International Journal of Theoretical Physics, Vol. 45, No. 7, July 2006 (© 2006)
DOI: 10.1007/s10773-006-9122-3

Bayesian Probabilities and the Histories Algebra

Thomas Marlow[1]

Received September 30, 2005; accepted March 5, 2006
Published Online: May 10, 2006

We attempt a justification of a generalisation of the consistent histories programme using a notion of probability that is valid for all complete sets of history propositions. This consists of introducing Cox's axioms of probability theory and showing that our candidate notion of probability obeys them. We also give a generalisation of Bayes' theorem and comment upon how Bayesianism should be useful for the quantum gravity/cosmology programmes.

KEY WORDS: Bayesian probability; consistent histories; linear positivity.

PACS: 02.50.Cw; 03.65.Ta; 04.60.-m.

During this paper we will introduce a novel notion of probability within the framework of a histories theory. Firstly, we shall introduce Cox's axioms of probability theory and then show that our proposed notion of probability obeys them. Secondly, we shall discuss the implications of such Bayesian probability assignments on the physical foundations of quantum history theories, quantum cosmology and relational gravity theories.

So, one can derive the standard rules of Bayesian probability theory from axioms that notions of probable inference should obey. This was shown by Cox (1961) who gave two simple axioms and derived probability theory from them. One of the axioms is that the probability that two propositions α and β are both true upon a given hypothesis I should depend only upon the probability of one of the propositions upon the same hypothesis and the probability of the other proposition upon the same hypothesis conjoined with the presumption that the former proposition is true. This can be written schematically as:

$$p(\alpha \wedge \beta \mid I) := F[p(\alpha \mid \beta I), p(\beta \mid I)] \tag{1}$$

where F is some function to be determined that is sufficiently well-behaved for our purposes.

Cox's other axiom is simply that the probability of the negation of a proposition upon a given hypothesis should depend only upon the probability of that

[1] School of Mathematical Sciences, University of Nottingham, UK, NG7 2RD; e-mail: pmxtm@nottingham.ac.uk.

proposition upon the same hypothesis. This can similarly be written as:

$$p(\neg\alpha \mid I) := G[p(\alpha \mid I)]. \tag{2}$$

Of course, there is an implicit zeroth axiom that probabilities should be represented by real numbers. Using such axioms Cox derived that the standard probability rules must be obeyed.

However, once one attempts to derive probability as a form of probable inference one is not wholly clear about the status of the zeroth axiom. Certain probabilities must be considered real when they are to be interpreted as relative frequencies but, in terms of a theory of probable inference, there is no *a priori* reason why probabilities should be real. In fact, we can split up the zeroth axiom into two further axioms (Jaynes, 2003). Firstly we can presume the transitivity of probability assignments:

$$\text{Axiom 0a:} \quad \text{If } p(\alpha \mid I) > p(\beta \mid I) \text{ and } p(\beta \mid I) > p(\gamma \mid I) \\ \text{then } p(\alpha \mid I) > p(\gamma \mid I) \tag{3}$$

where '$>$' is an ordering notion that is defined on the space we use to represent probabilities.

Secondly, we can presume what is called 'universal comparability':

$$\text{Axiom 0b:} \quad \text{For all } \alpha, \beta \text{ we have that either } p(\alpha \mid I) > p(\beta \mid I) \\ \text{or } p(\alpha \mid I) < p(\beta \mid I) \text{ or } p(\alpha \mid I) = p(\beta \mid I). \tag{4}$$

The combination of axioms 0a and 0b ensures that probability assignments can be real numbers. Obviously axioms 0a and 0b are restrictions upon the type of probability space we desire for notions of probable inference. However, we might easily not desire axiom 0b, especially in the light of quantum theory and special relativity. There are physical reasons why we might not be able to universally compare certain propositions probabilistically. For example, should we be able to compare two statements that involve spacelike separated regions, or that involve incompatible variables? A form of probability that does compare such statements might involve unjustified inference. So not only do we argue that probabilities need not be real numbers, we also argue that, in certain physical situations, they *should not* be real numbers. The reals simply might not have enough structure to represent a plausible notion of probability in certain situations. See also (Isham, 2003) for other reasons why we might choose not to use reals for all notions of probability.

This is a thesis argued rather cogently by Youssef (1994, 2001) who argues that complex numbers and quaternions could also be consistent with Cox's two axioms (or rather axioms analogous to Cox's). Such work attempts a derivation of the consistency of complex numbers with Cox's two axioms by presuming a distributive lattice of propositions. One can then derive quantum mechanical features for such probability theories. We shall not comment upon this work much more but rather we take a slightly different tack; we would like to discuss quantum history theories (Griffiths, 1984; Omnes, 1988; Gell-Mann and Hartle, 1990; Isham, 1994).

Following (Isham, 1994), we can define a homogeneous history as an ordered tensor product of Heisenberg picture projection operators:

$$\alpha := \hat{\alpha}_{t_n}(t_n) \otimes \hat{\alpha}_{t_{n-1}}(t_{n-1}) \otimes \ldots \hat{\alpha}_{t_2}(t_2) \otimes \hat{\alpha}_{t_1}(t_1) \tag{5}$$

such that $t_n > t_{n-1} > \ldots t_2 > t_1 > t_0$ and $\hat{\alpha}_{t_n}(t_n) = \hat{U}(t_n - t_0)\hat{\alpha}_{t_n}\hat{U}^\dagger(t_n - t_0)$ where $\hat{\alpha}_{t_n}$ is a standard Schrödinger picture projection operator and $\hat{U}$ is the standard unitary evolution operator. The ordered set of times upon which an homogeneous history is defined is called its temporal support. Inhomogeneous histories can then be defined using '$\vee$' or '$\neg$' operations that are naturally defined (Isham, 1994) to produce an algebra of history propositions. There are also natural notions of disjointness and exhaustivity. In the consistent histories programme one normally defines what is called the decoherence functional d which can then be used to define relative frequencies for some complete (disjoint and exhaustive) sets of histories—these are the same relative frequencies as predicted by the von Neumann collapse hypothesis. Sets in which the relative frequencies are well-defined are called d-consistent and not all complete sets are d-consistent. However, we do not wish to discuss relative frequencies *per se*; we would rather discuss the more general notion of Bayesian probabilities.

One can naturally define what is called the class operator of such a history (Isham, 1994):

$$C_\alpha := \hat{\alpha}_{t_n}(t_n)\hat{\alpha}_{t_{n-1}}(t_{n-1}) \ldots \hat{\alpha}_{t_2}(t_2)\hat{\alpha}_{t_1}(t_1). \tag{6}$$

Such class operators can also be defined for inhomogeneous combinations of homogeneous histories (Isham, 1994) in a natural manner. We have the property that, for disjoint homogeneous histories that are defined over the same temporal supports, the class operators just add:

$$C_{\alpha\vee\beta} = \hat{\alpha}_{t_n}(t_n) \ldots \hat{\alpha}_{t_1}(t_1) + \hat{\beta}_{t_n}(t_n) \ldots \hat{\beta}_{t_1}(t_1). \tag{7}$$

Instead of using the subset of the set of complete sets of history propositions that consists of d-consistent sets one can define a larger subset of sets that are Linearly Positive (LP)—this larger subset was introduced by Goldstein and Page (1995) as an alternative to the d-consistent subset which is used in the standard consistent histories programme. Recent work (Marlow, 2006) has shown that there is a certain amount of consistency between a real notion of Bayesian probability and the LP formalism. These LP probabilities seem to obey Bayesian reasoning whereas the standard notion of probability using the decoherence functional does not. Thus we have a 'good'—it seems to obey Cox's axioms—notion of probability for the LP subset of the space of history propositions. We do not, however, have an assignment that is good for all complete sets. All in all, it seems rather implausible that such a real probability can be found. We might instead be able to invoke a complex notion of probability. The decoherence functional gives complex numbers but it does not obey (2) and is designed specifically with the aim of giving real probabilities. Note that its form is derived specifically by presuming von Neumann collapse which gives a natural notion of relative frequency (Anastopoulos, 2004). So, von Neumann collapses are designed to give something real. Without presuming von Neumann collapse one is also released from discussing solely relative frequencies.

The most obvious complex candidate is simply:

$$p(\alpha \mid I) := \mathrm{tr}(C_\alpha \rho) \tag{8}$$

where ρ is the initial density matrix (defined at t_0) and α is either homogeneous as in (5) or is an inhomogeneous proposition defined by combining homogeneous histories using the natural '$\vee$' and '$\neg$' operations (Isham, 1994). The real part of this candidate behave like normal probabilities for the LP subset. Also, given the natural algebra of history propositions (Isham, 1994), (8) obeys (2):

$$p(\neg\alpha \mid I) = \mathrm{tr}((\mathbf{1} - C_\alpha)\rho) = \mathrm{tr}(\rho) - \mathrm{tr}(C_\alpha \rho) = G[p(\alpha \mid I)] \tag{9}$$

where **1** is the unit history proposition.

If Cox's other axiom (1) is to be obeyed in situations where we have the associativity of the '$\wedge$' operation such that:

$$\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma = \alpha \wedge \beta \wedge \gamma \tag{10}$$

then Bayes' rule should be valid (by Cox's proof (Cox, 1961)). Hence we should have that:

$$p(\alpha \mid \beta I) = \frac{\mathrm{tr}(C_{\alpha\wedge\beta}\rho)}{\mathrm{tr}(C_\beta \rho)} \tag{11}$$

which is well-defined as long as $\mathrm{tr}(C_\beta \rho) \neq 0$. Note that the '$\wedge$' is commutative for homogeneous histories defined upon the same temporal support such that $\alpha \wedge \beta = \beta \wedge \alpha$. Hence Cox's proof of the multiplication rule also remains valid:

$$p(\alpha \mid \beta I) p(\beta \mid I) = p(\beta \mid \alpha I) p(\alpha \mid I). \tag{12}$$

Now we must ask ourselves whether such a complex probability assignment is (a) consistent (b) unique and (c) useful? Since it obeys Cox's axioms then it has a certain amount of consistency as a probability assignment. Question (b) is quite hard to answer but we can certainly begin to tackle (c). Hartle (2004) has recently suggested (also see Feynman, 1987) that virtual probabilities—real but not within [0, 1]—might be useful as intermediate steps in any quantum analysis. In light of axiom 0b we would prefer to go the whole hog and discuss a different space for the probabilities altogether—Hartle's virtual probabilities are real so should be rejected unless one weakens one of the probability axioms, but axiom 0b is the most plausible to weaken and would give something other than the reals; thus it seems more natural to invoke the full complex probability. Also note the rather good analogy between the above and Youssef's work (Youssef, 1997). Youssef attempts to prove that complex and quarterion probabilities are compatible with Cox's axioms (or rather some axioms akin to Cox's) combined with the presumption that there exists a sublattice where the normal real probabilities are obtained. This is a necessary feature of such a theory since axiom 0b should presumably apply to a subset of propositions, just not universally. Such an argument adds weight to Youssef's derivation of quantum-like features—similarly for Caticha's recent pedagogical derivation of a quantum theory (Caticha, 1998). Similarly we could ask the same question for history theories—and Marlow (2006) has shown that there is a subset of histories where the standard Bayesian rules apply—but we already have a complex object that might obey Cox's axioms; namely (8).

If we are willing to accept a complex probability then note that the natural candidate behaves in a rather nice way. When histories α and β are homogeneous (and defined over the same temporal support) and disjoint ($\alpha \wedge \beta = \mathbf{0}$ where $\mathbf{0}$ is used to denote the null history) then we have that:

$$\mathrm{tr}(C_{\alpha\vee\beta}\rho) = \mathrm{tr}(C_\alpha\rho) + \mathrm{tr}(C_\beta\rho). \tag{13}$$

So, in such situations our complex probability assignment behaves in the completely standard manner:

$$p(\alpha \vee \beta \mid I) = p(\alpha \mid I) + p(\beta \mid I) - p(\alpha \wedge \beta \mid I). \tag{14}$$

Although this does not yet convince us that such an assignment is really useful conceptually, at least it behaves in a nice manner for such a large number of histories—it is valid *for all complete sets of homogeneous histories*.

Note that $\alpha \wedge \beta = \beta \wedge \alpha$ for homogeneous histories defined over the same temporal support regardless of commutation issues at each time point. In the simple case of two homogeneous histories defined over the same temporal support, where the projection operators at each time point commute, we have that:

$$\mathrm{tr}(C_{\alpha\vee\beta}\rho) = \mathrm{tr}(C_\alpha\rho) + \mathrm{tr}(C_\beta\rho) - \mathrm{tr}(C_{\alpha\wedge\beta}\rho). \tag{15}$$

For example, for two two-time histories in the HPO formulation $\alpha = \hat{\alpha}_{t_1} \otimes \hat{\alpha}_{t_2}$ and $\beta = \hat{\beta}_{t_1} \otimes \hat{\beta}_{t_2}$ such that $[\hat{\alpha}_{t_1}, \hat{\beta}_{t_1}] = 0$ and $[\hat{\alpha}_{t_2}, \hat{\beta}_{t_2}] = 0$ we have that:

$$\alpha \vee \beta = \alpha + \beta - \hat{\alpha}_{t_1}\hat{\beta}_{t_1} \otimes \hat{\alpha}_{t_2}\hat{\beta}_{t_2} \tag{16}$$

$$= \alpha + \beta - (\hat{\alpha}_{t_1} \wedge \hat{\beta}_{t_1}) \otimes (\hat{\alpha}_{t_2} \wedge \hat{\beta}_{t_2}). \tag{17}$$

So, at least for disjoint and commuting homogeneous histories, our probability assignment behaves in the standard manner.

The major use of this 'Bayesian Histories' formalism is its simplicity; and its ease of generalisation. What we have is a formalism that behaves as a standard probability theory for all complete sets of histories. The probabilities just happen to be complex. One might like to ask *why* we should use complex numbers and we can see that they do not obey axiom 0b. Youssef argues that, of the spaces that don't obey 0b, we should use complex or quaternion numbers because they can also obey Cox's other two axioms. Note that the natural partial order on complex numbers does obey axiom 0a. He uses a distributive lattice for his argument so we have yet to complete his argument fully for the non-distributive history algebra proper. We shall not attempt to do such a thing here.

Even given such complex probabilities, we always know that we can get standard real Bayesian probabilities out for a certain subset of the space of history propositions and, furthermore, we can get out relative frequencies by further defining consistency (note that consistency might not be the *only* way for us to get relative frequencies). The point is that we don't need all that, we know that we can get such things out in the end; for now we can just search for well-behaved Bayesian probabilities in domains not yet studied. Hence it is the simplicity and generality of such an unreal probability programme which means it might be useful in the quantum gravity domain. Note that what type of space of probability we use is dependent upon the space and logic of the history propositions we invoke; note also that we require notions of kinematics and dynamics before we can discuss our probability assignments (we use the Heisenberg picture). So perhaps quantum gravity will benefit from explicitly not using standard quantum theory (as is usually the case) and we can simply search for a consistent Bayesian probability assignment for whatever propositional space one ends up deriving by other means, say using causal sets (which, presuming some form of background independence, should include both kinematical and dynamical aspects of the theory). Youssef (2001) has shown that even for distributive logics one can derive many quantum mechanical features from just invoking a complex probability, so a general path is clear: try and find a natural proposition space to be derived prior to probabilistic notions, and then invoke Bayesian reasoning to get the quantum mechanical features from the theory.

So, weirdly enough, Bayesianism might be very useful in relational theories of gravity (where it is rarely invoked; see (Poulin, 2005; Caticha, 2005) for tentative proposals in this area). Here we would briefly like to discuss a few curious analogies between relationism and Bayesian philosophy. There are two basic principles of relationism according to Leibniz. Firstly there is the 'principle of sufficient reason', and secondly there is the 'principle of identifying the indiscernible' (see Smolin, 2005 for an accessible discussion of relationalism). These two principles are connected by the Bayesian 'principle of *in*sufficient reason' which states that if you do not have a rational reason for differentiating two statements you should assign them identical probabilities. If you do have a rational reason for differentiating statements then you should assign different probabilities according to rational rules. Thus, foundationally, Bayesian probability theory is wholly compatible with a relational philosophy. In fact, these two philosophical standpoints might be identified in the future. Bayesian probability is a way of representing such relational ideas via a probability space—the relational nature of gravity theories may help us, rather than hinder us, in searching for quantum probabilities. Hence why we believe this Bayesian histories programme may be useful for quantum gravity theories. We intend to investigate such a quantum gravity programme in future work.

Of course, acceptance of the above programme relies significantly upon a Bayesian view of probabilities. So, let us now briefly discuss why such a viewpoint is useful. Firstly note that Bayesian probabilities aren't incompatible with notions of relative frequency. Quite the opposite; Bayesian probabilities can incorporate most notions of relative frequencies within the literature, whereas theories of relative frequencies need to be *designed* for the problem at hand (Jaynes, 2003). Bayesian probability theory is an umbrella philosophy that can incorporate many different notions of relative frequency (using notions of independence, exchangability, and maximum entropy for example). Also, Bayesian probability is pedagogically useful because of the lack of philosophical presumptions that goes into the theory. One can show that any notion of probable inference that obeys completely transparent axioms must behave as we expect. Caticha has also recently shown that entropy formulae can also be derived in a particularly Bayesian way (Caticha, 2004) (and references therein). Entropy is normally considered a *physical* property of systems but Caticha shows that one can sometimes consistently take an opposing view. What better way to investigate a concept than to explicitly give the axioms by which we are allowed to consistently state it? Especially since there might be physical cases where such axioms are violated. This is the pedagogical power of the Bayesian programme. Since such axioms are clearly defined, the Bayesian programme is also ripe for generalisation because certain axioms can be relaxed if there are cogent reasons for doing so. This is not the case with relative frequencies which do not have such a clear pedagogical basis. In Marlow (2006b) we also argue that the use of Bayesian probability can provide a logical loophole in discussions about nonlocality and Bell's inequalities.

Often Bayesian probabilities are rejected because they are considered subjective. This, however, is the wrong way to look at it. It is only by considering probabilities as subjective that we can begin to understand the reasons we use the concept 'probability' as we do. Calling a probability 'objective', when we can never measure it directly, is not good ontology, especially since such 'objective' probabilities are usually invoked as relative frequencies which in turn are defined using an entirely unphysical notion of infinite ensembles. One can never measure 'entropy' or 'probability', one can only apply (either consistently or inconsistently) these concepts to the measurements we make (Mana, 2004). 'Entropy' and 'probability' are forms of reasoning we *use*, not things that are. We must work out *why* and *how* we use these concepts. Assuming that they are objective properties of systems does not help us in this enterprise as it only gives us the opportunity to accept them without thinking about why we use them in the way we do. So, it is better to derive such concepts from a consistent set of plausible axioms. Having to invoke some subjectivity in the notion of probability is thus not the result of the programme, it is the whole pedagogical basis of doing things in a Bayesian manner. One must first assume that probabilities are subjective in order to then begin to work out why we ought to assign a certain probability over another. It is also clear that these plausible axioms are not *a priori* philosophical axioms; we require them to be consistent with the physics we are doing, and the physics we are doing can justify further generalisations or axioms. So the term 'subjective' should not be considered synonymous with 'not physical' or 'arbitrary'.

Another reason Bayesian probabilities are sometimes rejected (especially by quantum cosmologists) is because of the notion of 'observers' that is often kept explicit. Perhaps this is a misconception however, as is shown by the pervasive use of Bayesian methods in the astrophysics community. Bayesian probability theory doesn't require that we have observers floating around and that we must model them—the observers need not be 'in' the theory nor 'outside' the system being discussed. Bayesianism is more about what 'we', as theorists, are allowed to consistently say about a theory. There is no measurement problem as soon as one accepts that it is *necessarily* 'us' who are interpreting a theory. We can either interpret a theory consistently or inconsistently; Bayesianism is an attempt to do so consistently. Thus we do not need to invoke 'observers'—one can do so for pedagogical reasons but it is not a necessary feature of Bayesian physics. One should rather use rational 'interpreters' instead of 'observers' but even that is just a pedagogical notion and could still be removed from the foundations of the theory. Interpreters are pedagogically invoked for the same reason that we need subjectivity; we must first presume that we could interpret things differently before we can begin to constrain how we ought to interpret things. No two equals are the same. So, such interpreters are not 'passive' because they are rational but nor are they 'invasive'—we don't have to assume that they ('we') are physically effecting the world around them ('us') through rationalising.

Note that implicit in our definition of probability is a notion of dynamics (we are using the Heisenberg picture). We also explicitly have a notion of initial state ρ. This may confuse some Bayesian practitioners because ρ is often considered subjective and we should be able to update any state assignment given further information. However, if we are to be discussing closed quantum systems then the dynamics *and* the initial state must first be postulated (Hartle, 1991). They could be postulated through Bayesian reasoning, but we do not discuss such a possibility here.

Clearly we are not presenting a complete programme. We still have to argue a direct path between weakening 0b and getting necessarily to complex numbers for the history algebra. This would presumably involve a physical justification for not probabilistically comparing those statements that a complex probability allows us not to compare universally. Note, however, the analogy between Bayesianism and Gleason's theorem: (1) and (2) are effectively non-contextuality assumptions. Hence we remain hopeful that a uniqueness theorem may be forthcoming. We are also still not clear how the notion of entropy should be generalised in such domains, but we do not rule out that it also might be complex. Note that the real part of (8) behaves like a normal Bayesian probability for the LP subset so the normal formula of Shannon entropy should also behave as normal for the LP subset (Marlow, 2006).

Note also that the conditional complex probabilities defined using Bayes' theorem (11) behave in a nice manner; we have that, for all complete sets of homogeneous histories defined over the same temporal support $\{\alpha^i\}_{i=1}^{N_\alpha}$ such that β is also homogeneous, on the same temporal support, and commutes with the α^i:

$$\sum_{i=1}^{N_\alpha} p(\alpha^i \mid \beta I) = 1. \tag{18}$$

This is because we have a certain amount of distributivity for such histories:

$$\sum_i (\alpha^i \wedge \beta) = \left(\sum_i \alpha^i\right) \wedge \beta. \tag{19}$$

For example, for two-time homogeneous histories in the HPO formulation (Isham, 1994) we have that:

$$\begin{aligned}(\alpha^1 + \alpha^2) \wedge \beta &= \left(\hat{\alpha}^1_{t_2}(t_2) \otimes \hat{\alpha}^1_{t_1}(t_1) + \hat{\alpha}^2_{t_2}(t_2) \otimes \hat{\alpha}^2_{t_1}(t_1)\right) \wedge (\hat{\beta}_{t_2}(t_2) \otimes \hat{\beta}_{t_1}(t_1)) \\ &= (\alpha^1 \wedge \beta) + (\alpha^2 \wedge \beta). \end{aligned} \tag{20}$$

This property arises because history propositions are projection operators on some larger histories Hilbert space; it is a standard result that distributivity is obeyed for mutually commuting projection operators. It passes across to the class operators such that:

$$\sum_{i=1}^{N_\alpha} \frac{\mathrm{tr}(C_{\alpha^i \wedge \beta}\rho)}{\mathrm{tr}(C_\beta \rho)} = \frac{\mathrm{tr}(C_{\mathbf{1} \wedge \beta}\rho)}{\mathrm{tr}(C_\beta \rho)} = \frac{\mathrm{tr}(C_\beta \rho)}{\mathrm{tr}(C_\beta \rho)} = 1. \tag{21}$$

Thus updating by Bayes' theorem in this manner gives *a posteriori* probabilities that behave in exactly the same manner as the *a priori* ones. For complete sets of homogeneous histories that are defined over the same temporal support, they add up to 1 and are additive (as long as the *a priori* history commutes with the *a posteriori* ones). This is, of course, exactly what Bayes' theorem is about; one takes probabilities and updates them to give something that are also good probabilities.

An opposing programme would be to keep probabilities real (presumably because they are to be interpreted as frequencies), but to try and give a pedagogical justification for a non-additive measure of relative frequency (Sorkin, 1995). Such a task would probably involve frequencies which don't converge to a single value (Aerts, 2002; Anastopoulos, 2005). For frequencies that don't converge to a single value, the most natural interpretation is that we are confusing contexts somehow—this is exactly the justification given in Marlow (2006) for the LP probabilities, albeit within a Bayesian framework. However, our Bayesian analysis need not be incompatible with notions of relative frequencies as such notions (or non-convergent generalisations) should be derivable from it as in the classical case.

So, although the presumption that probabilities are complex might initially seem patently absurd, there is a certain amount of internal consistency to the argument. Feynman has often cogently argued for the use of 'negative probabilities' in physics (Feynman, 1987) but these 'negative probabilities' don't *behave* like probabilities according to Cox's axioms. The added structure provided by using complex numbers is exactly what is required in order to make a 'good' notion of probability for all complete sets of homogeneous history propositions. The real part of our complex notion can give the standard real notion of Bayesian probabilities for LP history propositions and, furthermore, can give us the standard notion of relative frequency for the d-consistent subset.

Although not yet complete, this programme includes the two main previous quantum history theories in certain limits. This Bayesian histories programme also suggests where we should look in order to find a physical justification for quantum history theories; namely we point at axiom 0b. It is also foundationally compatible with a relational philosophy.

REFERENCES

Aerts, D. (2002). Reality and probability: Introducing a new type of probability calculus. In *Probing the Structure of Quantum Mechanics: Nonlocality, Computation and Axiomatics*, D. Aerts, M. Czachor and T. Durt eds. World Scientific, Singapore, preprint: quant-ph/0205165.

Anastopoulos, C. (2004). On the relation between quantum mechanical probabilities and event frequencies. *Annals Physics*, **313**, 368, preprint: quant-ph/0403207.

Anastopoulos, C. (2005). Classical vs quantum probability in sequential measurements. preprint: quant-ph/0509019.

Caticha, A. (1998). Consistency, amplitudes and probabilities in quantum theory. *Physical Review A*, **57**, 1572, preprint: quant-ph/9804012.

Caticha, A. (2004). Relative entropy and inductive inference. *AIP Conference Proceedings*, **707**, 75–96, preprint: physics/0311093.

Caticha, A. (2005). The information geometry of space and time. Preprint: gr-qc/0508108 v1.

Cox, R. T. (1961). *The Algebra of Probable Inference*, The Johns Hopkins University Press.

Feynman, R. P. (1987). Negative probabilities. In *Quantum Implications*, B. J. Hiley and F. David Peat, eds., Routledge and Kegan Paul, p. 235.

Gell-Mann, M. and Hartle, J. B. (1990). Quantum mechanics in the light of quantum cosmology. In *Proceedings of the Third International Symposium on the Foundations of Quantum Mechanics in the Light of New Technology*, Physical Society of Japan, Tokyo, Japan, pp. 321–343.

Goldstein, S. and Page, D. N. (1995). Linearly positive history propositions: Probabilities for a robust family of sequences of quantum events. *Physical Review Letters*, **74**, 3715–3719, preprint: gr-qc/9403055.

Griffiths, R. B. (1984). Consistent history propositions and the interpretation of quantum mechanics. *Journal of Statistical Physics*, **36**, 219–273.

Hartle, J. B. (1991). Excess baggage. In *Elementary Particles and the Universe*, J. Schwarz, ed., Cambridge University Press, updated preprint: gr-qc/0508001.

Hartle, J. B. (2004). Linear positivity and virtual probability. *Physical Review A*, **70**, 022104, preprint: quant-ph/0401108.

Isham, C. J. (2003). Some reflections on the status of conventional quantum theory when applied to quantum gravity. In *Proceedings of the Conference in Honour of Stephen Hawking's birthday*, G. Gibbons ed., Cambridge University Press, preprint: quant-ph/0206090 v1.

Isham, C. J. (1994). Quantum logic and the history propositions approach to quantum theory. *Journal of Mathematical Physics*, **35**, 2157–2185, preprint: gr-qc/9308006.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*, Cambridge University Press.

Mana, P. G. L. (2004). Consistency of the Shannon entropy in quantum experiments. *Physical Review A*, **69**, 062108, preprint: quant-ph/0302049 v5.

Marlow, T. (2006). A Bayesian account of quantum histories. *Annals of Physics*, **321**, 1103–1125, preprint: quant-ph/0509149.

Marlow, T. (2006b). Relationalism vs. Bayesianism. preprint: gr-qc/0603015 v1.

Omnés, R. (1988). Logical reformulation of quantum mechanics. I. Foundations. *Journal of Statistical Physics*, **53**, 933–955.

Poulin, D. (2005). Toy model for a relational formulation of quantum theory. preprint: quant-ph/0505081 v2.

Smolin, L. (2005). The case for background independence. preprint: hep-th/0507235.

Sorkin, R. D. (1997). Quantum measure theory and its interpretation. In *Quantum Classical Correspondence: Proceedings of the 4th Drexel Symposium on Quantum Nonintegrability*, D. H. Feng and B.L. Hu, eds., International Press, Cambridge Mass., preprint: gr-qc/9507057.

Youssef, S. (1994). Quantum mechanics as complex probability theory. *Mod. Physics Lett A*, **9**, 2571.

Youssef, S. (2001). Physics with exotic probability theory. preprint: hep-th/0110253 v2.

All preprints refer to the http://arxiv.org/ website