How can the appropriate objective and predictive probabilities get into non-collapse quantum mechanics?

Roberto H. Schonmann

Abstract It is proved that in non-collapse quantum mechanics the state of a closed system can always be expressed as a superposition of states all of which describe histories that conform to Born's probability rule. This theorem allows one to see Born probabilities in non-collapse quantum mechanics as an appropriate predictive tool, implied by the theory, provided an appropriate version of the superposition principle is included in its axioms
Key words and phrases: Non-collapse quantum mechanics, Everett, Born's rule, origin of probability in quantum mechanics.

1 Introduction

This is a shorter version of the paper [17], where the reader will find a much more detailed and thorough discussion of the relevance of the theorem introduced here, as well as further comparison of the role of probabilities in collapse and non-collapse quantum mechanics. This version is being written in memory of Vladas Sidoravicius, whose premature death shocked and saddened his friends and colleagues, and whose interests focused on probability theory not only in the abstract, but especially as it relates to physics. Vladas' passing happened close in time to that of Harry Kesten, good friend and mentor to both of us and to so many others. This paper is also dedicated to his memory.

For mathematicians who may need an introduction to quantum mechanics, I recommend the text [11]. (Chapters 1 and 3 suffice for the purposes of this paper.)

This paper deals with an important aspect of what is known as the "measurement problem in quantum mechanics". In standard quantum mechanics the state of a system (which is a vector in a Hilbert space) evolves in two distinct and incompatible fashions, and it is unclear when each one applies. When it is not being observed it

Dept. of Mathematics, University of California at Los Angeles, CA 90095, USA, e-mail: rhs@math.ucla.edu

evolves in a deterministic, continuous way, according to certain unitary transformations. But when observed, the system evolves in a probabilistic, discontinuous way (called a collapse, or reduction of the state), jumping to a new state according to a probabilistic prescription called Born's rule (we will refer to this as the collapse axiom, or the probability axiom). But what constitutes an "observation"? After all the "observers" (whether we are able to include in this class humans, other animals, robots, photographic plates, ...) should be considered as part of the system, so that "observations" should not have distinct physics. Non-collapse quantum mechanics (first introduced in [9]) proposes that the collapse axiom be eliminated from quantum mechanics, and claims that we would still have the same experiences that we predict from standard quantum mechanics. Instead of collapses happening, the system always evolves in the deterministic unitary fashion, and this implies that at the end of each experiment all the possible outcomes materialize, including one version of the "observers" (possibly humans) associated to each outcome, perceiving and recording that outcome and no other. This accounts for our observation of collapses as illusions, so to speak. But then, what accounts for them following Born's probability rule, rather than some other probability rule, or no probability rule at all? This is the focus of this paper (see below for references and some comments on the extensive work already available on this fundamental issue).

Before we proceed, a few words about terminology. We will use the expression "Born-rule collapse quantum mechanics" for the standard quantum mechanics theory, as presented in our textbooks, including the assumption that measurements lead to collapses of the state of the system according to Born's rule. "Collapse quantum mechanics" will be used for a broader set of theories, in which the collapses follow some probability distribution that may or not be the one given by Born's rule. And by "non-collapse quantum mechanics" we simply mean that we eliminate the assumption of collapse when measurements are performed. In non-collapse quantum mechanics, we do not include the words "measurement" or "observation" in the axioms of the theory, and use them only informally when applying the theory to explain and predict our experiences.

Readers who want an introduction to non-collapse quantum mechanics will benefit from the classic [8], where papers by those who first proposed and advertised it as a (better) alternative to collapse quantum mechanics are collected. The subject is not standard in textbooks geared to physicists, or mathematicians, but is standard in texts concerned with the philosophy of quantum mechanics; see, e.g., [2], [3], [21], [4] and [12]. For a positive appraisal of the theory, written for the general scientific public, see, e.g., [19]. For expositions for the general public, see, e.g., [5], [20] and [6]. For a recent collection of mostly philosophical discussions see [16]. And for some among the many research papers on the subject, see, e.g., [7], [1], [18] and [13], which also provide extensive additional references.

Our concern here is with the origin of our perception of Born-rule probabilities in a theory, non-collapse quantum mechanics, in which everything is deterministic and, in particular, no probabilities are introduced in its axioms. A great deal has been written about this problem, e.g., in the references cited in the last paragraph and references therein, with opinions ranging from "the problem is solved" (sometimes by the authors themselves), to "the problem is hopeless and the proposed solutions all flawed". This project was motivated by my dissatisfaction with the previously proposed solutions, especially with the current trend of treating the probabilities in non-collapse quantum mechanics as subjective ones ([7], [21], [18], [13]; see for instance Chapter 6 of [12] for a criticism). I hope nevertheless to convince the reader that the theorem stated and proved here provides a solution to this puzzle and explains how our perception of probabilities, as given by Born's rule, emerges in non-collapse quantum mechanics, if we include in its axioms an appropriate version of the superposition principle. I propose even that the puzzle be turned around: If collapses do happen, why do they happen precisely with the same rule that comes out of quantum mechanics without collapse?

In Section 2 we will state the theorem alluded to in the abstract, in a mathematically self-contained fashion, but without emphasizing the corresponding physics, which will then be briefly discussed in Section 3. (For a longer discussion the reader is referred to [17].) To help the reader keep in mind what is planned, we include next a few words of introduction on how the mathematical setting in Section 2 is motivated by collapse quantum mechanics.

We will be working in the Heisenberg picture (operators evolve in time, rather than states), as applied to a closed system (possibly the whole universe). Associated to the system there is a Hilbert space $\mathscr{H}$ (not assumed in this paper to be necessarily separable). The state of the system is given at any time by a non-null vector in $\mathscr{H}$ (with non-null scalar multiples of a vector corresponding to the same state). This state does not change with time except when there is a collapse. Collapses are associated with measurements and with their corresponding self-adjoint operators (which in the Heisenberg picture are time dependent). In each collapse, the state immediately after the collapse is a projection of the state immediately before the collapse on a subspace (a subset of $\mathscr{H}$ closed linearly and topologically) chosen at random, according to a specified probability law (in the standard case, Born's rule), from among the eigenspaces of that operator, one eigenspace for each possible outcome of the experiment. (To avoid unnecessary mathematical complications, and on physical grounds, we are assuming that every experiment can only have a finite number of possible outcomes.) To each subspace of $\mathscr{H}$ there is associated a projection operator (self-adjoint idempotent operator on $\mathscr{H}$) that projects on that subspace. If initially the state was a vector $\psi \in \mathscr{H}$, then immediately after a collapse the state can be expressed as $\mathrm{Proj}\,\psi$, where Proj is the composition of the projections that took place after each collapse, up to and including this last one.

It is natural to represent all the possible ways in which the system can evolve using a rooted (oriented) tree. The root vertex of the tree will correspond to the beginning of times for the system under study, and the other vertices will either correspond to collapse events, or be terminal vertices (vertices of degree 1) that signal that no further experiment is performed along a branch of the tree. (In the interesting cases the tree will be infinite. One can think of terminal vertices as uncommon in the tree, possibly even absent.) The projections associated to the possible outcomes in the collapses, as described at the end of the last paragraph, will then be indexed by the edges of the tree. The tree does not have to be homogeneous, as, e.g., decisions on what experiments to perform in a lab may depend on the outcomes of previous experiments. More interesting and dramatic examples of non-homogeneity of the tree occur if one thinks of some major human decisions being made by use of “quantum coins”, i.e., outcomes of experiments performed for this purpose (depending on these decisions the future of humanity may take quite different turns).

After stating the theorem in Section 2 and then briefly discussing its relevance in Section 3, we will prove it in Section 4.

2 The theorem

Let $(\mathbf{V},\mathbf{E})$ be a tree with vertex set $\mathbf{V}$, including a singled out vertex called the root vertex, and edge set $\mathbf{E}$. We assume that the root vertex has a single edge incident to it and call it the root edge. Such a tree will be called an edge-rooted tree. We orient the root edge from the root vertex to its other end, and we give an orientation to every edge in the tree, so that each vertex other than the root vertex has exactly one edge oriented towards it. If e is the edge oriented towards vertex v and $e_1,...,e_n$ are the edges incident to v and oriented away from it, we call $e_1,...,e_n$ the children of e, and we refer to $\{e_1,...e_n\}$ as a set of siblings and to e as their parent. (The advantage of using such “family” language, even if a bit funny, is that the terminology becomes easy to remember and easy to extend.) Childless edges will be called terminal edges, and the vertices to which terminal edges point will be called terminal vertices. Each edge belongs to a generation defined inductively by declaring the generation of the root edge as 1, and the generation of the children of the edges of generation i to be $i+1$. It will be convenient to declare that childless edges that belong to generation i also belong to generations $i+1$, $i+2$, ... A partial history is a finite sequence of edges $(e_1,e_2,...,e_n)$, where each e_i is a child of e_{i-1}, $i=2,...,n$. A complete history (or just a history) is either a partial history in which e_1 is the root edge and the last edge is a terminal edge, or an infinite sequence of edges $(e_1,e_2,...)$, where e_1 is the root edge and each e_i is a child of e_{i-1}, $i=2,....$

Definition 1. A tree-structured set of projections on a Hilbert space $\mathscr{H}$ is a collection of such projections, $\mathscr{P}=\{Proj_e : e\in\mathbf{E}\}$, where the index set $\mathbf{E}$ is the set of edges of an edge-rooted tree, and the following conditions are satisfied:

1. If e is the root edge, then $Proj_e$ is the identity operator.
2. If $e_1,...,e_n$ are the children of e, then $\sum_{i=1}^n Proj_{e_i} = Proj_e$.

We write $\mathscr{H}_e = Proj_e\mathscr{H}$ for the subspace associated to $Proj_e$. The first condition means that $\mathscr{H}_e=\mathscr{H}$ when e is the root edge, while the second one means that the subspaces $\mathscr{H}_{e_i}$ associated to a set of siblings $\{e_1,...,e_n\}$ are orthogonal to each other and their linear span is the subspace $\mathscr{H}_e$ associated to the parent e.

Implicit in the definition of a tree-structured set of projections $\mathscr{P}$ is the associated edge-rooted tree $(\mathbf{V},\mathbf{E})$. The set of histories on this tree, denoted Ω, is the sample space on which one defines Born's probabilities (and alternative ones) associated to $\mathscr{P}$. Recall that, informally speaking, an element $\omega \in \Omega$ is a sequence of edges starting from the root and having each of its elements succeeded by one of its children, either with no end, or ending at a terminal edge. Abusing notation, we will write $e \in \omega$ for the statement that the edge e is an element of the sequence ω. For each $e \in \mathbf{E}$, we define $\Omega_e = \{\omega : e \in \omega\}$, the set of histories that go through e. Unions of finitely many sets Ω_e define an algebra of sets (a class of sets that is closed with respect to complements, finite unions and finite intersections) that we denote by $\mathscr{A}$. (This statement requires a proof, which is easily obtained by noting that every set $A \in \mathscr{A}$ can be written as a union over sets Ω_e with all e in the same generation, and that A^c is then the union of the sets Ω_e over the other e belonging to this same generation. This shows closure under complements. Closure under unions is immediate and De Morgan's law then provides closure under intersections.) The smallest σ-algebra that contains $\mathscr{A}$ will be denoted by $\mathscr{B}$.

Born's probabilities are defined on the measure space $(\Omega, \mathscr{B})$ and, in addition to $\mathscr{P}$, depend on a vector $\psi \in \mathscr{H} \backslash \{0\}$. (In the theorem below, ψ is arbitrary, but in all our applications it will be the initial state of our system. In collapse quantum mechanics, ψ will be chosen as the state, in the Heisenberg picture, before collapses. In non-collapse quantum mechanics, ψ will be chosen as the unchanging state, in the Heisenberg picture.) Born's probability corresponding to ψ will be denoted by $\boldsymbol{P}_\psi$. It is described informally by imagining a walker that moves on the edges of the tree. The walker starts at the root vertex of the tree and then moves in the direction of the orientation, deciding at each vertex where to go in a probabilistic fashion, with edges chosen with probability proportional to norm-squared, i.e., when at a vertex that separates a parent e from its children, the walker chooses child e' with probability $||\mathrm{Proj}_{e'} \psi||^2 / ||\mathrm{Proj}_e \psi||^2$, independently of past choices. If ever at a terminal vertex, the walker stops. A simple inductive computation shows that this is equivalent to the statement

$$\boldsymbol{P}_\psi(\Omega_e) = \frac{||\mathrm{Proj}_e \psi||^2}{||\psi||^2}, \quad \text{for each } e \in \mathbf{E}. \tag{1}$$

It is standard to show that (1) extends in a unique fashion to $\mathscr{A}$ and then to $\mathscr{B}$, defining in this way a unique probability measure on $(\Omega, \mathscr{B})$. Actually, for our purposes it will be important to observe that this standard procedure yields even more. For each $\psi \in \mathscr{H}$, the extension of the probability measure is to a larger measure space, $(\Omega, \mathscr{M}_\psi)$, where $\mathscr{M}_\psi \supset \mathscr{B}$, completes $\mathscr{B}$ with respect to the measure $\boldsymbol{P}_\psi$, meaning in particular that if $A \in \mathscr{B}$, $\boldsymbol{P}_\psi(A) = 0$ and $B \subset A$, then also $B \in \mathscr{M}_\psi$ and $\boldsymbol{P}_\psi(B) = 0$. We should note that all that is needed to implement this extension is contained in two facts about the non-negative numbers $p_e = \boldsymbol{P}_\psi(\Omega_e)$, which are similar to conditions 1 and 2 in Definition 1: $p_e = 1$, when e is the root edge, and $\sum_{i=1,\ldots,n} p_{e_i} = p_e$, when $e_1, \ldots, e_n$ are the children of e. (In obtaining the extension of $\boldsymbol{P}_\psi$ to the algebra $\mathscr{A}$ as a premeasure, the only non-trivial claim that has to be checked is that if $A \in \mathscr{A}$ is described in two distinct ways as finite disjoint unions of sets Ω_e, then the sum of the p_e over these sets is the same for both descriptions. And this is not difficult, if one realizes that it is possible to compare both representations to a third one, in which all the sets Ω_e have all e in the same sufficiently large generation.

The extension from a premeasure on $\mathscr{A}$ to a measure on $\mathscr{M}_\psi$ is an application of Carathéodory's Extension Theorem; see Sections 1 and 2 of Chapter 12 in [15], or Section 4 of Chapter 1 in [10].)

Before stating our theorem, we need to introduce a few more definitions, which will play a fundamental role in this paper. Given $\phi \in \mathscr{H}$ and $\omega \in \Omega$, we say that ϕ persists on ω if for each $e \in \omega$, $\mathrm{Proj}_e \phi \neq 0$. Otherwise we say that ϕ terminates on ω. We set now

$$\Omega(\phi) = \{\omega \in \Omega : \phi \text{ persists on } \omega\}, \tag{2}$$

and

$$\Omega^c(\phi) = \Omega \backslash \Omega(\phi) = \{\omega \in \Omega : \phi \text{ terminates on } \omega\}. \tag{3}$$

Keep in mind that the choice of $\mathscr{P}$ is implicit in the definitions in the last two paragraphs. We omitted it from the notation, but should not forget that Ω, $\boldsymbol{P}_\psi$, $\Omega(\phi)$, etc, depend on this choice.

Theorem 1. *Let $\mathscr{H}$ be a Hilbert space and $\mathscr{P}$ be a tree-structured set of projections on $\mathscr{H}$. For any $\psi \in \mathscr{H} \backslash \{0\}$ and $A \subset \Omega$, the following are equivalent.*

(1) $\boldsymbol{P}_\psi(A) = 0$.
(2.i) There exist $\phi_1, \phi_2, \ldots$ orthogonal to each other, such that $\psi = \sum \phi_i$ and $\Omega(\phi_i) \subset A^c$, for each i.
(2.ii) There exist $\zeta_1, \zeta_2, \ldots$ such that $\zeta_n \to \psi$ and $\Omega(\zeta_n) \subset A^c$, for each n.

Note that we are not, a priori, making any assumption of measurability on A. But if we assume that one of (2.i), (2.ii) is true, then we learn from the theorem that $A \in \mathscr{M}_\psi$ (and $\boldsymbol{P}_\psi(A) = 0$). On the other hand, assuming that (1) holds means assuming that $A \in \mathscr{M}_\psi$ (and $\boldsymbol{P}_\psi(A) = 0$).

The first two propositions stated and proved in Section 4 will add mathematical structure to the content of Theorem 1, and allow it to be restated in a very compact form in display (9).

3 Relevance of the theorem

The theorem stated in the previous section holds for any Hilbert space $\mathscr{H}$ and any choice of tree-structured set of projections $\mathscr{P}$ on it. The arbitrariness of $\mathscr{P}$ should be kept in mind as our discussion returns to Physics. When we consider collapse quantum mechanics, there is a special choice of $\mathscr{P}$, namely the one described in the introduction: vertices (other than the root vertex and terminal ones) correspond to experiments and edges (other than the root edge) correspond to the possible outcomes in each experiment. In the case of collapse quantum mechanics, and with this choice of $\mathscr{P}$, Ω is the set of possible histories that could materialize from the collapses. And in the special case of Born-rule collapse quantum mechanics, statement

(1) in Theorem 1 means that event A is (probabilistically) precluded from happening. In the case of non-collapse quantum mechanics there is in principle no special choice of $\mathscr{P}$. But, as the reader may have anticipated, for the purpose of comparing non-collapse to collapse quantum mechanics, via Theorem 1, it is natural to choose precisely the same $\mathscr{P}$. We will observe below, as the reader may have also anticipated, that with this choice, the equivalence between statements (1) and (2.i) in the theorem implies (modulo plausible postulates on how the theories provide predictions) that quantum mechanics without collapse gives raise to the same predictions as Born-rule collapse quantum mechanics.

In collapse quantum mechanics only one history $\omega = (e_1, e_2, e_3, ...)$ materializes. In the Heisenberg picture that we are considering, the state of our system is initially some $\psi \in \mathscr{H}$, but it changes at each collapse, following the path $(\mathrm{Proj}_{e_1}\psi, \mathrm{Proj}_{e_2}\psi, \mathrm{Proj}_{e_3}\psi, ...) = (\psi, \mathrm{Proj}_{e_2}\psi, \mathrm{Proj}_{e_3}\psi, ...)$. In non-collapse quantum mechanics in the Heisenberg picture, ψ never changes. Everett [9] was the first to make the observation that this would still be compatible with our perception that collapses happen. As observers who are part of the system (otherwise we would not be able to interact with the experimental arrangement and observe it), the particles that form our bodies and in particular our brains must follow the same quantum mechanics that describes the rest of the system that we are observing. So that at the end of an experiment we can be described as being in a superposition of states, each one with a brain that encodes a different outcome for this experiment. All the possible outcomes materialize, and versions of the human observers, entangled to each possible experimental outcome, are included in this superposition.

The non-collapse view of quantum mechanics has the significant advantage of eliminating the mystery of collapse: How can systems behave differently when they are being "measured"? It yields a much simpler and consistent theory. One of the main hurdles that prevents its acceptance is probably psychological, as it affects substantially our sense of identity and of our reality. But other than this, probably the greatest obstacle to its acceptance is the issue addressed (once more) in this paper: even accepting Everett's observation that we will see collapses even if they do not happen, the question remains of why it is that we perceive them happening as if they were produced according to Born's probability rule. I will not discuss here the various previous approaches to this problem, and rather refer the reader to the recent papers [1], [18], [13], references therein and papers in the collection [16] for background and recent ideas. In [17] I argue at some length why Theorem 1 presents an answer. Here I will summarize the idea.

Stating that collapses do happen according to Born's probabilities can only have meaning if we add some postulate telling us how this leads to predictions. I assume that in collapse quantum mechanics, the predictive power derived from the collapse axiom is fully contained in the following postulate.

Prediction Postulate of Collapse Quantum Mechanics (PPCQM): *In making predictions in collapse quantum mechanics, events of probability 0 can be deemed as sure not to happen.*

If it is accepted that this postulate covers the full predictive power of the probability axiom in collapse quantum mechanics, and in particular of Born's rule in Born-rule collapse quantum mechanics, then Theorem 1 tells us that non-collapse quantum mechanics will yield the same predictions as Born-rule collapse quantum mechanics, provided we accept the following postulate for non-collapse quantum mechanics.

Prediction Postulate of Non-Collapse Quantum mechanics (PPNCQM): *In making predictions in non-collapse quantum mechanics, if the state of our system is a superposition of states all of which exclude a certain event (i.e., if (2.i) of Theorem 1 holds for this event A), then this event can be deemed as sure not to happen.*

This postulate can be seen as a version of the superposition principle of quantum mechanics, and does not include probabilities in its statement. Theorem 1 therefore provides an explanation of how probabilities emerge in non-collapse quantum mechanics, and why they are given by Born's rule.

The observation above provides an answer to the question in the title of this paper, but it also raises a fundamental question. Can one formulate non-collapse quantum mechanics in a precise and consistent fashion that provides a clear notion for what is reality in the theory (provides a precise ontology for the theory) and is compatible with the PPNCQM? In current work in progress I hope to provide an affirmative answer.

One important consequence of the observations above is that not only objective predictive probabilities emerge in non-collapse quantum mechanics (from the non-probabilistic PPNCQM), but that they are precisely the ones supported by experimental observation, namely, Born's rule probabilities. This point is discussed at some length in Section 4 of [17]. It implies that non-collapse quantum mechanics with the PPNCQM included would be falsified by data that indicated collapses with a (significantly) different probability law.

Theorem 1 and the discussion in this section help dismiss an old and important misconception associated to non-collapse quantum mechanics. That the "natural" probability distribution that it entails is some sort of "branch counting" or "uniform" one. (The quotation marks are used because in the case of non-homogeneous trees, there is ambiguity in such wording. In the case of a homogeneous tree in which each edge has b children, this probability distribution is well defined by setting $\boldsymbol{P}(\Omega_e) = b^{-n+1}$, when e is an edge in the n-th generation.) Typically this probability distribution will produce predictions at odds with those produced by using Born's rule. And this has been used as an argument against non-collapse quantum mechanics. But while the use of Born's rule to make predictions in non-collapse quantum mechanics is shown here to be equivalent to the PPNCQM, one cannot find any good reason why a "branch counting rule" would be the appropriate tool for this purpose. There is a tradition of saying something like: "since all branches are equally real, a branch counting probability distribution is implied". But this is a meaningless sentence. What would "real, but not equally real" mean? "Equally real" (whatever it may mean) does not imply equally likely in any predictive sense. For instance all the teams competing for a soccer World Cup are "equally real", but if we want to predict who will win the cup, there is no reason for using a uniform distribution.

4 Proof of the theorem

The orientation that was introduced on the tree $(\mathbf{V},\mathbf{E})$ induces a partial order on the set of edges: for any two edges we write $e' \leq e''$ if there is a partial history that starts with e' and ends with e''. We write $e' < e''$ if $e' \leq e''$ and $e' \neq e''$. If neither $e' \leq e''$, nor $e'' \leq e'$, then we say that e' and e'' are not comparable.

Definition 1 has some simple consequences. If e'' is a child of e', then $\mathscr{H}_{e''} \subset \mathscr{H}_{e'}$. By induction along a partial history line, this extends to:

$$\text{If } e' \leq e'', \text{ then } \mathscr{H}_{e''} \subset \mathscr{H}_{e'}. \tag{4}$$

In contrast, if e' and e'' are siblings, then $\mathscr{H}_{e'} \perp \mathscr{H}_{e''}$. By induction along partial history lines, this extends to:

$$\text{If } e' \text{ and } e'' \text{ are not comparable, then } \mathscr{H}_{e'} \perp \mathscr{H}_{e''}. \tag{5}$$

Proposition 1. *For every $\phi_1, \phi_2 \in \mathscr{H}$, $\Omega(\phi_1+\phi_2) \subset \Omega(\phi_1) \cup \Omega(\phi_2)$, or equivalently $\Omega^c(\phi_1) \cap \Omega^c(\phi_2) \subset \Omega^c(\phi_1+\phi_2)$.*

Proof. Suppose $\omega \in \Omega^c(\phi_1) \cap \Omega^c(\phi_2)$. Then there are $e_1, e_2 \in \omega$ such that $\mathrm{Proj}_{e_1}\phi_1 = \mathrm{Proj}_{e_2}\phi_2 = 0$. As ω is a history, e_1 and e_2 are comparable. Let e be the larger between e_1 and e_2. Also because ω is a history, for $i = 1,2$ we have now, from (4), $\mathscr{H}_e \subset \mathscr{H}_{e_i}$ and hence $\mathrm{Proj}_e\phi_i = 0$. Therefore $\mathrm{Proj}_e(\phi_1+\phi_2) = \mathrm{Proj}_e\phi_1 + \mathrm{Proj}_e\phi_2 = 0$, which means that $\omega \in \Omega^c(\phi_1+\phi_2)$. □

For each $A \subset \Omega$ we define the following two sets (T stands for "truth" and F for "falsehood"):

$$T(A) = \{\phi \in \mathscr{H} : \Omega(\phi) \subset A\}, \tag{6}$$

and

$$F(A) = T(A^c) = \{\phi \in \mathscr{H} : \Omega(\phi) \subset A^c\} = \{\phi \in \mathscr{H} : A \subset \Omega^c(\phi)\}. \tag{7}$$

Proposition 2. *For every $A \subset \Omega$, $T(A)$ and $F(A)$ are vector spaces.*

Proof. Since $F(A) = T(A^c)$, it suffices to prove the statement for $T(A)$. Suppose $\phi_1, \phi_2 \in T(A)$, a_1, a_2 scalars. Then, for $i = 1,2$, $\Omega(a_i\phi_i) = \Omega(\phi_i)$, if $a_i \neq 0$, and $\Omega(a_i\phi_i) = \emptyset$, if $a_i = 0$. In any case $\Omega(a_i\phi_i) \subset \Omega(\phi_i) \subset A$. From Proposition 1 we obtain $\Omega(a_1\phi_1 + a_2\phi_2) \subset \Omega(a_1\phi_1) \cup \Omega(a_2\phi_2) \subset A$, which means $a_1\phi_1 + a_2\phi_2 \in T(A)$. □

We can rephrase Statement (2.ii) in Theorem 1 as

$$\psi \in \overline{F(A)}, \tag{8}$$

where the bar denotes topological closure in the Hilbert space $\mathscr{H}$.

The equivalence of (2.ii) and the apparently stronger statement (2.i) in Theorem 1, can be obtained, in a standard fashion, by applying the Gram-Schmidt orthonormalization procedure (see p.46 of [14], or p.167 of [10]) to the vectors ζ_1, $\zeta_2 - \zeta_1$, $\zeta_3 - \zeta_2$, ... to produce an orthonormal system with the same span. Proposition 2 assures us that this orthonormal system will be contained in $F(A)$, since the ζ_i are. The vectors ϕ_1, ϕ_2, ϕ_3, ..., are then obtained by expanding ψ in this orthonormal system.

The proof of Theorem 1 is now reduced to showing that for any $\psi \in \mathscr{H} \setminus \{0\}$ and $A \subset \Omega$,

$$\boldsymbol{P}_\psi(A) = 0 \iff \psi \in \overline{F(A)}. \tag{9}$$

The class $\mathscr{A}_\sigma$ of subsets of Ω obtained by countable unions of elements of $\mathscr{A}$ will play a major role in the proof of (9). Every $A \in \mathscr{A}_\sigma$ is a union of sets in the countable class $\{\Omega_e \colon e \in \mathbf{E}\}$. But since $\Omega_{e''} \subset \Omega_{e'}$, whenever $e' \le e''$, we will avoid redundancies in this union by writing it as

$$A = \bigcup_{e \in \mathbf{E}(\mathbf{A})} \Omega_e, \tag{10}$$

where

$$\mathbf{E}(A) = \{e \in \mathbf{E} : \Omega_e \subset A \text{ and there is no } e' \in \mathbf{E} \text{ such that } e' < e \text{ and } \Omega_{e'} \subset A\}. \tag{11}$$

Any two distinct elements of $\mathbf{E}(A)$ are not comparable. And since $\Omega_{e'} \cap \Omega_{e''} = \emptyset$, whenever e' and e'' are not comparable, (10) is a disjoint union. Moreover, using (5) we see that $\{\mathscr{H}_e : e \in \mathbf{E}(\mathbf{A})\}$ is a countable collection of orthogonal subspaces of $\mathscr{H}$. We will associate to A their direct sum (the topological closure of the linear span of vectors in these $\mathscr{H}_e$), which we denote by

$$\mathscr{H}(A) = \bigoplus_{e \in \mathbf{E}(\mathbf{A})} \mathscr{H}_e. \tag{12}$$

If $\mathscr{S}$ is a subspace of $\mathscr{H}$ and $\phi \in \mathscr{H}$, we will use the notation $\mathrm{Proj}(\phi|\mathscr{S})$ to denote the projection of ϕ on $\mathscr{S}$. For instance $\mathrm{Proj}(\phi|\mathscr{H}_e) = \mathrm{Proj}_e \phi$.

Lemma 1. *For any $\phi \in \mathscr{H}$ and $A \in \mathscr{A}_\sigma$,*

(i) For any $e \in \mathbf{E}$, $\mathrm{Proj}_e \phi = 0 \iff \Omega_e \subset \Omega^c(\phi)$.
(ii) $\mathscr{H}^\perp(A) = F(A)$.
(iii) $\Omega^c(\phi) \in \mathscr{A}_\sigma$.
(iv) $\phi \in \mathscr{H}^\perp(\Omega^c(\phi))$.
(v) $||\mathrm{Proj}(\phi|\mathscr{H}(A))||^2 = ||\phi||^2 \boldsymbol{P}_\phi(A)$, if $\phi \neq 0$.

Proof. (i) The implication ($\Longrightarrow$) is clear. To prove ($\Longleftarrow$) suppose that $\mathrm{Proj}_e\phi \neq 0$. Then either e is a terminal edge, or it has a child e' with $\mathrm{Proj}_{e'}\phi \neq 0$. Repeating inductively this reasoning, we produce a history ω such that $e \in \omega$ and ϕ persists on ω. Hence $\Omega_e \not\subset \Omega^c(\phi)$.

(ii)

$$\begin{aligned}\mathscr{H}^{\perp}(A) &= \bigcap_{e\in\mathbf{E(A)}} \mathscr{H}_e^{\perp} = \bigcap_{e\in\mathbf{E(A)}} \{\phi \in \mathscr{H} : \Omega_e \subset \Omega^c(\phi)\} \\ &= \{\phi \in \mathscr{H} : A \subset \Omega^c(\phi)\} = F(A),\end{aligned}$$

where in the first equality we used the definition (12) of $\mathscr{H}(A)$, in the second equality we used part (i) of the lemma, in the third equality we used (10), and in the fourth equality we used (7)

(iii) $\Omega^c(\phi) = \cup\{\Omega_e : e \in \mathbf{E}, \mathrm{Proj}_{\mathbf{e}}\phi = \mathbf{0}\}$. And this set belongs to $\mathscr{A}_\sigma$, since this union is countable.

(iv) Thanks to part (iii) of the lemma, we can take $A = \Omega^c(\phi)$ in part (ii) of the lemma. Using then (7), we obtain

$$\mathscr{H}^{\perp}(\Omega^c(\phi)) = F(\Omega^c(\phi)) = \{\phi' \in \mathscr{H} : \Omega^c(\phi) \subset \Omega^c(\phi')\} \ni \phi.$$

(v)

$$\begin{aligned}||\mathrm{Proj}(\phi|\mathscr{H}(A))||^2 &= \sum_{e\in\mathbf{E(A)}} ||\mathrm{Proj}_e(\phi)||^2 = \sum_{e\in\mathbf{E(A)}} ||\phi||^2 \boldsymbol{P}_\phi(\Omega_e) \\ &= ||\phi||^2 \boldsymbol{P}_\phi(\cup_{e\in\mathbf{E(A)}}\Omega_e) \ = \ ||\phi||^2 \boldsymbol{P}_\phi(A),\end{aligned}$$

where in the first equality we used the definition (12) of $\mathscr{H}(A)$, in the second equality we used (1), in the third equality we used the disjointness of the sets involved, and in the fourth equality we used (10). $\square$

We will use some consequences of Carathéodory's theorem that extends the measure $\boldsymbol{P}_\psi$ from $\mathscr{A}$ to $\mathscr{M}_\psi$ (see Sections 1 and 2 of Chapter 12 in [15], or Section 4 of Chapter 1 in [10]). Given $\psi \in \mathscr{H}$, define the outer measure of any set $A \subset \Omega$ by

$$\boldsymbol{P}^*_\psi(A) = \inf\{\boldsymbol{P}_\psi(A') : A' \in \mathscr{A}_\sigma, A \subset A'\}, \tag{13}$$

and define also

$$\mathscr{M}_\psi \ = \ \{A \subset \Omega : \text{for all } S \subset \Omega, \boldsymbol{P}^*_\psi(A \cap S) + \boldsymbol{P}^*_\psi(A^c \cap S) = \boldsymbol{P}^*_\psi(S)\}. \tag{14}$$

Then it follows from Carathéodory's Extension Theorem that $\mathscr{M}_\psi$ is a σ-algebra that extends $\mathscr{B}$ and $\boldsymbol{P}^*_\psi(A) = \boldsymbol{P}_\psi(A)$ for every $A \in \mathscr{M}_\psi$, in particular for every $A \in \mathscr{B}$

and therefore for every $A \in \mathscr{A}_\sigma$. It also follows that $\boldsymbol{P}^*_\psi(A) = 0$ implies $A \in \mathscr{M}_\psi$ and is necessary and sufficient for $\boldsymbol{P}_\psi(A) = 0$.

The next two lemmas prove each one of the directions of the equivalence (9), completing the proof of Theorem 1.

Lemma 2. *For any $\psi \in \mathscr{H} \setminus \{0\}$ and $A \subset \Omega$,*

$$\psi \in \overline{F(A)} \implies \boldsymbol{P}_\psi(A) = 0.$$

Proof. If $\psi \in \overline{F(A)}$, there are $\zeta_n \in F(A)$ such that $\zeta_n \to \psi$. Set $B_n = \Omega^c(\zeta_n)$. From (7) and Lemma 1(iii) we have $A \subset B_n \in \mathscr{A}_\sigma$. Using (13) and Lemma 1(v), we obtain

$$0 \;\le\; \boldsymbol{P}^*_\psi(A) \;\le\; \boldsymbol{P}_\psi(B_n) \;=\; \frac{||\mathrm{Proj}(\psi|\mathscr{H}(B_n))||^2}{||\psi||^2}.$$

But since Lemma 1(iv) tells us that $\zeta_n \in \mathscr{H}^\perp(B_n)$, we can write

$$\begin{aligned} ||\mathrm{Proj}(\psi|\mathscr{H}(B_n))||^2 \;&=\; ||\mathrm{Proj}(\psi - \zeta_n|\mathscr{H}(B_n)) + \mathrm{Proj}(\zeta_n|\mathscr{H}(B_n))||^2 \\ &=\; ||\mathrm{Proj}(\psi - \zeta_n|\mathscr{H}(B_n))||^2 \;\le\; ||\psi - \zeta_n||^2. \end{aligned}$$

Since n is arbitrary, the two displays combined give

$$0 \;\le\; \boldsymbol{P}^*_\psi(A) \;\le\; \lim_{n\to\infty} \frac{||\psi - \zeta_n||^2}{||\psi||^2} \;=\; 0,$$

proving that $\boldsymbol{P}^*_\psi(A) = 0$ and hence $A \in \mathscr{M}_\psi$ and $\boldsymbol{P}_\psi(A) = 0$. □

Lemma 3. *For any $\psi \in \mathscr{H} \setminus \{0\}$ and $A \subset \Omega$,*

$$\boldsymbol{P}_\psi(A) = 0 \implies \psi \in \overline{F(A)}.$$

Proof. If $\boldsymbol{P}_\psi(A) = 0$, (13) tells us that there are $A_n \in \mathscr{A}_\sigma$ such that $A \subset A_n$ and $\boldsymbol{P}_\psi(A_n) \to 0$. Set $\xi_n = \mathrm{Proj}(\psi|\mathscr{H}^\perp(A_n))$. Then $\xi_n \in \mathscr{H}^\perp(A_n) = F(A_n) \subset F(A)$, where the equality is Lemma 1(ii), and in the last step we are using (7). Therefore, using Lemma 1(v), we obtain

$$||\xi_n - \psi||^2 \;=\; ||\mathrm{Proj}(\psi|\mathscr{H}(A_n))||^2 \;=\; ||\psi||^2\, \boldsymbol{P}_\psi(A_n) \;\to 0,$$

as $n \to \infty$. This shows that (ξ_n) is a sequence in $F(A)$ that converges to ψ and therefore $\psi \in \overline{F(A)}$. □

Acknowledgements I am grateful to Marek Biskup, Michael Gutperle, Ander Holroyd, Jim Ralston, Pierrre-François Rodriguez and Sheldon Smith for enlightening discussions. Special thanks go to Jim Ralston and Maria Eulalia Vares for their careful reading of the proof.

References

1. Aguirre, A. and Tegmark, M. (2011) Born in an infinite universe: a cosmological interpretation of quantum mechanics. *Physical Review D* **84**, 105002, 1-16.
2. Albert, D. (1992) *Quantum Mechanics and Experience* (Harvard University Press).
3. Barrett, J. (1999) *The Quantum Mechanics of Minds and Worlds* (Oxford University Press).
4. Bricmont, J. (2016) *Making Sense of Quantum Mechanics* (Springer).
5. Bruce, C.(2004) *Schrödinger Rabbits. The Many Worlds of the Quantum* (Joseph Henry Press).
6. Carroll, S. (2019) *Something Deeply Hidden* (Dutton).
7. Deutsch, D. (1999) Quantum theory of probability and decisions. *Proceedings of the Royal Society A* **455**, 3129-3137.
8. De Witt, B. and Grahan, N. (Eds., 1973) *The Many Worlds Interpretation of Quantum Mechanics* (Princeton University Press).
9. Everett, H. (1957) "Relative state" formulation of quantum mechanics. *Reviews of Modern Physics* **29**, 454-462.
10. Folland, G.B. (1984) *Real Analysis. Modern Techniques and Their Application* (John Wiley & Sons).
11. Hall, B.C. (2013) *Quantum Theory for Mathematicians* (Springer).
12. Maudlin, T. (2019) *Philosophy of Physics. Quantum Theory* (Princeton University Press).
13. McQueen, K.J. and Vaidman, L. (2019) In defense of the self-location uncertainty account of probability in the many-worlds interpretation. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **66**, 14-23.
14. Reed, M. and Simon, B. (1980) *Functional Analysis* (Revised and enlarged edition, Academic Press).
15. Royden, H.L. (1986) *Real Analysis* (3rd Ed. Macmillan Publishing Company).
16. Saunders, S., Barrett, J., Kent, A. and Wallace, D. (Eds., 2010) *Many Worlds?* (Oxford University Press).
17. Schonmann, R.H. (2019) A theorem and a remark with the purpose of comparing the role and origin of probabilities in non-collapse and in collapse quantum mechanics. *Preprint.*
18. Sebens, T. and Carroll, S.M. (2018) Self-locating uncertainty and the origin of probability in Everettian quantum mechanics. *The British Journal for the Philosophy of Science* **69(1)**, 25-74.
19. Tegmark, M. (2007) Many lives in many worlds. *Nature* **448(5)**, 23-24.
20. Tegmark, M. (2014) *Our Mathematical Universe. My Quest for the Ultimate Nature of Reality* (Doubleday).
21. Wallace, D. (2012) *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation* (Oxford University Press).