Information-geometric reconstruction of quantum theory

Philip Goyal*
Perimeter Institute, Waterloo, Canada

(Received 19 May 2008; revised manuscript received 13 October 2008; published 25 November 2008)

In this paper, we show how the framework of information geometry, the natural geometry of discrete probability distributions, can be used to derive the quantum formalism. The derivation rests upon a few elementary features of quantum phenomena, such as complementarity and global gauge invariance. When appropriately formulated within an information-geometric framework, and combined with a novel information-theoretic principle, these features lead to the abstract quantum formalism for finite-dimensional quantum systems, and the result of Wigner's theorem. By means of a correspondence principle, several correspondence rules of quantum theory, such as the canonical commutation relationships, are also systematically derived. The derivation suggests that information geometry is directly or indirectly responsible for many of the central structural features of the quantum formalism, such as the importance of square roots of probability and the occurrence of sinusoidal functions of phases in a pure quantum state. Global gauge invariance is shown to play a crucial role in the emergence of the formalism in its complex form.

DOI: 10.1103/PhysRevA.78.052120 PACS number(s): 03.65.Ta, 03.67.−a

I. INTRODUCTION

The unparalleled empirical success of quantum theory suggests that the mathematical formalism of quantum theory accurately captures fundamental aspects of the workings of the physical world. However, the elucidation of these aspects has been hampered by the abstract nature of the mathematical framework (complex Hilbert space) in whose language the quantum postulates are expressed. Over the past two decades, there has been growing interest in bringing these aspects into the open by formulating a set of postulates that are expressed in a simpler mathematical language that has a more transparent physical basis, and showing that the quantum formalism can be derived from these postulates [1–7]. A major motivation for such a reconstruction is the expectation that it might provide some indication as to the relative malleability of different parts of the quantum formalism, an understanding that could provide useful guidance when attempting to modify the quantum formalism (for example, to allow nonlinear continuous transformations [8–10]) or when attempting to build a theory of quantum gravity (see, for example, [11]).

Much of the recent effort in reconstructing the quantum formalism has been encouraged by belief in the hypothesis that the concept of information might be the key, hitherto missing, ingredient that, if appropriately conceptualized and formalized, might make such a reconstruction possible [1,2,5]. There have been several attempts to systematically explore the reconstruction of the quantum formalism from an informational starting point (for example, [2,4,12–21]). A number of these approaches take an operational approach, so that the statistical nature of measurements on quantum systems is taken as a given, and the framework of classical probability theory is accordingly taken as the starting point. Perhaps the earliest of these approaches is remarkable for showing that one can derive a correct, nontrivial physical prediction (Malus' law) concerning a quantum experiment from a simple information-theoretic principle [12]. However, this approach does not lead to the reconstruction of the quantum formalism itself, and other, more recent, approaches to understand the quantum formalism on the basis of simple information-theoretic principles within the framework of probability theory are similarly restricted in their scope [4,14–16]. In contrast, a number of recent approaches succeed in deriving a significant portion of the quantum formalism by using the concept of information [2,17–21], but they do so at the cost of employing frameworks that assume *a priori* the importance of the complex number field [35]. Such assumptions significantly limit the degree to which these reconstructions can elucidate the physical content of the quantum formalism since one of the most mysterious mathematical features of the quantum formalism is being assumed at the outset.

In this paper, we show that it is possible to reconstruct the majority of the quantum formalism within a classical probabilistic framework using the concept of information without relying on abstract assumptions that presuppose the importance of the complex number field. In particular, the complex Hilbert space structure of the quantum formalism is shown to arise naturally from simpler physical ideas.

We adopt an operational approach, and take the probabilistic nature of measurements as a given. Accordingly, the framework of classical probability theory is taken as a starting point. We equip the probabilistic framework with a metric, $ds^2=\frac{1}{4}\Sigma_i dp_i^2/p_i$, known as the *information metric* (or Fisher-Rao metric), which determines the distance between infinitesimally close probability distributions $\mathbf{p}=(p_1,\ldots,p_N)$ and $\mathbf{p}'=(p_1',\ldots,p_N')$. This distance can be thought of as quantifying the difficulty with which these distributions can be distinguished from one another given a finite number of samples from each. As we shall describe below, the metric can be seen as a natural consequence of introducing the concept of information into the probabilistic framework. Accordingly, we shall refer to this framework as the information geometric framework [22].

*pgoyal@perimeterinstitute.ca

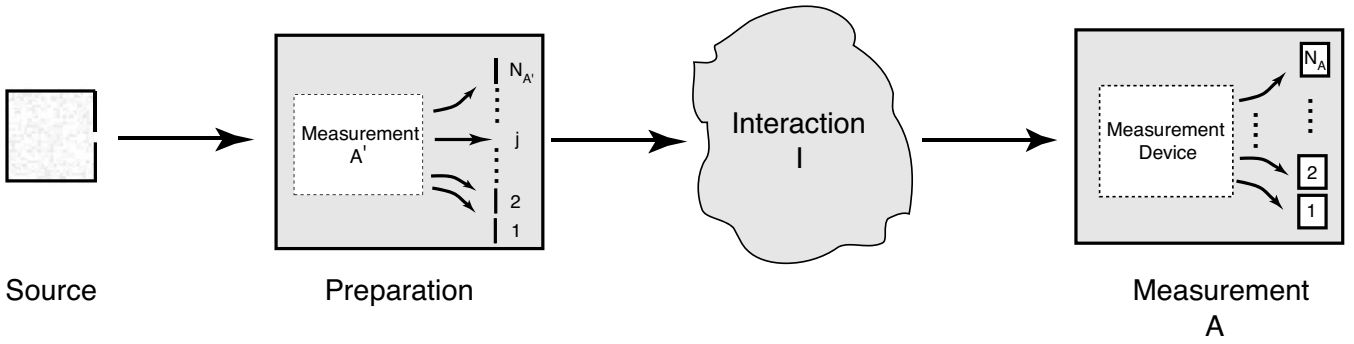

FIG. 1. An abstract experimental setup. In each run of the experiment, a physical system (such as a silver atom) is emitted from a source, passes a preparation step, undergoes an interaction, **I**, and is then subject to a measurement. The preparation is implemented as a measurement, $\mathbf{A}'$, which has $N_{A'}$ possible outcomes, followed by the selection of those systems that yield some outcome j ($j=1,2,\ldots,N_{A'}$). The measurement **A** has N_A possible outcomes.

Within this information-geometric framework, we formalize a few elementary features of quantum phenomena, such as complementarity and global gauge invariance. These features can be understood as assertions about the physical world quite apart from the setting of the quantum formalism within which they are usually encountered, and are sufficiently simple to be taken as primitives in the building up of quantum theory. To these features we add an information-theoretic principle, the principle of metric invariance. From these ingredients, we reconstruct the principal features of the quantum formalism.

This paper is organized as follows. In Sec. II, we present an experimental framework and a set of postulates that determine the theoretical model of a system in the context of this experimental framework. In Sec. III, we show how these postulates give rise to the finite-dimensional abstract quantum formalism, namely the von Neumann postulates for finite-dimensional systems (apart from the form of the temporal evolution operator) and the result of Wigner's theorem (that a one-to-one map over the state space of a system that represents a symmetry transformation of the system is either unitary or antiunitary) [23].

In Sec. IV, we formulate the average-value correspondence principle (AVCP), and use this to obtain the form of the temporal evolution operator and, taking the infinite-dimensional form of the quantum formalism as a given, derive several correspondence rules including the canonical commutation relationship $[\mathsf{x},\mathsf{p}_x]=i\hbar$.

In Sec. V, it is shown how many of the postulates can be regarded as reasonable generalizations of elementary experimental facts that are characteristic of quantum phenomena, are drawn from classical physics, or can be understood as novel physical assertions. We conclude in Sec. VI with a discussion of the insights provided by the derivation and of some of the questions it raises.

II. EXPERIMENTAL FRAMEWORK AND POSTULATES

A. Experimental framework

We shall be concerned with constructing a theoretical model of the abstract experimental setup shown in Fig. 1: (i) a *source* emits a copy of a physical system of interest, (ii) a *preparation step* then either selects or rejects the incoming system, (iii) the system undergoes an *interaction* with a physical apparatus, and (iv) the system undergoes a *measurement*.

A measurement is viewed as a process that acts upon an *input* system, and generates an *output* system together with an observed *outcome*. Measurements are idealized as follows: (a) measurements have a finite number of possible outcomes, (b) measurements are reproducible, so that immediate repetition of a measurement yields the same outcome with certainty, and (c) measurement outcomes obtained in many runs of the experiment are characterized by a probability distribution. A preparation is implemented by performing a measurement upon the incoming system, and then selecting the system (allowing it to pass) if a given outcome is obtained, and rejecting it (blocking it) otherwise. Consideration is restricted to interactions that are reversible and deterministic, and that preserve the integrity of the system upon which they act.

An example of this abstract experimental setup is an experiment where silver atoms emerge from a source (an evaporator), pass through a preparation step (implemented using a Stern-Gerlach measurement), undergo an interaction with a uniform magnetic field, and finally undergo a Stern-Gerlach measurement.

Experimental closure and experimental sets

In the above-mentioned Stern-Gerlach experiment, one finds experimentally that the probability distribution that characterizes the measurement data is *independent* of arbitrary interactions with the system prior to the preparation. That is, the experiment has been arranged in such a way that the measurement data are not influenced by conditions that are not under experimental control. We shall say that experimental setups of this kind are *closed* (or have the property of *closure*), and we shall restrict our consideration to such setups.

In the Stern-Gerlach experiment, given a particular preparation, the measurement Stern-Gerlach device can be rotated to perform a range of possible Stern-Gerlach measurements, and different uniform magnetic field interactions can occur between the preparation and measurement. In each of these

setups, the spin behavior of the system is being probed. In the case of an abstract experimental setup, we can use the concept of closure to formulate an operational procedure that delineates the set of all setups that probe the same behavioral aspect of a given system.

The procedure. The first step in the procedure is to identify the set of all measurements that probe precisely the same behavioral aspect as some given measurement, **A**. We refer to this set as the *measurement set*, $\mathcal{A}$, *generated* by **A**. Consider an experiment (Fig. 1) in which a system from a source is subject to a preparation consisting of measurement, $\mathbf{A}'$, with $N_{A'}$ possible outcomes, with outcome j selected ($j=1,\ldots,N_{A'}$), followed by measurement **A** (with N_A possible outcomes), *without* an interaction in the intervening time. Suppose that the data obtained in the experiment are characterized by a probabilistic source with N_A possible outcomes and with probability n-tuple $\vec{p}=(p_1,p_2,\ldots,p_{N_A})$, where p_i is the probability of the ith outcome ($i=1,2,\ldots,N_A$). If, for all j, $\vec{p}$ is independent of arbitrary pre-preparation interactions with the system, so that the setup is closed, the preparation will be said to be *complete* with respect to measurement **A**. If this completeness condition also holds true when **A** and $\mathbf{A}'$ are interchanged, then **A** and $\mathbf{A}'$ will be said to form a *measurement pair*. For example, any Stern-Gerlach preparation is complete with respect to a Stern-Gerlach measurement, and any two Stern-Gerlach measurements form a measurement pair.

The measurement set, $\mathcal{A}$, generated by **A** is then defined as the set of all measurements that (i) form a measurement pair with **A** and (ii) are not a composite of other measurements in $\mathcal{A}$. In the Stern-Gerlach example above, if **A** is any given Stern-Gerlach measurement, then it generates a measurement set, $\mathcal{A}$, consisting of all Stern-Gerlach measurements of the form $\mathbf{A}_{\theta,\phi}$, where (θ,ϕ) is the orientation of the Stern-Gerlach device. Measurements that are composed of two or more Stern-Gerlach measurements are excluded from $\mathcal{A}$ by definition.

In the second step, we determine the set of all interactions (the *interaction set*, $\mathcal{I}$) with the system that can occur between a preparation (implemented using a measurement in $\mathcal{A}$) and a measurement (chosen from $\mathcal{A}$). Suppose that, in the experiment of Fig. 1, an interaction, **I**, occurs between the preparation and measurement. If, for all $\mathbf{A},\mathbf{A}'\in\mathcal{A}$, the preparation remains complete with respect to the subsequent measurement, then **I** will be said to be *compatible* with $\mathcal{A}$. The set $\mathcal{I}$ is then defined as the set of all such compatible interactions.

In the case of the Stern-Gerlach experiment, if an interaction, $\mathbf{I}_{\theta_B,\phi_B,t,\Delta t}$, consisting of a uniform $\vec{B}$ field acting during the interval $[t,t+\Delta t]$ in some direction (θ_B,ϕ_B), occurs between the preparation and measurement, one finds experimentally that the completeness of the preparation with respect to the measurement is preserved; that is, the interaction is compatible with the set of all Stern-Gerlach measurements of the form $\mathbf{A}_{\theta,\phi}$. Hence, in this example, all interactions in which a uniform magnetic field acts between the preparation and measurement are in the interaction set, $\mathcal{I}$.

In terms of these definitions, a closed setup consists of a source of systems where each system is prepared using a measurement $\mathbf{A}'\in\mathcal{A}$, is subject to an interaction $\mathbf{I}\in\mathcal{I}$, and then undergoes a measurement $\mathbf{A}\in\mathcal{A}$. The set of all such setups constitutes an *experimental set*, and will be said to be *generated* by measurement **A**.

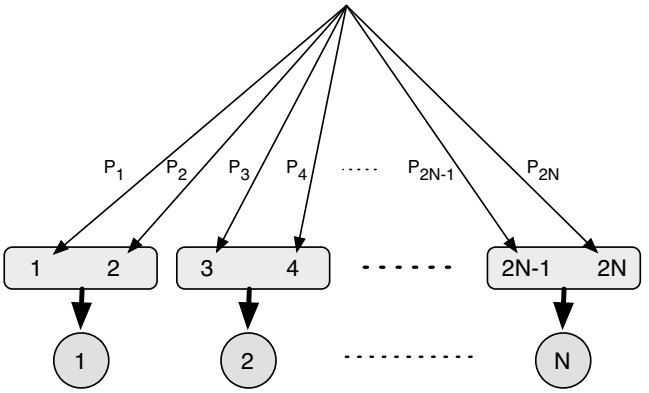

FIG. 2. *Complementarity postulate*. Probability tree showing the events that occur when measurement **A** is performed, and the corresponding outcomes (circled) that are observed. One of $2N$ possible events occurs, with probability $P_1,P_2,\ldots,P_{2N}$, respectively. The individual events are not resolved. Outcome i is obtained whenever either event $2i-1$ or $2i$ is realized ($i=1,\ldots,N$).

Finally, we provide an operational definition of the terms *subsystem* and *composite system*. Suppose that a physical system admits an experimental set generated by a measurement $\mathbf{A}^{(1)}$ in measurement set $\mathcal{A}^{(1)}$, and also admits an experimental set generated by measurement $\mathbf{A}^{(2)}$ in measurement set $\mathcal{A}^{(2)}$, where the sets $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ are disjoint. In that case, we shall say that each experimental set probes a different *subsystem* of the same physical system. Suppose further that the physical system admits an experimental set generated by measurement **A** that consists of measurement $\mathbf{A}^{(1)}$ followed by $\mathbf{A}^{(2)}$. We shall then say that this experimental set probes a system that is a *composite system*.

B. Postulates

The postulates, stated below, determine the abstract model of a physical system subject to any setup chosen from an experimental set. There are two key novel postulates. First, the Complementarity postulate (Postulate 2.1) expresses within a probabilistic framework the idea that, when a measurement is performed on a physical system, the data obtained only provide information about half of the experimentally accessible degrees of freedom of the state of the system. In particular, it is assumed that, when a measurement is performed, one of $2N$ possible *events* occurs, with a probability determined by the state of the system, but that only one of N possible *outcomes* is actually observed (each outcome *coarse-graining* over two of the events), so that the outcomes only provide information about half of the event probabilities (see Fig. 2). Second, the Metric Invariance postulate (Postulate 3.5), an information-theoretic postulate, rests upon the idea that the distinguishability of two states is a quantity intrinsic to this pair of states, and therefore does not depend upon the measurement used to distinguish between them. As described later (Sec. V C), this idea leads directly to the re-

quirement that the metric over state space (which is induced by the information metric, introduced through the Information Metric postulate, which is a measure of the distinguishability of nearby states) is invariant under mappings over state space.

Several of the postulates (2.3, 2.5, 3.6, 3.7, and 4) rest upon elementary properties of the classical Hamilton-Jacobi model of a particle (see Sec. V B). The global gauge invariance property of the Hamilton-Jacobi model is of key importance, and we generalize this property to an abstract physical system by assuming that, if the state of the system is represented as $(p_1,\ldots,p_N;\chi_1,\ldots,\chi_N)$ with respect to a measurement, where the p_i are the outcome probabilities of the measurement and the χ_i are real degrees of freedom, then any predictions made on the basis of the state are invariant under the transformation $\chi_i \rightarrow \chi_i+\chi_0$, for all $i=1,\ldots,N$, for any χ_0. The Measure Invariance (2.5) and the Gauge Invariance postulates (3.6) are both particular instances of this general assumption.

The remaining postulates are either drawn directly from the framework of classical physics (3.1-3.4) or from elementary experimental observations characteristic of quantum phenomena (1.1 and 1.2). The reader is referred to Sec. V for further discussion of the physical origin of the postulates.

Statement of the postulates

The abstract model of a physical system subject to an experimental setup where the measurements and interactions are chosen from a measurement set $\mathcal{A}$ and interaction set $\mathcal{I}$, respectively, satisfies the following postulates.

1. Measurements

1.1 *Outcomes.* When any measurement $\mathbf{A}\in\mathcal{A}$ is performed, one of N ($N\geqslant 2$) possible outcomes is observed.

1.2 *Measurement simulability.* For any given pair of measurements $\mathbf{A},\mathbf{A}'\in\mathcal{A}$, there exist interactions $\mathbf{I},\mathbf{I}'\in\mathcal{I}$ such that $\mathbf{A}'$ can, insofar as the outcome probabilities and output states of the measurement are concerned, be simulated by an arrangement where $\mathbf{I}$ is immediately followed by $\mathbf{A}$, which, in turn, is immediately followed by $\mathbf{I}'$.

2. States

2.1 *Complementarity.* When any given measurement $\mathbf{A}\in\mathcal{A}$ is performed on the system, one of $2N$ possible events occurs with probability $P_1,\ldots,P_{2N}$, respectively. The individual events are unobserved. Outcome i is observed whenever event $(2i-1)$ or event $2i$ occurs.

2.2 *States.* The state, $\mathbf{S}$, of the system with respect to measurement $\mathbf{A}\in\mathcal{A}$ is given by $\mathbf{Q}=(Q_1,Q_2,\ldots,Q_{2N})^{\mathsf{T}}$, where $Q_q\in[-1,1]$, $q=1,\ldots,2N$. The probability of event q is given by $P_q=Q_q^2$, and the variable, $\sigma_q=\mathrm{sgn}(Q_q)$, which is defined if $P_q\neq 0$, is a binary degree of freedom (a *polarity*) associated with event q.

2.3 *State representation.* The state, $\mathbf{S}$, of a system with respect to measurement $\mathbf{A}$ can be represented by the pair $(\vec{p},\vec{\chi})$, where $\vec{p}=(p_1,\ldots,p_N)$ and $\vec{\chi}=(\chi_1,\ldots,\chi_N)$ are real n-tuples, and where p_i is the probability that the ith outcome of measurement $\mathbf{A}$ is observed. In particular, the state is given by $\mathbf{Q}=(\sqrt{p_1}\cos\theta(\chi_1),\sqrt{p_1}\sin\theta(\chi_1),\ldots,\sqrt{p_N}\cos\theta(\chi_N),\sqrt{p_N}\sin\theta(\chi_N))$, where $\theta(\cdot)$ is a nonconstant differentiable function.

2.4 *Information metric.* The metric over $\mathbf{P}=(P_1,P_2,\ldots,P_{2N-1},P_{2N})$ is the information metric, $ds^2=\frac{1}{4}\sum_{q=1}^{2N}dP_q^2/P_q^2$.

Measure invariance. The measure, $\mu(p_1,\ldots,p_N;\chi_1,\ldots,\chi_N)$, over $(\vec{p},\vec{\chi})$ induced by the metric over $\mathbf{Q}$ satisfies the condition $\mu(p_1,\ldots,p_N;\chi_1,\ldots,\chi_N)=\mu(p_1,\ldots,p_N;\chi_1+\chi_0,\ldots,\chi_N+\chi_0)$ for all χ_0.

3. Transformations

3.1 *Mappings.* Any physical transformation of the system, whether active (due to temporal evolution) or passive (due to a change of frame of reference), is represented by a map, $\mathcal{M}$, over the state space, $\mathcal{S}$, of the system.

3.2 *One-to-one.* Every map, $\mathcal{M}$, that represents a physical transformation of the system is one-to-one.

3.3 *Continuity.* If a physical transformation is continuously dependent upon the real-valued parameter n-tuple $\boldsymbol{\pi}$, and is represented by the map $\mathcal{M}_{\boldsymbol{\pi}}$, then $\mathcal{M}_{\boldsymbol{\pi}}$ is continuously dependent upon $\boldsymbol{\pi}$.

3.4 *Continuous transformations.* If $\mathcal{M}_{\boldsymbol{\pi}}$ represents a continuous transformation, then, for some value of $\boldsymbol{\pi}$, $\mathcal{M}_{\boldsymbol{\pi}}$ reduces to the identity.

3.5 *Metric invariance.* The map $\mathcal{M}$ preserves the metric over the state space, $\mathcal{S}$, of the system.

3.6 *Gauge invariance.* The map $\mathcal{M}$ is such that, for any state $\mathbf{S}\in\mathcal{S}$, the probabilities, $p_1',p_2',\ldots,p_N'$, of the outcomes of measurement $\mathbf{A}\in\mathcal{A}$ performed upon a system in state $\mathbf{S}'=\mathcal{M}(\mathbf{S})$ are unaffected if, in any representation, $(p_i;\chi_i)\equiv(\vec{p},\vec{\chi})$, of the state $\mathbf{S}$ written down with respect to $\mathbf{A}$, an arbitrary real constant, χ_0, is added to each of the χ_i.

3.7 *Temporal evolution.* The map, $\mathcal{M}_{t,\Delta t}$, which represents temporal evolution of a system in a time-independent background during the interval $[t,t+\Delta t]$, is such that any state, $\mathbf{S}$, represented as $(p_i;\chi_i)$, of definite energy E, whose observable degrees of freedom are time-independent, evolves to $(p_i';\chi_i')=(p_i;\chi_i-E\Delta t/\alpha)$, where α is a nonzero constant with the dimensions of action.

The above postulates, together with the average-value correspondence principle (AVCP), which will be given in Sec. IV A, suffice to determine the form of the abstract quantum model. From the Outcomes postulate, it follows that, when any measurement in $\mathcal{A}$ is performed on the system, one of N possible outcomes is observed. Accordingly, we shall denote the abstract quantum model by $\mathbf{q}(N)$.

Finally, the Composite Systems postulate, below, is needed in order to obtain a rule for relating the quantum model of a composite system to the quantum models of its component systems:

4. Composite Systems. Suppose that a system admits a quantum model (model 1), with respect to the measurement set $\mathcal{A}^{(1)}$ whose measurements have N possible outcomes, and admits a quantum model (model 2) with respect to measurement set $\mathcal{A}^{(2)}$ whose measurements have N' possible outcomes, where the sets $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ are disjoint.

Consider the quantum model (model 3) of the system with respect to the measurement set $\mathcal{A}=\mathcal{A}^{(1)}\times\mathcal{A}^{(2)}$ that contains all possible composite measurements consisting of a measurement from $\mathcal{A}^{(1)}$ and a measurement from $\mathcal{A}^{(2)}$. If the states of the system in models 1 and 2 can be represented as $(p_i;\chi_i)$ $(i=1,2,\ldots,N)$ and $(p_j';\chi_j')$ $(j=1,2,\ldots,N')$, respectively, then the state of the system in model 3 can be repre-

sented as $(p_l'';\chi_l'')$ $(l=1,2,\ldots,NN')$, where $p_l''=p_ip_j'$ and $\chi_l''=\chi_i+\chi_j'$, where $l=N'(i-1)+j$.

III. DEDUCTION OF THE QUANTUM FORMALISM

A. States and dynamics in Q space

According to the States postulate, the state of a system prepared using any measurement in the measurement set $\mathcal{A}$ can be represented by $\mathbf{Q}=(Q_1,Q_2,\ldots,Q_{2N-1},Q_{2N})$. Now, the Information Metric postulate assigns the metric

$$ds^2=\frac{1}{4}\sum_{q=1}^{2N}\frac{dP_q^2}{P_q} \tag{1}$$

over $\mathbf{P}=(P_1,\ldots,P_{2N})$, which, from the relation $P_q=Q_q^2$ $(q=1,\ldots,2N)$, implies that the metric over $\mathbf{Q}$ is Euclidean, namely

$$ds^2=\sum_q dQ_q^2. \tag{2}$$

Hence, the state space of the system can be represented by the set of all unit vectors in a $2N$-dimensional real Euclidean space, which we will refer to as Q space or Q^{2N}.

By the Mappings and One-to-One postulates, any physical transformation is represented by a one-to-one map, $\mathcal{M}$, over state space, Q^{2N}. Furthermore, by the Metric Invariance postulate, it follows that $\mathcal{M}$ must be an orthogonal transformation of the unit hypersphere, S^{2N-1}, in Q^{2N}.

B. Global gauge invariance, and the complex form of states and dynamics

1. Determination of function $\theta(\cdot)$

According to the State Representation postulate, the state $\mathbf{Q}$ of a system with respect to some measurement $\mathbf{A}\in\mathcal{A}$ can be written

$$\mathbf{Q}=(\sqrt{p_1}\cos\,\theta(\chi_1),\sqrt{p_1}\sin\,\theta(\chi_1),\ldots,\sqrt{p_N}\cos\,\theta(\chi_N),\sqrt{p_N}\sin\,\theta(\chi_N)). \tag{3}$$

In order to determine the unknown function $\theta(\cdot)$, we first determine the metric over $\mathbf{Q}$ in terms of the p_i and χ_i.

Using Eq. (2),

$$\begin{aligned}ds^2&=dQ_1^2+dQ_2^2+\cdots+dQ_{2N-1}^2+dQ_{2N}^2\\&=\sum_{i=1}^{N}\frac{1}{4}\frac{dp_i^2}{p_i}+\sum_{i=1}^{N}p_i\theta'^2(\chi_i)d\chi_i^2.\end{aligned} \tag{4}$$

The measure over $(p_1,\ldots,p_N;\chi_1,\ldots,\chi_N)$ induced by this metric is proportional to the square root of the determinant of the metric, and so is given by

$$\mu(\vec{p},\vec{\chi})=\mu_0\prod_{i=1}^{N}|\theta'(\chi_i)|, \tag{5}$$

where μ_0 is a constant, which marginalizes to give

$$\mu_i(\chi_i)\equiv\int\cdots\int\mu(\vec{p},\vec{\chi})dV_i^-=c\mu_0|\theta'(\chi_i)|, \tag{6}$$

with $dV_i^-\equiv dp_1\cdots dp_Nd\chi_1\cdots d\chi_{i-1}d\chi_{i+1}\cdots d\chi_N$, as the measure over χ_i, where c is a constant.

Now, using the Measure Invariance postulate, $\mu_i(\chi_i+\chi_0)$ is given by

$$\begin{aligned}&\int\cdots\int\mu(\vec{p},\chi_1,\ldots,\chi_{i-1},\chi_i+\chi_0,\chi_{i+1},\ldots,\chi_N)dV_i^-\\&=\int\cdots\int\mu(\vec{p},\tilde{\chi}_1+\chi_0,\ldots,\tilde{\chi}_{i-1}+\chi_0,\chi_i+\chi_0,\\&\quad\tilde{\chi}_{i+1}+\chi_0,\ldots,\tilde{\chi}_N+\chi_0)d\tilde{V}_i^-\\&=\int\cdots\int\mu(\vec{p},\tilde{\chi}_1,\ldots,\tilde{\chi}_{i-1},\chi_i,\tilde{\chi}_{i+1},\ldots,\tilde{\chi}_N)d\tilde{V}_i^-=\mu_i(\chi_i),\end{aligned} \tag{7}$$

with $d\tilde{V}_i^-\equiv dp_1\cdots dp_Nd\tilde{\chi}_1\cdots d\tilde{\chi}_{i-1}d\tilde{\chi}_{i+1}\cdots d\tilde{\chi}_N$, where the variable substitution $\tilde{\chi}_j=\chi_j-\chi_0$ for $j\neq i$ has been used to obtain the second line. Hence, the measure $\mu_i(\chi_i)$ is independent of χ_i. Therefore, from Eq. (6), $\theta(\chi)=a\chi+b$, where a,b are constants, where $a\neq 0$ since, from the State Representation postulate, the function $\theta(\cdot)$ is not constant.

Defining $\phi_i=a\chi_i+b$, we can therefore, from Eq. (3), write the state of a system with respect to some measurement $\mathbf{A}\in\mathcal{A}$ as

$$\mathbf{Q}=(\sqrt{p_1}\cos\,\phi_1,\sqrt{p_1}\sin\,\phi_1,\ldots,\sqrt{p_N}\sin\,\phi_N)^{\mathsf{T}}. \tag{8}$$

2. Mappings

In this section, the general form of mappings that represent physical transformations of a system will be determined. The derivation proceeds in three steps.

Step 1: Imposition of the Gauge Invariance postulate. In this step, we show that the imposition of the Gauge Invariance postulate restricts $\mathcal{M}$ to a subset of the set of orthogonal transformations, and that these transformations can be recast as unitary or antiunitary transformations acting on a suitably defined complex vector space.

From Eq. (8), the Gauge Invariance postulate, and the relation $\phi_i=a\chi_i+b$ given above, it follows that the probabilities $p_1',p_2',\ldots,p_N'$ of the outcomes of measurement $\mathbf{A}$ performed on a system in state $\mathbf{Q}'=\mathcal{M}(\mathbf{Q})$ are unaffected if, in any state $\mathbf{Q}=(\sqrt{p_1}\cos\,\phi_1,\sqrt{p_1}\sin\,\phi_1,\ldots,\sqrt{p_N}\cos\,\phi_N,\sqrt{p_N}\sin\,\phi_N)^{\mathsf{T}}$ written down with respect to measurement $\mathbf{A}$, an arbitrary real constant, ϕ_0, is added to each of the ϕ_i. Our goal in this section is to determine the constraint imposed on $\mathcal{M}$ by this condition.

Since $\mathcal{M}$ is an orthogonal transformation (Sec. III A), it can be represented by the $2N$-dimensional orthogonal matrix, M. Under its action, the vector $\mathbf{Q}$ transforms as

$$\mathbf{Q}'=M\mathbf{Q}. \tag{9}$$

In order to determine the most general permissible form of M, it is suffices to consider two special cases.

First consider the case in which all but one of the p_i are zero. In the case in which $p_i=1$, one obtains

$$p'_k = \frac{1}{2}(\alpha_{ki}+\beta_{ki}) + \frac{1}{2}(\alpha_{ki}-\beta_{ki})\cos 2\phi_i + \gamma_{ki}\sin 2\phi_i \tag{10}$$

for $k=1,\ldots,N$, where

$$\alpha_{ki} = M^2_{2k-1,2i-1} + M^2_{2k,2i-1},$$

$$\beta_{ki} = M^2_{2k-1,2i} + M^2_{2k,2i},$$

$$\gamma_{ki} = M_{2k-1,2i-1}M_{2k-1,2i} + M_{2k,2i-1}M_{2k,2i}. \tag{11}$$

The invariance condition implies the conditions

$$\alpha_{ki} = \beta_{ki} \quad \text{and} \quad \gamma_{ki} = 0 \quad \text{for all } i,k, \tag{12}$$

which implies that M takes the form of an N-by-N array of two-by-two submatrices,

$$M = \begin{pmatrix} T^{(11)} & T^{(12)} & \cdots & T^{(1N)} \\ T^{(21)} & T^{(22)} & \cdots & T^{(2N)} \\ & & \cdots & \\ T^{(N1)} & T^{(N2)} & \cdots & T^{(NN)} \end{pmatrix}, \tag{13}$$

where

$$T^{(ij)} = \alpha_{ij}\begin{pmatrix} \cos\varphi_{ij} & -\sigma_{ij}\sin\varphi_{ij} \\ \sin\varphi_{ij} & \sigma_{ij}\cos\varphi_{ij} \end{pmatrix}$$

is a two-by-two matrix composed of a enlargement matrix (scale factor α_{ij}) and a rotation matrix if $\sigma_{ij}=1$ or a reflection-rotation matrix if $\sigma_{ij}=-1$, with rotation angle φ_{ij} in either case.

Before considering the second type of special case, we note that, due to the special form of M above, it is possible to rewrite Eq. (9) in a simpler way. The state $\mathbf{Q}$ can be faithfully represented by

$$\mathsf{v} = \begin{pmatrix} Q_1 + iQ_2 \\ Q_3 + iQ_4 \\ \cdots \\ Q_{2N-1} + iQ_{2N} \end{pmatrix}, \tag{14}$$

so that, from Eq. (8),

$$v_i = \sqrt{p_i}e^{i\phi_i}. \tag{15}$$

Define the complex matrix

$$\mathsf{W} = \begin{pmatrix} \alpha_{11}e^{i\varphi_{11}}K^{\sigma_{11}} & \cdots & \alpha_{1N}e^{i\varphi_{1N}}K^{\sigma_{1N}} \\ \alpha_{21}e^{i\varphi_{21}}K^{\sigma_{21}} & \cdots & \alpha_{2N}e^{i\varphi_{2N}}K^{\sigma_{2N}} \\ \cdots & \cdots & \cdots \\ \alpha_{N1}e^{i\varphi_{N1}}K^{\sigma_{N1}} & \cdots & \alpha_{NN}e^{i\varphi_{NN}}K^{\sigma_{NN}} \end{pmatrix}, \tag{16}$$

where K^σ is the conditional complex conjugation operation defined as

$$K^\sigma z = \begin{cases} z & \text{if } \sigma = 1 \\ z^* & \text{if } \sigma = -1, \end{cases} \tag{17}$$

where $z \in \mathbb{C}$. Then, Eq. (9) is equivalent to the equation

$$\mathsf{v}' = \mathsf{W}\mathsf{v}, \tag{18}$$

where v' is defined analogously to v.

Consider now the second special case, where two of the p_i, say p_i and p_j $(i \neq j)$, are set equal to $1/2$, and the remainder are set to zero. In the case in which $p_i=p_j=1/2$, with $i \neq j$, one obtains the expression

$$p'_k = \frac{1}{4}\{\alpha^2_{ki} + \alpha^2_{kj} + 2\alpha_{ki}\alpha_{kj}\cos[(\varphi_{ki}-\varphi_{kj}) + (\sigma_{ki}\phi_i - \sigma_{kj}\phi_j)]\}, \tag{19}$$

and the invariance condition implies that, for any k and any $i \neq j$, the values of σ_{ki} and σ_{kj} must be the same unless $\alpha_{ki}=0$ or $\alpha_{kj}=0$.

Since M represents the mapping $\mathcal{M}$, and, by the One-to-One postulate, $\mathcal{M}^{-1}$ exists, the matrix M^{-1} represents the mapping $\mathcal{M}^{-1}$. Hence, the matrix $M^{-1}=M^{\mathsf{T}}$ must also satisfy the Invariance postulate. Now, from Eq. (13), the matrix M^{T} takes the form

$$M^{\mathsf{T}} = \begin{pmatrix} (T^{(11)})^{\mathsf{T}} & (T^{(21)})^{\mathsf{T}} & \cdots & (T^{(N1)})^{\mathsf{T}} \\ (T^{(12)})^{\mathsf{T}} & (T^{(22)})^{\mathsf{T}} & \cdots & (T^{(N2)})^{\mathsf{T}} \\ & & \cdots & \\ (T^{(1N)})^{\mathsf{T}} & (T^{(2N)})^{\mathsf{T}} & \cdots & (T^{(NN)})^{\mathsf{T}} \end{pmatrix}, \tag{20}$$

and the corresponding complex matrix is

$$\tilde{W}_{ij} = \alpha_{ji}e^{-i\sigma_{ji}\varphi_{ji}}K^{\sigma_{ji}}. \tag{21}$$

Consider the transformation $\mathsf{v}'=\tilde{\mathsf{W}}\mathsf{v}$, with the above special case, namely $p_i=p_j=1/2$, where $i \neq j$. In this case, one obtains

$$p'_k = \frac{1}{4}\{\alpha^2_{ik} + \alpha^2_{jk} + 2\alpha_{ik}\alpha_{jk}\cos[(-\sigma_{ik}\varphi_{ik} + \sigma_{jk}\varphi_{jk}) + (\sigma_{ik}\phi_i - \sigma_{jk}\phi_j)]\}. \tag{22}$$

The invariance condition implies that, for any k and any $i \neq j$, the values of σ_{ik} and σ_{jk} must be the same unless $\alpha_{ik}=0$ or $\alpha_{jk}=0$. But this implies that, in W, all of the nonzero entries have the same value of σ_{ik}. Therefore, W is of one of two types,

$$\mathsf{W} = \mathsf{V} \quad \text{or} \quad \mathsf{W} = \mathsf{V}\mathsf{K}, \tag{23}$$

corresponding to the cases in which the $\sigma_{ij}=1$ and the $\sigma_{ij}=-1$, respectively, where

$$V_{ij} = \alpha_{ij}e^{i\varphi_{ij}}, \tag{24}$$

and K is the complex conjugation operator, $\mathsf{K}\mathsf{v}=\mathsf{v}^*$.

Now, since M is orthogonal, it follows that V is unitary. To see this, consider M with all $\sigma_{ij}=1$. In that case,

$$(M^{\mathsf{T}}M)_{[ij]} = \sum_k (T^{(ki)})^{\mathsf{T}} T^{(kj)} = \sum_k \alpha_{ki}\alpha_{kj} R(-\varphi_{ki}) R(\varphi_{kj})$$
$$= \sum_k \alpha_{ki}\alpha_{kj} R(\varphi_{kj} - \varphi_{ki}), \tag{25}$$

where $A_{[ij]}$ denotes the (i,j)th two-by-two submatrix of A, with A being an N by N array of two-by-two submatrices, and where $R(\varphi)$ is a two-by-two rotation matrix with rotation angle φ. Consider also

$$(\mathsf{V}^\dagger \mathsf{V})_{ij} = \sum_k \alpha_{ki}\alpha_{kj} e^{i(\varphi_{kj} - \varphi_{ki})}. \tag{26}$$

By inspection, the orthogonality condition $(M^{\mathsf{T}}M)_{[ij]} = \delta_{ij} I$, where I is the two-by-two identity matrix, is equivalent to the condition of unitarity, $(\mathsf{V}^\dagger \mathsf{V})_{ij} = \delta_{ij}$. Therefore, since M is orthogonal, V is unitary. Hence, from Eq. (23), matrix W is either unitary or antiunitary.

Finally, we must show that transformations of the types in Eq. (23) satisfy the Gauge Invariance postulate for any state $\mathbf{Q}$, not just for the special cases of $\mathbf{Q}$ considered above. This follows immediately: we note that the addition of ϕ_0 to each of the ϕ_i in the complex form of the state, v, generates the state $e^{i\phi_0}\mathsf{v}$, that is,

$$\mathsf{v} \xrightarrow{+\phi_0} e^{i\phi_0}\mathsf{v}. \tag{27}$$

As a result, the vector $\mathsf{v}' = \mathsf{V}\mathsf{v}$ transforms as

$$\mathsf{v}' \xrightarrow{+\phi_0} e^{i\phi_0}\mathsf{v}', \tag{28}$$

and the vector $\mathsf{v}' = \mathsf{V}\mathsf{K}\mathsf{v}$ transforms as

$$\mathsf{v}' \xrightarrow{+\phi_0} e^{-i\phi_0}\mathsf{v}'. \tag{29}$$

Since $p_i' = |v_i'|^2$, the p_i' are independent of the overall phase of v', so that, in both Eqs. (28) and (29), the p_i' remain unchanged by the addition of ϕ_0 to the ϕ_i. Therefore, the transformations V and $\mathsf{V}\mathsf{K}$ both satisfy the Gauge Invariance postulate for all $\mathbf{Q}$.

Step 2: General unitary and antiunitary transformations. We have shown so far that the imposition of the Gauge Invariance postulate restricts M to a subset of the set of orthogonal transformations, and that each transformation in this subset can be recast either as a unitary transformation or as an antiunitary transformation. However, we have not ruled out the possibility that there exist unitary or antiunitary transformations that are not equivalent to orthogonal transformations belonging to the above-mentioned subset. In this section, it shall be shown that, in fact, *any* N-dimensional unitary or antiunitary transformation satisfies the One-to-One, Metric Invariance, and Gauge Invariance postulates.

Consider first the arbitrary unitary transformation U, where $U_{ij} = \alpha_{ij} e^{i\varphi_{ij}}$. The transformation

$$\mathsf{v}' = \mathsf{U}\mathsf{v} \tag{30}$$

is equivalent to the transformation

$$\mathbf{Q}' = M\mathbf{Q}, \tag{31}$$

where $M_{[ij]} = \alpha_{ij} R(\varphi_{ij})$. Now, as we observed above, the condition of unitarity of U, namely $(\mathsf{U}^\dagger \mathsf{U})_{ij} = \delta_{ij}$, is equivalent to the orthogonality condition of M. Therefore, M is orthogonal.

Similarly, in the case of the arbitrary antiunitary transformation $\mathsf{U}\mathsf{K}$, where U is defined as above and K is the complex conjugation operator, the corresponding matrix M is given by $M_{[ij]} = \alpha_{ij} R(\varphi_{ij}) F$, where $F = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. In this case,

$$(M^{\mathsf{T}}M)_{[ij]} = \sum_k (M^{\mathsf{T}})_{[ik]} M_{[kj]} = \sum_k \alpha_{ki}\alpha_{kj} F R(-\varphi_{ki}) R(\varphi_{kj}) F$$
$$= \sum_k \alpha_{ki}\alpha_{kj} R(\varphi_{ki} - \varphi_{kj}), \tag{32}$$

which is $\delta_{ij} I$ due to the unitarity of U. Therefore, M is orthogonal in this case also.

Since M is an orthogonal matrix, it satisfies the One-to-One and Metric Invariance postulates. The invariance of the p_i' required by the Gauge Invariance postulate follows from the observation made previously that, under the addition of ϕ_0 to the ϕ_i in v, the transformed vectors $\mathsf{v}' = \mathsf{U}\mathsf{v}$ or $\mathsf{v}' = \mathsf{U}\mathsf{K}\mathsf{v}$ yield unchanged values of p_i'.

Hence, any unitary or antiunitary transformation satisfies the One-to-One, Metric Invariance, and Gauge Invariance postulates. In particular, we have obtained the result of Wigner's theorem that a one-to-one map over state space that represents a symmetry transformation of a system is either unitary or antiunitary.

Step 3: Physical transformations. By the Continuity postulate, a physical transformation (such as a reflection-rotation of a frame of reference) that depends continuously upon a real-valued parameter n-tuple $\boldsymbol{\pi}$ is represented by a map $\mathcal{M}_{\boldsymbol{\pi}}$, which depends continuously upon $\boldsymbol{\pi}$. From the above discussion, the set of mappings that represent physical transformations are the set of all unitary and antiunitary transformations. The set of all unitary transformations and the set of all antiunitary transformations are disjoint and it is not possible to continuously transform a unitary transformation into any antiunitary transformation. Therefore, $\mathcal{M}_{\boldsymbol{\pi}}$ is represented either by unitary transformations or by antiunitary transformations.

Furthermore, by the Continuous Transformations postulate, a *continuous* physical transformation that depends continuously upon a real-valued parameter n-tuple $\boldsymbol{\pi}$ is represented by a map $\mathcal{M}_{\boldsymbol{\pi}}$, which reduces to the identity map for some value of $\boldsymbol{\pi}$. However, only the set of unitary transformations include the identity. Therefore, a continuous physical transformation can only be represented by unitary transformations. In particular, since temporal evolution of a system is a continuous physical transformation, it must be represented by unitary transformations.

C. Representation of measurements

Suppose that a system undergoes measurement $\mathbf{A}$ and yields outcome j. By the assumption of reproducibility (Sec. II A), after $\mathbf{A}$ has been performed and outcome j obtained, immediate repetition yields the same outcome with certainty.

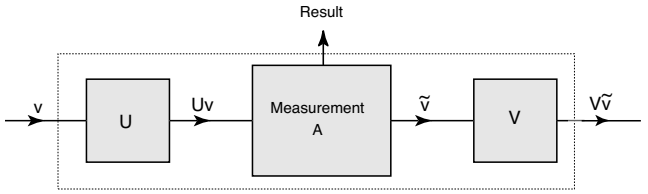

FIG. 3. Simulation of measurement $\mathbf{A}'$. A unitary transformation, U, transforms the input state, v, into Uv. Measurement $\mathbf{A}$ is performed on this state, and yields an outcome and the outgoing state, $\tilde{\mathsf{v}}$, which is then transformed by the unitary transformation V into the output state $\mathsf{V}\tilde{\mathsf{v}}$.

Therefore, for every outcome j there exists a corresponding state, v_j, such that the measurement $\mathbf{A}$ upon the system in state v_j yields outcome j with certainty. From Eq. (14), since $p_j=1$ and all the other p_j are zero, we have that

$$\mathsf{v}_j = (0, \ldots, e^{i\phi_j}, \ldots, 0)^{\mathsf{T}}, \tag{33}$$

where ϕ_j is undetermined.

By the Measurement Simulability postulate, measurement $\mathbf{A}'$ can be simulated by an arrangement consisting of a measurement $\mathbf{A}$ followed immediately before and after by suitable interactions. These interactions bring about continuous transformations of the system. From the results of the previous section, these interactions must, therefore, be represented by unitary transformations, which we shall denote U and V, respectively (see Fig. 3). In the following, we shall establish the form of these matrices, and then obtain an expression for the outcome probabilities for measurement $\mathbf{A}'$ performed on a system in state v.

First, from the Outcomes postulate and the assumption of reproducibility, there exist N states $\mathsf{v}'_1, \mathsf{v}'_2, \ldots, \mathsf{v}'_N$ such that measurement $\mathbf{A}'$ performed on a system in state v'_i yields outcome i with certainty. Hence, the arrangement in Fig. 3 must be such that $\mathbf{A}$ yields outcome i with certainty when the input state to the arrangement is v'_i. For this to be the case, U must transform v'_i to a state of the form $\mathsf{v}_i e^{i\xi_i}$, where ξ_i is arbitrary. That is, the matrix U must satisfy the relations

$$\mathsf{U}\mathsf{v}'_i = \mathsf{v}_i e^{i\xi_i}, \quad i = 1,2, \ldots, N. \tag{34}$$

Second, if outcome i is obtained from the arrangement, the output state of the arrangement must be of the form $\mathsf{v}'_i e^{i\xi'_i}$, where ξ'_i is arbitrary. But, immediately after measurement $\mathbf{A}$, the system is in state v_i up to an overall phase. Hence, the matrix V must satisfy the relations

$$\mathsf{V}\mathsf{v}_i = \mathsf{v}'_i e^{i\xi'_i} \quad i = 1,2, \ldots, N. \tag{35}$$

From Eq. (33), the v_i form an orthonormal basis for $\mathbb{C}^N$. From Eq. (34), $\mathsf{v}'_i = \mathsf{U}^\dagger \mathsf{v}_i e^{i\xi_i}$, which, since U is unitary, implies that the v'_i also form an orthonormal basis. Therefore, any state, v, can be expanded as $\Sigma_i c'_i \mathsf{v}'_i$, with $c'_i \in \mathbb{C}$, and the matrices U and V are determined by the relations in Eqs. (34) and (35) up to the ξ_i and the ξ'_i.

It is now possible to determine the measurement probabilities if a system in state v undergoes measurement $\mathbf{A}'$. Using Eq. (34) and the expansion $\mathsf{v}=\Sigma_i c'_i \mathsf{v}'_i$, the first interaction of the arrangement transforms v into

$$\mathsf{U}\left(\sum_i c'_i \mathsf{v}'_i\right) = \sum_i c'_i \mathsf{v}_i e^{i\xi_i}. \tag{36}$$

Using Eq. (15), the probability that measurement $\mathbf{A}$ in the arrangement yields outcome i is therefore $|c'_i e^{i\xi_i}|^2 = |c'_i|^2$. Hence, measurement $\mathbf{A}'$ performed on the state v yields outcome i with probability $|c'_i|^2$. If outcome i is obtained, the outgoing state of measurement $\mathbf{A}$ is v_i, so that the output state of the arrangement is $\mathsf{V}\mathsf{v}_i$ which is v'_i up to an overall phase.

In summary, every measurement, $\mathbf{A}' \in A$, has an associated orthonormal basis, $\{\mathsf{v}'_1, \mathsf{v}'_2, \ldots, \mathsf{v}'_N\}$. If measurement $\mathbf{A}'$ is performed upon a system in state v, the probability, p'_i, of obtaining outcome i and corresponding output state v'_i is $|c'_i|^2$, where c'_i is determined by the relation $\mathsf{v}=\Sigma_i c'_i \mathsf{v}'_i$, which is the Born rule.

Expected values

If the ith outcome of measurement $\mathbf{A}'$ has an associated real value a'_i, the expected value obtained in an experiment in which a system in state v undergoes measurement $\mathbf{A}'$ is defined as

$$\langle \mathsf{A}' \rangle = \sum_i a'_i p'_i. \tag{37}$$

Since $p'_i = |c'_i|^2$ and $c'_i = \mathsf{v}'^{\dagger}_i \mathsf{v}$, this expression can be also written as

$$\langle \mathsf{A}' \rangle = \sum_i \mathsf{v}^\dagger (\mathsf{v}'_i a'_i \mathsf{v}'^{\dagger}_i) \mathsf{v} = \mathsf{v}^\dagger \left(\sum_i \mathsf{v}'_i a'_i \mathsf{v}'^{\dagger}_i \right) \mathsf{v} = \mathsf{v}^\dagger \mathsf{A}' \mathsf{v}, \tag{38}$$

where the matrix $\mathsf{A}' \equiv \Sigma_i \mathsf{v}'_i a'_i \mathsf{v}'^{\dagger}_i$ is Hermitian since the a'_i are real.

D. Composite systems

1. Tensor product rule

In order to derive the tensor product rule, we shall apply the Composite Systems postulate to the case of a composite system with two subsystems with abstract models $\mathbf{q}(N)$ and $\mathbf{q}(N')$, respectively, where the composite system has the abstract model $\mathbf{q}(N'')$.

Suppose that the subsystems are in states represented by $(p_i; \chi_i)$ and $(p'_j; \chi'_j)$, respectively, and the state of the composite system is represented by $(p''_l; \chi''_l)$ $(l=1, \ldots, N'')$ in the manner defined by the Composite Systems postulate. If we write the states of the subsystems in complex form,

$$\mathsf{v}^{(1)} = (\sqrt{p_1} e^{i\phi_1}, \sqrt{p_2} e^{i\phi_2}, \ldots, \sqrt{p_N} e^{i\phi_N})^{\mathsf{T}}$$

and

$$\mathsf{v}^{(2)} = (\sqrt{p'_1} e^{i\phi'_1}, \sqrt{p'_2} e^{i\phi'_2}, \ldots, \sqrt{p'_{N'}} e^{i\phi'_{N'}})^{\mathsf{T}},$$

respectively, where $\phi_i = a\chi_i + b$ and $\phi'_j = a\chi'_j + b$, and, similarly, we write the state of the composite system as

$$\mathsf{v} = (\sqrt{p''_1} e^{i\phi''_1}, \sqrt{p''_2} e^{i\phi''_2}, \ldots, \sqrt{p''_{N''}} e^{i\phi''_{N''}})^{\mathsf{T}},$$

where $\phi''_l = a\chi''_l + b$, then it follows from the Composite Systems postulate that v can simply be written as $\mathsf{v}^{(1)} \otimes \mathsf{v}^{(2)}$.

More generally, consider a composite system with d subsystems, numbered $1,2,\ldots,d$, in states $\mathsf{v}^{(1)},\mathsf{v}^{(2)},\ldots,\mathsf{v}^{(d)}$, respectively. We can regard subsystems 1 and 2 as comprising a bipartite composite system, system $1'$, which, according to the above result, is in state $\mathsf{v}^{(1)}\otimes\mathsf{v}^{(2)}$. Next, we can regard system $1'$ and subsystem 3 as comprising a bipartite composite system, system $2'$, which is therefore in state $(\mathsf{v}^{(1)}\otimes\mathsf{v}^{(2)})\otimes\mathsf{v}^{(3)}$. Continuing in this way, we can see that the state of the composite system with d subsystems has the state $\mathsf{v}=\mathsf{v}^{(1)}\otimes\mathsf{v}^{(2)}\otimes\ldots\otimes\mathsf{v}^{(d)}$.

2. Representation of subsystem measurements

Suppose that measurement $\mathbf{A}^{(1)}\in\mathcal{A}^{(1)}$, represented by N-dimensional Hermitian operator $\mathsf{A}^{(1)}$, with eigenstates $\mathsf{v}_i^{(1)}$ and eigenvalues a_i, respectively, is performed on subsystem 1 of dimension N of a bipartite composite system of dimension $N''=NN'$, where N' is the dimension of subsystem 2. With respect to the abstract quantum model $\mathbf{q}(N'')$ of the composite system, measurement $\mathbf{A}^{(1)}$ is *not* in the measurement set $\mathcal{A}$ of the composite system. However, it is convenient to be able to describe measurement $\mathbf{A}^{(1)}$ as an N''-dimensional operator A, in the framework of $\mathbf{q}(N'')$.

To determine the form of A, it is sufficient to consider the effect of A on product states of the form $\mathsf{v}_i^{(1)}\otimes\mathsf{v}^{(2)}$ of the composite system, where $\mathsf{A}^{(1)}\mathsf{v}_i^{(1)}=a_i\mathsf{v}_i^{(1)}$. If the composite system is in such a state, then subsystem 1 is in state $\mathsf{v}_i^{(1)}$. Therefore, when measurement $\mathbf{A}^{(1)}$ is performed, result a_i is obtained with certainty, and the state of subsystem 1 is unchanged (up to an irrelevant overall phase). Therefore, the state of the composite system remains unchanged. If we require that A has eigenvectors $\mathsf{v}_i^{(1)}\otimes\mathsf{v}^{(2)}$, with respective eigenvalues a_i, it follows that A can be taken to be $\mathsf{A}^{(1)}\otimes\mathsf{I}^{(2)}$, where $\mathsf{I}^{(2)}$ is the identity matrix in the model of subsystem 2, with the only freedom being a physically irrelevant overall phase in each of the eigenstates of $\mathsf{A}^{(1)}$. This result generalizes trivially to the case of a measurement performed on one or more subsystems of a composite system consisting of d subsystems. We shall refer to such a measurement as a *subsystem measurement*.

IV. TEMPORAL EVOLUTION AND THE CORRESPONDENCE RULES OF QUANTUM THEORY

We begin by stating the average-value correspondence principle (AVCP) and then using it to derive several *generalized operator rules* that connect the average values of operators at different times. The commonly used operator rules of quantum theory follow as a special case. Using the AVCP and the Temporal Evolution postulate, we then derive the explicit form of the temporal evolution operator. Finally, we illustrate how the AVCP can be used to derive commutation relations by deriving $[\mathsf{x},\mathsf{p}_x]=i\hbar$.

A. Average-value correspondence

Suppose that, in a classical experiment, a measurement of observable A on a physical system can be implemented by an arrangement where a number of other measurements of observables $A',A'',\ldots$ are performed on the same system (perhaps performed at a different time to the measurement of A), such that the outcome value of the measurement of A can be calculated as a function of the outcome values of measurements of $A',A'',\ldots$. For example, a measurement of the position of a particle of mass m at time $t+\delta t$ can be implemented by measuring its position, x, and momentum, p_x, at time t, and computing the function $x+p_x\delta t/m$. The AVCP, stated below, then asserts that, in this case, providing certain conditions are met, a corresponding quantum experiment can be constructed where quantum measurements $\mathbf{A},\mathbf{A}',\mathbf{A}'',\ldots$ (which may be subsystem measurements), corresponding to classical measurements of $A,A',A'',\ldots$, are performed on copies of the physical system, such that the same functional relation holds *on average* between the outcome values of the quantum measurements, the average being taken over infinitely many trials of the quantum experiment. See Sec. V for discussion of the origin of the principle.

Average-value correspondence principle

Consider a classical idealized experiment in which a system (possibly a composite system) is prepared in some state at time t_0, and is allowed to evolve in a given background (see the example in Fig. 4). Suppose that a measurement of $A^{(m)}(m\geqslant 2)$, performed on the system at time t_2 with value $a^{(m)}$, can be implemented by an arrangement where measurements of $A^{(1)},A^{(2)},\ldots,A^{(m-1)}$ are performed upon one copy of the system at time t_1, and the values of their respective outcomes, denoted $a^{(1)},a^{(2)},\ldots,a^{(m-1)}$, are then used to compute the output $f(a^{(1)},a^{(2)},\ldots,a^{(m-1)})$, where f is an analytic function, so that the relation

$$a^{(m)}=f(a^{(1)},a^{(2)},\ldots,a^{(m-1)}) \qquad (*)$$

holds for all initial (classical) states of the system.

Consider the case in which the quantum measurements $\mathbf{A}^{(1)},\mathbf{A}^{(2)},\ldots,\mathbf{A}^{(m)}$ (which may be subsystem measurements), with operators $\mathsf{A}^{(1)},\mathsf{A}^{(2)},\ldots,\mathsf{A}^{(m)}$, represent the measurements of $A^{(1)},A^{(2)},\ldots,A^{(m)}$, respectively. Then, consider the following idealized quantum experimental arrangement consisting of several setups, each consisting of identical sources and backgrounds, where, in each setup, a copy of the system is prepared in the same initial state, v_0, at time t_0.

In one setup, only measurement $\mathbf{A}^{(m)}$ is performed (at time t_2) and, for any i,j with $i\neq j$ and $i,j\leqslant m-1$, the measurements $\mathbf{A}^{(i)},\mathbf{A}^{(j)}$ are performed (at time t_1) in (i) the *same* setup if $[\mathsf{A}^{(i)},\mathsf{A}^{(j)}]=0$, and (ii) *different* setups if $[\mathsf{A}^{(i)},\mathsf{A}^{(j)}]\neq 0$. Let the values of the outcomes of the measurements $\mathbf{A}^{(1)},\ldots,\mathbf{A}^{(m)}$ in any given run of the experimental arrangement be denoted $a^{(1)},\ldots,a^{(m)}$, respectively. The function $f(a^{(1)},a^{(2)},\ldots,a^{(m-1)})$ is defined as *simple* provided that its polynomial expansion contains no terms involving a product of eigenvalues belonging to measurements whose operators do not commute. If f is simple, then $(*)$ holds on average, the average being taken over an infinitely large number of runs of the experiment.

B. Generalized operator rules

We will now apply the AVCP to derive operator relations that hold when the function f takes various useful forms. In

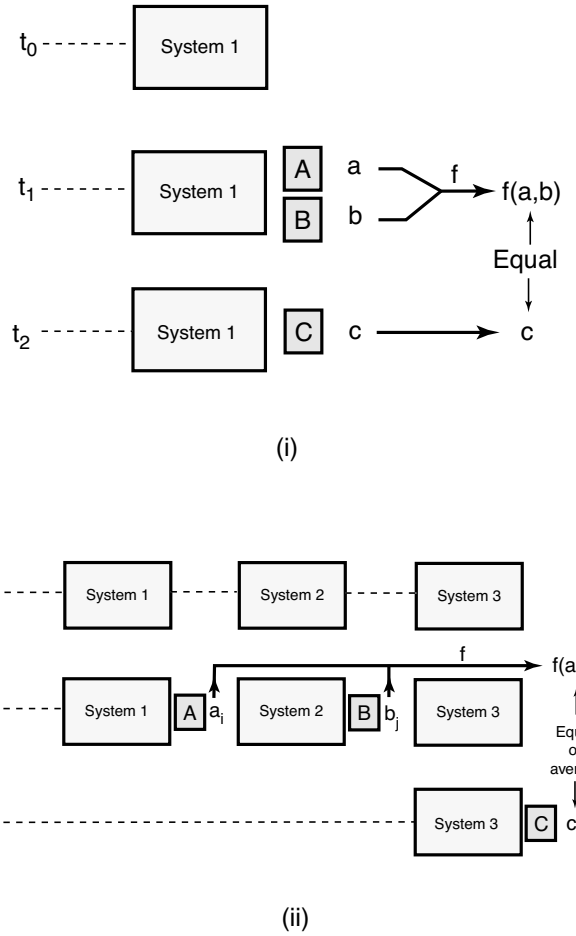

FIG. 4. An example of the application of the AVCP. (i) A classical experiment showing the measurements of A, B, and C performed at times t_A, t_B, and t_C, respectively, with values denoted as a, b, and c, respectively. Here, $t_A=t_B=t_1$ and $t_C=t_2$. Suppose that one finds that the relation $c=f(a,b)$ holds for all initial states of the system. (ii) The corresponding quantum experiment. Three copies of the system are prepared in the same initial state, v_0, at time t_0, and are placed in identical backgrounds. In this example, it is assumed that the operators A and B do not commute. Hence, by the AVCP, measurements **A** and **B** are performed on different copies of the system. Measurement **C** is performed on a separate copy of the system. In any given run of the experiment, the probabilities that measurements **A**, **B**, and **C** yield values a_i, b_j, and c_k $(i,j,k=1,\ldots,N)$ are p_i, p'_j, and p''_k, respectively. The AVCP then asserts that, provided the polynomial expansion of $f(a,b)$ contains no product terms involving a and b, the relation $\bar{c}=f(\bar{a},\bar{b})$ holds for all initial states, v_0, of the system, where the average is taken over an infinite number of runs of the experiment; that is, $\Sigma_k c_k p''_k = \Sigma_{ij} f(a_i,b_j) p_i p'_j$ for all $\mathbf{v}_0$.

each instance of f, we shall first derive a *generalized operator rule* that relates the expected values of the relevant operators at *different times*. Then, taking the special case in which the expected values are computed at the same time, we obtain the corresponding operator rule, which relates the operators themselves.

For simplicity, we shall consider the simple case of a classical experiment where a system is subject to measurements of A and B at time t_1, and to a measurement of C at time t_2. The generalization to experiments involving more than two measurements at t_1 is trivial. We shall suppose that a measurement of C, with value c, can be implemented by an arrangement in which the measurements of A and B are performed, with respective values a and b, and the function $f(a,b)$ then computed, so that the relation

$$c = f(a,b) \tag{39}$$

holds for all initial states of the system.

In a quantum model of the appropriate experimental arrangement, let the operators that represent these measurements be denoted A, B, and C, respectively. To simplify the presentation, we shall only consider the case in which these operators have finite dimension, N; the results obtained below can be readily shown to apply in the infinite-dimensional case. Let the elements of orthonormal sets of eigenvectors of A, B, and C be denoted v_i, v'_j, and v''_k, respectively $(i,j,k=1,2,\ldots,N)$, let the corresponding eigenvalues be denoted a_i, b_j, and c_k, and let the probabilities of the ith, jth, and kth outcomes of measurements **A**, **B**, and **C** in any given experimental arrangement be denoted by p_i, p'_j, and p''_k, respectively.

Case 1. f is a function of a only

In this case, according to the AVCP, the quantum experiment simply consists of two setups, involving two copies of the system, where **A** is performed on one copy at time t_1 and **C** on the other copy at time t_2. Consequently, we obtain the relation

$$\langle \mathsf{C} \rangle_{t_2} = \langle f(\mathsf{A}) \rangle_{t_1}, \tag{40}$$

which holds for all v_0. We can summarize the above result in the form of the *generalized function rule*,

$$c(t_2) = f(a(t_1)) \mapsto \langle \mathsf{C} \rangle_{t_2} = \langle f(\mathsf{A}) \rangle_{t_1} \quad \forall\ \mathsf{v}_0, \tag{41}$$

where, for clarity, the times at which the outcomes are obtained has been explicitly indicated. In the special case in which $t=t_1=t_2$, we obtain the usual operator rule, the *function rule*,

$$c = f(a) \mapsto \mathsf{C} = f(\mathsf{A}). \tag{42}$$

Case 2. $f(a,b)=f_1(a)+f_2(b)$

In this case, by straightforward application of the AVCP, one finds that, whether or not the operators A, B commute, one obtains the *generalized sum rule*,

$$\begin{aligned} c(t_2) &= f_1(a(t_1)) + f_2(b(t_1)) \mapsto \langle \mathsf{C} \rangle_{t_2} \\ &= \langle f_1(\mathsf{A}) \rangle_{t_1} + \langle f_2(\mathsf{B}) \rangle_{t_1} \quad \forall\ \mathsf{v}_0. \end{aligned} \tag{43}$$

In the special case in which $t=t_1=t_2$, we obtain the *sum rule*,

$$c = f_1(a) + f_2(b) \mapsto \mathsf{C} = f_1(\mathsf{A}) + f_2(\mathsf{B}). \tag{44}$$

Case 3. $f(a,b)=f_1(a)f_2(b)$

In this case, application of the AVCP yields the *generalized product rule*,

$$c(t_2) = f_1(a(t_1))f_2(b(t_1)) \mapsto \langle \mathsf{C} \rangle_{t_2}$$
$$= \langle f_1(\mathsf{A}) f_2(\mathsf{B}) \rangle_{t_1} \quad \forall\ \mathsf{v}_0 \quad \text{if } [\mathsf{A},\mathsf{B}] = 0. \tag{45}$$

In the special case in which $t_1 = t_2$, we obtain the *product rule*,

$$c = f_1(a)f_2(b) \mapsto \mathsf{C} = f_1(\mathsf{A})f_2(\mathsf{B}) \quad \text{if } [\mathsf{A},\mathsf{B}] = 0. \tag{46}$$

Note that there is no rule in the case in which A and B do not commute.

C. Temporal evolution

In this section, we will use the AVCP, together with the Temporal Evolution postulate, to derive the explicit form of the temporal evolution operator for a system in a time-dependent background.

From results derived previously, over the course of the interval $[t, t+\Delta t]$, the state $\mathsf{v}(t)$ evolves as

$$\mathsf{v}(t + \Delta t) = \mathsf{U}_t(\Delta t)\mathsf{v}(t), \tag{47}$$

where $\mathsf{U}_t(\Delta t)$ is the unitary matrix that represents temporal evolution of the system during $[t, t+\Delta t]$. Suppose now that the background of the system is time-independent during this interval. Then we shall write $\mathsf{U}_t(\Delta t)$ as $\mathsf{V}_t(\Delta t)$. Now, for $0 \leqslant \Delta t_1 + \Delta t_2 \leqslant \Delta t$, and $\Delta t_1, \Delta t_2$ both positive, we have

$$\mathsf{V}_t(\Delta t_1 + \Delta t_2) = \mathsf{V}_{t+\Delta t_1}(\Delta t_2)\mathsf{V}_t(\Delta t_1). \tag{48}$$

But the time independence of the background implies that $\mathsf{V}_{t+\Delta t_1}(\Delta t_2) = \mathsf{V}_t(\Delta t_2)$. Therefore,

$$\mathsf{V}_t(\Delta t_1 + \Delta t_2) = \mathsf{V}_t(\Delta t_2)\mathsf{V}_t(\Delta t_1), \tag{49}$$

which can be solved to yield

$$\mathsf{V}_t(\Delta t_1) = \exp(-i\mathsf{K}_t \Delta t_1), \tag{50}$$

where K_t is a Hermitian matrix.

To determine the nature of K_t, we proceed as follows. In the classical model of a physical system in a time-independent interval $[t, t+\Delta t]$, the classical Hamiltonian, H, for the system is not explicitly dependent upon time during this interval. On the assumption that H is a simple function of the observables of the system, and that the measurements of these observables are represented by quantum measurements, it follows from the AVCP that the corresponding Hamiltonian operator is also not explicitly dependent upon time during this interval. More precisely, suppose that $H = f(A_1, \ldots, A_M)$ is a simple function, f, of the observables $A_1, \ldots, A_M$, and is not explicitly a function of time, t, during the interval $[t, t+\Delta t]$. If the operators $\mathsf{A}_1, \ldots, \mathsf{A}_M$ represent measurements of the observables $A_1, \ldots, A_M$, respectively, then, by the AVCP, there exists an operator $\mathsf{H} = f(\mathsf{A}_1, \ldots, \mathsf{A}_M)$, which represents the classical Hamiltonian, H. Since t does not appear explicitly in H during the above-mentioned interval, it cannot appear explicitly in H during this interval. Therefore, in particular,

$$\mathsf{H}_t = \mathsf{H}_{t+\Delta t}, \tag{51}$$

where H_t denotes the Hamiltonian operator at time t. In addition, in the classical model, the total energy of the system is constant during this interval for all states of the system. Therefore, by the generalized function rule, the relation

$$\langle \mathsf{H}_t \rangle_t = \langle \mathsf{H}_{t+\Delta t} \rangle_{t+\Delta t} \tag{52}$$

holds for any state v. Hence, from Eqs. (51) and (52), it follows that

$$\langle \mathsf{H}_t \rangle_t = \langle \mathsf{H}_t \rangle_{t+\Delta t} \tag{53}$$

for all v. But

$$\langle \mathsf{H}_t \rangle_{t+\Delta t} = \langle \mathsf{H}_t \rangle_t + i\langle [\mathsf{K}_t, \mathsf{H}_t] \rangle_t \Delta t + \mathsf{O}(\Delta t^2). \tag{54}$$

Therefore, the commutator $[\mathsf{K}_t, \mathsf{H}_t] = 0$, which implies that there exist N mutually orthogonal eigenvectors, $\mathsf{v}_1, \ldots, \mathsf{v}_N$, which K_t and H_t share in common. In particular, the state $\mathsf{v}_j (j = 1, 2, \ldots, N)$ is an eigenvector of H_t, with some eigenvalue E_j.

Now, if the system is in an eigenstate, v_j, with eigenvalue k_j, of K_t, at time t, the state evolves as

$$\mathsf{v}(t + \Delta t) = e^{-ik_j \Delta t}\mathsf{v}(t). \tag{55}$$

Therefore, during the interval $[t, t+\Delta t]$, this state remains an eigenstate of H_t, and is therefore a state of constant energy, E_j, during this interval. In addition, since evolution only affects the overall phase of the state, the observable degrees of freedom of the state are time-independent during this interval. But, by the Temporal Evolution postulate, recalling that $\phi_i = a\chi_i + b$, the state $\mathsf{v}(t)$, representing a system in a time-independent background of definite energy, E_j, whose observable degrees of freedom are time-independent, evolves as

$$\mathsf{v}(t + \Delta t) = e^{-iE_j \Delta t/\alpha'}\mathsf{v}(t), \tag{56}$$

where $\alpha' = \alpha/a$. By comparison of Eqs. (55) and (56), we find that $k_j = E_j/\alpha'$ holds for $j = 1, 2, \ldots, N$, which implies that $\mathsf{K}_t = \mathsf{H}_t/\alpha'$. Hence, in a time-independent background, any state v evolves as

$$\mathsf{v}(t + \Delta t) = \exp(-i\mathsf{H}_t \Delta t/\alpha')\mathsf{v}(t). \tag{57}$$

In order to generalize to the case of temporal evolution in a time-dependent background, we split the interval $[t, t+\Delta t]$ into intervals of duration ϵ, approximate the evolution during each of these intervals assuming that the background is time-independent, and then take the limit as $\epsilon \to 0$,

$$\mathsf{U}_t(\Delta t) = \mathsf{U}_{t+\Delta t-\epsilon}(\epsilon) \cdots \mathsf{U}_{t+\epsilon}(\epsilon)\mathsf{U}_t(\epsilon)$$
$$= \lim_{\epsilon \to 0} \{\mathsf{V}_{t+\Delta t-\epsilon}(\epsilon) \cdots \mathsf{V}_{t+\epsilon}(\epsilon)\mathsf{V}_t(\epsilon)\}, \tag{58}$$

which, upon expansion, yields

$$\mathsf{U}_t(\Delta t) = I - \frac{i}{\alpha'}\mathsf{H}_t \Delta t + \mathsf{O}(\Delta t^2), \tag{59}$$

so that

$$i\alpha' \frac{d\mathsf{v}(t)}{dt} = \mathsf{H}_t \mathsf{v}(t). \tag{60}$$

The value of the constant α' is determined to be $\hbar$ by comparison with experiment.

D. Commutation relations

We have shown above that the operator rules (the function, sum, and product rules) follow as special cases of the generalized operator rules. Remarkably, these generalized operator rules also enable the derivation of other types of correspondence rule, such as the measurement commutation relations (including Dirac's Poisson-Bracket rule) and measurement-transformation commutation relations [24]. Here, taking the infinite-dimensional form of abstract quantum formalism as a given, we shall give one example, namely the derivation of the measurement commutation relation $[\mathsf{x},\mathsf{p}_x]=i\hbar$.

The position-momentum commutation relationship

We consider the special case of structureless and massless particle moving in the $+x$ direction, where measurements of the x component of position, the x component of momentum, and the energy are represented by the operators $\mathsf{x},\mathsf{p}_x,\mathsf{H}$, respectively.

First, to determine the relationship of H to the operators x and p_x, we make use of the fact that the relation $H=cp_x$, where c is the speed of light, holds for all classical states (x,p_x) of the system, so that, from the function rule, it follows that $\mathsf{H}=c\mathsf{p}_x$.

Next, to obtain a relation between x and p_x, we make use of the fact that, in the quantum model, the expected value of x at time $t+\delta t$ can be calculated in two separate ways. First, from the definition of $\langle x\rangle_t$, the relation

$$\begin{aligned}\langle \mathsf{x}\rangle_{t+\delta t} &= \langle \mathsf{U}_t^\dagger(\delta t)\mathsf{x}\mathsf{U}_t(\delta t)\rangle_t \\ &= \left\langle \left(1+\frac{i\mathsf{H}\delta t}{\hbar}\right)\mathsf{x}\left(1-\frac{i\mathsf{H}\delta t}{\hbar}\right)\right\rangle_t + O(\delta t^2) \\ &= \langle \mathsf{x}\rangle_t + \frac{i}{\hbar}\delta t\langle \mathsf{H}\mathsf{x}-\mathsf{x}\mathsf{H}\rangle_t + O(\delta t^2) \\ &= \langle \mathsf{x}\rangle_t - \frac{ic}{\hbar}\delta t\langle [\mathsf{x},\mathsf{p}_x]\rangle_t + O(\delta t^2) \end{aligned} \tag{61}$$

holds for all states, v, of the system. Second, using the generalized function rule, it follows from the classical relation $x(t+\delta t)=x(t)+c\,\delta t+O(\delta t^2)$ that the relation

$$\langle \mathsf{x}\rangle_{t+\delta t} = \langle \mathsf{x}\rangle_t + c\,\delta t + O(\delta t^2) \tag{62}$$

holds for all v.

Equating the above two expressions for $\langle x\rangle_{t+\delta t}$, we obtain

$$\mathsf{v}^\dagger[\mathsf{x},\mathsf{p}_x]\mathsf{v} = i\hbar \tag{63}$$

for all v, which implies that

$$[\mathsf{x},\mathsf{p}_x] = i\hbar. \tag{64}$$

V. PHYSICAL COMPREHENSIBILITY OF THE POSTULATES

A. Postulates based upon experimental facts

The Outcomes and Measurement Simulability postulates can be viewed as a direct generalization of experimental observations made on Stern-Gerlach experiments performed on silver atoms.

Consider an experiment in which Stern-Gerlach preparations and measurements are performed upon silver atoms, and where the set $\mathcal{A}$ consists of the elements $\mathbf{A}_{\theta,\phi}$ representing Stern-Gerlach measurements in the direction (θ,ϕ). In this experimental setup, which is closed in the sense defined earlier, we find that each measurement yields one of two possible outcomes. The Outcomes postulate generalizes this finding by asserting that, when performed on a given system, all measurements in a measurement set yield the same number, N, of possible outcomes.

If, in the above experiment, an interaction consisting of a uniform magnetic field acts between the preparation and measurement, one finds that the probabilities of the measurement outcomes are the same as those obtained if a different Stern-Gerlach measurement is performed with the interaction absent. Using this observation, one finds that it is possible to simulate measurement $\mathbf{A}_{\theta,\phi}$ using any given measurement $\tilde{\mathbf{A}}\in\mathcal{A}$ if followed immediately before and after by suitable interactions. The simulation behaves precisely as $\mathbf{A}_{\theta,\phi}$ insofar as the probabilities of the measurement outcomes and the corresponding output states are concerned. The Measurement Simulability postulate can be regarded as a direct generalization of this observation.

B. Postulates drawn from classical physics

The Mappings and One-to-One postulates correspond to the classical assumptions that physical transformations are represented by mappings over state space, and that these mappings are one-to-one. The Continuity and Continuous Transformations postulates are identical in form to the corresponding classical assumptions.

Postulates obtained through classical-quantum correspondence

A general guiding principle in building up a quantum model of a physical system is that, in an appropriate limit, the predictions of the quantum model of the system stand in some one-to-one correspondence with those of a classical model of the system. The key idea we shall employ here is that a quantum model of a particle corresponds, in the appropriate classical limit, to the Hamilton-Jacobi ensemble model of the particle.

In the Hamilton-Jacobi model of particle of mass m moving along the x axis, the state of the ensemble is given by $(P(x,t),S(x,t))$, where $P(x,t)$ is the probability density function over position, and $S(x,t)$ is the action function, and the governing equations are

$$\begin{aligned}&\frac{\partial P}{\partial t}+\frac{\partial}{\partial x}\left(P\frac{1}{m}\frac{\partial S}{\partial x}\right)=0,\\ &\frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2+V(x,t)=-\frac{\partial S}{\partial t},\end{aligned} \tag{65}$$

where $V(x,t)$ is the potential. The observables are the probability density function, $P(x,t)$, the local momentum, $\partial S/\partial x$, and the total energy, $-\partial S/\partial t$ (see [25], for example). We con-

sider coarse position measurements, and discretize the Hamilton-Jacobi state $(P(x,t),S(x,t))$, as $(p_i^{(\mathrm{CM})};S_i)$, where $p_i^{(\mathrm{CM})}$ is the probability that the position measurement yields outcome i $(i=1,\ldots,N)$, and S_i is the action associated with position i. Therefore, on the assumption that the quantum state of the particle can be put into one-to-one correspondence with the classical state in the classical limit, it follows that the quantum state of the particle must be expressible in the form $(p_i;\chi_i)$, where, again, the p_i are the probabilities of the position measurement, and where the χ_i are N real degrees of freedom. In the limit, we assume $p_i=p_i^{(\mathrm{CM})}$ and $S_i=\alpha\chi_i$, where α has the dimensions of action.

Now, on the basis of empirical observations, it is reasonable to infer that, in the limit of infinite precision, a coarse position preparation is complete (in the sense defined in Sec. II) with respect to a coarse position measurement. Therefore, a direct generalization of the above observation is that the state of a system described by the abstract model $\mathbf{q}(N)$ with respect to some measurement $\mathbf{A}$ can be represented by $(\vec{p},\vec{\chi})$, where $\vec{p}=(p_1,\ldots,p_N)$ are the outcome probabilities of measurement $\mathbf{A}$. As will be discussed below, this assumption provides the motivation for the State Representation postulate.

Using this correspondence, we can transpose a number of the elementary features of the Hamilton-Jacobi model across to the quantum model of the particle, which leads to several of the postulates.

Global gauge invariance

In the continuum Hamilton-Jacobi model, the observables associated with $S(x,t)$ for a system in state $(P(x,t),S(x,t))$ are $\partial S/\partial x$ and $-\partial S/\partial t$. Furthermore, the Hamilton-Jacobi equations only depend upon $S(x,t)$ through these observables. Hence, the transformation $S(x,t)\rightarrow S(x,t)+S_0$ gives rise to no change in the values of the observables of the system at time t or any other time, and is therefore a global gauge transformation of the model. Therefore, the discretized form of the model has a global gauge invariance property, namely that, for a system with state $(p_i^{(\mathrm{CM})};S_i)$, the transformation $S_i\rightarrow S_i+S_0$ for $i=1,\ldots,N$ and for any S_0 is a global gauge transformation, thus leaving invariant all physical predictions made on the basis of the state.

From this property of the Hamilton-Jacobi model, using the above classical-quantum correspondence, we assume, in the quantum model of a particle, and, even more generally for the abstract quantum model $\mathbf{q}(N)$, that the transformation

$$(p_i;\chi_i)\rightarrow(p_i;\chi_i+\chi_0), \tag{66}$$

where $\chi_0\in\mathbb{R}$, is also a gauge transformation. From this assumption, we now shall draw two postulates.

Postulate 3.6: Gauge Invariance

First, we note that, as a direct result of this global gauge invariance assumption, it follows that a transformation (representing passive or active physical transformation of the system) of the state (p_i,χ_i) to the state $(p_i';\chi_i')$ is such that the p_i' are unchanged if an arbitrary real constant, χ_0, is added to each of the χ_i. This is the content of the Gauge Invariance postulate, which may be regarded as a specific example of the global gauge invariance property.

Postulate 2.5: Measure Invariance

Second, we impose the requirement that the measure (or, in the language of Bayesian probability theory, the *prior*) over $(p_1,\ldots,p_N,\chi_1,\ldots,\chi_N)$ induced by the metric over state space (which metric arises from the Information Metric postulate) is compatible with the global gauge invariance property, and therefore satisfies the relation

$$\mu(p_i;\chi_i)=\mu(p_i;\chi_i+\chi_0) \tag{67}$$

for any χ_0, which is the content of the Measure Invariance postulate.

The requirement of the consistency of the measure with the global gauge invariance property can be understood as follows. Suppose that one is performing Bayesian inference on the quantum system, and one uses the measure as one's prior over state space. If one's knowledge about the quantum system includes the fact that it has a global gauge invariance property, then the prior over the p_i and χ_i that one employs should reflect this fact. Otherwise, one's inference will sometimes lead to predictions that are not consistent with the global gauge invariance property.

Postulate 3.7: Temporal Evolution

Consider the special case of a system in a time-independent background whose observable degrees of freedom are time-independent. According to the Hamilton-Jacobi equations, the state of such a system evolves in time Δt as

$$(P(x,t+\Delta t),S(x,t+\Delta t))=(P(x,t),S(x,t)-E\Delta t),$$

where E is the energy of the ensemble. That is, the temporal rate of change of the unobservable degree of freedom encodes the total energy of the system.

Using the above classical-quantum correspondence, we assume that the quantum model of a particle in a time-independent background that is in a state of definite energy with time-independent observable degrees of freedom evolves as $(p_i;\chi_i)\rightarrow(p_i;\chi_i-E/\alpha\Delta t)$ during the interval $[t,t+\Delta t]$. The Temporal postulate directly transposes this assumption to the quantum model $\mathbf{q}(N)$.

Postulate 4: Composite Systems

Consider a composite system, described in the model $\mathbf{q}(N'')$, consisting of two subsystems that are known to be in states represented by $(p_i;\chi_i)$ and $(p_j';\chi_j')$, respectively, with $i=1,\ldots,N$ and $j=1,\ldots,N'$, and $N''=NN'$. The composite system is in a state represented by (p_l'',χ_l''), where $l=N'(i-1)+j$, and we assume that

$$p_l''=p_ip_j', \tag{68a}$$

$$\chi_l''=g(\chi_i,\chi_j'), \tag{68b}$$

where g is a function, symmetric in its arguments, to be determined.

Suppose that the first subsystem undergoes the gauge transformation $(p_i;\chi_i)\rightarrow(p_i;\chi_i+\chi_0)$. We require that this transformation leads to a gauge transformation of the composite system, so that

$$(p_l'',\chi_l'') \rightarrow (p_l'',\chi_l''+h(\chi_0)), \tag{69}$$

where h is some function to be determined. Together with Eq. (68b), this implies that g is linear in its first argument. Applying the same argument to the second subsystem, one obtains that $g(\chi_i,\chi_j')=c\chi_i+d\chi_j'+e$. Imposing symmetry, and setting $e=0$ without loss of generality, we obtain $g(\chi_i,\chi_j')=c(\chi_i+\chi_j')$.

To determine c, we apply the Temporal Evolution postulate. We require that, if the energies of the subsystems are E and E', respectively, the energy of the composite system is $E+E'$. From the Temporal Evolution postulate, it follows at once that $c=1$. Hence, we obtain $g(\chi_i,\chi_j')=\chi_i+\chi_j$. Therefore, the state of the composite system $(p_l'',\chi_l'')=(p_ip_j';\chi_i+\chi_j')$, which is the content of the Composite Systems postulate.

Alternatively, one can obtain this result more directly from the Hamilton-Jacobi model. We note that, if, with respect to position measurements along the x and y axes, the discretized Hamilton-Jacobi state of a particle is $(p_i;S_i)$ and $(p_j';S_j')$, respectively, where $i=1,\ldots,N$ and $j=1,\ldots,N'$, then, with respect to (x,y)-position measurements, its state is $(p_l'';S_l'')=(p_ip_j';S_i+S_j')$, where $l=N'(i-1)+j$. By the classical-quantum correspondence, and then generalizing to the abstract quantum model, $\mathbf{q}(N)$, we obtain the same result.

C. Novel postulates

Postulate 2.1: Complementarity

According to the discussion of correspondence above, the state $S(t)$, written with respect to some measurement $\mathbf{A}\in\mathcal{A}$, can be represented by the pair $(\vec{p},\vec{\chi})$, where $\vec{p}$ contains the outcome probabilities of the measurement, and $\vec{\chi}$ is an ordered set of real-valued degrees of freedom. Hence, the state consists of a mixture of probabilities and degrees of freedom unconnected to probabilities, and measurement $\mathbf{A}$ yields information about the p_i but not the χ_i. The Complementarity postulate is motivated by the aesthetic desideratum that the quantum state, as far as possible, should consist of probabilities of events rather than being such a mixture, and aims to express the restriction on measurement $\mathbf{A}$ as a restriction on the ability of measurement $\mathbf{A}$ to completely resolve the events that occur when it is performed.

In particular, we hypothesize that, when measurement $\mathbf{A}$ is performed, there are, in fact, $2N$ possible events, with respective probabilities $P_1,\ldots,P_{2N}$, and that outcome i is observed whenever either event $2i-1$ or event $2i$ is realized. We note that similar assumptions have been made in a toy model of quantum theory in order to give concrete expression to complementarity [26]. In the Discussion, we provide a tentative explanation for the unobservability of the individual events.

Postulate 2.2: States

The States postulate asserts that the state of a system with respect to measurement $\mathbf{A}\in\mathcal{A}$ is given by $\mathbf{Q}=(Q_1,\ldots,Q_{2N})^{\mathsf{T}}$, with $Q_q\in[-1,1]$, where $P_q=Q_q^2$. Hence, in addition to the P_q, this postulate asserts that there is an additional, binary degree of freedom, σ_q, associated with each event q, where $\sigma_q=\mathrm{sgn}(Q_q)$, which is defined whenever $P_q\neq 0$.

The motivation for the introduction of the σ_q is the following. If one takes the $\mathbf{P}$ themselves as the state space of the system, one finds that nontrivial one-to-one transformations of the state space that preserve the metric over the state space (as we shall require in the Metric Invariance postulate, described below) are not possible. A simple way to allow the existence of such transformations is to take all $\mathbf{Q}$ as the state space of the system, for then there exist transformations, namely orthogonal transformations of the $\mathbf{Q}$, which preserve the metric over the $\mathbf{Q}$. Such an extension of the state space is an assumption that, although formally rather natural, presently awaits a clear physical basis.

Postulate 2.3: State Representation

The State Representation postulate connects together the hypothesis, expressed by the States postulate, that the state of a system is given by $\mathbf{Q}=(Q_1,\ldots,Q_{2N})$, and the above assertion that the state can be represented by $(\vec{p},\vec{\chi})$.

First, note that, since $p_i=Q_{2i-1}^2+Q_{2i}^2$, given $p_1,\ldots,p_N$, the state $\mathbf{Q}$ can, without loss of generality, be written as

$$Q_{2i-1}=\sqrt{p_i}\cos\theta_i,$$

$$Q_{2i}=\sqrt{p_i}\sin\theta_i, \tag{70}$$

where $\theta_1,\theta_2,\ldots,\theta_N$ are N real degrees of freedom.

The State Representation postulate now connects the θ_i together with the χ_i by asserting that $\theta_i=\theta(\chi_i)$, where $\theta(\cdot)$ is a differentiable function, not a constant, to be determined.

Postulate 2.4: Information Metric

The Information Metric postulate asserts that the metric over the space of probability distributions $\mathbf{P}$ is the information metric. The following argument shows explicitly how the metric arises from the concept of information.

Suppose that Alice has two coins, A and B, characterized by the probability distributions $\mathbf{p}=(p_1,p_2)$ and $\mathbf{p}'=(p_1',p_2')$, respectively, and suppose that she chooses coin A, tosses it n times, and then sends the data to Bob, without disclosing which coin she chose. If Bob knows $\mathbf{p}$ and $\mathbf{p}'$, how much information does the data provide him about which coin was tossed?

Using Bayes' theorem and Stirling's approximation for the case in which n is large, on the assumption that coins A and B are *a priori* equally likely to be chosen, one finds that $P_A/P_B=\exp[n\Sigma_{i=1}^2 p_i\ln(p_i/p_i')]$, where P_A is the probability that the tossed coin is A given the data, and likewise for P_B. When the probability distributions are close, so that $\mathbf{p}'=\mathbf{p}+d\mathbf{p}$, the argument of the exponent can be expanded in the dp_i to give $P_A/P_B=\exp(2nds^2)$, where $ds^2=\frac{1}{4}\Sigma_i dp_i^2/p_i$.

Now, the information gained by Bob, ΔI, is defined as $\Delta I\equiv U(1/2,1/2)-U(P_A,P_B)$, with U being an entropy (uncertainty) function such as the Shannon entropy. But, since

$P_A+P_B=1$ and since P_A/P_B is determined by ds, once U is selected, ΔI is determined by ds. This result immediately generalizes to the case in which $\mathbf{p}$ and $\mathbf{p}'$ are M-dimensional probability distributions ($M\geqslant 2$). Hence, from an informational viewpoint, it is natural to endow the space of discrete probability distributions with the information metric, $ds^2=\frac{1}{4}\Sigma_i dp_i^2/p_i$.

We remark that the information metric is the infinitesimal form of the statistical distance, $\mathrm{d}_S=\cos^{-1}(\Sigma_i\sqrt{p_i p_i'})$, between two probability distributions, $\mathbf{p}=(p_1,\ldots,p_N)$ and $\mathbf{p}'=(p_1',\ldots,p_N')$, introduced by Wootters [13] (see also the discussion in [27]).

Postulate 3.5: Metric Invariance

The Metric Invariance postulate accords the metric over state space a fundamental place in the theoretical framework: any transformation of state space is required to preserve the distance between any pair of nearby states.

The postulate is based on the intuitive physical idea that the distance (which is a measure of the degree of distinguishability) between any pair of nearby states is a quantity that is intrinsic to this pair of states, and is therefore the same irrespective of the measurement from whose perspective the states are observed. More precisely, suppose that there are two nearby states, $\mathbf{Q}=(Q_1,\ldots,Q_{2N})$ and $\mathbf{Q}'=(Q_1',\ldots,Q_{2N}')$, written down with respect to measurement $\mathbf{A}$, a distance $\mathrm{d}(\mathbf{Q},\mathbf{Q}')$ apart with respect to the metric over state space. If these states are now observed with respect to measurement $\mathbf{A}'$, then, by the Measurement Simulability postulate, this is equivalent to observing the states $\tilde{\mathbf{Q}}=\mathcal{M}\mathbf{Q}$ and $\tilde{\mathbf{Q}}'=\mathcal{M}\mathbf{Q}'$, respectively, using measurement $\mathbf{A}$, where $\mathcal{M}$ represents the interaction $\mathbf{I}$ and is known to exist from the Measurement Simulability postulate. By the above assumption of the invariance of distance with change of measurement perspective, we require that $\mathrm{d}(\mathbf{Q},\mathbf{Q}')=\mathrm{d}(\tilde{\mathbf{Q}},\tilde{\mathbf{Q}}')$ for all nearby states, which implies that $\mathrm{d}(\mathbf{Q},\mathbf{Q}')=\mathrm{d}(\mathcal{M}\mathbf{Q},\mathcal{M}\mathbf{Q}')$ for all nearby $\mathbf{Q},\mathbf{Q}'$. By assuming that this condition must hold also in the case in which $\mathcal{M}$ represents a discontinuous transformation, we obtain the Metric Invariance postulate.

Average-value correspondence principle

The general notion of average-value correspondence between the behavior of a physical system when modeled in the classical and quantum frameworks, respectively, is familiar in elementary quantum mechanics through Ehrenfest's theorem [28]. However, the possibility that such a correspondence might serve as the basis for a constructive principle that allows the correspondence rules of quantum theory to be determined by appropriately chosen classical relations does not appear to have been widely explored [36].

The intention behind the AVCP is to formulate the notion of average-value correspondence in a form that is sufficiently exact that, roughly speaking, it allows the logic of Ehrenfest's argument to be reversed, enabling correspondence rules of quantum theory to be derived in a systematic manner from relations known to hold in classical physics.

The precise form of the principle is arrived at by considering a sequence of particular examples (see Ref. [24] for details). The key difficulty is what the quantum experiment must be in the case in which the operators involved are noncommuting. For example, suppose that, in the classical framework, a measurement of C at time t can be implemented by performing measurements of A and B at time t, and suppose that it is known that operators A and B represent measurements of A and B, respectively, and that $[\mathsf{A},\mathsf{B}]\neq 0$. In this case, the AVCP asserts that the quantum measurements are performed on separate copies of the system, thereby sidestepping the problem of measurement ordering.

The application of the AVCP is restricted to particular functions $f(a^{(1)},a^{(2)},\ldots,a^{(m-1)})$ in order that the case of measurements of products of classical observables whose quantum operators are noncommuting does not need to be addressed. However, this restriction does not prevent the derivation of the correspondence rules that one commonly requires when applying the quantum formalism.

VI. DISCUSSION

A. Discussion concerning individual events and their polarities

The Complementarity and States postulates assert that, when measurement $\mathbf{A}$ is performed, there are $2N$ possible events, each with an associated polarity (a binary degree of freedom), but that, when events $2i-1$ or $2i$ occur (with their respective polarities), only the outcome i is registered. The above derivation of the quantum formalism lends support to the plausibility of these assertions, but raises the immediate question as to why the measurement does not (or cannot) resolve the individual events and their associated polarities. A tentative answer to this question is as follows.

First, the overall phase ϕ of a system in an eigenstate of energy E changes at the rate $E/\hbar$. Consequently, the probabilities P_{2i-1} and P_{2i} of events $2i-1$ and $2i$ are oscillating at frequency $2E/h$, and the polarities of these events are switching at frequency E/h. Now, it seems reasonable to suppose that, if one wishes to observe the realization of events $2i-1$ or $2i$ and their polarities, the measurement performed must have a temporal resolution $\Delta t\ll h/2E$. Conversely, if the measurement does not have such resolution, it seems reasonable to suppose that one event and its polarity will not be cleanly realized, but rather *both* events and their polarities will be realized over the duration of the measurement, leading to the situation in which only the property that both events hold in common, namely the outcome i, can be observed, with the other properties being "smeared out."

Now, according to the energy-time uncertainty relation $\Delta E\Delta t\geqslant\hbar/2$ [37], the energy associated with the interaction used to implement a measurement with the temporal resolution needed to observe the individual events and polarities has uncertainty $\Delta E\geqslant\frac{1}{2}\hbar/\Delta t$, so that $\Delta E\gg E/2\pi$. From $E=mc^2$, it then follows that ΔE must be of the order of the rest energy of the system. A measurement of such energy would therefore probably not preserve the integrity of the system, thereby violating the assumption made in Sec. II A that interactions preserve the integrity of the system. Hence, a measurement with the requisite temporal resolution cannot be consistently described within the quantum formalism. Conversely, a measurement that, with high probability, preserves

the integrity of the system, will have insufficient temporal resolution to resolve the individual events and their polarities.

B. Some implications of the deduction

1. Role of information in quantum theory

From the point of view afforded by the derivation, it appears that the concept of information, via the information metric, plays a substantial role in giving rise to various structural features of the quantum formalism. First, we have seen that, if the space of probability distributions is endowed with the information metric, a transformation to the space where the distributions are parametrized by the square roots of probability is endowed with a Euclidean metric. Hence, even at the level of classical probability theory, it appears, just as one sees in the quantum formalism, that the square roots of probability have a rather fundamental significance

Second, once a quantum state is represented as a unit vector, $\mathbf{Q}=(\sqrt{p_1}\cos\theta(\chi_1),\sqrt{p_1}\sin\theta(\chi_1),\ldots,\sqrt{p_N}\sin\theta(\chi_N))^{\mathsf{T}}$, in a $2N$-dimension real Euclidean space, the imposition of the Metric Invariance postulate (which rests on the global gauge condition) leads immediately to the function $\theta(\chi)=a\chi+b$, with the measure over χ being uniform. Hence, the sinusoidal functions into which the phases, $\phi_i=a\chi_i+b$, in a quantum state, $\mathsf{v}=(\sqrt{p_1}e^{i\phi_1},\ldots,\sqrt{p_N}e^{i\phi_N})^{\mathsf{T}}$, enter can be directly traced to the concept of information.

Third, the requirement that transformations of the state space preserve the metric leads to the conclusion that the transformations must be orthogonal. As we have shown, it is then only necessary to impose the Gauge Invariance and Measure Invariance postulates (both of which rest on the global gauge condition) to obtain the result that the set of allowed transformations is in one-to-one correspondence with the set of unitary and antiunitary transformations of a suitably defined complex vector space.

2. Insights into the quantum formalism

The derivation provides a number of significant insights into the quantum formalism, of which we mention only a few. First, we note that, since the development of the quantum formalism, there has been some uncertainty as to whether the formalism is the most general formalism for the description of quantum phenomena in flat space-time. Various possibilities have been suggested for the generalization of the formalism that, from a purely mathematical point of view, seem to be plausible, and which may have interesting physical consequences. For example, the possibility of nonunitary temporal evolution has been considered by several authors [8–10]. The derivation given above gives rise to a mathematical structure that is neither more nor less general than the finite-dimensional abstract quantum formalism, and thereby lends support to the view that the quantum formalism is the most general formalism for the description of quantum phenomena in flat space-time.

Second, the prevalence of complex numbers in the quantum formalism is perhaps its most mysterious mathematical features. In the derivation, one can see that the emergence of the complex numbers is a direct consequence of the imposition of the global gauge invariance condition. Prior to the imposition of the two postulates based on this condition (namely, the Gauge Invariance and the Measure Invariance postulates), the state space is S^{2N-1} in a $2N$-dimensional real space, and the set of all possible transformations is the set of orthogonal transformations. However, these two postulates restrict the set of all possible transformations of Q^{2N} to a subset of the orthogonal transformations, and thereby allow the set of all possible transformations to be represented by the set of all unitary and antiunitary transformations of a suitably defined complex vector space. The critical importance of the global gauge invariance condition is striking: the physical irrelevance of the overall phase of a pure state is usually regarded as being a minor mathematical feature of the quantum formalism, but the reconstruction presented here suggests that this simple-looking feature has a central role to play.

Finally, it has been suggested that the quantum formalism may owe at least a significant part of its structure to the fact that quantum theory permits nonlocality and no-signaling to "peacefully coexist" [3], and a number of recent reconstructive approaches ([7,21,29], for example) rely upon postulates that concern the behavior of physically separated subsystems to derive the mathematical features of the quantum formalism that are necessary to describe a *single* physical system. However, the above derivation, which makes no reference to subsystems in deriving the quantum formalism for a single system, suggests that such considerations are not, in fact, necessary.

ACKNOWLEDGMENTS

I am indebted to Steve Gull and Mike Payne for their support, and to Yiton Fu and Tetsuo Amaya for extensive critical comments and suggestions. I am also indebted to Harvey Brown, Ariel Caticha, Matthew Donald, Chris Fuchs, Suguru Furuta, Lucien Hardy, Lane Hughston, Kevin Knuth, Rüdiger Schach, John Skilling, Lee Smolin, Rob Spekkens, William Wootters, and many others for discussions and invaluable comments and suggestions. I would like to thank the Cavendish Laboratory; Trinity College, Cambridge; Wolfson College, Cambridge; and Perimeter Institute for institutional and financial support. Research at Perimeter Institute is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MEDT.

[1] J. A. Wheeler, *Proceedings of the 3rd International Symposium on the Foundations of Quantum Mechanics, Tokyo* (Physical Society of Japan, Tokyo, 1989).
[2] C. Rovelli, Int. J. Theor. Phys. **35**, 1637 (1996).
[3] S. Popescu and D. Rohrlich, e-print arXiv:quant-ph/9709026.
[4] J. Summhammer, Int. J. Theor. Phys. **33**, 171 (1994).
[5] A. Zeilinger, Found. Phys. **29**, 631 (1999).
[6] C. A. Fuchs e-print arXiv:quant-ph/0205039.
[7] L. Hardy, e-print arXiv:quant-ph/0101012.
[8] S. Weinberg, Phys. Rev. Lett. **62**, 485 (1989).
[9] S. Weinberg, Ann. Phys. (N.Y.) **194**, 336 (1989).
[10] N. Herbert, Found. Phys. **12**, 1171 (1982).
[11] C. J. Isham, e-print arXiv:quant-ph/0206090.
[12] W. K. Wootters, Ph.D. thesis, University of Texas at Austin, 1980.
[13] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).
[14] Č. Brukner and A. Zeilinger, Phys. Rev. Lett. **83**, 3354 (1999).
[15] Č. Brukner and A. Zeilinger, Phys. Rev. A **63**, 022113 (2001).
[16] Č. Brukner and A. Zeilinger, in *Time, Quantum, and Information*, edited by L. Castell and O. Ischebeck (Springer, Berlin, 2002).
[17] A. Grinbaum, Int. J. Quantum Inf. **1**, 289 (2003).
[18] A. Grinbaum, Ph.D. thesis, Ecole Polytechnique, Paris, 2004.
[19] A. Caticha, Phys. Rev. A **57**, 1572 (1998).
[20] A. Caticha, Found. Phys. **30**, 227 (2000).
[21] R. Clifton, J. Bub, and H. Halvorson, Found. Phys. **33**, 1561 (2003).
[22] S. Amari, *Differential-Geometrical Methods in Statistics* (Springer-Verlag, Berlin, 1985).
[23] E. Wigner, *Group Theory* (Academic, New York, 1959).
[24] P. Goyal, e-print arXiv:0805.2765v1.
[25] J. Butterfield, in *Quo Vadis Quantum Mechanics?*, edited by A. Elitzur and N. Kolenda (Springer, Berlin, 2005), http://philsci-archive.pitt.edu/archive/00001193.
[26] R. W. Spekkens, Phys. Rev. A **75**, 032110 (2007).
[27] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States* (Cambridge University Press, Cambridge, UK, 2006).
[28] P. Ehrenfest, Z. Phys. **45**, 455 (1927).
[29] L. Hardy, e-print arXiv:quant-ph/0111068v1.
[30] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955).
[31] H. J. Groenewold, Physica (Amsterdam) **12**, 405 (1946).
[32] D. Bohm, *Quantum Theory* (Dover, New York, 1989).
[33] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1995).
[34] J. Oppenheim, Ph.D. thesis, University of Toronto, 1999.
[35] For example, (a) Caticha [19], in a derivation of Feynman's rules for combining probability amplitudes, assumes that a complex number can be associated with every experimental setup; (b) Grinbaum [18], in Axiom VII, makes specific assumptions regarding the applicable number fields; and (c) in [21], it is assumed that a physical theory can be accommodated within a C^*-algebraic framework, which is based on the complex number field.
[36] For example, Refs. [30–32] mention the general idea of average-value correspondence in their discussion of the operator rules of quantum theory. For example, Groenewold [31] [Eqs. (1.32)–(1.34)] remarks that the sum rule is equivalent to a condition on the expectations of the respective operators, but the idea is not formulated in a manner that is sufficiently systematic to derive the operator rules, and no attempt is made to derive any of the other types of correspondence rule (such as the measurement commutation relations) using average-value correspondence. Bohm [32] clearly articulates the idea that average-value correspondence can be used as a constraint on quantum theory, and uses it to determine particular instances of the function rule (Secs. 9.5–9.21) and to determine the Hamiltonian operator that represents a nonrelativistic particle (Secs. 9.24–9.26). However, the idea is not systematically formulated and applied beyond these special cases.
[37] We shall regard $\Delta E \Delta t \geqslant \hbar/2$ as being a consequence of the classical result $\Delta\omega\Delta t \geqslant 1/2$ (relating the uncertainty in the duration and angular frequency of a wave) and the photon energy-frequency relationship $E=\hbar\omega$. However, the validity and meaning of the energy-time uncertainty relation, and of the inferences that can legitimately by drawn from it, have been, and continue to be, the subject of debate (see, for example, Ref. [33], Sec. 12.8, and Ref. [34]). The argument given in the text leading to $\Delta E \geqslant E/2$ should, accordingly, only be regarded as suggestive insofar as it relies on a particular interpretation of the energy-time uncertainty relation.