Valid PACS appear here

The Role of Information in the Probabilistic Reconstruction of Quantum Theory

Philip Goyal

Cavendish Laboratory
University of Cambridge

Abstract. In this paper, we explore the possibility that the concept of information may enable a derivation of the quantum formalism from a set of physically comprehensible postulates. Taking the probabilistic nature of measurements as a given, we introduce the concept of information via a novel invariance principle, the Principle of Information Gain. Using this principle, we then show that it is possible to deduce the abstract quantum formalism for finite-dimensional quantum systems from a set of postulates, of which one is a novel physical assumption, and the remainder are based on experimental facts characteristic of quantum phenomena or are drawn from classical physics. The concept of information plays a key role in the derivation, and gives rise to some of the central structural features of the quantum formalism.

Keywords: Foundations of Physics, Quantum Theory, Information Theory

INTRODUCTION

Since its formulation in the mid-1920s, quantum theory has been successfully applied to an ever broadening range of physical phenomena. However, a fundamental question has remained largely unanswered: what does quantum theory tell us about the way that nature operates? One of the main obstacles in formulating a comprehensive answer to this question is that, owing to the fact that the quantum formalism was obtained using a significant amount of mathematical guesswork, the formalism has many features (such as its use of complex numbers) whose physical origin is obscure. Consequently, it is not at all clear what constitutes the actual *physical content* of the theory, a situation that has led to a plethora of often conflicting answers to the above question.

Over the last two decades, a number of authors have expressed the view that our efforts to develop an understanding of quantum theory would be significantly aided by a systematic derivation of the formalism from a set of physically comprehensible assumptions [1, 2, 3]. Furthermore, several authors have proposed that the concept of information may be the key, hitherto missing, ingredient which, if appropriately applied and formalized, might make such a derivation possible [4, 1, 5, 6, 3].

The proposal that information might enable a derivation of the quantum formalism rests, to a significant degree, upon the recognition that the concept of information plays a new and fundamental role in quantum physics. One way to see this is as follows. In classical physics, an experimenter presented with a system in an unknown state can, in principle, perform a ideal measurement upon the system which gives perfect knowledge about the state of the system. Hence, there is no fundamental distinction between the state of the system on the one hand, and an ideal experimenter's *knowledge* of the state on the other. However, if one takes the probabilistic nature of measurement as suggested by quantum theory at face value, so that the state of a system only determines the outcome probabilities of measurements performed upon it, it follows that an ideal measurement (or even a finite number of such measurements performed upon an ensemble of identically-prepared systems) provides only partial knowledge about the unknown state of the state of the system. Hence, in sharp contrast to the situation in classical physics, a fundamental distinction is drawn between the state of the system and the knowledge that the experimenter can possibly have of the state. It is then natural to attempt to *quantitatively* relate the two, and so to ask: 'How much information has been obtained by the experimenter about the state?', whereby the concept of information assumes a fundamental role.

One of the earliest attempts to explore the possible role of information in determining the quantum formalism is due to Wootters [7], who showed that Malus' law can be derived from an intuitively plausible information-theoretical principle using the standard inferential methods of probability theory and the standard Shannon information measure. More recently, other attempts [8, 9, 10, 5] have been made to examine and quantify the gain of information in the measurement process, and which differ in various ways from Wootters' approach, but which lead to Malus' law. However, none of these approaches are able to generalize their results in a physically motivated way to obtain a significant part of the quantum formalism.

In contrast, several other recent approaches [1, 11, 12, 13, 14, 15] which involve the concept of information succeed in deriving a significant fraction of the quantum formalism, but make abstract assumptions of key importance which are given no physical interpretation, and which thereby significantly detract from the understanding of the physical origin of the quantum formalism that can be obtained.

In [16, 17], we have attempted to build upon the insights provided by Wootters' approach, and to derive the abstract quantum formalism, together with the correspondence rules of quantum theory, from assumptions that can be clearly understood as assertions about the physical world. The key information-theoretic postulate is the *Principle of Information Gain*, which expresses the idea that, although different measurements (such as different Stern-Gerlach measurements on a spin) yield different information about the state of a system, they nonetheless provide the same *amount* of information about the state. That is, although different measurements provide different perspectives on a system, none is informationally privileged with respect to any other. In this paper, we outline our approach, and, for reasons of space, focus on the postulates and their physical origin, describing the derivation only briefly. The reader is referred to [16, 17] for a more thorough discussion.

EXPERIMENTAL SET-UP AND POSTULATES

Abstract Experimental Set-up

Consider an experimental set-up where, in each run, a system (from a source of identical systems) undergoes a preparation and an interaction, and is then subject to

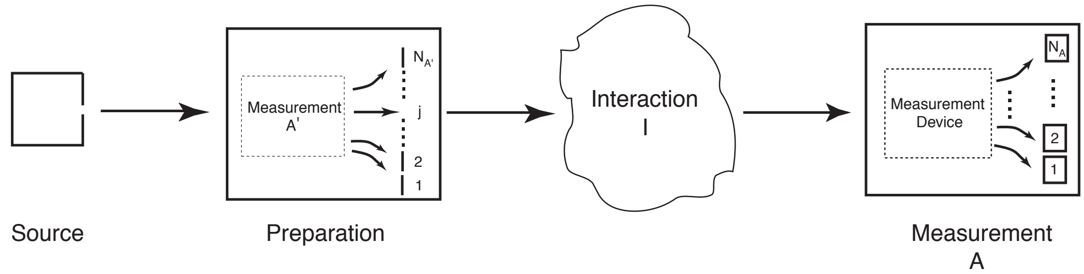

FIGURE 1. An abstract experimental-set up. In each run of the experiment, a physical system (such as a silver atom) is emitted from a source, passes a preparation step, and is then subject to a measurement. The preparation is implemented as a measurement, $\mathbf{A}'$, which has $N_{A'}$ possible outcomes,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

a measurement. Suppose that the purpose of the set-up is to allow the experimenter to investigate how some property (or observable), such as position or spin, of the system is affected by the interaction with the system.

Ideally, in such an experiment, the measurement data obtained should be independent of interactions with the system that occur prior to the preparation. For, otherwise, the measurement data could be influenced by conditions that are not under experimental control. In our derivation of the quantum formalism, we shall take such set-ups, which we shall say are *closed* (or have the property of *closure*), as our starting point.

On the experimentally-founded assumptions that the possible outcomes of a measurement performed upon a system are finite in number and the outcome probabilities are determined by the state of the system, it follows that, in a given experimental set-up, the measurement data obtained in repeated trials is theoretically characterized by a finite set of probabilities. Therefore, the closure condition, when applied in the context of these assumptions, requires that these outcome probabilities are independent of the pre-preparation history of the system.

Accordingly, in the abstract experimental set-up illustrated in Fig. 1, we restrict the measurements and interactions that are permitted to those which result in a closed set-up. Further, we define the measurement set $\mathscr{A}$ and interaction set $\mathscr{I}$ such that, for any pair of measurements $\mathbf{A}, \mathbf{A}' \in \mathscr{A}$ and any interaction $\mathbf{I} \in \mathscr{I}$, the resulting set-up is closed.

Statement of the Postulates.

Consider the idealized experiment illustrated in Fig. 1 in which a system passes a preparation step that employs a measurement $\mathbf{A}'$ in measurement set $\mathscr{A}$, undergoes an interaction, $\mathbf{I}$ in the interaction set $\mathscr{I}$, and is then subject to a measurement, $\mathbf{A}$, in $\mathscr{A}$. The abstract theoretical model that describes this set-up satisfies the following postulates.

1. **Measurements**
   1.1 *Outcome Number.* When any measurement $\mathbf{A} \in \mathscr{A}$ is performed, one of N ($N \geq 2$) possible outcomes are observed.

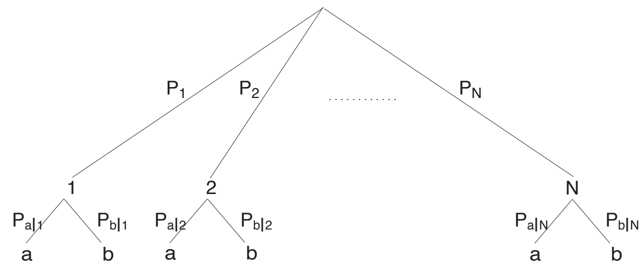

FIGURE 2. (*State Interpretation Postulate*) A probability tree showing the outcomes of measurement **A** when performed on a system in state **S** given by $(\vec{P}, \vec{\chi})$. Outcome i is observed with probability P_i, and outcome a or b is realized with probability $P_{a|i}$ or $P_{b|i}$, respectively. In addition, outcome $+$ or $-$ is realized, determined by the sign of $Q_{a|i}$ if a has been realized, or $Q_{b|i}$ if b has been realized. The outcomes a, b, and $+, -$ cannot be observed by the experimenter.

1.2 *Measurement Representation*. For any given pair of measurements $\mathbf{A}, \mathbf{A}' \in \mathscr{A}$, there exist interactions $\mathbf{I}, \mathbf{I}' \in \mathscr{I}$ such that $\mathbf{A}'$ can, insofar as probabilities of the observed outcomes and insofar as the possible output states of the measurement are concerned, be represented 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 *State Representation*. With respect to any given measurement $\mathbf{A} \in \mathscr{A}$, the state, $\mathbf{S}(t)$, of a quantum system at time t is given by $(\vec{P}, \vec{\chi})$, where $\vec{P} = (P_1, P_2, \ldots, P_N)$ and $\vec{\chi} = (\chi_1, \chi_2, \ldots, \chi_N)$ are real n-tuples, and where P_i is the probability that the ith outcome of measurement $\mathbf{A}$ is observed.

2.2 *State Interpretation*. When measurement $\mathbf{A} \in \mathscr{A}$ is performed on a system in state $\mathbf{S}(t)$ and the outcome i is observed, there are additional outcomes that are objectively realized but unobserved (see Fig. 2):

(i) one of two outcomes, labeled a and b, which are obtained with respective probabilities $P_{a|i} = Q^2_{a|i}$ and $P_{b|i} = Q^2_{b|i}$, where $Q_{a|i} = f(\chi_i)$ and $Q_{b|i} = \tilde{f}(\chi_i)$, where f is not a constant function and $f, \tilde{f}$ have range $[-1, 1]$, and

(ii) one of two possible outcomes, with values labeled $+$ and $-$, which is determined by the sign of $Q_{a|i}$ if a has been realized, or $Q_{b|i}$ if b has been realized.

2.3 *Information Gain*. When measurement $\mathbf{A} \in \mathscr{A}$ is performed on a system in any given unknown state $\mathbf{S}(t)$, the amount of Shannon-Jaynes information provided by the observed outcomes and the outcomes a and b about $\mathbf{S}(t)$ in n runs of the experiment is independent of $\mathbf{S}(t)$ in the limit as $n \to \infty$.

2.4 *Prior Probabilities*. The prior probability $\Pr(\chi_i|\mathrm{I})$ is uniform for $i = 1, \ldots, N$, where I is the background knowledge of the experimenter prior to performing the experiment.

3. **Transformations**

3.1 *One-to-one*. If a physical transformation of the system is represented by a map, $\mathscr{M}$, over the state space, $\mathscr{S}$, of the system, then $\mathscr{M}$ is one-to-one.

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

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

3.4 *Invariance*. The map $\mathscr{M}$ is such that, for any state $\mathbf{S} \in \mathscr{S}$, the observed outcome probabilities, $P_1', P_2', \ldots, P_N'$, of measurement $\mathbf{A} \in \mathscr{A}$ performed upon a system in state $\mathbf{S}' = \mathscr{M}(\mathbf{S})$ are unaffected if, in any representation, $(\vec{P}, \vec{\chi}) = (P_i; \chi_i)$, of the state $\mathbf{S}$ written down with respect to $\mathbf{A}$, any arbitrary real constant, χ_0, is added to each of the χ_i.

3.5 *Consistency* The posterior probability distributions over $\mathscr{S}$ that result from the following two processes coincide in the limit as $n \to \infty$:
   (i) inferring a posterior over $\mathscr{S}$ based upon the objectively realized outcomes when the measurement $\mathbf{A} \in \mathscr{A}$ is performed upon n copies of a system in state $\mathbf{S}$, and then transforming the posterior using $\mathscr{M}$, or
   (ii) inferring a posterior over $\mathscr{S}$ based upon the objectively realized outcomes when the measurement $\mathbf{A} \in \mathscr{A}$ is performed upon n copies of a system in state $\mathscr{M}(\mathbf{S})$,

3.6 *Temporal Evolution*. The map, $\mathscr{M}_{t,\Delta t}$, 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 non-zero constant with the dimensions of action.

The above postulates, together with the Average-Value Correspondence Principle (AVCP), not discussed here, suffice to determine the form of the abstract quantum model for the abstract set-up. From the Outcome Number Postulate, it follows that, when any measurement in $\mathscr{A}$ is performed on the system, one of N possible outcomes is observed. Accordingly, we shall denote the abstract quantum model of such a set-up by $\mathbf{q}(N)$.

Finally, we shall need a further postulate (Composite Systems postulate), which is not discussed here, in order to obtain the tensor product rule that allows one to relate the quantum model of a composite system to the quantum models of its component systems.

ORIGIN OF THE POSTULATES

In our discussion below, we shall divide the postulates according to their origin as follows:

1. *Based on Experimental Observations*. Postulates obtained through direct generalization of experimental facts that are taken to be characteristic of quantum phenomena.

2. *Drawn from Classical Physics*.
   (i) Postulates adopted unchanged from the theoretical framework of classical physics.
   (ii) Postulates obtained from classical physics via a classical–quantum correspondence argument.
3. *Novel Postulates*. Postulates (one informational and one physical postulate) that are based on novel theoretical principles or ideas which cannot obviously be traced to classical physics or experimental observations.

1. Postulates based on Experimental Observations

In experiments where Stern-Gerlach measurements are performed upon microscopic particles such as silver atoms, one observes that:

1. Every Stern-Gerlach measurement yields the same number of possible outcomes (for example, two outcomes in the case of silver atoms), and
2. Any Stern-Gerlach measurement can be implemented using a given Stern-Gerlach measurements flanked either side with suitable magnetic fields.

The Outcome Number and the Measurement Representation postulates can be regarded as direct generalization of these observations.

2. Postulates drawn from Classical Physics

Postulates adopted from Classical Framework

The framework of classical physics assumes that physical transformations of a system are represented by maps over the state space of the system, and that these maps have the properties expressed in the One-to-One, Continuity, and Continuous Transformations Postulates. We adopt these properties unchanged in the quantum framework.

Finally, we note that the classical framework is consistent in sense that, if the framework allows one to make a prediction about a given measurement outcome by following two or more calculational paths, then these paths must yield the same result. This elementary consistency requirement, adopted and applied in a probabilistic framework, leads directly to the Consistency postulate.

Postulates obtained by a Correspondence Argument

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. In particular, we shall consider a quantum model of a particle and its classical counterpart, and shall require that these models are in one-to-one correspondence in the classical limit as the mass, m, of the particle tends to macroscopic values. Using this correspondence, we shall transpose several elementary properties of the classical model into the quantum model of the particle, and thence, by generalization, into the quantum model of a arbitrary physical system.

Based on experimental data, one finds that, in a quantum experiment, if one prepares the particle using a coarse position measurement, and then performs a coarse position measurement on the particle, the set-up is closed (up to experimental precision) provided that the position measurements used are of sufficiently high resolution. Now, in the classical limit, we expect that the system will behave classically between the preparation and measurement, and one can readily establish that the relevant classical mode in this situation is a particle ensemble model, the Hamilton-Jacobi model.

In the Hamilton-Jacobi model, the state of the ensemble is given by $(P(\vec{r},t), S(\vec{r},t))$, where $P(\vec{r},t)$ is the spatial probability density function, and S is the action function. In the case of coarse position measurements with N possible outcomes, we shall use the discretized form of the Hamilton-Jacobi state, $(P_1^{(\mathrm{CM})}, \ldots, P_N^{(\mathrm{CM})}, S_1, \ldots, S_N) = (P_i^{(\mathrm{CM})}; S_i)$, where $P_i^{(\mathrm{CM})}$ is the probability that the position measurement yields a detection at the ith measurement location, and S_i is the classical action at the ith measurement location.

In order that the predictions of the quantum and classical models agree in the classical limit, the quantum state $\mathbf{S}(t)$ $(t > t_0)$ must contain degrees of freedom which encode N quantities, which we shall denote $S_1^{(\mathrm{QM})}, \ldots, S_N^{(\mathrm{QM})}$, which, in the classical limit, are equal to the S_i. Equivalently, we shall assume that $\mathbf{S}$ contains N dimensionless real quantities, $\chi_1, \ldots, \chi_N$, such that $S_i^{(\mathrm{QM})} = \alpha\chi_i$, where α is a non-zero constant with the dimensions of action.

From the above discussion, in the quantum model of the particle, which we shall denote $\mathbf{q}^*(N)$, the state, $\mathbf{S}$, is given by $(\vec{P}, \vec{\chi})$, where $\vec{\chi} = (\chi_1, \ldots, \chi_N)$. The State Representation postulate directly generalizes this statement to the abstract model $\mathbf{q}(N)$.

We now observe that the Hamilton-Jacobi model has the following properties, which can be readily verified from the Hamilton-Jacobi equations:

1. *Invariance*. The evolution of the state $(P_i^{(\mathrm{CM})}(t_1); S_i(t_1))$ to the state $(P_i^{(\mathrm{CM})}(t_2); S_i(t_2))$ is such that the $P_i^{(\mathrm{CM})}(t_2)$ are unchanged if an arbitrary real constant, S_0, is added to each of the $S_i(t_1)$.
2. *Temporal Evolution*. In a time-independent background, a state, $(P_i^{(\mathrm{CM})}(t); S_i(t))$ whose observable degrees of freedom are time-independent, evolves in time Δt to the state $(P_i^{(\mathrm{CM})}(t); S_i(t) - E\Delta t)$, where E is the total energy of the system.

Furthermore, from the first property, since the zero-value of the S_i is conventional and therefore has no physical correlate, it follows that, for any $S_1, \ldots, S_N$ and any S_0,

$$\Pr(S_1, \ldots, S_N | \mathrm{I}) = \Pr(S_1 + S_0, \ldots, S_N + S_0 | \mathrm{I}), \tag{1}$$

from which, by marginalization, if follows that, for any S_0, S_1,

$$\Pr(S_1 | \mathrm{I}) = \Pr(S_1 + S_0 | \mathrm{I}), \tag{2}$$

where I represents the state of knowledge of the experimenter prior to performing measurements on the system. Therefore, the prior $\Pr(S_i|\mathrm{I})$ is uniform, which, for convenience, we shall list as a third property:

3. *Prior Probabilities*. The prior $\Pr(S_i|\mathrm{I})$ is uniform ($i = 1, 2, \ldots, N$), where I represents the state of knowledge of the experimenter prior to performing a measurement on the system.

On the assumption of the above correspondence between the Hamilton-Jacobi model and the model $\mathbf{q}^*(N)$, it is now possible to transpose these properties to the model $\mathbf{q}^*(N)$ in the classical limit, and then to generalize the abstract quantum model $\mathbf{q}^N$. This transposition immediately yields the Invariance, Temporal Evolution, and Prior Probabilities postulates.

3. Novel Postulates

The two novel postulates consist of a physical postulate (the State Interpretation postulate), and an informational postulate (the Information Gain postulate).

Postulate 2.2: State Interpretation.

According to the State Representation postulate, the state $S(t)$, written with respect to some measurement $\mathbf{A} \in \mathscr{A}$, consists of the pair $(\vec{P}, \vec{\chi})$, where $\vec{P}$ contains the probabilities of the observed outcomes, 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. The State Interpretation postulate is motivated by the aesthetical desideratum that a quantum state consist, as far as possible, of probabilities of events, rather than being such a mixture.

Accordingly, we postulate that χ_i encodes the probabilities of some events, labeled a and b. Hence, when measurement $\mathbf{A}$ is performed on the system, one of $2N$ possible outcomes is obtained, with probabilities determined by the state of the system. Since, by the Outcome Number postulate, an experimenter observes only one of N possible outcomes upon performing measurement $\mathbf{A}$, we are forced to postulate that, for some reason to be investigated later, the outcomes a and b are not observed by the experimenter.

Now, we make the reasonable assumption that the abstract quantum framework being developed is capable of modeling the behavior of a photon when subject to polarization measurements, and that this model will agree with the predictions of electromagnetism under a particle interpretation. Now, an electromagnetic plane wave of constant amplitude moving along the $+z$-direction is described by the vector-valued function $\vec{E} = E_0(\cos\theta\,\vec{i} + \sin\theta\,\vec{j})$, and the information about the polarization of the wave is contained in $(\cos\theta, \sin\theta)$ with respect to polarization measurements in the xy-plane. In the particle interpretation, the probability that a photon will pass through a polarizer whose axis points along the x-axis or y-axis is given by $\cos^2\theta$ or $\sin^2\theta$, respectively. The key feature which we wish to abstract from this example is that, since the map from $(\cos\theta, \sin\theta)$ (the 'state-level') to $(\cos^2\theta, \sin^2\theta)$ (the 'probability-level') is many-to-one, the computed probabilities are *not* the fundamental quantities when describing the state of the photon. Rather, the more fundamental quantities are $\cos\theta$ and $\sin\theta$, which we can regard as square roots of probability in the range $[-1, 1]$, which are squared to obtain probabilities.

To incorporate this two-layered feature into the abstract quantum model, we assume that, following the realization of outcome a or b, one of two outcomes, labeled $+$ and $-$, is obtained. This ensures that one binary-valued degree of freedom is associated with each of the $2N$ possible probabilistically-determined outcomes. Furthermore, we assume that the value of χ_i *determines* whether $+$ or $-$ is obtained via the sign of either $Q_{a|i}$ or $Q_{b|i}$, depending upon whether a or b was obtained, where $P_{a|i} = Q^2_{a|i}$ and $P_{b|i} = Q^2_{b|i}$. In summary, the quantum state consists of the N probabilities $P_1, \ldots, P_N$ and the $2N$ quantities $Q_{a|1}, Q_{b|1}, \ldots, Q_{a|N}, Q_{b|N}$ which encode the probabilities $P_{a|1}, P_{b|1}, \ldots, P_{a|N}, P_{b|N}$ and encode the values of the $2N$ binary-valued degrees of freedom.

In Sec. **??**, we sketch some ideas which help to provide a better physical understanding of this postulate.

Postulate 2.3: Information Gain.

Suppose that, in trial 1 of n runs of an experiment, a measurement $\mathbf{A}$ is performed on a system in state $\mathbf{S}(t)$, and suppose that trial 2 is identical to trial 1 except that measurement $\mathbf{A}'$ is performed instead of $\mathbf{A}$. The data obtained in trials 1 and 2 provides information (via the Shannon-Jaynes entropy functional) about $S(t)$. If, in one of the two trials 1 and 2, the data obtained yields more information about the state $S(t)$ than in the other trial, this means that one of the two measurements $\mathbf{A}$ and $\mathbf{A}'$ is privileged compared to the other insofar as the amount of information that it yields about $\mathbf{S}(t)$. Although this possibility cannot be ruled out *a priori*, we make the intuitively plausible assertion that, although these different measurements provide different perspectives on the system, these perspectives are not informationally privileged.

Now, by the Measurement Representation postulate, trial 2 is equivalent (insofar as the probabilities of the probabilistically-determined outcomes are concerned) to trial $2'$ consisting of n runs of an experiment where a system in state $\mathbf{S}(t)$ is sent through an arrangement consisting of a suitable physical interaction with the system, represented by map $\mathscr{M}$, followed by measurement $\mathbf{A}$, followed by another physical interaction. Since the data obtained in trials 2 and $2'$ is statistically identical (as ensured by the Measurement Representation postulate), the amount of information obtained about $S(t)$ in trial 2 is asymptotically equal to the amount of information obtained about $S'(t) = \mathscr{M}(S(t))$ in trial $2'$

But, our assertion that the amount of information obtain about $S(t)$ in trials 1 and 2 is the same therefore implies that the amount of information obtained in trial 1 about $S(t)$ and in trial $2'$ about $S'(t)$ is the same. Hence, we are led to the Information Gain principle. That is, the postulate can be understood as arising from the requirement that no measurement in the measurement set provides an informationally privileged perspective on the system.

DEDUCTION OF THE QUANTUM FORMALISM

In this section, we shall indicate the broad steps which lead to the deduction of the quantum formalism from the above postulates. The reader is referred to [16] for details.

- **Step 1.** Obtain State Space using the States Postulates (2.1–2.4).
  - Postulates 2.1–2.3 imply that:
    - States are represented as unit vectors, $\vec{Q}$, in a $2N$-dimensional Euclidean space.
    - The prior over the unit hypersphere sphere is uniform..
    - After n runs of experiment, the posterior is symmetric Gaussian over one orthant with standard deviation $1/2\sqrt{n}$, and is zero in other orthants.
  - Postulate 2.4 (Prior Probabilities) then implies that the functions $f(\chi_i)$ and $\tilde{f}(\chi_i)$ in the State Interpretation postulate can be taken, without loss of generality, to be $\cos\chi_i$ and $\sin\chi_i$, respectively.
- **Step 2.** Obtain the set of allowable Transformations using the Transformations Postulates (3.1–3.6).
  - The One-to-one and Consistency Postulates imply that transformations are orthogonal transformations of the unit hypersphere.
  - The Invariance postulate implies that the transformations are a particular subset of the orthogonal transformations.
  - This subset of the orthogonal transformations can equivalently be represented by the set of all unitary and antiunitary transformations of a suitably-defined N-dimensional complex vector space.
  - The Parameterized Transformations and Continuous Transformations postulates then imply that continuously-parameterized transformations are either unitary or antiunitary, and that continuous transformation are unitary.
  - The Temporal Evolution postulate, together the the Average-Value Correspondence Principle, gives the explicit form of the temporal evolution operator.
- **Step 3.** Obtain the Hermitian operator representation of measurements using the Measurement Representation postulate (1.2).

DISCUSSION

Owing to the transparency of the assumptions used in the derivation, it is apparent from the derivation that the concept of information plays a substantial role in giving rise to various structural features of the quantum formalism, and can reasonably be said to provide the backbone of the quantum formalism. In particular, the information gain condition directly leads to Q-space, which introduces square-roots of probability, or *real* amplitudes and, via the State Interpretation postulate, leads to a $2N$-dimensional Q-space. Furthermore, in conjunction with the Prior Probabilities postulate, the Information Gain postulate leads to the function $f(\chi_i) = \cos(\chi_i)$. Hence, the sinusoidal functions into which the phases in a quantum state enter can be directly traced to the concept of information. Finally, the prior over the unit hypersphere in Q^{2N}-space induced by the imposition of the Information Gain postulate leads, via the Consistency postulate, to the strong constraint that physical transformations can only be represented by orthogonal transformations of the unit hypersphere.

The derivation provides a number of insights into the quantum formalism. First, 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 use of complex numbers in the quantum formalism is perhaps one of its most mysterious mathematical features. The emergence of complex numbers in the derivation depends on most of the postulates, and so is not easy to unravel, but the role of the Invariance postulate is perhaps the most obvious: its imposition leads to the possibility of representing the set of all possible orthogonal transformations of Q^{2N} by the set of all unitary or antiunitary transformations of a suitably defined complex vector space. This is particularly interesting since it suggests that the importance of complex numbers is directly tied to the set of possible transformations.

Discussion of the State Interpretation postulate

The above derivation rests upon two key postulates, the Information Gain postulate and the State Interpretation postulate, the remainder of the postulates being traceable to elementary experimental observations and to classical physics. Of these two key postulates, the State Interpretation postulate invites the most immediate questions: (i) when a measurement is performed on the system, why is the experimenter able to observe outcomes $1,\ldots,N$, but not the outcomes a,b and $+,-$, and (ii) how one can more intuitively understand the physical meaning of these outcomes. A tentative answer to these questions is as follows.

First, for a system in an eigenstate of energy E, the overall phase of its quantum state (in the complex representation) changes at the rate $E/\hbar$. Consequently, the probabilities $P_{a|i}$ and $P_{b|i}$ are oscillating at frequency $2E/h$, and the signs of $Q_{a|i}$ and $Q_{b|i}$ are switching at frequency E/h. Now, it is reasonable to expect that, if one wishes to observe the actualization of one of the possible outcomes a,b and $+,-$, the measurement performed must have a temporal resolution $\Delta t \ll h/2E$. If the measurement does not have such resolution, it seems plausible to suppose that none of the possible outcomes will be cleanly actualized, leading to the situation where each measurement performed on the system effectively yields *all* of the possible outcomes, forming an unchanging, smeared 'background' to the outcomes $1,\ldots,N$, so that one would remain unaware that other outcomes (other than $1,\ldots,N$) are being actualized.

Now, according to the energy-time uncertainty relation $\Delta E \Delta t \geq \hbar/2$, the energy associated with the interaction used to implement a measurement with the temporal resolution needed to observe the additional outcomes has uncertainty $\Delta E \geq \frac{1}{2}\hbar/\Delta t$, 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 identity of the system. Hence, a measurement with the requisite temporal resolution cannot be consistently described within the quantum formalism. Conversely, a measurement which, with high probability, preserves the identity of the system, will have insufficient temporal resolution to resolve the outcomes a, b and $+, -$.

Second, the outcomes a, b and $+, -$ can be graphically represented as follows. Consider a unit circle in Euclidean space, with orthogonal axes labeled a and b, respectively, and consider a unit vector at angle χ_i to the a-axis. Since the prior $\Pr(\chi_i|I)$ is uniform, the vector is equally likely to be pointing in any direction. When the measurement **A** is performed, the outcomes a and b register which axis the vector is found to be pointing along, and the additional $+$ and $-$ outcomes determine whether the vector is pointing along the positive or negative direction along the respective axis. From repeated trials, one obtains information about the probabilities $P_{a|i}$ and $P_{b|i}$, from which one can infer that the angle χ_i is one of four possibilities; the signs $+$ or $-$ associated with a and b determine which of the four possibilities is the correct one. Under temporal evolution, the vector rotates at angular frequency $E/\hbar$ if the system is in an eigenstate of energy E. This picture appears to be closely related to the idea that massive particles can be regarded as energy that is 'trapped' in a region of space and is undergoing rapid, to-and-fro motion (an idea that was shown by Einstein to quantitatively account for the inertia of mass), and related to Hestenes' [18, 19] contention that there is a localized circular motion which accounts for electron spin.

REFERENCES

1. C. Rovelli, *Int. J. Theor. Phys.* **35**, 1637–1678 (1996), see quant-ph/9609002v2.
2. S. Popescu, and D. Rohrlich (1997), see quant-ph/9709026v2.
3. C. A. Fuchs (2002), see quant-ph/0205039.
4. J. A. Wheeler, "It from bit," in *Proceedings of the 3rd international symposium on the foundations of quantum mechanics, Tokyo*, 1989.
5. J. Summhammer (1999), see quant-ph/9910039.
6. A. Zeilinger, *Found. Phys.* **29**, 631 (1999).
7. W. K. Wootters, *The acquisition of information from quantum measurements*, Ph.D. thesis, University of Texas at Austin (1980).
8. Č. Brukner, and A. Zeilinger, *Phys. Rev. Lett.* **83**, 3354–3357 (1999).
9. Č. Brukner, and A. Zeilinger, *Phys. Rev. A* **63** (2001).
10. Č. Brukner, and A. Zeilinger (2002), see quant-ph/0212084v1.
11. A. Caticha, *Phys. Rev. A* **57**, 1572 (1998).
12. A. Caticha (1999), see quant-ph/9810074v2.
13. R. Clifton, J. Bub, and H. Halvorson (2003), see quant-ph/0211089v2.
14. A. Grinbaum, *Int. J. Quant. Inf.* **1**, 289–300 (2003), see also quant-ph/0306079.
15. A. Grinbaum, *The Significance of Information in Quantum Theory*, Ph.D. thesis, Ecole Polytechnique, Paris (2004), see quant-ph/0410071.
16. P. Goyal (2007), see quant-ph/0702124.
17. P. Goyal (2007), see quant-ph/0702149.
18. D. Hestenes, *Found. Phys.* **15**, 63–87 (1985).
19. D. Hestenes, *Found. Phys.* **20**, 1213–1232 (1990).