}\right) \cup \left(\hat{p}_{k_0'} \to \hat{p}_{k_1'} \to \cdots \to \bigcup_{k_L=1}^{l_L} \hat{p}_{k_L}\right).$$

Notice that, in contrast to the usual route in probability theory,[8] we have not defined $\{E(C)\}$ to be the Cartesian product space of the σ-algebras $\{E_0(C_0)\} \otimes \{E_1(C_1)\} \otimes \cdots \otimes \{E_L(C_L)\}$. Such a choice is not the most general one since it requires that set operations in $\{E(C)\}$ be defined in terms of set operations in $\{E_i(C_i)\}$. For our definition of $\{E(C)\}$ we see that C does not form a σ-algebra since it contains no unions of members of C. However, by choosing an equivalence relation between members of C and the compliment of C in $\{E(C)\}$, one can "induce" a σ-algebra on C. As we shall see in the next section, such a choice is empirical and seems necessary in order to produce the stochastic process.

PROBABILITY ON $\{E(C)\}$

We now turn our attention to probability functions on $\{E(C)\}$ and, in particular, conditional probabilities. We will assume in the following discussions that the environment for the sequence $\{M_i\}_{i=1,L}$ is fixed and described by $\hat{q}$. We will tacitly require that all probability functions on $\{E(C)\}$ be conditioned by $\hat{q}$.

The unit norm condition for P on $\{E_i(C_i)\}$ is given by

$$P(C_i) = 1, \quad (2)$$

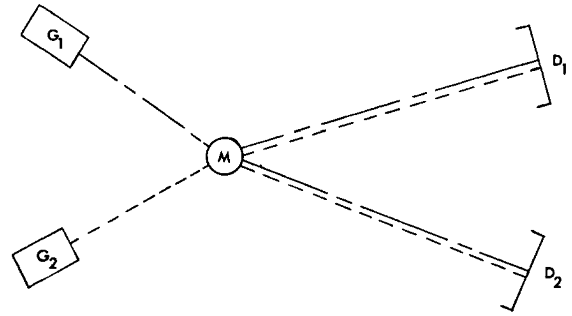

FIG. 1. Electron gun apparatus.

which was interpreted as the probability for some event to occur during M_i. In view of this, it would seem reasonable that, for the sequence of measurements,

$$P(C_0 \to C_1 \to \cdots \to C_L) = 1 \tag{3}$$

and, for the simple paths $\{S_n\}_{n=1,2,\ldots}$ in $\{E(C)\}$,

$$P\left(\bigcup_{n=1}^{\infty} S_n\right) = 1, \tag{4}$$

which is interpreted as *some simple path* must occur. In order for (3) and (4) to be true, we must postulate the following relation:

If

$$E_0 \in \{E_0(C_0)\}, \quad E_1 \in \{E_1(C_1)\}, \cdots,$$

both

$$E_i' \;\&\; E_i'' \in \{E_i(C_i)\}, \cdots, E_L \in \{E_L(C_L)\},$$

then

$$\begin{aligned}(E_0 \to E_1 \to \cdots \to E_i' \cup E_i'' \to \cdots \to E_L) \\ = (E_0 \to E_1 \to \cdots \to E_i' \to \cdots \to E_L) \\ \cup (E_0 \to E_1 \to \cdots \to E_i'' \to \cdots \to E_L).\end{aligned} \tag{5}$$

We will also require the class of measurements that we are investigating to obey

$$\begin{aligned}P((\hat{p}_{k_0} \to \hat{p}_{k_1} \to \cdots \to \hat{p}_{k_i} \to \cdots \to \hat{p}_{k_L}) \\ \cap (\hat{p}_{k_0} \to \hat{p}_{k_1} \to \cdots \to \hat{p}_{k_i'} \to \cdots \to \hat{p}_{k_L})) \\ = \begin{cases} 0, & k_i' \neq k_i \\ P(\hat{p}_{k_0} \to \hat{p}_{k_1} \to \cdots \to \hat{p}_{k_i} \to \cdots \to \hat{p}_{k_L}), & k_i' = k_i, \end{cases}\end{aligned} \tag{6}$$

which simply states that only one configuration may be obtained as the result of a measurement. Statements (4), (5), and (6) must be *a posteriori* in nature, not derivable from any *a priori* consideration. To clarify this point, consider the following measurement situation.

The schematic in Fig. 1 describes two electron guns G_1 and G_2 firing at a fixed target M. These electrons are scattered from M and detected at D_1 or D_2. The entire apparatus is placed in a cloud chamber so that the track of each electron can be monitored if desired. Such a device will serve to examine the generalized σ-algebra $\{E(C)\}$ and Eqs. (5) and (6).

Let M_0 denote detection of the firing of the guns, M_1 denote detection of scattering from the target, and M_2 denote detection at D_1 or D_2. We may now build $\{E(C)\}$ for the sequence $\{M_0 \to M_1 \to M_2\}$. The σ-algebras $\{E_0(C_0)\}$, $\{E_1(C_1)\}$, and $\{E_2(C_2)\}$ are given by

$$\begin{aligned}\{E_0(C_0)\} &= \{(G_1), (G_2), (G_1 \cup G_2), (G_1 \cap G_2), \varnothing\}, \\ \{E_1(C_1)\} &= \{M, \varnothing\}, \\ \{E_2(C_2)\} &= \{(D_1), (D_2), (D_1 \cup D_2), (D_1 \cap D_2), \varnothing\}.\end{aligned} \tag{7}$$

C as defined earlier is given by the Cartesian product space $\{E_0(C_0)\} \otimes \{E_1(C_1)\} \otimes \{E_2(C_2)\}$, and $\{E(C)\}$, the event space for $\{M_0 \to M_1 \to M_2\}$, is the σ-algebra of subsets of C.

If we form $\{E(C)\}$ by the prescription given above, we see that $\{E(C)\}$ contains events such as $(G_1 \to M \to D_1)$, $(G_1 \to M \to D_2)$, $(G_2 \to M \to D_1)$, and $(G_2 \to M \to D_2)$, the union of these $[(G_1 \to M \to D_1) \cup (G_1 \to M \to D_2) \cup (G_2 \to M \to D_1) \cup (G_2 \to M \to D_2)]$, and $(G_1 \cup G_2 \to M \to D_1 \cup D_2)$. It is quite natural to interpret each of the events in the collection $\{(G_i \to M \to D_j)\}$ as the event for a certain *simple path* to be observed in the cloud chamber. The union of these simple paths would, of course, be interpreted as the event for one or another of the *simple paths* to occur. However, the event $(G_1 \cup G_2 \to M \to D_1 \cup D_2)$ would appear to have no simple interpretation as an event independent of the events for simple paths. [The event $(G_1 \cup G_2 \to M \to D_1 \cup D_2)$ seems a likely candidate for a "superposition" event defined by Jauch[10] if the σ-algebraic structure of $\{E(C)\}$ is modified. This investigation will constitute another paper.]

We do see, however, that Eqs. (3) and (4) can be satisfied for $\{E(C)\}$ only in case Eq. (5) is valid on $\{E(C)\}$. Equation (5) defines the event $(G_1 \cup G_2 \to M \to D_1 \cup D_2)$ in terms of the simple paths in $\{E(C)\}$, i.e., by Eq. (5)

$$\begin{aligned}(G_1 \cup G_2 \to M \to D_1 \cup D_2) \\ = (G_1 \to M \to D_1 \cup D_2) \cup (G_2 \to M \to D_1 \cup D_2) \\ = (G_1 \to M \to D_1) \cup (G_1 \to M \to D_2) \\ \cup (G_2 \to M \to D_1 \cup D_2) \\ = (G_1 \to M \to D_1) \cup (G_1 \to M \to D_2) \\ \cup (G_2 \to M \to D_1) \cup (G_2 \to M \to D_2),\end{aligned} \tag{8}$$

and therefore the requirement for $\{E(C)\}$ that $P(C)$ is unity is consistent with Eqs. (3) and (4).

We will call Eq. (5) the *generalized distributive relation* of the set operation $\bullet$ with respect to the ordering operation $\to$. We see that this relation is *a posteriori* in nature, i.e., it is *not* required by the structure of $\{E(C)\}$. Only when we require Eq. (3) or Eq. (4) to be valid must we require the generalized distributive relation. The validity of Eq. (4) can be tested only if each of the simple paths are observable; thus the generalized distributive relation is ultimately *a posteriori* in nature.

It should be evident that Eq. (5) "induces" a σ-algebra on C and thus reduces $\{E(C)\}$ to the conventional σ-algebra of simple paths. [If Eqs. (3) and (4) are to be consistent with the requirement $P(C) = 1$, then the generalized distributive relation must be valid for both union and intersection with respect to ordering.] We will see, however, that the generalized notation obtained from generalizing $\{E(C)\}$ leads to some new notions in stochastic processes.

Let us return to the experiment of Fig. 1, assuming that the generalized distributive relation is valid for this experiment. We see, in general, that the probability for $G_1 \cap G_2$ and $D_1 \cap D_2$ is nonzero. However, if we suppose that G_1 and G_2 never fire simultaneously and that D_1 and D_2 never detect simultaneously, then Eq. (6) is satisfied; thus we see that Eq. (6) is a requirement motivated by *a posteriori* knowledge.

From Eqs. (6) and (8) and the additive property of P, we see that

$$\begin{aligned} &P(G_1 \cup G_2 \to M \to D_1 \cup D_2) \\ &\quad = P(G_1 \to M \to D_1) + P(G_1 \to M \to D_2) \\ &\qquad + P(G_2 \to M \to D_1) + P(G_2 \to M \to D_2), \end{aligned} \tag{9}$$

from which we conclude that

$$\begin{aligned} &P(G_1 \cup G_2 \to M \to D_j) \\ &\quad = P(G_1 \to M \to D_j) + P(G_2 \to M \to D_j); \end{aligned} \tag{10}$$

thus we are provided with the definition

$$P(D_j) \triangleq P(C_0 \to C_1 \to D_j) = \sum_{i=1,2} P(G_i \to C_1 \to D_j) \tag{11}$$

for the *unconditional probability* to detect a particle at D_j. This definition may be generalized to an L-term measurement sequence, i.e., for the L-term measurement sequence $\{M_0 \to M_1 \to M_2 \to \cdots \to M_L\}$, the *unconditional probability* for a result $\hat{p}_{k_i}$ during M_i, $0 \le i \le L$, is given by

$$\begin{aligned} P(\hat{p}_{k_i}) &\triangleq P(C_0 \to C_1 \to \cdots \to \hat{p}_{k_i} \to C_{i+1} \to \cdots \to C_L) \\ &= P\Big(\bigcup_{k_0=1}^{\infty} \hat{p}_{k_0} \to \bigcup_{k_1=1}^{\infty} \hat{p}_{k_1} \to \cdots \to \hat{p}_{k_i} \\ &\qquad \to \bigcup_{k_{i+1}=1}^{\infty} \hat{p}_{k_{i+1}} \to \cdots \to \bigcup_{k_L=1}^{\infty} \hat{p}_{k_L}\Big). \end{aligned} \tag{12}$$

Thus, using the generalized distributive relation, the disjointness of the simple paths in $\{E(C)\}$, and the σ-additivity of P, we see that Eq. (12) may be written in the more familiar form

$$\begin{aligned} P(\hat{p}_{k_i}) &= \sum_{k_0=1}^{\infty} \sum_{k_1=1}^{\infty} \cdots \sum_{k_{i-1}=1}^{\infty} \sum_{k_{i+1}=1}^{\infty} \cdots \\ &\qquad \sum_{k_L=1}^{\infty} P(\hat{p}_{k_0} \to \hat{p}_{k_1} \to \cdots \to \hat{p}_{k_i} \to \cdots \to \hat{p}_{k_L}), \end{aligned} \tag{13}$$

that is, the unconditional probability for $\hat{p}_{k_i}$ is the sum of the probabilities of all simple paths containing $\hat{p}_{k_i}$.

With a suitable definition of *conditional probability*, Eq. (12) provides the general mathematical structure for a stochastic process. *Conditional probability* on $\{E(C)\}$ may be defined by analogy with the traditional definition. Conventionally, the probability for "E_i^l is observed if E_i^k is observed" is given by

$$P_C(E_i^l \mid E_i^k) \triangleq P(E_i^l \cap E_i^k)/P(E_i^k). \tag{14}$$

For the conventional event space, such a definition suffers from causal ambiguities; however, for the time-ordered event space such ambiguities disappear.

In addition to the simultaneous events of Eq. (14), we wish to consider the conditional probability for the time-separated events E_i^l and E_j^k, $i \neq j$. By analogy with the conventional definition (14), we define

$$\begin{aligned} P(E_j^l \mid E_i^k) &\triangleq P(C_0 \to C_1 \to \cdots \to E_j^l \to C_{j+1} \to \cdots \to C_L \mid C_0 \to C_1 \to \cdots \to E_i^k \to C_{i+1} \to \cdots \to C_L) \\ &= \frac{P((C_0 \to C_1 \to \cdots \to E_j^l \to C_{j+1} \to \cdots \to C_L) \cap (C_0 \to C_1 \to \cdots \to E_i^k \to C_{i+1} \to \cdots \to C_L))}{P(C_0 \to C_1 \to \cdots \to E_i^k \to C_{i+1} \to \cdots \to C_L)}; \end{aligned} \tag{15}$$

we see that this is well defined, independent of the magnitude of i with respect to j. Let us examine this definition for the case where $i < j$ and the case where $i = j$.

When $i < j$, Eq. (15) becomes

$$P(E_j^l \mid E_i^k) = \frac{\begin{aligned} &P((C_0 \to C_1 \to \cdots \to C_i \to \cdots \to E_j^l \to C_{j+1} \to \cdots \to C_L) \\ &\qquad \cap (C_0 \to C_1 \to \cdots \to E_i^k \to C_{i+1} \to \cdots \to C_j \to \cdots \to C_L)) \end{aligned}}{P(C_0 \to C_1 \to \cdots \to E_i^k \to C_{i+1} \to \cdots \to C_j \to \cdots \to C_L)}; \tag{16}$$

thus $P(E_j^l \mid E_i^k)$; $i < j$ has the obvious interpretation "the conditional probability for the event E_j^l to occur at time t_j if E_i^k is known to have occurred at an earlier time t_i."

Now $P(E_i^k \mid E_j^l)$ is also well defined by Eq. (15). Let us examine the nature of this conditional probability. Equation (15) yields

$$P(E_i^k \mid E_j^l) = \frac{P((C_0 \to C_1 \to \cdots \to E_i^k \to C_{i+1} \to \cdots \to C_j \to \cdots \to C_L) \cap (C_0 \to C_1 \to \cdots \to E_j^l \to C_{j+1} \to \cdots \to C_L))}{P(C_0 \to C_1 \to \cdots \to C_i \to \cdots \to E_j^l \to C_{j+1} \to \cdots \to C_L)}, \tag{17}$$

which, in view of the nature of the sequenced event space, can only be interpreted as "the conditional probability for the event E_i^k to occur at t_i if E_j^l is known to have occurred at a later time t_j."

In case $i = j$, we see from Eqs. (16) and (17) that our definition of Eq. (15) is the analog of the conventional definition given by Eq. (14).

It is our claim, and we discuss this more fully in the sections to follow, that the sequenced formalism clearly distinguishes and defines both "types" of conditional probabilities as given in Eqs. (16) and (17). We will demonstrate that the conditional probability of Eq. (17) can be the "inverse" or "time-reversed" form of the conditional probability of Eq. (16) only in case the system follows a deterministic path through the measurement sequence. We also will see that $P(E_i^k \mid E_j^l)$, $i < j$, is definable only because of the *a posteriori* nature of the data from a measurement sequence.

We will postpone this discussion until we have more fully developed the stochastic equations describing the measurement sequence.

THE RANDOM WALK

Now that we have developed the definitions for conditional probability and unconditional probability, we are able to consider the measurement sequence as a generalized random walk problem. We will, in this section, develop the random walk equation which determines the probability for the statement, "the simple event $\hat{p}_{k_i}$ is the outcome of M_i, regardless of the outcomes of the rest of the measurements in the sequence," in terms of the conditional probabilities of $\hat{p}_{k_i}$ with respect to the outcomes of other measurements in the sequence.

We accomplish this by beginning with the definition in Eq. (12) of the unconditional probability. From this we may write

$$P(\hat{p}_{kj}) = \sum_{ki} P(C_0 \to C_1 \to \cdots \to \hat{p}_{ki} \to C_{i+1} \to \cdots \to \hat{p}_{kj} \to C_{j+1} \to \cdots \to C_L), \quad j > i. \tag{18}$$

Since the conditional probability is defined for each member of $\{E(C)\}$, we may write, from Eq. (15),

$$P(\hat{p}_{kj} \mid \hat{p}_{ki}) = \frac{P((C_0 \to C_1 \to \cdots \to \hat{p}_{kj} \to C_{j+1} \to \cdots \to C_L) \cap (C_0 \to C_1 \to \cdots \to \hat{p}_{ki} \to C_{i+1} \to \cdots \to C_L))}{P(C_0 \to C_1 \to \cdots \to \hat{p}_{ki} \to C_{i+1} \to \cdots \to C_L)}. \tag{19}$$

Using the generalized distributive relation, we may reduce the numerator of Eq. (19) so that Eq. (19) becomes

$$P(\hat{p}_{kj} \mid \hat{p}_{ki}) = \frac{P(C_0 \to C_1 \to \cdots \to \hat{p}_{ki} \to C_{i+1} \to \cdots \to \hat{p}_{kj} \to C_{j+1} \to \cdots \to C_L)}{P(C_0 \to C_1 \to \cdots \to \hat{p}_{ki} \to C_{i+1} \to \cdots \to C_L)}. \tag{20}$$

Since the numerator of Eq. (20) is exactly the term inside the sum of Eq. (18), we may employ Eq. (20) to write Eq. (18) as

$$P(\hat{p}_{kj}) = \sum_{ki} P(\hat{p}_{kj} \mid \hat{p}_{ki}) P(C_0 \to C_1 \to \cdots \to \hat{p}_{ki} \to C_{i+1} \to \cdots \to C_L). \tag{21}$$

Before we "expose" this as the random walk equation, let us consider the unconditional probability for $\hat{p}_{k_i}$. From Eq. (12), we may write

$$P(\hat{p}_{ki}) = \sum_{kj} P(C_0 \to C_1 \to \cdots \to \hat{p}_{ki} \to C_{i+1} \to \cdots \to \hat{p}_{kj} \to C_{j+1} \to \cdots \to C_L), \quad j > i, \tag{22}$$

and, as we saw in the development of Eq. (20), we may write from Eq. (15)

$$P(\hat{p}_{k_i} \mid \hat{p}_{k_j}) = \frac{P(C_0 \to C_1 \to \cdots \to \hat{p}_{k_i} \to \cdots \to \hat{p}_{k_j} \to C_{j+1} \to \cdots \to C_L)}{P(C_0 \to C_1 \to \cdots \to \hat{p}_{k_j} \to C_{j+1} \to \cdots \to C_L)}, \quad j > i, \tag{23}$$

which allows us to write Eq. (22) as

$$P(\hat{p}_{k_i}) = \sum_{k_j} P(\hat{p}_{k_i} \mid \hat{p}_{k_j}) P(C_0 \to C_1 \to \cdots \to \hat{p}_{k_j} \to C_{i+1} \to \cdots \to C_L), \quad j > i. \tag{24}$$

In the simplified notation provided by the definition of unconditional probability, Eq. (21) may be written as

$$P(\hat{p}_{k_j}) = \sum_{k_i} P(\hat{p}_{k_j} \mid \hat{p}_{k_i}) P(\hat{p}_{k_i}), \quad j > i, \tag{25}$$

and Eq. (24) may be written

$$P(\hat{p}_{k_i}) = \sum_{k_j} P(\hat{p}_{k_i} \mid \hat{p}_{k_j}) P(\hat{p}_{k_j}), \quad j > i, \tag{26}$$

which we will name the *predictive random walk equation* and the *retrodictive random walk equation*, respectively. This is an obvious choice of terminology since Eq. (25) calculates probability distributions for events occurring at t_j in terms of the probability distributions for events occurring at an *earlier* time t_i and since Eq. (26) calculates probability distributions for events occurring at t_i in terms of the probability distributions for events occurring at a *later* time t_j.

We may go a step further in adapting our notation to the standard notation by defining the *predictive transition probability*

$$\begin{aligned} T_{k_jk_i} &\triangleq P(\hat{p}_{k_j} \mid \hat{p}_{k_i}) \\ &= P(\hat{p}_{k_j} \mid C_0 \to C_1 \to \cdots \to \hat{p}_{k_i} \\ &\qquad \to C_{i+1} \to \cdots \to C_j \to \cdots \to C_L) \end{aligned} \tag{27}$$

and the *retrodictive transition probability*

$$\begin{aligned} T'_{k_ik_j} &\triangleq P(\hat{p}_{k_i} \mid \hat{p}_{k_j}) \\ &= P(\hat{p}_{k_i} \mid C_0 \to C_1 \to \cdots \to C_i \\ &\qquad \to \cdots \to \hat{p}_{k_j} \to C_{j+1} \to \cdots \to C_L), \end{aligned} \tag{28}$$

so that the predictive random walk equation becomes

$$P(\hat{p}_{k_j}) = \sum_{k_i} T_{k_jk_i} P(\hat{p}_{k_i}) \tag{29}$$

and the retrodictive random walk equation becomes

$$P(\hat{p}_{k_i}) = \sum_{k_j} T'_{k_ik_j} P(\hat{p}_{k_j}). \tag{30}$$

We see from the preceding analysis that Eq. (29) is a generalized form of the conventional Markoff random walk equation. It is generalized in the sense that $T_{k_jk_i}$ is not Markoffian.

We also see that Eq. (30) is not at all conventional since it implies that if we know the probability set $\{P(\hat{p}_{k_j})\}$ at t_j and the set of retrodictive transition probabilities $\{T'_{k_jk_i}\}$, then we may calculate the probability set $\{P(\hat{p}_{k_i})\}$ even when $t_i < t_j$. Such a result is completely consistent with the *a posteriori* nature of data. We will discuss this property of data in the conclusion section of this paper.

PROPERTIES OF THE STOCHASTIC PROCESS

In this section we will examine the temporal behavior of the stochastic process in terms of prediction and retrodiction. This examination will clarify the relationship between the predictive dynamics and the retrodictive dynamics and will provide a foundation for our examination of the $\mathcal{H}$ representation of stochastic processes.

Each measurement pair M_i, M_j, $i < j$, in the measurement sequence $\{M_0 \to M_1 \to \cdots \to M_L\}$ defines a collection of predictive transition probabilities $\{T_{k_jk_i}\}$, a collection of retrodictive transition probabilities $\{T'_{k_ik_j}\}$, and a collection of *simultaneous conditional probabilities* $\{T_{k_ik_i'}\}$.

$\mathsf{T}(j, i)$ is the *predictive transition matrix* for the measurement pair M_i, M_j, $i < j$, means that $\mathsf{T}(j, i)$ is a matrix such that $T_{k_jk_i}$ is the k_jth-row and the k_ith-column element of $\mathsf{T}(j, i)$.

$\mathsf{T}'(i, j)$ is the *retrodictive transition matrix* for the measurement pair M_i, M_j, $i < j$, means that $\mathsf{T}'(i, j)$ is a matrix such that $T'_{k_ik_j}$ is the k_ith-row and the k_jth-column element of $\mathsf{T}'(i, j)$.

$\mathsf{T}(i, i)$ is the *simultaneous conditional probability matrix* for the measurement M_i means that $\mathsf{T}(i, i)$ is a matrix such that $T_{k_ik_i'}$ is the k_ith-row and the k_i'th-column element of $\mathsf{T}(i, i)$.

We see then that an L-term measurement sequence defines $\frac{1}{2}L(L+1)$ measurement pairs M_i, M_j, $i < j$, and thus defines $\frac{1}{2}L(L+1)$ retrodictive transition matrices, $\frac{1}{2}L(L+1)$ predictive transition matrices, and L simultaneous conditional probability matrices.

Let $\{\mathsf{T}(j, i)\}$ denote the collection of predictive transition matrices, $\{\mathsf{T}'(i, j)\}$ denote the collection of retrodictive transition matrices, and $\{\mathsf{T}(i, i)\}$ denote the collection of simultaneous conditional probability matrices. Let $\{\mathcal{T}(j, i)\}$ denote the collection of members of $\{\mathsf{T}(j, i)\}$, $\{\mathsf{T}'(i, j)\}$, and $\{\mathsf{T}(i, i)\}$.

We will now investigate the conditions, if any, for the collections $\{\mathsf{T}(j,i)\}$, $\{\mathsf{T}'(i,j)\}$, and $\{\mathcal{T}(j,i)\}$ to form either groups or semigroups with respect to matrix multiplication.

First, we note that Eq. (6) requires that the collection $\{\mathsf{T}(i,i)\}$ be the collection of unit matrices $\{\mathsf{I}_i\}$. In general, each member of $\{\mathsf{I}_i\}$ is of a different dimension, depending on the spectrum of M_i. In this investigation, we will assume that each spectrum is countably infinite, and thus each member of $\{\mathsf{I}_i\}$ will be of the same dimension.

It is not difficult to see that matrix multiplication between certain members of $\{\mathsf{T}(j,i)\}$ produces a transition matrix in $\{\mathsf{T}(j,i)\}$. To show this, we simply use Eq. (29) to write the following equation set:

$$\begin{aligned}
P(\hat{p}_{k_1}) &= \sum_{k_0} T_{k_1k_0} P(\hat{p}_{k_0}),\\
P(\hat{p}_{k_2}) &= \sum_{k_1} T_{k_2k_1} P(\hat{p}_{k_1}) = \sum_{k_0} T_{k_2k_0} P(\hat{p}_{k_0})\\
&\cdot\\
&\cdot\\
&\cdot\\
P(\hat{p}_{k_L}) &= \sum_{k_{L-1}} T_{k_Lk_{L-1}} P(\hat{p}_{k_{L-1}}) = \sum_{k_{L-2}} T_{k_Lk_{L-2}} P(\hat{p}_{k_{L-2}})\\
&= \cdots = \sum_{k_1} T_{k_Lk_1} P(\hat{p}_{k_1}) = \sum_{k_0} T_{k_Lk_0} P(\hat{p}_{k_0}).
\end{aligned} \tag{31}$$

Substituting the first equation of the set (31) into the second equation in the set, we obtain

$$P(\hat{p}_{k_2}) = \sum_{k_0} P(\hat{p}_{k_0}) \sum_{k_1} T_{k_2k_1} T_{k_1k_0} = \sum_{k_0} P(\hat{p}_{k_0}) T_{k_2k_0}, \tag{32}$$

which implies by comparison the Chapman–Kolmogorov relation[11]

$$T_{k_2k_0} = \sum_{k_1} T_{k_2k_1} T_{k_1k_0}. \tag{33}$$

This procedure may be repeated for the entire set (31) to obtain

$$T_{k_Lk_0} = \sum_{k_{L-1}} \sum_{k_{L-2}} \cdots \sum_{k_1} T_{k_Lk_{L-1}} T_{k_{L-1}k_{L-2}} \cdots T_{k_1k_0}. \tag{34}$$

Since $T_{k_Lk_0}$ is the k_Lth row and k_0th column of $\mathsf{T}(L,0)$, we see that Eq. (34) provides a *multiplication theorem* for transition matrices,

$$\mathsf{T}(L,0) = \mathsf{T}(L,L-1)\mathsf{T}(L-1,L-2)\cdots\mathsf{T}(1,0). \tag{35}$$

From the retrodictive equation (30), we may write an equation set similar to the equation set (31) and derive the multiplication theorem for the retrodictive transition matrices

$$\mathsf{T}'(0,L) = \mathsf{T}'(0,1)\mathsf{T}'(1,2)\cdots\mathsf{T}'(L-1,L). \tag{36}$$

In addition, Eqs. (29) and (30) can be combined for various integers i and j so that multiplication is defined between members of $\{\mathsf{T}(j,i)\}$ and $\{\mathsf{T}'(i,j)\}$. For example, consider the integers q, s, and t such that $0 \le q < s < t \le L$. Equations (29) and (30) then define the products

$$\begin{aligned}
\mathsf{T}'(q,s)\mathsf{T}'(s,t) &= \mathsf{T}'(q,t),\\
\mathsf{T}'(q,t)\mathsf{T}(t,s) &= \mathsf{T}'(q,s),\\
\mathsf{T}(s,q)\mathsf{T}'(q,t) &= \mathsf{T}'(s,t),\\
\mathsf{T}'(s,t)\mathsf{T}(t,q) &= \mathsf{T}(s,q),\\
\mathsf{T}(t,q)\mathsf{T}'(q,s) &= \mathsf{T}(t,s).
\end{aligned} \tag{37}$$

However, we also obtain from this process

$$\begin{aligned}
P(\hat{p}_{k_s}) &= \sum_{k_s'} P(\hat{p}_{k_s'}) \sum_{k_t'} T_{k_sk_t'} T_{k_t'k_s'}\\
&= \sum_{k_s'} M_{k_sk_s'} P(\hat{p}_{k_s'}),\\
P(\hat{p}_{k_t}) &= \sum_{k_t'} P(\hat{p}_{k_t'}) \sum_{k_s'} T_{k_tk_s'} T'_{k_s'k_t'}\\
&= \sum_{k_t'} M_{k_tk_t'} P(\hat{p}_{k_t'}).
\end{aligned} \tag{38}$$

Equations (38) define the matrix products

$$\begin{aligned}
\mathsf{M}(s,s) &= \mathsf{T}'(s,t)\mathsf{T}(t,s),\\
\mathsf{M}(t,t) &= \mathsf{T}(t,s)\mathsf{T}'(s,t).
\end{aligned} \tag{39}$$

The immediate inclination is to identify the collection $\{\mathsf{M}(i,i)\}_{i=0,L}$ as the collection $\{\mathsf{T}(i,i)\}$ of simultaneous conditional probability matrices. However, such an identification would require that

$$\begin{aligned}
\mathsf{M}(s,s) &= \mathsf{I}_s = \mathsf{T}'(s,t)\mathsf{T}(t,s),\\
\mathsf{M}(t,t) &= \mathsf{I}_t = \mathsf{T}(t,s)\mathsf{T}'(s,t),
\end{aligned} \tag{40}$$

and, if the dimension of I_s is the dimension of I_t, then Eqs. (40) imply that

$$\mathsf{T}'(s,t) = [\mathsf{T}(t,s)]^{-1}. \tag{41}$$

Wu[12] has shown, however, that since each member of $\mathsf{T}(t,s)$ is positive, then its inverse transition matrix $[\mathsf{T}(t,s)]^{-1}$ must have at least one *negative* member, unless, of course, $\mathsf{T}(t,s)$ has only one nonzero member. Since $\mathsf{T}'(s,t)$ is itself a transition matrix, Eq. (41) and thus Eqs. (40) can be satisfied only in case $\mathsf{T}(t,s)$ has only one nonzero member. [In this case $\mathsf{T}(t,s)$ would describe a *deterministic process*.] Thus we see that, in general, $\mathsf{M}(s,s)$ cannot be identified as the matrix $\mathsf{T}(s,s)$ of simultaneous conditional probabilities.

With multiplication defined in $\{\mathsf{T}(j,i)\}$ and $\{\mathsf{T}'(i,j)\}$, we may proceed to examine these collections as groups or semigroups.

Since $\{\mathsf{T}(j, i)\}$ can form a group only in case each member $\mathsf{T}(r, q) \in \{\mathsf{T}(j, i)\}$ has an inverse $[\mathsf{T}(r, q)]^{-1} \in \{\mathsf{T}(j, i)\}$, we see from the preceding arguments that neither $\{\mathsf{T}(j, i)\}$ nor $\{\mathsf{T}'(i, j)\}$ can form a group.

We also see that the collection $\{\mathcal{T}(j, i)\}$ cannot form a semigroup since the product M given by Eqs. (40) is not a member of $\{\mathcal{T}(j, i)\}$ unless, for each positive integer i such that $i \leq L$,

$$\mathsf{M}(i, i) = \mathsf{T}(i, i) = \mathsf{I}, \qquad (42)$$

which, as we argued, is possible only for a deterministic system.

Let us now examine the conditions for $\{\mathsf{T}(j, i)\}$ and $\{\mathsf{T}'(i, j)\}$ to form semigroups. Suppose $\{\mathsf{T}(j, i)\}$ forms a semigroup. In this case, closed associative multiplication must be defined between each pair in $\{\mathsf{T}(j, i)\}$. We see from Eq. (35) that left multiplication of $\mathsf{T}(r, s)$ by $\mathsf{T}(q, r)$ yields $\mathsf{T}(q, s)$; thus the product $\mathsf{T}(q, r)\mathsf{T}(r, s)$ is a member of $\{\mathsf{T}(j, i)\}$ and the multiplication is closed. Since this multiplication is matrix multiplication, it is also associative.

We see, however, that multiplication of the two matrices $\mathsf{T}(p, q)\mathsf{T}(r, s)$ produces a transition matrix in $\{\mathsf{T}(j, i)\}$ only in case $q = r$ or $p = s$. This fact motivates us to define the following notion: Two transition matrices $\mathsf{T}(p, q)$ and $\mathsf{T}(r, s)$ are *adjacent* means either $p = s$ or $q = r$. It is clear, then, that if each pair of matrices in $\{\mathsf{T}(j, i)\}$ can be made adjacent, then $\{\mathsf{T}(j, i)\}$ will form a semigroup.

If each member of $\{\mathsf{T}(j, i)\}$ has the property that

$$\mathsf{T}(l, k) = \mathsf{T}(x, y) \quad \text{in case} \quad |l - k| = |x - y|, \qquad (43)$$

then any two matrices $\mathsf{T}(p, q) \in \{\mathsf{T}(j, i)\}$ and $\mathsf{T}(r, s) \in \{\mathsf{T}(j, i)\}$ can be made adjacent simply by relabeling $\mathsf{T}(r, s)$ as $\mathsf{T}(q, s')$, where $\mathsf{T}(q, s') \in \{\mathsf{T}(j, i)\}$ and $|q - s'| = |r - s|$ so that

$$\mathsf{T}(p, q)\mathsf{T}(r, s) = \mathsf{T}(p, q)\mathsf{T}(q, s') = \mathsf{T}(p, s'). \qquad (44)$$

Thus the collection $\{\mathsf{T}(j, i)\}$ can form a semigroup in case the matrices in the collection are all conformable and Eq. (43) is satisfied for each matrix in the collection. The same argument applies for the collection $\{\mathsf{T}'(i, j)\}$. If, in addition, we include the collection $\{\mathsf{T}(i, i)\}$ in $\{\mathsf{T}(j, i)\}$, we see that $\{\mathsf{T}(j, i)\}$ can form a *monoid* semigroup. The same argument applies for $\{\mathsf{T}'(i, j)\}$.

We see then that the predictive collection $\{\mathsf{T}(j, i)\}$ and the retrodictive collection $\{\mathsf{T}'(i, j)\}$ can each form a group only in case each member in $\{\mathsf{T}(j, i)\}$ and each member in $\{\mathsf{T}'(i, j)\}$ describes a deterministic system. However, each of $\{\mathsf{T}(j, i)\}$ and $\{\mathsf{T}'(i, j)\}$ can form a semigroup in case each member of $\{\mathsf{T}(j, i)\}$ and each member of $\{\mathsf{T}'(i, j)\}$ satisfies Eq. (43). Physically, Eq. (43) restricts the transition probabilities to be a function only of the number of measurements between M_i and M_j; this requires that each $\mathsf{T}(j, i)$ be a function only of the relative time difference between M_i and M_j. Thus Eq. (43) is analogous to the *quantum* requirement that $\mathsf{U}(t_2, t_1)$ be a function only of $|t_2 - t_1|$ if U is to be a member of the unitary group.

We also demonstrated that a retrodictive transition matrix is not the inverse of the corresponding *predictive* transition matrix. However, the equations resulting from the sequenced event space clearly define and distinguish between *retrodiction and prediction* and show that one may always predict or retrodict the stochastic process.

PROBABILITY FUNCTIONS IN l^2

In this section we will demonstrate that probabilities for simple paths in $\{E(C)\}$ may be represented as products of complex functions in l^2, the space of square summable sequences. From the isomorphism of l^2 to a separable Hilbert space $\mathcal{H}$, we deduce the existence of a continuous linear operator in $\mathcal{H}$ which corresponds to the transition probability of Eq. (27). Hilbert space representations for probabilities of simple paths in $\{E(C)\}$ are shown to be possible because of the positive-definite, unit norm and σ-additive properties of P.

Since $P(\hat{p}_{k_j})$ is positive definite, there exists a complex function α_{k_j} such that for each $\hat{p}_{k_j}$

$$P(\hat{p}_{k_j}) = \alpha_{k_j}^* \alpha_{k_j} \qquad (45)$$

and the phase of α_{k_j} is *arbitrary*.

Using the unit norm property and the generalized distributive relation, we see that

$$\sum_{k_j} P(\hat{p}_{k_j}) = 1 = \sum_{k_j=1}^{\infty} \alpha_{k_j}^* \alpha_{k_j}. \qquad (46)$$

Thus the sequence $\{\alpha_{k_j}\}_{k_j=1,2,\cdots}$ is square summable and is a member of l^2. If we now consider the vector $|\alpha(j)\rangle$ defined by

$$|\alpha(j)\rangle = \sum_{k_j=1}^{\infty} C_{k_j} |k_j\rangle, \qquad (47)$$

where $\{|k_j\rangle\}_{k_j=1,2,\cdots}$ is an orthonormal basis for a separable Hilbert space $\mathcal{H}$, then $|\alpha(j)\rangle \in \mathcal{H}$ only in case $\{C_{k_j}\}$ is a square-summable sequence.[13] Thus, if we define C_{k_j} as

$$C_{k_j} \triangleq \alpha_{k_j} (\langle \alpha(j) \mid \alpha(j) \rangle)^{\frac{1}{2}} \qquad (48)$$

we see that $\{C_{k_j}\}$ is square summable; therefore, $|\alpha(j)\rangle$ defined by

$$|\alpha(j)\rangle = (\langle \alpha(j) \mid \alpha(j) \rangle)^{\frac{1}{2}} \sum_{k_j} \alpha_{k_j} |k_j\rangle \qquad (49)$$

is a member of $\mathcal{H}$. Thus we see that for each square-summable sequence $\{\alpha_{k_j}\}$ there exists a vector $|\alpha(j)\rangle \in \mathcal{H}$ such that each member of $\{\alpha_{k_j}\}$ has a representation in $\mathcal{H}$ given by

$$\alpha_{k_j} = \langle k_j \mid \alpha(j)\rangle / (\langle \alpha(j) \mid \alpha(j)\rangle)^{\frac{1}{2}}. \tag{50}$$

Thus we have established an $\mathcal{H}$ representation for each member in the collection $\{\alpha_{k_j}\}$ and therefore for $\{P(\hat{p}_{k_j})\}$.

Now let us examine the transition probability $T_{k_jk_i}$. Since $T_{k_jk_i}$ is positive, there exists a complex function for each k_j and k_i such that

$$T_{k_jk_i} = K^*_{k_jk_i} K_{k_jk_i}, \tag{51}$$

and, since $\{T_{k_jk_i}\}$ is singly stochastic, the sequence $\{K_{k_jk_i}\}_{k_j=1,2,\cdots}$ is square summable for each k_i. Therefore, there exists a countable orthonormal basis $\{|k_j\rangle\}$ and a member $|Q_{k_i}\rangle \in \mathcal{H}$ such that for each k_i

$$K_{k_jk_i} = \langle k_j \mid Q_{k_i}\rangle / (\langle Q_{k_i} \mid Q_{k_i}\rangle)^{\frac{1}{2}}. \tag{52}$$

We see from (51) and (52) that, for a given basis $\{|k_j\rangle\}$, each member of the countable collection $\{|Q_{k_i}\rangle\}$ is determined only to within a phase.

$K_{k_jk_i}$ may be written in a different form since we may associate with the collection $\{|Q_{k_i}\rangle\}$ an orthonormal basis $\{|k_i\rangle\}$ in $\mathcal{H}$ by an operator $\mathsf{K}(j, i)$ mapping $\mathcal{H}_i$ onto $\mathcal{H}_i$, i.e., for each k_i

$$|Q_{k_i}\rangle = \mathsf{K}(j, i)\,|k_i\rangle; \tag{53}$$

thus we may write (52) as

$$K_{k_jk_i} = \langle k_j|\,\mathsf{K}(j, i)\,|k_i\rangle / (\langle k_i|\,\mathsf{K}^+\mathsf{K}\,|k_i\rangle)^{\frac{1}{2}}. \tag{54}$$

With these representations for $T_{k_jk_i}$ and P, we may write the $\mathcal{H}$ representation for the predictive random-walk equation as

$$\frac{\langle k_j \mid \alpha(j)\rangle \langle \alpha(j) \mid k_j\rangle}{\langle \alpha(j) \mid \alpha(j)\rangle} = \sum_{k_i} \frac{\langle k_j|\,\mathsf{K}(j, i)\,|k_i\rangle \langle k_i|\,\mathsf{K}^+(j, i)\,|k_j\rangle \langle k_i \mid \alpha(i)\rangle \langle \alpha(i) \mid k_i\rangle}{\langle k_i|\,\mathsf{K}^+(j, i)\mathsf{K}(j, i)\,|k_i\rangle \langle \alpha(i) \mid \alpha(i)\rangle}; \tag{55}$$

clearly, from this development, an $\mathcal{H}$ representation can be generated for the retrodictive equation (30). This equation would be given by

$$\frac{\langle k_i \mid \alpha(i)\rangle \langle \alpha(i) \mid k_i\rangle}{\langle \alpha(i) \mid \alpha(i)\rangle} = \sum_{k_j} \frac{\langle k_i|\,\mathsf{K}'(i, j)\,|k_j\rangle \langle k_j|\,\mathsf{K}'^+(i, j)\,|k_i\rangle \langle k_j \mid \alpha(j)\rangle \langle \alpha(j) \mid k_j\rangle}{\langle k_j|\,\mathsf{K}'^+(i, j)\mathsf{K}(i, j)\,|k_j\rangle \langle \alpha(j) \mid \alpha(j)\rangle}, \tag{56}$$

where the operator $\mathsf{K}'(i, j)$ is constructed so that

$$T'_{k_ik_j} = \frac{\langle k_i \mid Q_{k_j}\rangle \langle Q_{k_j} \mid k_i\rangle}{\langle Q_{k_j} \mid Q_{k_j}\rangle} = \frac{\langle k_i|\,\mathsf{K}'(i, j)\,|k_j\rangle \langle k_j|\,\mathsf{K}'^+(i, j)\,|k_i\rangle}{\langle k_j|\,\mathsf{K}'^+(i, j)\mathsf{K}'(i, j)\,|k_j\rangle}, \tag{57}$$

the retrodictive transition probability, is reproduced. Thus we have established $\mathcal{H}$ representations for both the retrodictive and predictive random-walk equations.

RANDOM WALK AND TIME EVOLUTION IN $\mathcal{H}$

Now that we have established an $\mathcal{H}$ representation for the random walk equation, we may employ a phase choice theorem established in a previous paper[1] to establish another $\mathcal{H}$ representation for the random walk equation which will allow us to compare the dynamics of stochastic and quantum theory.

This theorem demonstrates the existence of choices for the phases of the sequence of products

$$\{K_{k_jk_i}\alpha_{k_i}\}_{k_i=1,2,\cdots}$$

such that Eq. (55) factors to yield (see Appendix A for this theorem and its connection here)

$$\frac{\langle k_j \mid \alpha(j)\rangle}{(\langle \alpha(j) \mid \alpha(j)\rangle)^{\frac{1}{2}}} = \sum_{k_i} \frac{\langle k_j|\,\mathsf{K}(j, i)\,|k_i\rangle \langle k_i \mid \alpha(i)\rangle}{\langle k_i|\,\mathsf{K}^+\mathsf{K}\,|k_i\rangle^{\frac{1}{2}} (\langle \alpha(i) \mid \alpha(i)\rangle)^{\frac{1}{2}}}. \tag{58}$$

Equation (58) provides a very simple representation in $\mathcal{H}$ for the dynamics of classical probability theory; i.e., Eq. (58) may be written

$$|\alpha'(j)\rangle = \sum_{k_i} \mathsf{K}(j, i)\,\frac{|k_i\rangle\langle k_i|}{a^j_{k_i}}\,|\alpha'(i)\rangle, \tag{59}$$

where

$$a^j_{k_i} \triangleq (\langle k_i|\,\mathsf{K}^+(j, i)\mathsf{K}(j, i)\,|k_i\rangle)^{\frac{1}{2}},$$
$$|\alpha'(j)\rangle = |\alpha(j)\rangle / (\langle \alpha(j) \mid \alpha(j)\rangle)^{\frac{1}{2}}. \tag{60}$$

We can further simplify by defining the operator $\mathsf{S}(j, i)$ as

$$\mathsf{S}(j, i) \triangleq \sum_{k_i} \mathsf{K}(j, i)\,\frac{|k_i\rangle\langle k_i|}{a^j_{k_i}} \tag{61}$$

so that Eq. (59) becomes

$$|\alpha'(j)\rangle = \mathsf{S}(j, i)\,|\alpha'(i)\rangle, \tag{62}$$

and we see that in a similar manner we may construct this representation for the retrodictive case

$$|\alpha'(i)\rangle = \mathsf{S}'(i, j)\,|\alpha'(j)\rangle. \tag{63}$$

Equation (62) is similar in form to the evolution equation of quantum theory, although, as we will see in the discussion to follow, the stochastic operator $\mathsf{S}(j,i)$ differs strikingly from the quantum evolution operator $\mathsf{U}(t_j, t_i)$. In addition to Eq. (62), we have Eq. (63), the retrodictive evolution equation. No such formalism appears in conventional quantum theory.

Thus we see that, for the measurement sequence $\{M_0 \to M_1 \to \cdots \to M_L\}$, there exists a collection $\{\mathsf{S}(j,i)\}$ of $\frac{1}{2}L(L+1)$ predictive stochastic operators and a collection $\{\mathsf{S}'(i,j)\}$ of $\frac{1}{2}L(L+1)$ retrodictive stochastic operators. Let us now examine the properties of $\{\mathsf{S}'(i,j)\}$ and $\{\mathsf{S}(j,i)\}$.

First we see from Eq. (61) that

$$\langle k_j|\, \mathsf{S}(j,i)\, |k_i\rangle = \frac{\langle k_j|\, \mathsf{K}(j,i)\, |k_i\rangle}{(\langle k_i|\, \mathsf{K}^+(j,i)\mathsf{K}(j,i)\, |k_i\rangle)^{\frac{1}{2}}}. \quad (64)$$

If we multiply Eq. (64) by its complex conjugate and sum over all $|k_i\rangle$, then we obtain the isometric property for S,

$$\mathsf{S}^+(j,i)\mathsf{S}(j,i) = \mathsf{I}. \quad (65)$$

However, multiplying Eq. (64) from the right by its complex conjugate, we see that S is unitary ($\mathsf{S}^+\mathsf{S} = \mathsf{S}\mathsf{S}^+ = \mathsf{I}$) only in case K is unitary. Thus we see that S is automatically isometric by construction, but can be unitary only if K is unitary. This relationship of S to K, as we shall see, has important physical implications. In order to see these implications, we must explore the properties of the collections $\{\mathsf{S}(j,i)\}$ and $\{\mathsf{S}'(i,j)\}$.

The approach to the examination of $\{\mathsf{S}(j,i)\}$ and $\{\mathsf{S}'(i,j)\}$ will be almost identical to our earlier approach when we examined the collections $\{\mathsf{T}(j,i)\}$ and $\{\mathsf{T}'(i,j)\}$, and, not surprisingly, the results will be almost identical. The complex analogs to Eqs. (31) are by the phase choice theorem

$$\begin{aligned}
\alpha_{k_1} &= \sum_{k_0} \langle k_1|\, \mathsf{S}(1,0)\, |k_0\rangle\, \alpha_{k_0}, \\
\alpha_{k_2} &= \sum_{k_1} \langle k_2|\, \mathsf{S}(2,1)\, |k_1\rangle\, \alpha_{k_1} = \sum_{k_0} \langle k_2|\, \mathsf{S}(2,0)\, |k_0\rangle\, \alpha_{k_0} \\
&\ \ \vdots \\
\alpha_{k_L} &= \sum_{k_{L-1}} \langle k_L|\, \mathsf{S}(L, L-1)\, |k_{L-1}\rangle\, \alpha_{k_{L-1}} \\
&= \sum_{k_{L-2}} \langle k_L|\, \mathsf{S}(L, L-2)\, |k_{L-2}\rangle\, \alpha_{k_{L-2}} \\
&= \cdots = \sum_{k_0} \langle k_L|\, \mathsf{S}(L,0)\, |k_0\rangle\, \alpha_{k_0}.
\end{aligned} \quad (66)$$

Substituting the first of Eqs. (66) into the second equation in the set and comparing, we obtain

$$\sum_{k_0} \alpha_{k_0} \sum_{k_1} \langle k_2|\, \mathsf{S}(2,1)\, |k_1\rangle \langle k_1|\, \mathsf{S}(1,0)\, |k_0\rangle = \sum_{k_0} \langle k_2|\, \mathsf{S}(2,0)\, |k_0\rangle\, \alpha_{k_0}, \quad (67)$$

so that we obtain the $\mathcal{K}$ representation of Eq. (33),

$$\langle k_2|\, \mathsf{S}(2,0)\, |k_0\rangle = \sum_{k_1} \langle k_2|\, \mathsf{S}(2,1)\, |k_1\rangle \langle k_1|\, \mathsf{S}(1,0)\, |k_0\rangle, \quad (68)$$

which implies the multiplication theorem

$$\mathsf{S}(2,0) = \mathsf{S}(2,1)\mathsf{S}(1,0). \quad (69)$$

This procedure may be repeated for the entire set (66) to obtain the general multiplication theorem for the stochastic operator set $\{\mathsf{S}(j;i)\}$, i.e.,

$$\mathsf{S}(L,0) = \mathsf{S}(L, L-1) \times \mathsf{S}(L-1, L-2) \cdots \mathsf{S}(2,1)\mathsf{S}(1,0) \quad (70)$$

and similarly for the retrodictive set:

$$\mathsf{S}'(0,L) = \mathsf{S}'(0,1)\mathsf{S}'(1,2) \cdots \mathsf{S}'(L-2, L-1) \times \mathsf{S}'(L-1, L). \quad (71)$$

In addition, we have the set $\{\mathsf{S}(i,i)\}$, which by Eq. (64) and the definition of $\{\mathsf{T}(i,i)\}$ is given by

$$\{\mathsf{S}(i,i)\} = \{\mathsf{I}_i\}. \quad (72)$$

Suppose $\{\mathsf{S}(j,i)\}$ forms a subset of a group. It must be true then that each member of $\{\mathsf{S}(j,i)\}$ has an inverse. We show in Appendix B that, in case $\mathsf{S}^{-1}(j,i)$ exists, then

$$P(\hat{p}_{k_i}) = \delta_{k_i k_i'}, \quad (73)$$

that is, the state of the system at M_i must be precisely determined. Consider the predictive random-walk equation in case $\mathsf{S}^{-1}(j,i)$ exists for each measurement pair in the sequence:

$$P(\hat{p}_{k_l}) = \sum_{k_{L-1}} \sum_{k_{L-2}} \cdots \sum_{k_0} T_{k_L k_{L-1}} T_{k_{L-1} k_{L-2}} \cdots T_{k_1 k_0} P(\hat{p}_{k_0}), \quad (74)$$

which by (73) must reduce to

$$P(\hat{p}_{k_l}) = T_{k_L k_{L-1}} \delta_{k_{L-1} k_{L-1}'} \delta_{k_{L-2} k_{L-2}'} \cdots \delta_{k_0 k_0'}. \quad (75)$$

Equation (75) is the random walk equation for a system which is *deterministic* from M_0 through M_{L-1}. We see from this that, in case $\{\mathsf{S}(j,i)\}$ is a subset of a group, then the members of $\{\mathsf{S}(j,i)\}$ cannot describe the most general class of stochastic processes. The same argument applies for $\{\mathsf{S}'(i,j)\}$.

Let $\{\mathcal{S}(j,i)\}$ denote the collection of members of $\{\mathsf{S}(j,i)\}$, $\{\mathsf{S}'(i,j)\}$, and $\{\mathsf{S}(i,i)\}$. As we did for the transition matrices, we may define multiplication between members of $\{\mathsf{S}(j,i)\}$ and $\{\mathsf{S}'(i,j)\}$ and show that, for t and s each a positive integer such that $t > s$,

$$\begin{aligned}
\alpha_{k_t} &= \sum_{k_t'} \langle k_t|\, \mathsf{S}(t,s)\mathsf{S}'(s,t)\, |k_t'\rangle\, \alpha_{k_t'}, \\
\alpha_{k_s} &= \sum_{k_s'} \langle k_s|\, \mathsf{S}'(s,t)\mathsf{S}(t,s)\, |k_s'\rangle\, \alpha_{k_s'}.
\end{aligned} \quad (76)$$

Equations (76) are satisfied in case

$$\mathsf{S}(t, s)\mathsf{S}'(s, t) = \mathsf{S}'(t, t) = \mathsf{I},$$
$$\mathsf{S}'(s, t)\mathsf{S}(t, s) = \mathsf{S}(s, s) = \mathsf{I}, \tag{77}$$

but can be satisfied, as could Eqs. (38), without the conditions imposed by Eqs. (77). In fact, if Eqs. (77) are required of each $\mathsf{S}(j, i)$ and each $\mathsf{S}'(i, j)$, then the system described by the collection $\{\mathcal{S}(j, i)\}$ would, by Eq. (75), be completely deterministic. In addition, we see that if $\{\mathcal{S}(j, i)\}$ is to form a semigroup, then Eqs. (77) must be satisfied if multiplication between $\mathsf{S}(j, i)$ and $\mathsf{S}'(j, i)$ is to be closed in $\{\mathcal{S}(j, i)\}$. Therefore, if $\{\mathcal{S}(j, i)\}$ forms a semigroup, then it must form a group, and this group must be a *unitary group* since each $\mathcal{S} \in \{\mathcal{S}(j, i)\}$ is isometric and has an inverse.

Now suppose that $\{\mathsf{S}(j, i)\}$ forms a semigroup. As with $\{\mathsf{T}(j, i)\}$, we must require that

$$\mathsf{S}(l, k) = \mathsf{S}(x, y), \quad |l - k| = |x - y|, \tag{78}$$

that is, $\{\mathsf{S}(j, i)\}$ can form a semigroup only if each $\mathsf{S} \in \{\mathsf{S}(j, i)\}$ is a function of the relative time. The same argument applies for $\{\mathsf{S}'(i, j)\}$.

We are now in a position to fully appreciate the difference between stochastic dynamics and quantum dynamics. First we note that the stochastic evolution operator S is, in general, isometric while the quantum evolution operator U is always unitary.

We see that in case the collection of stochastic operators $\{\mathcal{S}(j, i)\}$ for the measurement sequence $\{M_0 \to M_1 \to \cdots \to M_L\}$ forms a unitary group, then a system must follow a *deterministic* path through the measurement sequence. We also see from Appendix B that in case each member of the collection $\{\mathsf{S}(j, i)\}$ has an inverse, then $\{\mathsf{S}(j, i)\}$ is a unitary collection and Eq. (75) implies that each measurement in the sequence, except the last, yields a unique result.

Since the quantum evolution operator U always has an inverse, we see that the quantum evolution equation, when subjected to the phase choice of Appendix A, can only describe evolution corresponding to Eq. (75). In case the quantum evolution operators form a unitary group, then unitary evolution in $\mathcal{H}$ can only describe a *deterministic stochastic process* when the phase choice is imposed. Thus we see that quantum dynamics, i.e., unitary evolution in $\mathcal{H}$, can never reproduce the random walk structure of stochastic processes.

QUANTUM AND STOCHASTIC DYNAMICS IN A SINGLE $\mathcal{H}$ REPRESENTATION

From the preceding section we see that quantum dynamics and stochastic dynamics in $\mathcal{H}$ are identical only in case the quantum evolution equation is subject to the phase choice of Appendix A and the stochastic operator S is unitary. However, if the quantum evolution equation is subject to the phase choice, then the peculiar probability structure produced by the "square" of this equation disappears; on the other hand, if the stochastic operator S is unitary, then the more general singly stochastic structure of the transition matrices of stochastic processes is restricted to the doubly stochastic structure of quantum theory. Furthermore, if the phase choice is imposed on unitary evolution in $\mathcal{H}$, then the ensuing dynamical model in $\mathcal{H}$ can reproduce only a special case, given by Eq. (75), of the random walk equation (29).

In view of this, it is interesting to note that Nelson[14,15] has derived the time-dependent Schrödinger equation from the diffusion equation. However, one may readily see from Chandrasekhar's[16] derivation of the diffusion equation that the diffusion format follows from the random walk equation (29) only in case $\mathsf{T}(j, i)$ is *doubly stochastic*.

Such a result emphasizes the peculiarity of the doubly stochastic "transition" matrix of quantum theory. The quantum "transition" matrix is clearly doubly stochastic since its elements are given by

$$T_{k_j k_i} \triangleq |\langle k_j|\, \mathsf{U}(t_j, t_i)\, |k_i\rangle|^2, \tag{79}$$

and we see from this equation that since U is unitary,

$$\sum_{k_j} T_{k_j k_i} = \sum_{k_i} T_{k_j k_i} = 1. \tag{80}$$

However, the stochastic representation with elements

$$T_{k_j k_i} = |\langle k_j|\, \mathsf{S}(j, i)\, |k_i\rangle|^2 \tag{81}$$

is in general not doubly stochastic since in general S is only isometric and not unitary.

The above properties of the evolution equations and the transition matrices of quantum and stochastic dynamics provide the motivation for a more general mathematical structure in $\mathcal{H}$ which will include both stochastic and quantum dynamics as a special case. To do this, we simply hypothesize that each "state" of a physical system has a representation by a member of a separable Hilbert space $\mathcal{H}$ and that the dynamical evolution of the system is described by

$$|\alpha(t)\rangle = \mathsf{S}(t, t_0)\, |\alpha(t_0)\rangle, \tag{82}$$

where S is, in general, isometric. The quantum dynamical description is given by a unitary S, and the stochastic dynamical description is given by applying the phase choice theorem to Eq. (82). In this way, we encompass both the peculiar probability structure provided by quantum theory and the singly stochastic transition matrix of classical stochastic theory.

CONCLUSION

We have discussed in this paper a novel formulation for the σ-algebra of stochastic chains and have seen how the sequenced event space leads to the notions of both prediction and retrodiction in stochastic theory. We have shown also that the equations for stochastic dynamics have a representation in a separable Hilbert space $\mathcal{H}$ which, in general, is distinct from the conventional quantum representation in $\mathcal{H}$. The stochastic picture in $\mathcal{H}$ suggests a more general evolution picture in $\mathcal{H}$ which includes quantum evolution and stochastic evolution as special cases.

That retrodiction in stochastic theory is possible is not surprising and, in fact, is necessary when one considers the definitions upon which stochastic theory is built. For example, consider the measuring sequence $\{M_0 \to M_1 \to \cdots \to M_L\}$. Suppose we let N systems pass this sequence one at a time, so that a moving picture camera may record the configurations assigned to a system as it passes through the sequence. Let the ith frame on the film record the result of M_i. Then the passage of a single system through the L-term measurement sequence will be recorded on an L-frame strip of film, each frame containing the result of one measurement. Suppose we record each system's passage through the sequence until we obtain N L-frame strips of motion picture film. Suppose we mark the first frame of each strip to identify the direction of time passage for each strip. We may now place the N strips into a box and shuffle them. If the configuration of the environment is fixed for the N systems, then we may operationally define the unconditional probability, for some $\hat{p}_{k_i}$ during M_i, as the number of strips $n(\hat{p}_{k_i})$ which have the configuration $\hat{p}_{k_i}$ on the ith frame divided by the total number of strips, N, i.e.,

$$P(\hat{p}_{k_i}) = n(\hat{p}_{k_i})/N. \tag{83}$$

The unconditional probability for the sequence $(C_0 \to C_1 \to \cdots \to \hat{p}_{k_i} \to C_{i+1} \to \cdots \to \hat{p}_{k_j} \to C_{j+1} \to \cdots \to C_L)$ then is simply

$$P(\hat{p}_{k_i} \to \hat{p}_{k_j}) = n(\hat{p}_{k_i} \to \hat{p}_{k_j})/N, \tag{84}$$

and the predictive conditional probability is given by

$$P(\hat{p}_{k_j} \mid \hat{p}_{k_i}) = \frac{P(\hat{p}_{k_i} \to \hat{p}_{k_j})}{P(\hat{p}_{k_i})} = \frac{n(\hat{p}_{k_i} \to \hat{p}_{k_j})}{n(\hat{p}_{k_i})}, \quad i < j. \tag{85}$$

With these operational definitions, it is then absolutely reasonable to define the retrodictive conditional probability

$$P(\hat{p}_{k_i} \mid \hat{p}_{k_j}) = \frac{P(\hat{p}_{k_i} \to \hat{p}_{k_j})}{P(\hat{p}_{k_j})} = \frac{n(\hat{p}_{k_i} \to \hat{p}_{k_j})}{n(\hat{p}_{k_j})}, \tag{86}$$

which, as we see from our example, is not anticausal in nature but is a simple result of the *a posteriori* nature of the film data.

From the above example, we see that we may interpret the predictive and the retrodictive random-walk equations in the following way: The predictive random-walk equation will describe the diffusion of a drop of cream placed in a cup of coffee. If we film this process, then the retrodictive random-walk equation will describe the "reverse diffusion process" as it appears on a projection screen when the film is run in reverse. We saw, however, from the analysis of the transition matrices, that the retrodictive transition matrix is the inverse of the predictive transition matrix only for deterministic systems.

When the stochastic equations were cast into their respective $\mathcal{H}$ representations, we saw that the predictive evolution operator S and the retrodictive evolution operator S' defined predictive and retrodictive evolution in $\mathcal{H}$. We saw that S and S' are isometric, but that S' is S^{-1} only for deterministic systems. Furthermore, we saw that, in contrast to conventional quantum theory, S is unitary only for systems described by Eq. (75). Thus we saw that the stochastic $\mathcal{H}$ representation is distinct from the quantum representation so that stochastic processes cannot be considered as a special case of quantum evolution.

We then postulated a mathematical structure [Eq. (82)] in $\mathcal{H}$ which would include both quantum evolution and stochastic evolution as special cases. No basis was given for such a structure, but it is envisioned that a more general definition of the event space $\{E(C)\}$ might well produce the more general postulated structure. Recall that we required $\{E(C)\}$ to be a σ-algebra and further imposed the generalized distributive relation on $\{E(C)\}$. It is hoped that a removal of the generalized distributive requirement, or a mathematical generalization of the σ-algebraic structure of $\{E(C)\}$, or both, will produce the more general evolution picture in $\mathcal{H}$.

APPENDIX A

Suppose that each of $\{T_{k_jk_i}\}_{k_i=1,2,\cdots}$ and

$$\{P(\hat{p}_{k_i})\}_{k_i=1,2,\cdots}$$

is a sequence of positive real numbers and that there exists a real number $P(\hat{p}_{k_j})$ such that

$$P(\hat{p}_{k_j}) = \sum_{k_i=1}^{\infty} T_{k_jk_i} P(\hat{p}_{k_i}). \tag{A1}$$

Then there exists a sequence of complex numbers $\{K_{k_jk_i}\}_{k_i=1,2,\cdots}$, a sequence of complex numbers $\{\alpha_{k_i}\}_{k_i=1,2,\cdots}$, and a complex number α'_{k_j} such that the following equations are consistent:

$$\alpha_{k_j} = \sum_{k_i=1}^{\infty} K_{k_jk_i}\alpha_{k_i}, \tag{A2}$$

$$P(\hat{p}_{k_j}) = \alpha^*_{k_j}\alpha_{k_j} \tag{A3}$$

and, for each positive integer k_i,

$$T_{k_jk_i} = K^*_{k_jk_i}K_{k_jk_i}, \tag{A4}$$

$$P(\hat{p}_{k_i}) = \alpha^*_{k_i}\alpha_{k_i}. \tag{A5}$$

This theorem thus states that phases for the sequence $\{K^*_{k_jk_i}\alpha_{k_i}\}_{k_i=1,2,\cdots}$ can be found such that the double sum formed from the square of equation (A2) reduces to a single sum of real numbers. Equations (55) and (58) are nothing more than the $\mathcal{H}$ representations of Eqs. (A1) and (A2), respectively.

APPENDIX B

Theorem: Suppose that S is a linear continuous operator such that S^{-1} exists and I is a collection of positive integers such that k_i belongs to I only in case $\alpha_{k_i} \neq 0$. Then the equations

$$\alpha^*_{k_j}\alpha_{k_j} = \sum_{k_i} \langle k_j| S(j, i) |k_i\rangle \langle k_i| S^+(j, i) |k_j\rangle \alpha^*_{k_i}\alpha_{k_i} \tag{B1}$$

and

$$\alpha_{k_j} = \sum \langle k_j| S(j, i) |k_i\rangle \alpha_{k_i} \tag{B2}$$

are consistent only in case I has only one member, i.e.,

$$\alpha_{k_i} = \delta_{k_ik_i'}. \tag{B3}$$

Proof: Substitution of (B2) into (B1) for α_{k_j} produces

$$\alpha^*_{k_j} \sum_{k_i\in I} \langle k_j| S(j, i) |k_i\rangle \alpha_{k_i} = \sum_{k_i\in I} \langle k_j| S(j, i) |k_i\rangle \langle k_i| S^+(j, i) |k_j\rangle \alpha^*_{k_i}\alpha_{k_i}. \tag{B4}$$

Rearranging, we obtain

$$\sum_{k_i\in I} (\alpha^*_{k_j} - \langle k_i| S^+(j, i) |k_j\rangle \alpha^*_{k_i})\alpha_{k_i} \langle k_j| S(j, i) |k_i\rangle = 0, \tag{B5}$$

which may be written

$$\sum_{k_i\in I} (\alpha^*_{k_j} - \langle k_i| S\ (j, i) |k_j\rangle \alpha^*_{k_i})S(j, i) = |0\rangle. \tag{B6}$$

If S^{-1} exists, the collection $\{S\ |k_i\rangle\}_{k_i\in I}$ is a linearly independent set so that (B6) is satisfied only in case

$$(\alpha^*_{k_j} - \langle k_i| S^+(j, i) |k_j\rangle \alpha^*_{k_i})\alpha_{k_i} = 0, \quad k_i \in I. \tag{B7}$$

Since α_{k_i} is nonzero for each k_i in I, (B7) is satisfied only in case

$$\alpha^*_{k_j} = \langle k_i| S^+(j, i) |k_j\rangle \alpha^*_{k_i}, \quad k_i \in I, \tag{B8}$$

or

$$\alpha_{k_j} = \langle k_j| S(j, i) |k_i\rangle \alpha_{k_j}, \quad k_i \in I. \tag{B9}$$

Thus we see that (B9) is consistent with (B2) only in case the set $\{\alpha_{k_i}\}$ has only one nonzero member, i.e., I has only one member.

Let us examine the implications of this in terms of probabilities. Since

$$P(\hat{p}_{k_i}) = |\langle k_i \mid \alpha(i)\rangle|^2 = \alpha^*_{k_i}\alpha_{k_i} = \delta_{k_ik_i'}, \tag{B10}$$

we see that the system must be in some initial state of M. We also see then that the random walk equation yields

$$P(\hat{p}_{k_j}) = P(\hat{p}_{k_j} \mid \hat{p}'_{k_i'})\delta_{k_ik_i'}. \tag{B11}$$

* Supported in part by a Frederick-Gardner Cottrell Grant-in-Aid from Research Corporation.

[1] R. E. Collins, Phys. Rev. **183**, 5,1081 (1969).
[2] R. E. Collins, in *Critical Review of Thermodynamics*, E. B. Stuart, Benjamin Gal-Or, and A. J. Brainard, Eds. (Mono, Baltimore, Md., 1970).
[3] R. E. Collins, Phys.Rev. D 1, 2, 379 (1970).
[4] R. E. Collins, Phys.Rev. D 1, 4, 658 (1970).
[5] R. E. Collins, Phys.Rev. D 1, 4, 666 (1970).
[6] R. E. Collins and F. G. Hall, *Proceedings of the International Conference on Thermodynamics*, P. T. Landsberg, Ed., University College, Cardiff, Wales, (April 1970), in press.
[7] P. R. Halmos, *Measure Theory* (Van Nostrand, New York, 1950).
[8] E. Feller, *Introduction to Probability Theory and Its Applications* (Wiley, New York, 1950), Vol. 2.
[9] M. Loeve, *Probability Theory* (Van Nostrand, New York, 1955).
[10] J. M. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, Reading, Mass., 1968).
[11] K. L. Chung, *Markov Chains with Stationary Transition Probabilities* (Springer-Verlag, Berlin, 1960).
[12] T.-Y. Wu, *Kenetic Equations of Gases and Plasmas* (Addison-Wesley, Reading, Mass., 1966).
[13] S. K. Berberian, *Introduction to Hilbert Space* (Oxford U.P., London, 1961).
[14] E. Nelson, Phys. Rev. **150**, 4 (1966).
[15] E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton U.P., Princeton, N.J., 1967).
[16] S. Chandrasekhar, Rev. Mod. Phys. **15**, 17 (1943).