ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/248526472

Quantum probabilities as degrees of belief

Article *in* Studies In History and Philosophy of Science Part B Studies In History and Philosophy of Modern Physics · June 2007

DOI: 10.1016/j.shpsb.2006.09.002

CITATIONS
18

READS
48

1 author:

Jeffrey Bub
University of Maryland, College Park

102 PUBLICATIONS 2,504 CITATIONS

SEE PROFILE

All content following this page was uploaded by Jeffrey Bub on 06 March 2016.

The user has requested enhancement of the downloaded file. All in-text references underlined in blue are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.

Provided for non-commercial research and educational use only.
Not for reproduction or distribution or commercial use.

VOLUME 38 Number 2 June 2007 ISSN 1355-2198

ELSEVIER

Studies in History and Philosophy of Science Part B

Studies in History and Philosophy of Modern Physics

SPECIAL ISSUE—PROBABILITIES IN QUANTUM MECHANICS

Guest Editors:
Roman Frigg, Stephan Hartmann

Available online at www.sciencedirect.com
ScienceDirect

1355-2198(200706)38:2;1-T

This article was originally published in a journal published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues that you know, and providing a copy to your institution's administrator.

All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

http://www.elsevier.com/locate/permissionusematerial

ELSEVIER

Studies in History and Philosophy of Modern Physics 38 (2007) 232–254

Studies in History and Philosophy of Modern Physics

www.elsevier.com/locate/shpsb

Quantum probabilities as degrees of belief

Jeffrey Bub

Department of Philosophy, University of Maryland, College Park, MD 20742, USA

Author's personal copy

Abstract

I outline an argument for a subjective Bayesian interpretation of quantum probabilities as degrees of belief distributed subject to consistency constraints on a quantum rather than a classical event space. I show that the projection postulate of quantum mechanics can be understood as a noncommutative generalization of the classical Bayesian rule for updating an initial probability distribution on new information, and I contrast the Bayesian interpretation of quantum probabilities sketched here with an alternative approach defended by Chris Fuchs.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Quantum probabilities; Quantum information; Quantum measurement; Quantum foundations; Projection postulate; Bayes' rule

1. Introduction

In this paper I outline an argument for a subjective Bayesian interpretation of quantum probabilities as degrees of belief distributed subject to consistency constraints on a quantum rather than a classical event space. I begin with a brief review of quantum mechanics as a noncommutative modification of classical mechanics in Section 2. In Section 3, I discuss what Pitowsky (2007) has called two 'dogmas' about quantum mechanics. The first dogma is Bell's (1990) assertion that measurement should always be given a dynamical analysis and never introduced as a primitive in a fundamental theory of mechanics. The second dogma is the view that the quantum state is a (perhaps incomplete) representation of physical reality. I argue that both dogmas are called into question by a 'no cloning' principle that distinguishes quantum information from classical information. In Section 4, I distinguish two measurement problems: a problem about individual events,

E-mail address: jbub@umd.edu.

1355-2198/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.shpsb.2006.09.002

which I characterize as a pseudo-problem, and a tractable problem about probabilities, which finds a solution in the phenomenon of decoherence. Section 5 develops an earlier argument of mine (Bub, 1977), where I show that the projection postulate of quantum mechanics can be understood as a noncommutative generalization of the classical Bayesian rule for updating an initial probability distribution on new information, and Section 6 is a critical analysis of Chris Fuchs' alternative treatment of the projection postulate as Bayesian updating. Finally, in Section 7, I consider whether the Bayesian interpretation of quantum probabilities sketched here, following Ramsey rather than de Finetti—see Howson (1990, 2003), Galavotti (1991), Ramsey (1931), de Finetti (1964)—is instrumentalist, and I contrast this with the approach defended by Fuchs (2001a, 2001b, 2002a, 2002b).

2. From classical to quantum mechanics

Quantum mechanics first appeared as a noncommutative modification of classical mechanics in the form of Heisenberg's matrix mechanics in 1925, following the 'old quantum theory'; a patchwork of *ad hoc* modifications of classical mechanics to accommodate Planck's quantum postulate. Shortly afterwards, Schrödinger developed a wave mechanical version of quantum mechanics and proved the formal equivalence of his theory and Heisenberg's. It used to be common to understand the significance of the transition from classical to quantum mechanics in terms of 'wave-particle duality,' the idea that a quantum system like an electron manifests itself as a wave under certain circumstances and as a particle under other circumstances. This picture obscures far more than it illuminates. We can see more clearly what is going on conceptually if we consider the implications of Heisenberg's move for the way we think about objects and their properties in the most general sense.

Heisenberg replaced the commutative algebra of dynamical variables of classical mechanics—position, momentum, angular momentum, energy, etc.—with a noncommutative algebra. Some of these dynamical variables take the values 0 and 1 only and correspond to properties. For example, we can represent the property of a particle being in a certain region of space by a dynamical variable that takes the value 1 when the particle is in the region and 0 otherwise. A dynamical variable like position corresponds to a set of such two-valued dynamical variables or physical properties. In the case of the position of a particle, these are the properties associated with the particle being in region R, for all regions R. If, for all regions R, you know whether or not the particle is in that region, you know the position of the particle, and conversely. The two-valued dynamical variables or properties of a classical system form a Boolean algebra, a subalgebra of the commutative algebra of dynamical variables. Equivalently, we could refer to events: the instantiation a property can be associated with the occurrence of the corresponding event.

Replacing the commutative algebra of dynamical variables with a noncommutative algebra is equivalent to replacing the Boolean algebra of two-valued dynamical variables representing properties or events with a non-Boolean algebra. The really essential thing about the classical mode of representation of physical systems in relation to quantum mechanics is that the possible properties of classical systems, or the possible events that can occur to classical systems, are represented as having the structure of a Boolean algebra. Every Boolean algebra is isomorphic to a Boolean algebra of subsets of a set. To say that the possible properties of a classical system or the corresponding events form a Boolean

Author's personal copy

algebra is to say that they can be represented as the subsets of a set, the phase space or state space of classical mechanics. To say that a physical system has a certain property or that a certain event occurs is to associate the system with a certain set in a representation space where the elements of the space—the points of the set—represent all possible states of the system. A state picks out a collection of sets, the sets to which the point representing the state belongs, as the properties of the system in that state, or the events concerning the system that occur in that state. The dynamics of classical mechanics is described in terms of a law of motion describing how the state moves in the state space. As the state changes with time, the set of properties or events selected by the state changes.

The transition from classical to quantum mechanics involves replacing the Boolean event space with the representation of possible events as a certain sort of non-Boolean algebra: essentially, the non-Boolean algebra obtained by 'pasting together'[1] various Boolean algebras in a certain way, so that the total set of elements cannot be represented as a single Boolean algebra. Dirac and von Neumann developed Schrödinger's equivalence proof into a representation theory for the properties of quantum systems as subspaces in a linear vector space over the complex numbers: Hilbert space. The non-Boolean algebra in question is the algebra of subspaces of this space. Instead of representing events as the subsets of a set, quantum mechanics represents events as the subspaces of a linear space—as lines, or planes, or hyperplanes, i.e., as a projective geometry. Algebraically, this is the central structural change in the transition from classical to quantum mechanics.

A given property is instantiated or a given event occurs if and only if the corresponding proposition is true. So we could talk equivalently in terms of propositions. In a Boolean propositional structure, there exist two-valued homomorphisms on the structure that correspond to truth-value assignments to the propositions. Each point in phase space—representing a classical state—defines a truth-value assignment to the subsets representing the propositions: each subset to which the point belongs represents a true proposition or a property that is instantiated by the system or an event that occurs, and each subset to which the point does not belong represents a false proposition or a property that is not instantiated by the system or an event that does not occur. So a classical state corresponds to a complete assignment of truth values to the propositions, or a maximal consistent catalogue of properties of the system or a maximal set of co-occurring events, and all possible states correspond to all possible maximal consistent catalogues or maximal sets of co-occurring events.

Probabilities can be introduced on a classical event space as measures on the subsets representing the events. Since each phase space point defines a truth-value assignment, the probability of an event is the measure of the set of truth-value assignments that assign a 1 ('true') to the event—in effect, we 'count' (in the measure-theoretic sense) the relative number of state descriptions in which the event occurs (or the corresponding proposition is true), and this number represents the probability of the event. So it makes sense to interpret the probability of an event as a measure of our ignorance as to whether or not the event occurs. Probability distributions over classical states represented as phase space points are sometimes referred to as 'mixed states,' in which case states corresponding to phase space points are distinguished as 'pure states.'

[1]Two elements of two different Boolean algebras are 'pasted together' in the sense that they are identified as the same element in the non-Boolean algebra.

Author's personal copy

Author's personal copy

The problem for a quantum event space arises because two-valued homomorphisms do not exist on these structures (except in the special case of a two-dimensional Hilbert space). If we take the subspace structure of Hilbert space seriously as the structural feature of quantum mechanics corresponding to the Boolean event space or propositional structure of classical mechanics, then the non-existence of two-valued homomorphisms on the algebra of subspaces of a Hilbert space means that there is no partition of the totality of events of the associated quantum system into two sets: the events that occur, and the events that do not occur; i.e., there is no partition of the totality of propositions into true propositions and false propositions. (Of course, other ways of associating propositions with features of a Hilbert space are possible, and other ways of assigning truth values, including multi-valued truth-value assignments and contextual truth-value assignments. Ultimately, the issue here concerns what we take as the salient structural change involved in the transition from classical to quantum mechanics, and this is reflected in the identification of quantum propositions or events that take the same probabilities for all quantum states.)

It might appear that, on the standard interpretation, a pure quantum state represented by a one-dimensional subspace in Hilbert space—a minimal element in the subspace structure—defines a truth-value assignment to quantum propositions in an analogous sense to the truth-value assignment to classical propositions defined by a pure classical state. Specifically, on the standard interpretation, a pure quantum state selects the propositions represented by subspaces containing the state as true, and the propositions represented by subspaces orthogonal to the state as false.[2]

There is, however, an important difference between the two situations. In the case of a classical state, every possible event represented by a phase space subset is selected as either occurring or not; equivalently, every proposition is either true or false. For a quantum state, the events represented by Hilbert space subspaces are not partitioned into two such mutually exclusive and collectively exhaustive sets: some propositions are assigned no truth value. Only propositions represented by subspaces that contain the state are assigned the value 'true,' and only propositions represented by subspaces orthogonal to the state are assigned the value 'false.' This means that propositions represented by subspaces that are at some non-zero or nonorthogonal angle to the ray representing the quantum state are not assigned any truth value in the state, and the corresponding events must be regarded as *indeterminate* or *indefinite*: according to the theory, there *can be no fact of the matter* about whether these events occur or not.

For Hilbert spaces of three or more dimensions, the possible assignments of probabilities to quantum events, i.e., weights that satisfy the usual Kolmogorov axioms for a probability measure on Boolean subalgebras of the non-Boolean algebra of quantum events, are uniquely characterized by Gleason's theorem (Gleason, 1957). An event represented by a subspace associated with a projection operator P can only be assigned the probability $\mathrm{Tr}(\rho P)$, for some density operator ρ on the Hilbert space. Density operators are pure or mixed, where mixed density operators correspond to classical probability distributions over quantum pure states: convex combinations of pure states represented by rays (one-dimensional subspaces) or vectors in Hilbert space. If ρ is a pure state, $\rho = |\psi\rangle\langle\psi|$, the probability of the event represented by P is $|\langle\psi_P|\psi\rangle|^2$, where $|\psi_P\rangle$ is the

[2]Orthogonality is the analogue of set-complement, or negation, in the subspace structure; the set-theoretical complement of a subspace is not in general a subspace.

normalized orthogonal projection of $|\psi\rangle$ onto the subspace P, i.e., the probability of the event is the square of the cosine of the angle between the vector $|\psi\rangle$ and its orthogonal projection in the subspace P (the Born rule). This means that events represented by subspaces containing the state are assigned probability 1, events represented by subspaces orthogonal to the state are assigned probability 0, and all other events, represented by subspaces at a non-zero or nonorthogonal angle to the state are assigned a probability between 0 and 1. So quantum probabilities are not represented as measures over truth-value assignments and cannot be given an ignorance interpretation in the obvious way.

3. Two dogmas

Pitowsky (2007) has pointed out that there are two assumptions or dogmas that characterize debates about the foundations of quantum mechanics. The first dogma is Bell's assertion (defended in Bell, 1990) that measurement should never be introduced as a primitive in a fundamental mechanical theory like classical or quantum mechanics, but should always be open to a complete dynamical analysis in principle. The second dogma is the view that the quantum state has an ontological significance analogous to the ontological significance of the classical state (which specifies a complete catalogue of a system's properties), i.e., that the quantum state is a (perhaps incomplete) representation of physical reality.

The second dogma leads immediately to what Pitowsky calls the 'big measurement problem,' the problem of reconciling the individual outcome of a quantum measurement with a dynamical account of how the quantum state changes in a measurement interaction. Von Neumann proposed two modes of dynamical evolution for the quantum state of a system: a deterministic, unitary evolution when a system is not measured, and a stochastic, non-unitary evolution when a system undergoes measurement, described as the 'collapse' of the wave function in configuration space and referred to more generally as the 'projection postulate.' In the light of the first dogma, this is hardly a solution to the measurement problem. A solution is provided by a theory like Bohm's hidden variable theory (Bohm, 1952; Goldstein, 2001), according to which quantum mechanics is incomplete, there is no wave-function collapse, but there is always an 'effective collapse' corresponding to one definite measurement outcome that depends on the values of the hidden variables, or the GRW theory (Ghirardi, 2002), according to which quantum mechanics is only approximately true, there is a unified stochastic dynamics in terms of which the wave function sometimes undergoes a real collapse in configuration space, and there is always one definite measurement outcome that depends on the collapse, or an Everettian many worlds interpretation of quantum mechanics (Everett, 1957; Wallace & Brown, 2005), according to which quantum mechanics is exactly true, there is no collapse, and all measurement outcomes occur in some indexical sense relative to different 'worlds' corresponding to the different terms in the linear superposition representing the quantum state.

Now, the first dogma is called into question if, as a contingent matter of fact, there is a limitation on copying information—if the dynamical implementation of a universal cloning machine is in principle excluded by structural features of events. In a 'no cloning' world, as I will show below, no complete dynamical account of a measurement process is possible in general: ultimately, a measuring instrument in a quantum measurement process simply acts as a source of classical information, i.e., it produces a probability distribution over distinguishable measurement outcomes, and how the individual outcomes come about is not subject to further dynamical analysis.

Author's personal copy

The 'no cloning' principle can be formulated quite generally as follows:

> There is no universal cloning machine.

The principle asserts that it is impossible to construct a cloning machine that will clone the output of an arbitrary information source. By contrast, note that a universal computing machine is possible: a Turing machine can simulate any other Turing machine and hence compute any computable function.

A 'universal cloning machine' is defined in the following way:

> Given an arbitrary information source, s, producing outputs e_i with probabilities p_i, a universal cloning machine is a device that can be coupled to s, where the outputs of s are inputs to the device, so that the compound device is a new information source s^* that produces outputs f_i with the same probabilities p_i, where each f_i consists of the original output e_i and a copy of e_i. A copy of an output e_i is an output e_i^* that conveys the same information as e_i, in the sense that the information source $\{p_i, e_i^*\}$ is statistically indistinguishable from the information source $\{p_i, e_i\}$ by any possible measurements on the outputs.

To be precise, the notion of 'cloning' defined here should more properly be called 'broadcasting'. The process of taking a probability distribution over an event space to a new probability distribution over a product space of events, where the marginal probability distribution over each factor space are the same as the original distribution, is called broadcasting. (This is what happens, for example, when there are corresponding events at different radio receivers, all reacting in the same way to the same varying signal, so that the marginal probability distributions of events at each receiver is the same as the joint probability of corresponding events.) I shall continue to use the term 'cloning' rather than 'broadcasting' because it is more intuitive and more familiar. But the reader should bear in mind that the process we are talking about concerns copying or cloning the outputs of an information source, not the information source itself (defined by the probability distribution).

Some comments are relevant here about the notion of an information source, and the difference between classical and quantum information as this is usually understood.

The concept of information in the physical sense was first clearly formulated by Shannon (1948). In Shannon's theory, information is a quantifiable resource associated with a (suitably idealized) stochastic source of output states, where the physical nature of the systems embodying these output states is irrelevant to the amount of information associated with the source. The quantity of information associated with an information source is defined by its optimal compressibility as given by the Shannon entropy. The fact that the output of an information source can be optimally compressed is, ultimately, what justifies the attribution of a quantifiable resource to the source.

Information is represented physically in the states of physical systems. The essential difference between classical and quantum information is associated with the different distinguishability properties of classical and quantum states. Only sets of orthogonal quantum states are reliably distinguishable (i.e., with zero probability of error), as are sets of different (pure) classical states (which are represented by disjoint singleton subsets in a phase space, and so are formally orthogonal as subsets of phase space in a sense analogous to the orthogonality of subspaces of a Hilbert space).

Classical information, then, is understood as that sort of information represented in a probability distribution over distinguishable states—pure states of classical systems, or

Author's personal copy

orthogonal quantum states—and so can be regarded as a subcategory of quantum information, where the states may or may not be distinguishable. The theory of quantum information extends Shannon's notion of compressibility to sources of quantum states, for which the von Neumann entropy turns out to be a suitable measure of information. The indistinguishability of nonorthogonal quantum states is associated with the non-unique decomposition of mixed states into specific mixtures of pure states. So the statistics of a quantum information source that produces a probability distribution of quantum states is completely defined by the density operator of the mixture, which represents a mixed state that is decomposable, in general, into different mixtures of quantum states, all statistically indistinguishable by any possible measurement.

To see that a cloning machine for all classical information sources is possible, in principle, consider first a classical information source that produces the bits[3] 0 or 1 with probabilities $p(0)$, $p(1)$. A cloning device can be constructed as follows: the device implements a 'controlled-not' Boolean gate on the contents of two 1-bit input registers, the 'control' register and the 'target' register: if the control bit is 0, the target bit keeps the same value; if the control bit is 1, the target bit is flipped. The control and target output registers contain the content of the input control register (unchanged) and the content of the (possibly altered) target register, respectively. So the transformation is as follows (where the first bit in the sequence represents the control and the second bit the target):

$$
\begin{aligned}
&00 \rightarrow 00 \\
&01 \rightarrow 01 \\
&10 \rightarrow 11 \\
&11 \rightarrow 10
\end{aligned} \tag{1}
$$

To clone the output of a classical information source, the output of the source is fed to the control input register of a controlled-not gate. The target input register is set to 0. Then, depending on the output of the information source, 0 or 1, the cloning device performs the required cloning transformation:

$$
\begin{aligned}
&00 \rightarrow 00 \\
&10 \rightarrow 11
\end{aligned} \tag{2}
$$

If the information source produces n-bit sequences, the cloning device is constructed with n-bit control and target input registers, and n 1-bit controlled-not gates which act sequentially on the n individual bits of the target input register, initially set to $00 \cdots 0$, in terms of the corresponding bit of the control input register. No principle prohibits the construction of such a cloning device, for any n-bit sequences. Think of the content of the target register as the resource used up in the copying (cf. the paper in a copying machine), and the sequence of controlled-not gates as the copying engine. In principle, the distinguishable output states of any classical information source can be represented, to arbitrary accuracy, by n-bit sequences for any n. (So, to be precise, the claim is that, for any n, there is a universal cloning machine that will clone the outputs of an arbitrary classical

[3]The term 'bit' (for 'binary digit') is used to refer to the basic unit of classical information in terms of Shannon entropy, and to an elementary two-state classical system considered as representing the possible outputs of a classical information source.

Author's personal copy

information source that produces m-bit sequences, for $m \leqslant n$, as outputs—that, for any n, there is a universal broadcasting machine for classical information sources defined by probability distributions over m-bit sequences, for $m \leqslant n$.)

To see why a cloning machine for all quantum information sources is impossible, in principle, note that a quantum controlled-not gate for qubits[4] will similarly copy the output of an information source that produces orthogonal qubit states $|0\rangle$ and $|1\rangle$, but the device will fail for quantum states that are linear superpositions of $|0\rangle$ and $|1\rangle$ states. Since any quantum gate implements a unitary, hence linear, transformation, it will produce an entangled state in the output registers for a superposition in the input control register:

$$|0\rangle|0\rangle \stackrel{U}{\rightarrow} |0\rangle|0\rangle, \tag{3}$$

$$|1\rangle|0\rangle \stackrel{U}{\rightarrow} |1\rangle|1\rangle \tag{4}$$

from which it follows that

$$|\psi\rangle|0\rangle = (c_1|0\rangle + c_2|1\rangle)|0\rangle \stackrel{U}{\rightarrow} c_1|0\rangle|0\rangle + c_2|1\rangle|1\rangle \neq |\psi\rangle|\psi\rangle \tag{5}$$

unless $c_i = 0$ or 1.

The class of classical information sources can be defined as the class of information sources for which there exists a cloning machine (i.e., one machine for the whole class). Then, for example, a Stern-Gerlach device set to measure spin in the x-direction, where the inputs are quantum states $|z+\rangle$, will produce outputs $|x+\rangle$ (0) and $|x-\rangle$ (1) with probabilities $\frac{1}{2}$, and so this would be a clonable hence classical information source. But if the $|x+\rangle$ output is sent to a Stern–Gerlach device oriented in the y-direction, the whole set-up—$|z+\rangle$ inputs and two coupled Stern-Gerlach devices, producing outputs $|x+\rangle$ with probability $\frac{1}{2}$ and $|y+\rangle, |y-\rangle$ each with probability $\frac{1}{4}$—would be a nonclassical quantum information source that cannot be cloned.

The rejection of both dogmas, as I will argue in the following section, leads to an information-theoretic interpretation of quantum mechanics. On this interpretation, the structure of Hilbert space, i.e., the non-Boolean algebra of Hilbert space subspaces, defines the structure of a quantum event space, just as, classically, a Boolean algebra, the subsets of a set (phase space), defines the structure of a classical event space. Gleason's theorem then determines all possible probability measures on this structure as given by quantum states (pure and mixed) according to the trace rule, where the probabilities are interpreted as degrees of belief or measures of uncertainty about events in the Bayesian sense.

4. The measurement problem

Richard Feynman is often quoted as saying that nobody understands quantum mechanics (Feynman, 1967, p. 129):

> There was a time when the newspapers said that only twelve men understood the theory of relativity. I do not believe there ever was such a time. ... On the other hand,

[4]By analogy with the term 'bit,' a 'qubit' refers to the basic unit of quantum information in terms of the von Neumann entropy, and to an elementary two-state quantum system considered as representing the possible outputs of a quantum information source.

Author's personal copy

> I think I can safely say that nobody understands quantum mechanics. ... Do not keep saying to yourself, if you can possibly avoid it, 'But how can it be like that?' because you will get 'down the drain,' into a blind alley from which nobody has yet escaped. Nobody knows how it can be like that.

What is it about quantum mechanics, as opposed to the theory of relativity, that raises a special problem of intelligibility?

The salient and surprising discovery leading to the special theory of relativity was that the velocity of light is independent of the motion of the source—that our world is the sort of world in which there is 'no overtaking of light by light,' as Bondi (1980) puts it. The special theory of relativity is a revision of the conceptual framework of classical mechanics that stems from Einstein's recognition of the foundational significance of this fact.

Newtonian mechanics makes a fundamental distinction between inertial motion and accelerated motion, not between motion and rest as Aristotle's theory had done. An inertial frame is a reference frame in which Newton's laws hold, i.e., bodies not under the action of forces move in straight lines with constant velocity, accelerations are proportional to applied forces and in the same direction, and all forces come in action–reaction pairs. Any system of bodies defining a reference frame at rest or moving uniformly and rectilinearly with respect to an inertial frame is also an inertial frame.

There is a contradiction between this Galilean principle of relativity characterizing the symmetry group of Newtonian spacetime, that 'velocity doesn't matter,' as Bondi (1980) puts it, and the light postulate, 'no overtaking of light by light.' Lorentz resolved the contradiction by appealing, in effect, to 'hidden variables': mechanical distortions in electromagnetic and intermolecular forces as a result of motion through the ether. On Lorentz's view, the apparent invariance of the velocity of light for all inertial frames is explained on the assumption that measuring devices moving relative to the ether contract as a result of these distorting forces in a way that conforms to the Lorentz transformation, the symmetry group of Maxwell's electrodynamics. For Lorentz, Einstein's derivation of the Lorentz transformation from the light postulate is simply an alternative, but hardly more fundamental, resolution of the contradiction.

The revolutionary significance of special relativity lies in Einstein's recognition that the concept of an inertial system in Newtonian mechanics involved definitions of space and time that were constitutive of a conceptual framework in terms of which relative motion and causal interaction were physically intelligible, and this entire conceptual framework was called into question by unexpected facts about the propagation of light. What was required, then, was a definition of simultaneity that took account of the invariance of the velocity of light as the basis for a new conceptual framework. The new definition of simultaneity proposed by Einstein is constitutive of the new conception of time and space introduced by special relativity.

Einstein's conceptual analysis, like Newton's, is in effect a transcendental argument for the spatio-temporal concepts presupposed by physics. By contrast, Lorentz's assumption about the effect of motion through the ether on electrodynamic and intermolecular forces involves an ontological hypothesis about the nature of physical reality rather than a precondition for physical inquiry. This hypothesis implicitly depends on a notion of simultaneity, and hence a spatio-temporal structure, that is not well-defined physically if the light postulate is true. Ultimately, Lorentz's rejection of Einstein's approach depends on a subjective distaste for a physical explanation that does not conform to

Author's personal copy

this hypothesis, rather than a principled objection to the conceptual framework of special relativity. For an illuminating analysis of space–time theories along these lines, see DiSalle (2006).

The salient and surprising discovery underlying the quantum theory is that there is a limitation on copying information—that our world is the sort of world in which there does not exist a cloning machine that will clone the outputs of an arbitrary information source (more precisely: there are information sources that cannot be broadcast). The quantum theory is the revised conceptual framework for mechanics that stems from the recognition of the foundational significance of this fact.

There are, in principle if not in fact, nonclassical information sources that are not quantum information sources—e.g., the information sources defined by the nonlocal boxes of Popescu and Rhorlich (2005). In part, the answer to Feynman's question will involve identifying (suitably informative) principles, in addition to 'no cloning,' that distinguish quantum information sources from other nonclassical information sources and from classical information sources. The remaining part, understanding 'how it can be like that,' is the measurement problem.

The 'no cloning' principle was first explicitly recognized by Dieks (1982) and Wootters and Zurek (1982) in 1982, almost 60 years after the emergence of quantum mechanics in the 1920s, so this claim about the quantum theory in relation to the theory of relativity would seem to be oddly anachronistic. Nevertheless, the debate between Einstein, Bohr, Heisenberg, and others about the interpretation of quantum mechanics ultimately had to do with features of quantum measurement that are puzzling just because of the conflict between the 'no cloning' principle and the conceptual framework of classical mechanics. Of course, it would be more true to historical fact to say that the quantum theory is the revised conceptual framework for mechanics that stems from the recognition of the fundamental significance of Planck's quantum postulate. But the concern here is with a conceptual rather than a historical analysis, and the relevance of the quantum postulate for a theory of mechanics is precisely a limitation on cloning that is directly associated with the noncommutativity of quantum mechanics.

There are two ways of understanding the 'no cloning' principle in quantum mechanics. Either (i) the principle is not fundamental and unrestricted cloning is possible in principle, and what prevents cloning in certain cases is some feature of the dynamics, so that there is a dynamical explanation for the fact that cloning in these cases is, as a matter of fact, impossible or practically impossible, or (ii) unrestricted cloning is impossible *in principle*. In the latter case it follows that no complete dynamical account of a measurement process is possible in general, and there must always be some system involved in the process, taken as the ultimate measuring instrument, that functions simply as a classical information source in Shannon's sense, where the occurrence of the outputs of this measuring instrument is not analyzed dynamically.

This can be seen from the following line of reasoning: suppose there exists a measurement device that functions dynamically in such a way as to identify (with certainty) the output of an arbitrary information source. That is, the device distinguishes a given output from every other possible output by undergoing a dynamical transformation that results in a state that represents a distinguishable record of the output. If we assume that any known state can be prepared from a standard reference state by some dynamical evolution, then the possibility of identifying the output of an arbitrary information source means that it must be possible to clone the output of such a source. So if unrestricted

Author's personal copy

cloning is impossible in principle, it follows that such a measurement device must be impossible in principle.

Now this entails that there exist information sources that produce outputs that cannot be distinguished in this way through measurement interactions. Since we understand a measurement device as something that produces *distinguishable* measurement outcomes ('pointer readings' or records), it follows that in such cases a measurement device must produce outcomes stochastically, over a range of measurement outcomes that is the same for different outputs of the information source, or else the outputs of the information source would be distinguishable. That is, such a measurement device itself acts as a *classical* information source that produces a range of distinguishable outputs, where *only the probabilities can depend on what is being measured*. This, then, is what we mean by a measurement device in a quantum event space.

Position (i) corresponds to Bohm's position, or Einstein's interpretation of quantum mechanics as incomplete. Position (ii) corresponds to a Bohrian position, but reformulated information-theoretically. From the impossibility of cloning, in principle, it follows that the structure of events must be non-Boolean, or else cloning would be possible—there exists a universal cloning machine for the outputs of an arbitrary information source if the outputs of all possible information sources can be represented as events in a common Boolean algebra.

The 'big measurement problem' is the problem of explaining how *individual measurement outcomes* come about dynamically. The problem arises if the quantum state is understood as representing physical reality in an analogous sense in which the classical state represents physical reality—if the quantum state is understood as the 'truthmaker' with respect to what is true and what is false about the physical world at a particular time. Then we would expect the quantum state of a system plus measuring instrument in a quantum measurement process to evolve dynamically during the course of the measurement process in such a way that, at the end of the process, the state selects a particular measurement outcome as actually occurring, i.e., the corresponding proposition is assigned the value 'true' by the state. But we have seen that this is inconsistent with the 'no cloning' principle—there can be no dynamical account of this sort for the same reason that nonorthogonal quantum states cannot be cloned, i.e., the argument here is precisely the argument of Eqs. (3)–(5), with the controlled-not gate interpreted as a measuring device rather than a cloning device.

Putting it differently, a solution to the big measurement problem, say along the lines of Bohm's hidden variable theory, is simply an attempt to provide a dynamical explanation for 'no cloning.' It is analogous to Lorentz's attempt to provide a dynamical explanation for length contraction in terms of distortions that occur to bodies as they move through the ether. But, as Einstein saw, the significance of the surprising discovery that the velocity of light is independent of the velocity of the light source is that there is simply no relation of absolute simultaneity in the world—it is not that there is a relation of absolute simultaneity, but there is some dynamical reason why events that are not absolutely simultaneous but appear simultaneous to an observer depend on the observer's state of motion. Similarly, it is only if unrestricted cloning is regarded as possible in principle that it makes sense to look for a dynamical explanation of our inability to clone nonorthogonal quantum states, having to do, say in Bohm's theory, with the form of the equation of motion and the fact that the distribution of hidden variables is assumed to have reached equilibrium (so cloning would have been possible in the early universe).

Author's personal copy

The ‘small measurement problem’ is the problem of accounting for the possibility of classical information sources in a world in which cloning is in principle impossible. The solution is provided by the phenomenon of decoherence: a particular sort of interaction between macrosystems, such as our macroscopic measuring instruments, and the environment. Decoherence effectively selects a preferred basis dynamically, i.e., a preferred Boolean subalgebra in the quantum event space. It is with respect to *this Boolean subalgebra* that the Gleason measures can be interpreted as classical probabilities. Formally, the von Neumann entropy of quantum information (of the reduced density matrix of the measured system plus measuring instrument) becomes very rapidly equivalent to the Shannon entropy of classical information in the basis in which the density matrix is diagonal, and this is effectively the decoherence basis. So we have a dynamical justification for the interpretation of the Gleason measures as *classical probability distributions on the Boolean subalgebras* of the quantum event space.

To sum up: if ‘no cloning’ is accepted as a fundamental principle, then our world must be such that there is no dynamical account of the individual occurrence of the outcome of a quantum measurement, which is to say that the world is ‘irreducibly statistical.’ But the impossibility of a dynamical account here does not entail that there can be no actually occurring measurement outcomes or actually occurring events. Rather, we *begin* with a space of possible events that the quantum theory represents as structured in a particular (non-Boolean) way. In this ‘no cloning’ world of quantum events, a particular Boolean subalgebra is dynamically selected by decoherence, and the Gleason measures are interpreted as probabilities with respect to this Boolean subalgebra. So the set of possibly occurring events to which probabilities are assigned is relativized to this Boolean subalgebra—just as, in special relativity, the set of simultaneous events is relativized to a space-like hypersurface in Minkowski space defined by an inertial system. In a quantum event space, an event, in the sense of something to which a probability is assigned, occurs or does not occur with respect to a particular Boolean subalgebra selected by decoherence.

Now, the (effective) selection of a Boolean subalgebra and the emergence of classical probability distributions through the phenomenon of decoherence depends on certain contingent dynamical features of the quantum event space that apply to a particular way (relevant to us) of ‘carving up’ the space into microsystem (localized atomic nucleus, electron, etc.), macroscopic measuring instrument (extended metallic body with a registration and recording device, such as a computer), and the environment (gases, electromagnetic radiation, etc.), and averaging over (‘tracing out’) the environmental degrees of freedom. But this requires that we reject the second dogma that the quantum state has an ontological significance analogous to the significance of the classical state in specifying which events occur and which events do not occur, and interpret the probabilities defined by the quantum state as degrees of belief subject to consistency constraints on a quantum rather than a classical event space. Then there is no inconsistency in taking an event, in the sense of something we would be prepared to bet on, as relativized to a Boolean subalgebra selected by decoherence in the non-Boolean quantum event structure. And no inconsistency in the quantum description of events is involved in ignoring (tracing out) certain information in the environment that is in practice impossible to keep track of. If we were able to keep track of more of this information, we would derive our degrees of belief from constraints on a richer non-Boolean structure of

Author's personal copy

quantum events in a similar way, but there would be no inconsistency with the 'coarser' probability assignment with respect to our expectations about the outcomes of possible measurements.

5. The projection postulate as Bayesian updating

Pitowsky (2002, 2007) has formulated an explicit subjective Bayesian interpretation of the quantum probability calculus as a logic of partial belief in terms of 'quantum gambles' defined by consistency constraints on a quantum event structure. See also Schack, Brun, and Caves (2001) and Caves, Fuchs, and Schack (2002). Here I want to show that von Neumann's projection postulate for the change induced by measurement on a quantum state, or the more general Lüders version, is in fact just a non-Boolean or noncommutative version of the classical Bayesian rule for updating an initial probability distribution on new information.[5]

For simplicity, consider a countable classical probability space $(X, \mathscr{F}, \mu)$, with atomic or elementary events $x_1, x_2, \ldots$. These are associated with singleton subsets $X_1, X_2, \ldots$ and characteristic functions $\chi_1, \chi_2, \ldots$. Denote other, possibly non-atomic events by $a, b, \ldots$.

For any probability measure μ defined by an assignment of probabilities p_i to the elementary events x_i, it is possible to introduce a density operator $\rho = \sum_i p_i \chi_i$ (where $\sum_i p_i = 1, p_i \geqslant 0$, for all i) in terms of which the probability of an event a can be represented as

$$p_\mu(a) = \sum_j \left(\sum_i p_i \chi_i(x_j) \right) \chi_a(x_j) \tag{6}$$

$$= \sum_j \rho(x_j) \chi_a(x_j) \tag{7}$$

$$= \mu(X_a). \tag{8}$$

Writing $p_\rho(a)$ for $p_\mu(a)$, we have

$$p_\rho(a) = \sum \rho \chi_a, \tag{9}$$

where a summation sign without an index is understood as summing over all the atomic events.

In terms of this density operator ρ, the conditional probability of an event b, given an event a, can be represented as

$$p_\rho(b|a) = \frac{\sum_j \rho(x_j) \chi_a(x_j) \chi_b(x_j)}{\sum_j \rho(x_j) \chi_a(x_j)} \tag{10}$$

$$= \frac{\sum \rho \chi_a \chi_b}{\sum \rho \chi_a}. \tag{11}$$

[5]The analysis can be extended to the general case of measurements represented by positive operator valued measures (POVMs), see Henderson (2007). A general measurement represented by a POVM on a system $S \in \mathscr{H}_S$ is equivalent to a projection-valued measurement on a larger Hilbert space: specifically, a projective measurement on an ancilla system $E \in \mathscr{H}_E$ suitably entangled with S. An analogous equivalence holds for classical systems. For an account of such general measurements, see the section on measurement in Nielsen & Chuang (2000) or Bub (2006).

Author's personal copy

To see this, simply notice that

$$p_\mu(b|a) = \frac{\mu(X_a \cap X_b)}{\mu(X_a)} \tag{12}$$

$$= \frac{\sum \rho\chi_a\chi_b}{\sum \rho\chi_a}. \tag{13}$$

The transition

$$\mu \to \mu', \tag{14}$$

where μ' is defined for any event b by

$$\mu'(X_b) = \frac{\mu(X_a \cap X_b)}{\mu(X_a)} \tag{15}$$

represents the classical Bayesian rule for updating an initial probability distribution on new information a. It can be justified in terms of consistency constraints by a Dutch book argument. The rule can be represented in terms of the density operator ρ as the transition:

$$\rho \to \rho' = \frac{\rho\chi_a}{\sum \rho\chi_a} \tag{16}$$

or, equivalently, in the symmetrized form

$$\rho \to \rho' = \frac{\chi_a\rho\chi_a}{\sum \chi_a\rho\chi_a} \tag{17}$$

so that

$$p_\rho(b|a) = \sum \rho'\chi_b. \tag{18}$$

Now, Eq. (17) is just the classical analogue of the von Neumann–Lüders projection postulate in quantum mechanics! Consider the case of a quantum system in a state represented by a density matrix:

$$\rho = \sum_i p_i P_i, \tag{19}$$

where the P_i are projection operators onto one-dimensional subspaces spanned by the vectors $|x_i\rangle$ representing atomic events, and the p_i are probabilities. So ρ here is a convex combination of atomic projection operators of the system, just as in the classical case the density operator is a convex combination of characteristic functions (which are projection operators in the commutative algebra of classical dynamical variables).

In terms of the density operator, the probability of an event a is given by

$$p_\rho(a) = \mathrm{Tr}(\rho P_a). \tag{20}$$

Note that the trace of an operator O is just the sum of the eigenvalues of O, i.e., the sum of the possible values of O at each atom in the Boolean subalgebra defined by O. So $\mathrm{Tr}(\rho P_a)$ is the noncommutative analogue of $\sum \rho\chi_a$.

After a measurement of an observable A with outcome a, the conditional probability of an event b, relative to an initial probability assignment given by ρ, is

$$p_\rho(b|a) = \mathrm{Tr}(\rho' P_b), \tag{21}$$

Author's personal copy

where

$$\rho' = \frac{P_a \rho P_a}{\mathrm{Tr}(P_a \rho P_a)}. \tag{22}$$

That is, since the projection operators P_a are the noncommutative analogues of the characteristic functions χ_a, the transition $\rho \to \rho'$ in (22), which is the quantum projection postulate, is just the Bayesian rule (17) for updating a probability distribution on new information. It can be justified in the quantum case as a rule for updating beliefs on new information on the basis of consistency constraints by an analysis in terms of quantum gambles.

Note that if the prior probability assignment of an observer about to make a measurement on a quantum system S is taken as $\rho_S = I_S/\mathrm{Tr}_S(I_S)$, where I_S is the identity on the Hilbert space of S, then updating ρ_S on the new information a obtained in the measurement yields the transition:

$$\rho_S \to \rho'_S = \frac{P_a \rho_S P_a}{\mathrm{Tr}_S(P_a \rho_S P_a)} \tag{23}$$

$$= \frac{P_a}{\mathrm{Tr}_S(P_a)}. \tag{24}$$

If a, b are atomic events, so $P_a = |a\rangle\langle a|, P_b = |b\rangle\langle b|$ are projection operators onto one-dimensional subspaces spanned by the vectors $|a\rangle, |b\rangle$, respectively, then the probability of the event b conditional on the event a, where b may be incompatible with a and so belong to a different Boolean subalgebra in the non-Boolean event space, is computed according to the rule

$$p_{\rho_S}(b|a) = p_{\rho'_S}(b) \tag{25}$$

$$= \mathrm{Tr}_S(P_a P_b) \tag{26}$$

$$= |\langle b|a\rangle|^2. \tag{27}$$

So the 'transition probability' between the pure states $|a\rangle$ and $|b\rangle$ can be interpreted as the conditional probability of the event b occurring in the Boolean subalgebra selected by decoherence in a suitable measurement process on a system about which we have the updated information that the event a has occurred (or, equivalently, the conditional probability of a given b), relative to a prior probability assignment given by the equiprobable density operator $\rho_S = I_S/\mathrm{Tr}_S(I_S)$.

Note that if S is obtained by placing a 'cut' between the system of interest and the rest of the environment E, and $S + E$ is in some pure entangled state $|\Psi\rangle$, it follows from a result by Popescu, Short, and Winter (2005)[6] that $\rho_S = \mathrm{Tr}_E(|\Psi\rangle\langle\Psi|) \approx I_S/\mathrm{Tr}_S(I_S)$, for any pure state $|\Psi\rangle$.

6. Fuchs on the projection postulate as Bayesian updating

In Fuchs (2002b), Chris Fuchs presents a different analysis of the status of the projection postulate as Bayesian updating, associated with a very different account of quantum probabilities as degrees of belief than the view I have sketched above.

[6]See also Goldstein, Lebowitz, Tumulka, & Zanghi (2005).

Author's personal copy

He begins by presenting a general formulation of the state change following a quantum measurement of a POVM $\{E_d\}$ with outcome d on a system S as

$$\rho \rightarrow \rho_d = \frac{1}{\mathrm{Tr}(\rho E_d)} \sum_i A_{di} \rho A_{di}^{\dagger}, \tag{28}$$

where

$$\sum_i A_{di}^{\dagger} A_{di} = E_d. \tag{29}$$

As Fuchs points out, this is completely general: there is no constraint on the number of indices i and the operators A_{di} need not even be Hermitian.

As a special case, of course, the index i could take a single value and E_d could be a projection operator P_d (in which case $A_d = P_d$). Then ρ_d reduces to $(1/\mathrm{Tr}(\rho P_a))P_a \rho P_a$.

Fuchs now considers this special case and comments (Fuchs, 2002b, pp. 29–30):

> Let us take a moment to think about this special case in isolation. What is distinct about it is that it captures in the extreme a common folklore associated with the measurement process. For it tends to convey the image that measurement is a kind of gut-wrenching violence: In one moment the state is $\rho = |\psi\rangle\langle\psi|$, while in the very next it is a $\Pi_i = |i\rangle\langle i|$. Moreover, such a wild transition need depend on no details of $|\psi\rangle$ and $|i\rangle$; in particular the two states may even be almost orthogonal to each other. In density-operator language, there is no sense in which Π_i is contained in ρ: the two states are in distinct places of the operator space. That is,
>
> $$\rho \neq \sum_i p(i) \Pi_i. \tag{30}$$
>
> Contrast this with the description of information gathering that arises in Bayesian probability theory. There, an initial state of belief is captured by a probability distribution $p(h)$ for some hypothesis H. The way gathering a piece of data d is taken into account in assigning one's new state of belief is through Bayes' conditionalization rule. That is to say, one expands $p(h)$ in terms of the relevant joint probability distribution and picks off the appropriate term:
>
> $$p(h) = \sum_d p(h, d) \tag{31}$$
>
> $$= \sum_d p(d) p(h|d) \tag{32}$$
>
> $$p(h) \xrightarrow{d} p(h|d), \tag{33}$$
>
> where $p(h|d)$ satisfies the tautology
>
> $$p(h|d) = \frac{p(h, d)}{p(d)}. \tag{34}$$

He asks (Fuchs, 2002b, p. 30): 'Why does quantum collapse not look more like Bayes' rule? Is quantum collapse really a more violent kind of change, or might it be an artifact of a problematic representation?' His answer to this question is as follows:

Author's personal copy

If we suppress the index i, then

$$\rho_d = \frac{1}{p(d)} A_d \rho A_d^\dagger, \tag{35}$$

$$E_d = A_d^\dagger A_d. \tag{36}$$

Now, in general,

$$\rho \neq \sum_d p(d) \rho_d \tag{37}$$

while classically we have

$$p(h) = \sum_d p(d) p(h|d) \tag{38}$$

So Fuchs writes

$$\rho = \sum_d p(d) \tilde{\rho}_d, \tag{39}$$

where

$$\tilde{\rho}_d = \frac{1}{p(d)} \rho^{1/2} E_d \rho^{1/2}. \tag{40}$$

Note that ρ_d and $\tilde{\rho}_d$ have the same eigenvalues.

Now, Fuchs points out (Fuchs, 2002b, p. 34) that the state change following a quantum measurement of a POVM $\{E_d\}$ can be presented as a two-stage process:

> First one imagines an observer refining his initial state of belief and simply plucking out a term corresponding to the 'data' collected:
>
> $$\rho = \sum_d p(d) \tilde{\rho}_d, \tag{41}$$
>
> $$\rho \xrightarrow{d} \tilde{\rho}_d. \tag{42}$$
>
> Finally, there may be a further 'mental readjustment' of the observer's beliefs, which takes into account details both of the measurement interaction and the observer's initial quantum state. This is enacted via some (formal) unitary operator V_d:
>
> $$\rho_d \rightarrow \tilde{\rho}_d = V_d \rho_d V_d^\dagger. \tag{43}$$

Fuchs contrasts this representation of the two-stage process with the following alternative representation. Since one can write

$$\rho_d = \frac{1}{p(d)} A_d \rho A_d^\dagger \tag{44}$$

$$= \frac{1}{p(d)} U_d E_d^{1/2} \rho E_d^{1/2} U_d^\dagger \tag{45}$$

for some unitary operator U_d (using the polar decomposition theorem; see Nielsen & Chuang, 2000, Theorem 2.3, p. 78), the state transformation in a measurement of a POVM

Author's personal copy

$\{E_d\}$ with outcome d can be represented as a 'raw collapse':

$$\rho \to \sigma_d = \frac{1}{p(d)} E_d^{1/2} \rho E_d^{1/2} \tag{46}$$

followed by a further 'back action' or 'feedback' of the measuring instrument on the measured system:

$$\sigma_d \to \rho_d = U_d \sigma_d U_d^{\dagger}. \tag{47}$$

He notes that these different representations of the measurement transformation are purely conceptual.

Now, Fuchs' claim (Fuchs, 2002b, p. 34) is that on his formulation of the two-stage process, quantum collapse can be seen as 'not such a violent state of affairs after all,' but rather as nothing more than a refinement and readjustment of one's initial state of belief. He considers two limiting cases as support for this claim.

In the first type of case, the observer's initial state of belief is maximal, represented by a pure state

$$\rho = |\psi\rangle\langle\psi| \tag{48}$$

for the system. Here no refinement is possible because for any $\{E_d\}$, $\rho_d = \rho^{1/2} E_d \rho^{1/2} = p(d)|\psi\rangle\langle\psi|$ and the only state change can be a 'mental readjustment.' We can learn nothing new from a measurement—we only change what we can predict as a consequence of our experimental intervention. Here the measurement is solely disturbance, resulting in a state transition of the 'violent' sort. In the case of a projective measurement $\{E_d\} = \{\Pi_d = |d\rangle\langle d|\}$, where the Π_d are projection operators, the state change on measurement with outcome d is a collapse corresponding to a readjustment by some unitary operator U_d that takes $|\psi\rangle$ to $|d\rangle$.[7]

The second type of case involves a measurement on one of a pair of separated systems, A and B, in an entangled state. A measurement on A leads to a change in the state of B that is purely a Bayesian updating with no further readjustment. Fuchs considers a pure state of the system $A+B$ which takes the Schmidt decomposition:

$$|\psi^{AB}\rangle = \sum_i \sqrt{\lambda_i}|a_i\rangle|b_i\rangle. \tag{49}$$

A measurement of the POVM $\{E_d = A_d A_d^{\dagger}\}$ on A yields

$$|\psi^{AB}\rangle\langle\psi^{AB}| \to \rho_d = \frac{1}{p(d)}(A_d \otimes I)|\psi^{AB}\rangle\langle\psi^{AB}|(A_d^{\dagger} \otimes I). \tag{50}$$

Tracing over A yields

$$\mathrm{Tr}_A(\rho_d) = \frac{1}{p(d)} \rho^{1/2}(UA_d^{\dagger}A_d U^{\dagger})^{\mathrm{T}} \rho^{1/2}, \tag{51}$$

where ρ is the initial state of B, U is the unitary operator connecting the $\{|a_i\rangle\}$ basis with the $\{|b_i\rangle\}$ basis, and $^{\mathrm{T}}$ represents the transpose with respect to the $\{|a_i\rangle\}$

[7] Fuchs writes $U_d = |d\rangle\langle\psi|$, see Fuchs, 2002b, Eq. (34), but this operator is not unitary, as Palge & Konrad, 2006 point out.

Author's personal copy

basis. Since

$$(UA_d^{\dagger}A_dU^{\dagger})^{\mathrm{T}} = F_d, \tag{52}$$

where $\{F_d\}$ is a POVM, the result follows (cf. (40)).

Fuchs concludes from this analysis that quantum collapse can be regarded as a noncommutative version of Bayes's rule. Now, granted, the purely 'gentle' selection from an initial density operator of a term corresponding to the outcome of a measurement is just like classical Bayesian conditionalization, when we update an initial probability distribution on new information. But the 'violent' transition cannot be attributed to a mechanical disturbance, as a result of a dynamical interaction between a system and a measuring instrument. A dynamical interaction leads to an entangled state, not to the state obtained by 'mental readjustment.' The disturbance is 'uncontrollable,' to use Bohr's terminology, and reflects the constraints imposed on conditionalization in a quantum event structure. So the 'violent' collapse transition described by the projection postulate too is an effect of conditionalization.

Consider the following example of a measurement on one of a pair of systems, A and B, in an entangled state $|\psi^{AB}\rangle$, where the Hilbert space of B is two-dimensional and the Hilbert space of A is three-dimensional. I shall refer to A as the ancilla system, because projection-valued measurements on the ancilla A alone correspond to POVM measurements (with three possible values) on B. Consider a case where the Schmidt decomposition (49) takes the form

$$|\psi^{AB}\rangle = \frac{1}{\sqrt{2}}\left(\frac{2|a_1\rangle - |a_2\rangle - |a_3\rangle}{\sqrt{6}}|b_1\rangle + \frac{|a_2\rangle - |a_3\rangle}{\sqrt{2}}|b_2\rangle\right), \tag{53}$$

where $\{|a_1\rangle, |a_2\rangle, |a_3\rangle\}$ is an orthonormal basis in $\mathscr{H}_A$ and $|b_1\rangle, |b_2\rangle$ is an orthonormal basis in $\mathscr{H}_B$.

It is easy to see that $|\psi^{AB}\rangle$ can be expressed as

$$|\psi^{AB}\rangle = \frac{1}{\sqrt{3}}(|a_1\rangle|b_1\rangle + |a_2\rangle|c\rangle + |a_3\rangle|d\rangle), \tag{54}$$

where

$$|c\rangle = -\frac{1}{2}|b_1\rangle + \frac{\sqrt{3}}{2}|b_2\rangle, \tag{55}$$

$$|d\rangle = -\frac{1}{2}|b_1\rangle - \frac{\sqrt{3}}{2}|b_2\rangle. \tag{56}$$

The states $|b_1\rangle, |c\rangle, |d\rangle$ are non-orthogonal states in $\mathscr{H}_B$. Note that

$$\frac{1}{3}|b_1\rangle\langle b_1| + \frac{1}{3}|c\rangle\langle c| + \frac{1}{3}|d\rangle\langle d| = \frac{I_B}{2} \tag{57}$$

i.e., the state of B (obtained by tracing over $\mathscr{H}_A$) is the completely mixed state $\rho_B = \frac{1}{2}I_B$.

Similarly, $|\psi^{AB}\rangle$ can be expressed as

$$|\psi^{AB}\rangle = \frac{1}{\sqrt{3}}(|a_1'\rangle|b_2\rangle + |a_2'\rangle|e\rangle + |a_3'\rangle|f\rangle), \tag{58}$$

Author's personal copy

where

$$|e\rangle = \frac{\sqrt{3}}{2}|b_1\rangle - \frac{1}{2}|b_2\rangle, \tag{59}$$

$$|f\rangle = -\frac{\sqrt{3}}{2}|b_1\rangle - \frac{1}{2}|b_2\rangle \tag{60}$$

and

$$|a'_1\rangle = \tfrac{1}{3}|a_1\rangle + (1+\sqrt{3})|a_2\rangle + (1-\sqrt{3})|a_3\rangle, \tag{61}$$

$$|a'_2\rangle = \tfrac{1}{3}(1+\sqrt{3})|a_1\rangle + (1-\sqrt{3})|a_2\rangle + |a_3\rangle, \tag{62}$$

$$|a'_3\rangle = \tfrac{1}{3}(1-\sqrt{3})|a_1\rangle + |a_2\rangle + (1+\sqrt{3})|a_3\rangle \tag{63}$$

is an orthonormal basis in $\mathscr{H}_A$. The states $|b_2\rangle, |e\rangle, |f\rangle$ are non-orthogonal states in $\mathscr{H}_B$ and, of course,

$$\frac{1}{3}|b_2\rangle\langle b_2| + \frac{1}{3}|e\rangle\langle e| + \frac{1}{3}|f\rangle\langle f| = \frac{I_B}{2}. \tag{64}$$

Now suppose the observable with eigenstates $\{|a_1\rangle, |a_2\rangle, |a_3\rangle\}$ is measured on the ancilla A. Depending on the outcome, the system B will be left in one of the states $|b_1\rangle, |c\rangle, |d\rangle$, and since the completely mixed state $\rho_B = \frac{1}{2}I_B$ can be regarded as an equal weight mixture of $|b_1\rangle, |c\rangle, |d\rangle$, the change in the state of B as a result of the measurement will be of the 'gentle' sort, representing a pure refinement of the observer's beliefs. Similarly, if the observable with eigenstates $\{|a'_1\rangle, |a'_2\rangle, |a'_3\rangle\}$ is measured on A, the system B will be left in one of the states $|b_2\rangle, |e\rangle, |f\rangle$, and since the completely mixed state can equivalently be regarded as an equal weight mixture of $|b_2\rangle, |e\rangle, |f\rangle$, the change in the state of B as a result of the measurement will again be a pure refinement of the observer's beliefs. This is simply because the mixed state $\rho_B = \frac{1}{2}I_B$ corresponds to an infinite variety of mixtures of states in $\mathscr{H}_B$ (not necessarily equal weight mixtures, of course). Any one of these mixtures can be obtained by a suitable measurement on an ancilla system entangled with B. This is the content of the Hughston–Jozsa–Wootters theorem (Hughston, Jozsa, & Wootters, 1993). It is what Schrödinger called 'remote steering' and is the basis of quantum teleportation, quantum dense coding, and other peculiarities of quantum information, including the impossibility of unconditionally secure quantum bit commitment (see Bub, 2006).

By contrast, consider the effect of the measurement on A. The initial state of A is obtained by tracing $|\psi^{AB}\rangle$ over B. This yields the mixed state:

$$\rho_A = \tfrac{1}{2}|g\rangle\langle g| + \tfrac{1}{2}|h\rangle\langle h|, \tag{65}$$

where

$$|g\rangle = \frac{2|a_1\rangle - |a_2\rangle - |a_3\rangle}{\sqrt{6}}, \tag{66}$$

$$|h\rangle = \frac{|a_2\rangle - |a_3\rangle}{\sqrt{2}}. \tag{67}$$

Note that $\rho_A \neq \frac{1}{3}I_A$. In fact, ρ_A has support on a two-dimensional subspace in the three-dimensional Hilbert space $\mathscr{H}_A$: the plane spanned by $|g\rangle$ and $|h\rangle$. A measurement of either

Author's personal copy

Author's personal copy

of the maximal (non-degenerate) observables with eigenvectors $\{|a_1\rangle, |a_2\rangle, |a_3\rangle\}$ or $\{|a'_1\rangle, |a'_2\rangle, |a'_3\rangle\}$ will result in a state that has a component outside this plane. So the state change on measurement will necessarily be partly of the 'violent' sort, i.e., it will involve at least in part a readjustment of beliefs and cannot be a simple refinement.

The above example simply *illustrates* certain features of the projection postulate in quantum mechanics. It does not show how quantum collapse can be understood as arising from a non-Boolean or noncommutative version of Bayes' rule for updating states of belief. All that Fuchs' analysis in terms of POVMs shows is that the relevant features of my example are quite general. What is not explained is how the 'uncontrollable' disturbance in a quantum measurement process can be attributed to the constraints on Bayesian updating in a non-Boolean event structure.

7. Instrumentalism?

On the view proposed here, an objective feature of the world, the fact that events are structured in a non-Boolean way, underlies a limitation on copying information, formulated as the 'no cloning' principle. Consistency constraints on how partial beliefs are distributed on a quantum event structure yield a nonclassical probability theory. This accords with Ramsey's 'big idea' (see (Howson, 1990)) that the laws of probability are laws of consistency.

Is this information-theoretic interpretation of quantum mechanics simply instrumentalist? It is no more instrumentalist than the special theory of relativity is instrumentalist relative to Lorentz's theory of the electron. If the rejection of the 'big measurement problem' as a pseudo-problem is instrumentalist, then it is a principled instrumentalism, entailed by the acceptance of 'no cloning' as a fundamental principle, just as the rejection of absolute simultaneity is entailed by accepting the light postulate as a fundamental principle. If it is indeed the case that there is no universal cloning machine—and, of course, we may ultimately be mistaken about this—then a measurement cannot be the sort of process we thought it was in classical physics, just as simultaneity cannot be the sort of relation we thought it was in classical physics if there is no overtaking of light by light.

On Fuchs' view, quantum states represent subjective degrees of belief, and quantum collapse involves Bayesian conditionalization as the straightforward refinement of prior degrees of belief in the usual sense, together with a readjustment of the observer's beliefs required because, as Fuchs puts it (Fuchs, 2002b, p. 8): 'The world is sensitive to our touch.' It seems that for Fuchs, as for de Finetti (see Galavotti, 1991), physics is an extension of common sense and can only be relevant to the extra-logical and context-dependent *evaluation* of probabilities. But then, if the constraints on the distribution of partial belief about quantum events do not reflect structural features of the world that we take to be characteristic of a quantum event space, it seems that what is doing the work in Fuchs' Bayesian analysis of quantum collapse is ultimately an instrumentalist interpretation of quantum probabilities.

Acknowledgments

This paper represents the results of research undertaken during the tenure of a University of Maryland General Research Board semester award in 2005, and as a long-term visiting researcher at the Perimeter Institute for Theoretical Physics in Waterloo, Canada, in 2006. Discussions with Itamar Pitowsky have been extremely helpful in formulating the Bayesian view developed here.

References

Bell, J. S. (1990). Against measurement. *Physics World*, *8*, 33–40 [reprinted In Arthur Miller (Ed.), Sixty-two years of uncertainty: Historical, philosophical and physical inquiries into the foundations of quantum mechanics, New York: Plenum, pp. 17–31].

Bohm, D. (1952). A suggested interpretation of quantum theory in terms of 'hidden' variables. i and ii. *Physical Review*, *85*, 166–193.

Bondi, H. (1980). *Relativity and common sense*. New York: Dover Publications.

Bub, J. (1977). Von Neumann's projection postulate as a probability conditionalization rule in quantum mechanics. *Journal of Philosophical Logic*, *6*, 381–390.

Bub, J. (2006). Quantum information and computation. In *Handbook of philosophy of physics*. Amsterdam: North-Holland, arXiv e-print quant-ph/0512125.

Caves, C. M., Fuchs, C. A., & Schack, R. (2002). Quantum probabilities as Bayesian probabilities. *Physical Review A*, *65*, 022305.

de Finetti, B. (1964). Foresight: It's logical laws its subjective sources. In H. Kyburg, & H. Smokler (Eds.), *Studies in subjective probability* (pp. 93–159). New York: Wiley.

Dieks, D. (1982). Communication by EPR devices. *Physics Letters A*, *92*, 271–272.

DiSalle, R. (2006). *Understanding space– time*. Cambridge: Cambridge University Press.

Everett, H. (1957). 'Relative state' formulation of quantum mechanics. *Reviews of Modern Physics*, *29*, 454–462.

Feynman, R. P. (1967). *The character of physical law*. Cambridge, MA: MIT Press [first published in 1965 by the British Broadcasting Corporation].

Fuchs, C. A. (2001a). Notes on a Paulian idea: Foundational, historical, anecdotal and forward-looking thoughts on the quantum. arXiv e-print quant-ph/0105039.

Fuchs, C. A. (2001b). Quantum foundations in the light of quantum information. In A. Gonis (Ed.). *Proceedings of the NATO advanced research workshop on decoherence and its implications in quantum computation and information transfer*. arXiv e-print quant-ph/0106166.

Fuchs, C. A. (2002a). *The anti-Växjo interpretation of quantum mechanics*. arXiv e-print quant-ph/0204146.

Fuchs, C. A. (2002b). *Quantum mechanics as quantum information* (*and only a little more*). arXiv e-print quant-ph/0205039.

Galavotti, M. C. (1991). The notion of subjective probability in the work of Ramsey and de Finetti. *Theoria*, *LVII*(3), 239–259.

Ghirardi, G. -C. (2002). Collapse theories. In E. Zalta (Ed.). The Stanford encyclopedia of philosophy. ⟨http://plato.stanford.edu/entries/qm-collapse/⟩.

Gleason, A. (1957). Measures on the closed sub-spaces of hilbert spaces. *Journal of Mathematics and Mechanics*, *6*, 885–893.

Goldstein, S. (2001). Bohmian mechanics. In E. Zalta (Ed.). The Stanford encyclopedia of philosophy. ⟨http://plato.stanford.edu/entries/qm-bohm⟩.

Goldstein, S., Lebowitz, J. L., Tumulka, R., & Zanghi, N. (2005). *Canonical typicality*. arXiv e-print quant-ph/0511091.

Henderson, L. (2007). Quantum measurement as Bayesian updating. In A. Bokulich (Ed.). *Foundations of quantum information and entanglement*.

Howson, C. (1990). Ramsey's big idea. In D. H. Mellor (Ed.). F.P. Ramsey: Philosophical papers. Cambridge: Cambridge University Press.

Howson, C. (2003). Probability and logic. *Journal of Applied Logic*, *1*, 151–165.

Hughston, L. P., Jozsa, R., & Wootters, W. K. (1993). A complete classification of quantum ensembles having a given density matrix. *Physics Letters A*, *183*, 14–18.

Nielsen, M. A., & Chuang, I. (2000). *Quantum computation and quantum information*. Cambridge: Cambridge University Press.

Palge, V., & Konrad, T. (2006). *A remark on Fuchs' Bayesian interpretation of quantum mechanics*. arXiv e-print.

Pitowsky, I. (2002). *Betting on the outcomes of measurements*: *A Bayesian theory of quantum probability*. arXiv e-print quant-ph/0208121.

Author's personal copy

Pitowsky, I. (2007). Quantum mechanics as a theory of probability. In W. Demopoulos, & I. Pitowsky (Eds.). *Festschrift in honor of Jeffrey Bub*. Western Ontario Series in Philosophy of Science. New York: Springer, arXiv e-print quantph/0510095.

Popescu, S., & Rhorlich, D. (2005). Nonlocality as an axiom. *Foundations of Physics*, *24*, 379.

Popescu, S., Short, A. J., & Winter, A. (2005). *Entanglement and the foundations of statistical mechanics*. arXiv e-print quant-ph/0511225.

Ramsey, F. P. (1931). Truth and probability. In *The foundations of mathematics and other logical essays*. London: Routledge and Kegan Paul (pp. 156–199).

Schack, R., Brun, T. A., & Caves, C. M. (2001). Quantum Bayes rule. *Physical Review A*, *64*, 014305–014308.

Shannon, C. E. (1948). A mathematical theory of computation. *Bell System Technical Journal* 27, 379–423, 623–656.

Wallace, D., & Brown, H. (2005). Solving the measurement problem: de-Broglie–Bohm loses out to Everett. *Foundations of Physics*, *35*, 517–540.

Wootters, W. K., & Zurek, W. H. (1982). A single quantum cannot be cloned. *Nature*, *299*, 802–803.

Author's personal copy

View publication stats