From information geometry to quantum theory

To cite this article: Philip Goyal 2010 New J. Phys. 12 023012

View the article online for updates and enhancements.

Related content
- Quasi-probability representations of quantum theory with applications to quantum information science
  Christopher Ferrie
- Three-dimensionality of space and the quantum bit: an information-theoretic approach
  Markus P Müller and Lluís Masanes
- Is the local linearity of space-time inherited from the linearity of probabilities?
  Markus P Müller, Sylvain Carrozza and Philipp A Höhn

Recent citations
- Quantum theory from questions
  Philipp Andres Höhn and Christopher S. P. Wever
- Reflections on the information paradigm in quantum and gravitational physics
  Philipp Andres Höhn
- Michael J. W. Hall and Marcel Reginatto

IOP eBooks
Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.
Start exploring the collection - download the first chapter of every title for free.
Abstract. In this paper, we show how information geometry, the natural geometry of discrete probability distributions, can be used to derive the quantum formalism. The derivation rests upon three elementary features of quantum phenomena, namely complementarity, measurement simulability, and global gauge invariance. When these features are appropriately formalized within an information geometric framework, and combined with a novel information-theoretic principle, the central features of the finite-dimensional quantum formalism can be reconstructed.
physical intuition cannot directly relate. Over the last two decades, there has been growing interest in elucidating these aspects by expressing, in a less abstract mathematical language, what quantum theory might be telling us about how nature works, and trying to derive, or reconstruct, quantum theory on this basis [1], [6]–[10].

Much of the recent effort in reconstructing the quantum formalism is motivated by the hypothesis that the concept of information might be the key, hitherto missing, ingredient that may enable a reconstruction, and several attempts have been made to systematically explore the reconstruction of the quantum formalism from an informational starting point (for example, [7], [11]–[18]). Although these approaches have yielded significant insights, they are either incomplete (for example, [11, 12, 14]) or employ abstract assumptions that involve the assumption of the complex number field (for example, [16]–[18]). Such assumptions significantly limit the degree to which the physical content of the quantum formalism can be elucidated, since one of the most mysterious mathematical features of the quantum formalism is being assumed at the outset. In this paper, we show that the principal mathematical features of quantum theory can be reconstructed using the concept of information without employing such assumptions.

Our approach develops intimate connections, known to exist for some time, between structures that arise naturally in classical probability theory on the one hand, and the quantum formalism for pure states on the other [19]–[22]. For example, Wootters [19] has shown in the framework of classical probability theory that one can quantify the degree to which two discrete probability distributions, \( p = (p_1, \ldots, p_N) \) and \( p' = (p'_1, \ldots, p'_N) \), can be distinguished, given the same number of samples from each by means of the statistical distance, \( d_S(p, p') = \cos^{-1} \left( \sum_i \sqrt{p_i p'_i} \right) \), between them. If one considers the statistical distance, \( d_S(p, p') \), between the probability distributions \( p \) and \( p' \), which characterize the results of projective measurement \( A \) when performed upon two \( N \)-dimensional pure states \( u \) and \( v \), respectively, and if one chooses \( A \) such that \( d_S \) is maximized, Wootters shows that \( d_S \) is equal to the Hilbert space distance, \( d_H(u, v) = \cos^{-1} |u^\dagger v| \), between \( u \) and \( v \) [19]. The existence of such a connection is remarkable, and suggests that the usual formalism of quantum theory might owe at least some of its structure to the notion of distinguishability that arises naturally in a purely classical probabilistic setting.

Following Wootters, we adopt an operational approach, and so take the probabilistic nature of measurements as a given. Accordingly, the framework of classical probability theory is taken as a starting point. We equip this framework with a metric, \( d_S^2 = \frac{1}{2} \sum_i dP_i^2/p_i \), the information metric (or Fisher–Rao metric), the infinitesimal form of the statistical distance, rather than the statistical distance itself, as this suffices for the purposes of the reconstruction. This metric determines the distance between infinitesimally close probability distributions \( p = (p_1, \ldots, p_N) \) and \( p' = (p'_1, \ldots, p'_N) \). As we shall describe below, the information metric can be understood as a natural consequence of the introduction of the concept of information into the probabilistic framework. Accordingly, we shall refer to this framework as the information geometric framework [23].

Within this framework, we formalize three elementary features of quantum phenomena, namely complementarity, global gauge invariance and measurement simulability, detailed below. 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 [24], 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 finite-dimensional quantum
formalism, namely that pure states are represented by complex vectors, physical transformations are represented by unitary or antiunitary transformations, and the outcome probabilities (and the corresponding output states) of measurements are given by the Born rule. The present paper provides a streamlined derivation of the key parts of the finite-dimensional quantum formalism, focusing on the essential ideas. The reader is referred to [24, 25] for a more detailed discussion of the underlying ideas and methodology, as well as a derivation of the remainder of the finite-dimensional quantum formalism.

1. Information metric

We begin by giving a simple argument that shows how the information metric arises in a classical probabilistic setting from the concept of information. Suppose that Alice has two coins, A and B, characterized by the probability distributions \( p = (p_1, p_2) \) and \( p' = (p'_1, p'_2) \), respectively. Suppose that she chooses coin A, tosses it \( n \) times, and then sends the data to Bob, without disclosing to him which coin she chose. If Bob knows \( p \) and \( p' \), how much information does the data provide him with about which coin was tossed? Intuitively, the more information the data provides, the more sharply the distributions are distinguished.

Using Bayes’ theorem and Stirling’s approximation for the case where \( n \) is large, on the assumption that coins A and B are \( a \ priori \) equally likely to be chosen, one finds that

\[ \frac{P_A}{P_B} = \exp \left( n \sum_{i=1}^2 p_i \ln \frac{p_i}{p'_i} \right), \tag{1} \]

where \( P_A \) is the probability that the tossed coin is A given the data, and likewise for \( P_B \).\(^1\) When the probability distributions are close, so that \( p' = p + dp \), the argument of the exponent can be expanded in the \( dp_i \) to give

\[ \frac{P_A}{P_B} = \exp (2n ds^2), \tag{2} \]

where \( ds^2 = \frac{1}{4} \sum_i dp_i^2 / p_i \) is the information metric.

Now, the information gained by Bob, \( \Delta I \), is the reduction in his uncertainty, and is therefore defined as

\[ \Delta I \equiv U(1/2, 1/2) - U(P_A, P_B), \tag{3} \]

with \( U \) being an entropy (uncertainty) function such as the Shannon entropy. But, since \( P_A + P_B = 1 \) and \( P_A / P_B \) is determined by \( ds \), once \( U \) is selected, \( \Delta I \) is determined by \( ds \). For example, if \( U \) is chosen to be the Shannon entropy \( U(\pi_1, \pi_2) = -\sum_i \pi_i \ln \pi_i \), one finds that

\[ \Delta I = \frac{1}{2} (n ds^2)^2. \tag{4} \]

This result immediately generalizes to the case where \( p \) and \( p' \) are \( M \)-dimensional probability distributions \((M \geq 2)\). Hence, from an informational viewpoint, it is natural to endow the space of discrete probability distributions with the information metric.

\(^1\) See [27], chapter 4, for a discussion of such applications of Bayes’ theorem.
Parenthetically, we remark that Wootters’ statistical distance, \( d_5(p, p') = \cos^{-1}\left(\sum_i \sqrt{p_i p'_i}\right) \), between the probability distributions \( p \) and \( p' \) is the minimum distance between \( p \) and \( p' \) with respect to the information metric\(^2\). We do not, however, make use of this result in what follows.

2. Derivation

2.1. Construction of state space

Measurement is idealized as a process that (i) when performed upon some physical system, yields one of \( N \) possible outcomes, with probabilities, \( p_1, \ldots, p_N \), that are determined by the state of the system immediately prior to the measurement, and (ii) is reproducible, so that, upon immediate repetition of the measurement, the same outcome is obtained with certainty.

2.1.1. Formalizing complementarity. We take the first feature, complementarity, to consist of the general idea that, when a measurement is performed upon a system in some state, the measurement outcome only yields information about half of the experimentally accessible degrees of freedom of the state. In the above classical probabilistic model of measurement, we can express this idea in a very simple way as follows.

Postulate 1. Complementarity. When measurement \( A \) is performed, one of \( 2N \) possible events occur; but they are not individually observed. Outcome \( i \) is observed \((i = 1, \ldots, N)\) whenever either event \( 2i - 1 \) or event \( 2i \) is realized. The events \( 1, \ldots, 2N \) are assumed to occur with probabilities \( P_1, \ldots, P_{2N} \), respectively, so that

\[
p_i = P_{2i-1} + P_{2i},
\]

where \( p_i \) is the probability of outcome \( i \).

The \( P_q (q = 1, \ldots, 2N) \) can be summarized by the probability \( n \)-tuple \( P = (P_1, \ldots, P_{2N}) \). As a result, of the \( 2N - 1 \) degrees of freedom of \( P \), the measurement outcome only yields information about the \( p_i \), which constitute \( N - 1 \) degrees of freedom. We shall shortly impose an additional constraint (global gauge invariance), which implies that only \( 2(N - 1) \) of the \( 2N - 1 \) degrees of freedom of \( P \) are physically relevant. Hence, the measurement yields information about exactly one half of the experimentally accessible degrees of freedom in \( P \).

Intuitively, performing the measurement brings about the realization of one of \( 2N \) possible events but the observed outcomes coarse-grain over these events; when event \( 2i - 1 \) or \( 2i \) occurs, the measurement is (for some reason to be investigated) unable to resolve the individual events, so that only outcome \( i \) is registered. This is a novel hypothesis, which, at this point in the derivation, is recommended by its simplicity, and remains to be judged by its explanatory power (namely its capacity to support a derivation of the quantum formalism)\(^3\).

\(^2\) This can most easily be seen by making a change of coordinates, so that \( q_i = \sqrt{p_i} \). In terms of these coordinates, the information metric becomes Euclidean, \( d_5^2 = \sum_i dq_i^2 \), and the probability simplex becomes the positive orthant of the unit hypersphere in \( q \)-space. The minimum distance between the vectors \( q = (\sqrt{p_1}, \ldots, \sqrt{p_N}) \) and \( q' = (\sqrt{p'_1}, \ldots, \sqrt{p'_N}) \) is then simply \( \cos^{-1}(q, q') \), which is \( d_5(p, p') \).

\(^3\) A similar hypothesis (the ‘knowledge balance principle’) has been made in a recent toy model of quantum theory in order to give concrete expression to complementarity [28]. The insights provided by this toy model provide additional reason to explore the complementarity hypothesis given here. See also the discussion section of [24].
2.1.2. **Imposing the information metric.** Next, we endow the space of probability distributions $\mathbf{P}$ with the information metric, $d^2 = \frac{1}{2} \sum_i dP_i^2 / P_i$, where $q = 1, \ldots, 2N$. It is convenient to define $Q_q = \sqrt{P_q}$, where $Q_q \in [0, 1]$, since the metric over the $Q_q$ is then simply the Euclidean metric, $d^2 = dQ_1^2 + \cdots + dQ_{2N}^2$, so that $\mathbf{Q} = (Q_1, Q_2, \ldots, Q_{2N})^T$ is a unit vector that lies on the positive orthant of the unit hypersphere $S^{2N-1}$ in a $2N$-dimensional Euclidean space.

2.1.3. **Representing physical transformations.** We now consider transformations of state space that represent physical transformations of the system. We postulate that transformations of the state space, assumed one-to-one, preserve the metric over state space—that is, the *information distance*, $d(\mathbf{Q}, \mathbf{Q}')$, between any pair of infinitesimally close states, $\mathbf{Q}, \mathbf{Q}'$, where $d(\cdot)$ denotes distance with respect to the metric over state space, is preserved. The essential idea here is that the discriminability of any pair of nearby states is a quantity that is intrinsic to this pair of states, and therefore should remain invariant under reversible and deterministic transformations of the system.

Now, if one takes the $\mathbf{Q}$ themselves as the state space of the system, one immediately finds that continuous one-to-one transformations of the state space that preserve the information metric are not possible. A simple way to allow the existence of such transformations is to take the entire unit hypersphere, $S^{2N-1}$, as the state space of the system. That is, we take the state of the system as given by a unit vector $\mathbf{Q} = (Q_1, Q_2, \ldots, Q_{2N})^T$, with $Q_q \in [-1, 1]$, where the probabilities $P_q$ are given by $P_q = Q_q^2$. From the information metric over $\mathbf{P}$, it follows from the relation $P_q = Q_q^2$ that the metric over $\mathbf{Q}$ is Euclidean,

$$d^2 = dQ_1^2 + dQ_2^2 + \cdots + dQ_{2N}^2. \quad (6)$$

We can summarize the above requirements as follows.

**Postulate 2. Metric invariance.** The state of the system is given by the unit vector $\mathbf{Q} = (Q_1, Q_2, \ldots, Q_{2N})^T$, with $Q_q \in [-1, 1]$, where the probabilities $P_q$ are given by $P_q = Q_q^2$. The metric over $\mathbf{Q}$ is Euclidean, $d^2 = dQ_1^2 + dQ_2^2 + \cdots + dQ_{2N}^2$, which any transformation, $\mathcal{M}$, of state space must preserve.

It follows from this postulate that $\mathbf{Q}$ lies on the unit hypersphere, $S^{2N-1}$, in a $2N$-dimensional real Euclidean space. From the requirement of metric preservation, it follows that $\mathcal{M}$ is an orthogonal transformation of $S^{2N-1}$, so that every transformation can be expressed as $\mathbf{Q}' = M \mathbf{Q}$, where $M$ is a $2N$-dimensional real orthogonal matrix.

The above extension of the state space from the positive orthant of $S^{2N-1}$ to the entire hypersphere is an assumption that, although formally quite natural, presently awaits a clear physical basis.

2.2. **Global gauge invariance**

The second feature, *global gauge invariance*, consists of the idea that one can find a representation of the state of a system such that, if one displaces a subset of the degrees of freedom, the information distance between any pair of nearby states should be the same irrespective of the measurement from whose perspective the states are observed.

---

4 Additionally, using the measurement simulability postulate given in section 2.3, this postulate can be grounded in the idea that the information distance between any pair of nearby states should be the same irrespective of the measurement from whose perspective the states are observed.
freedom of the state by the same amount, any physical predictions based on the state are left invariant. To formalize this feature, we begin by making a change of variables by expressing the state $Q$ in terms of the probabilities $p_1, p_2, \ldots, p_N$, and $N$ additional real degrees of freedom, $\theta_1, \theta_2, \ldots, \theta_N$, so that, without loss of generality,

$$Q_{2i-1} = \sqrt{p_i} \cos \theta_i, \quad Q_{2i} = \sqrt{p_i} \sin \theta_i. \quad (7)$$

Only the $\theta_i$ can be subject to displacement since a displacement involving any of the $p_i$ would be experimentally detectable. Accordingly, we formalize the idea of global gauge invariance by requiring that $\theta_i = \theta(\chi_i)$, where $\theta(\cdot)$ is an unknown, non-constant, differentiable function to be determined, and that the transformation $\chi_i \to \chi_i + \chi_0$ for $i = 1, \ldots, N$ brings about no predictive changes for any $\chi_0 \in \mathbb{R}$. From this global gauge condition, we immediately draw the following postulate.

**Postulate 3. Gauge invariance.** The map $M$ is such that, for any state $Q \in S^{2N-1}$, the probabilities, $p_1', p_2', \ldots, p_N'$, of the outcomes of measurement $A$ performed upon a system in state $Q' = M(Q)$ are unaffected if, in any representation, $(p_i'; \chi_i)$, of the state $Q$, an arbitrary real constant, $\chi_0$, is added to each of the $\chi_i$.

Additionally, we draw the requirement that the measure, $\mu(p_i; \chi_i)$, over $p_1, \ldots, p_N$, $\chi_1, \ldots, \chi_N$ induced by the metric over $S^{2N-1}$ is consistent with the global gauge condition. This requirement is necessary so that probabilistic inference using the measure as a prior over state space is consistent with our physical knowledge of the system. This requirement yields the following postulate.

**Postulate 4. Measure invariance.** The measure $\mu(p_i; \chi_i)$ induced by the metric over state space 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 any $\chi_0$.

### 2.2.1. Determining the function $\theta(\cdot)$.

From equations (5)–(7),

$$ds^2 = \frac{1}{4} \sum_{i=1}^{N} \frac{dp_i^2}{p_i} + \sum_{i=1}^{N} p_i \theta'^2(\chi_i) d\chi_i^2. \quad (8)$$

The measure, $\mu(p_i; \chi_i)$, 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 marginalizes to give

$$\mu_i(\chi_i) = c|\theta'(\chi_i)| \quad (9)$$

as the measure over $\chi_i$, where $c$ is a constant.

Now, from the measure invariance postulate, it follows by marginalization that the measure $\mu_i(\chi_i)$ satisfies the relation $\mu_i(\chi_i + \chi_0) = \mu_i(\chi_i)$ for all $\chi_0$, and is therefore independent of $\chi_i$. Hence, from equation (9), $\theta(\chi) = a\chi + b$, where $a, b$ are constants, where $a \neq 0$ since, by assumption, the function $\theta(\cdot)$ is not constant. We can therefore write

$$Q = (\sqrt{p_1} \cos \theta_1, \sqrt{p_1} \sin \theta_1, \ldots, \sqrt{p_N} \sin \theta_N). \quad (10)$$
2.2.2. Implementing gauge invariance, and the emergence of complex vector space. From equation (10), the gauge invariance postulate and the relation $\theta_i = a\chi_i + b$ given above, one can show that $M$ is restricted to one of two types: $M$ has the general form

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

(11)

where $T^{(ij)}$ has the form

$$T^{(ij)} = \alpha_{ij} \begin{pmatrix} \cos \varphi_{ij} & -\sin \varphi_{ij} \\ \sin \varphi_{ij} & \cos \varphi_{ij} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^\beta,$$

and where either $\beta = 0$ (type 1), in which case $T^{(ij)}$ is a scale–rotation matrix, or $\beta = 1$ (type 2), in which case $T^{(ij)}$ is a scale–rotation–reflection matrix, with scale factor $\alpha_{ij}$ and rotation angle $\varphi_{ij}$ in either case.$^5$

Now, state $Q$ can be faithfully represented by the complex unit vector

$$v \equiv (Q_1 + iQ_2, \ldots, Q_{2N-1} + iQ_{2N})^T = (\sqrt{p_1}e^{i\theta_1}, \ldots, \sqrt{p_N}e^{i\theta_N})^T.$$  

(12)

Remarkably, one can then show that every transformation $M$ of type 1 corresponds one-to-one with the set of unitary transformations of $v$, and that every transformation $M$ of type 2 corresponds one-to-one with the set of antiunitary transformations of $v$. In particular, on the assumption that a parameterized transformation that represents a continuous physical transformation must reduce to the identity for some value of the parameters, it follows that a continuous transformation must be represented by unitary transformations.

2.3. Representation of measurements

The third feature, measurement simulability, can be stated as follows.

**Postulate 5. Measurement simulability.** Any reproducible measurement, $A'$, describable in the formalism can, insofar as its outcome probabilities and associated output states are concerned, be simulated by an arrangement consisting of measurement $A$ flanked by suitable interactions with the system.

Given the results derived above, this postulate immediately implies that $A'$ can be simulated by the arrangement shown in figure 1, where $U$ and $V$ are unitary transformations representing interactions with the system.

The reproducibility of measurement $A$ implies that the state of a system immediately after $A$ has yielded outcome $i$ is given by $v_i = (0, \ldots, e^{i\phi_i}, \ldots, 0)^T$, where $\phi_i$ is undetermined. Hence, the input state $v'_i = U^{-1}v_i$ will yield outcome $i$. For the arrangement to behave like a reproducible measurement, the output state must be $v'_i$ up to an overall phase, so that it suffices to choose $Vv_i = v'_i$ for $i = 1, \ldots, N$, which implies that $V = U^{-1}$.

Since the $v_i$ form an orthonormal basis, it follows from $v'_i = U^{-1}v_i$ that the $v'_i$ also form an orthonormal basis. Therefore, any state $v$ can be expanded as $\sum_i c_i v'_i$, where $c_i = v'_i v$. With the

$^5$ See [24], section V B 2.
input state $v$, the state measured by measurement $A$ in the arrangement is $Uv = \sum_i c'_i v_i$. From equation (12), the probabilities, $p_1, \ldots, p_N$, of the outcomes of measurement $A$ performed on state $v = (v_1, \ldots, v_N)$ are given by $p_i = |v_i|^2$. Therefore, in this case, the measurement yields outcome $i$, together with output state $v'_i$, with probability $|c'_i|^2 = |v'_i|^2$, which is the Born rule.

3. Discussion

The physical irrelevance of the overall phase of a pure state is usually regarded as being a minor mathematical feature of the quantum formalism and of little physical importance. From this standpoint, one of the most surprising findings in the derivation is that the global gauge condition (which expresses in a more general way the physical irrelevance of the overall phase) is sufficiently strong to transform a $2^N$-dimensional real formalism (where states are real unit vectors, and transformations are orthogonal transformations) into the familiar $N$-dimensional complex vector formalism of quantum theory (where states are complex unit vectors, and transformations are unitary and antiunitary transformations). In particular, the fact that the set of possible transformations one obtains is precisely the set of all unitary and antiunitary transformations (and neither more nor less) is not something that could, a priori, have been reasonably anticipated.

The derivation provides a number of other important insights into the structure of the quantum formalism. From the perspective of the derivation, it is clear that the use of complex numbers in the quantum formalism is directly tied to the set of possible transformations of state space. For example, if the set of all orthogonal transformations were allowed, then the complex form of the formalism, while still possible to write down, would involve nonlinear continuous transformations and would therefore not appear mathematically natural. The derivation also suggests that information geometry is directly or indirectly responsible for many of its key mathematical features (such as the importance of square roots of probability, and the sinusoidal functions that appear in a quantum state), thereby providing significant new support for the hypothesis that information plays a fundamental role in determining the structure of quantum theory.

Finally, the derivation illuminates a previous partial reconstruction of quantum theory due to Stueckelberg [26]. Stueckelberg makes an assumption similar to the complementarity postulate to arrive at the idea that the state of a system is given by a $2^N$-dimensional probability distribution that can be written as a unit vector in a $2^N$-dimensional ‘square root of probability space’, as we have done. He then asserts that the allowable transformations of the state space are orthogonal transformations, and shows that if the transformations are
restricted by a superselection rule, then the set of restricted transformations is equivalent to the set of unitary transformations acting on a suitably defined $N$-dimensional complex state space. The present derivation shows that Stueckelberg’s assertion that the allowable transformations are orthogonal transformations can be naturally accounted for in terms of the information metric over the probability simplex via the metric invariance postulate. The derivation also shows that Stueckelberg’s superselection rule can be replaced by the global gauge invariance postulate.

Acknowledgments

I thank Harvey Brown, Steve Flammia, Yiton Fu, Chris Fuchs, Lucien Hardy, Gerard’t Hooft, Lane Hughston and Lee Smolin for discussions and invaluable comments and suggestions. I also thank an anonymous reviewer for helpful suggestions. Research at Perimeter Institute is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MEDT.

References

[1] Fuchs C A 2002 arXiv:quant-ph/0205039
[2] Weinberg S 1989a Phys. Rev. Lett. 62 485
[3] Weinberg S 1989b Ann. Phys., NY 194 336
[4] Herbert N 1982 Found. Phys. 12 1171
[5] Isham C J 2002 arXiv:quant-ph/0206090
[6] Wheeler J A 1989 Proc. 3rd Int. Symp. on the Foundations of Quantum Mechanics, Tokyo
[7] Rovelli C 1996 Int. J. Theor. Phys. 35 1637 (arXiv:quant-ph/9609002v2)
[8] Popescu S and Rohrlich D 1997 Causality and locality in modern physics and astronomy: open questions and possible solutions (arXiv:quant-ph/9709026)
[9] Zeilinger A 1999 Found. Phys. 29 631
[10] Hardy L 2001 arXiv:quant-ph/0101012
[11] Wootters W K 1980 PhD Thesis, University of Texas at Austin
[12] Summhammer J 1994 Int. J. Theor. Phys. 33 171 (arXiv:quant-ph/9910039)
[13] Brukner Č and Zeilinger A 1999 Phys. Rev. Lett. 83 3354
[14] Brukner Č and Zeilinger A 2002 Time, Quantum, and Information ed L Castell and O Ischebeck (Berlin: Springer) (arXiv:quant-ph/0212084v1)
[15] Grinbaum A 2003 Int. J. Quantum Inf. 1 289 (arXiv:quant-ph/0306079)
[16] Grinbaum A 2004 PhD Thesis Ecole Polytechnique, Paris (arXiv:quant-ph/0410071)
[17] Catcha A 2000 Found. Phys. 30 227 (arXiv:quant-ph/9810074v2)
[18] Clifton R, Bub J and Halvorson H 2003 Found. Phys. 33 1561
[19] Wootters W K 1981 Phys. Rev. D 23 357–362
[20] Brody D C and Hughston L P 1996 Phys. Rev. Lett. 77 2851
[21] Brody D C and Hughston L P 1998 Proc. R. Soc. A 454 2445
[22] Mehrafarin M 2005 Int. J. Theor. Phys. 44 429 (arXiv:quant-ph/0402153)
[23] Amari S 1985 Differential-Geometrical Methods in Statistics (Berlin: Springer)
[24] Goyal P 2008a arXiv:0805.2761v1
[25] Goyal P 2008b arXiv:0805.2765v1
[26] Stueckelberg E C G 1960 Helv. Phys. Acta. 33 727
[27] Sivia D S 1996 Data Analysis: A Bayesian Tutorial (New York: Oxford University Press)
[28] Spekkens R W 2007 Phys. Rev. A 75 032110

New Journal of Physics 12 (2010) 023012 (http://www.njp.org/)