Finite frequentism explains quantum probability[§]

Simon Saunders[*]

Abstract: I show that frequentism, as an explanation of probability in classical statistical mechanics, can be extended in a natural way to a decoherent quantum history space, the analogue of a classical phase space. The result is further a form of *finite* frequentism, in which Gibbs' concept of an infinite ensemble of gases is replaced by the total quantum state expressed in terms of the decoherence basis, as defined by the history space. It is a form of finite and *actual* frequentism (as opposed to hypothetical frequentism), insofar as all the microstates exist, in keeping with the decoherence-based Everett interpretation, and some versions of pilot-wave theory.

1. Introduction

Frequentism in philosophy of probability holds that the probability of an event is its relative frequency of occurrence in a suitable ensemble. Of its many champions, perhaps the most prominent was von Mises. For him the ensembles were finite and actual, but as such they were vulnerable to the reference class problem -- the problem of *which* ensemble – as generally there are either too many, or too few. Von Mises recommended pluralism: there is no such thing as the probability of an event, only probabilities relative to different references classes – and for too rare an event, it may be, no probability at all (as lacking a reference class). But that proved unpersuasive. Finite frequentism as a philosophy of probability is today routinely dismissed as unworkable.[1]

However, the relevant ensembles in von Mises' sense were collections of events in a classical world. In the real word we find the quantum. Quantum mechanics is the most prolific, most accurate, and most unificatory theory in all the long history of physics, but also, astonishingly, now after more than a century, the most controversial. It is seemingly incapable of a realist reading -- hence the repeated attempts to add to or modify the formalism, despite its empirical success. The main difficulty is the quantum measurement problem, which concerns the way probability is usually introduced into quantum mechanics using the Born rule: a rule that puzzlingly seems always to involve the notion of 'observer', however indirectly.

The one clear realist reading of standard quantum mechanics takes seriously the idea that the superposition principle has unlimited validity and that the Schrödinger equation can be applied to macroscopic bodies. It takes serious the reality of quantum states, including when entering into superpositions. It is otherwise conservative, leaving quantum

---

[§] To appear in the *British Journal for the Philosophy of Science* (2024).

[*] University of Oxford and Merton College, Oxford.

[1] As in Rowbottom (2015 p.112), La Caze 2016), Myrvold (2016). A particularly influential critique is Hájek (1996). For a dissenting voice, see Hubert (2021). Here I put to one side Humean best-system accounts of probability, which are sometimes considered sophisticated forms of actual frequentism.

mechanics unchanged. The result is realism about superpositions of states that differ macroscopically, as arise in quantum measurements, given only the Schrödinger equation. With no theoretical resource for picking out one rather than another outcome as actual, all the outcomes are actual. Welcome to many worlds, otherwise known as the Everett interpretation of quantum mechanics, first introduced in Everett (1957).

No wonder, if this is on anything like the right lines, the controversy over quantum mechanics has been unending; and very likely, if on the right lines, it *will* be unending: for the worlds remain unobservable. Once formed, short of super-technology or Poincaré recurrence, the states that arise on measurement, and the states that arise from them in turn on further measurements, will never interfere with one other. But this is the only way states *can* interact with one another, in a linear theory; so once formed, each state evolves as if all the others were not there. This explains the appearance of quantum jumps and collapsing wave-packets.

The picture is challenging – there is no helping that – but the question is whether it succeeds on its own terms, and nowhere more so than when it comes to probability. The Born rule involves the squared amplitudes of macrostates produced by measurements: It is not obvious how differences in these amplitudes yield differences in probabilities, with all else unchanged. Attention in recent years to the decision-theory approach of Deutsch (1999) and Wallace (2010) has if anything only reinforced the worry: for if the Born rule is explained in terms of rational agency – if it is, indeed, rationally forced, as claimed, for agents knowledgeable of the superpositions of macrostates produced by measurements -- then it may be that *entirely* accounts for probability, with no residue to be explained.[2] Or rather, there remains one small residue: the relative frequencies of macrostates (experimental outcomes) on repeated measurements. But the rule based on these, as set out in Graham (1973), contradicts the Born rule, and so impacts negatively, and had to be explained away rather than used.

The situation, however, is quite different when branches are defined in decoherence theory.[3] Branching, in roughly Everett's sense, for macroscopic bodies, now takes place at the level of differences in molecular configurations and momenta. At ordinary temperatures and pressures branching is everywhere, and extremely rapid (much faster than thermal relaxation). It is taking place in an experiment even if the apparatus just sits there and doesn't measure anything -- reflected only in all those minute differences in molecular motions.

From this perspective, the old branch-counting rule is simply an aberration: the point is to count the relative numbers of microstates for the various experimental outcomes, not the numbers of outcomes. And yet, until very recently, no new rule was forthcoming based on the statistics of decohering microstates. To the contrary, standard wisdom has it that decoherence is an inherently messy, approximate affair, and that branches themselves are

---

[2] According to Deutsch (2016), Brown and Porath (2020), this is a virtue of the approach, not an objection.

[3] Folllowing Joos and Zeh (1985), Tegmark (1993), Saunders (1993); it was hinted in Gell-Mann and Hartle (1990).

emergent entitites, to the point where no boundaries can be drawn – so that ultimately, branch number makes no physical sense.[4]

Yet this consensus may be challenged. The theory of decoherent histories, as introduced in Gell-Mann and Hartle (1990)), provides a common language for decoherence theory, including master equations for the reduced density matrix. [5] It further *does* permit a precisification of microstates, cunningly shifting the approximation away from the states and onto the no-interference condition: branches may be well-defined, yet only approximately decohere (and residual but tiny interference is just fine for Everettians, as may usefully lead to observable effects). The upshot: there may be no point at which a unique notion of branch number can be defined, but well away from that point, *ratios* in branch numbers may yet make physical sense.

It is more than a hope. As shown in Saunders (2021), when the decoherent history space involves a continuous dynamical variable (like position or momentum), a branch-counting rule can be defined that for finite ensembles approximately agrees with the Born-rule (up to 'rounding errors'). The rule counts microstates parameterised by histories of events, up to time t; it counts the total number, up to that time, and it counts those that contain an event E, at some earlier time. The ratio of the two is the probability (relative frequency) of E. This rule has a ready extension to an infinite ensemble, by taking the limit $t \to \infty$ (assuming branching is unending). In that limit the relative frequency of E equals the Born rule probability of E, to which relative frequencies for finite times are approximations ('rounding errors'). This rule is, moreover, traceable to Boltzmann's microstate counting rule, so it is scarcely ad hoc, replacing Boltzmann's use of Liouville measure by the Hilbert-space norm.

There are, however, certain negatives to this approach. Its dependence on cosmology is troubling: does quantum probability make no sense in a bounded universe, with Poincaré recurrence? Yet it is not obvious, absent the infinite time limit, what 'approximation' (and 'rounding error') really mean -- or not in terms of frequentism. And the parallel between Liouville measure and the Hilbert-space norm is hardly new; it was first drawn in Everett (1957).

Here I propose a different strategy: a purely synchronic microstate-counting rule, restricted to finite ensembles throughout.[6] Decoherence is still essential (and remains an intrinsically diachronic concept) but for our purposes its role can be limited, roughly, to the definition of the parameter space (the 'decoherence basis'). The comparison with the methods used in classical statistical mechanics is now direct – and shows that the relevant connection is not so much to Boltzmann's method, passing from Liouville measure to the Hilbert space norm, but to Gibbs' method of defining probability, which precisely and sharply distinguishes the two

Boltzmann's method suggests rather a new and *different* branch-counting rule, that we will also consider – and ultimately dismiss, as failing to provide a consistent account of

[4] As argued, among others, by me, in Saunders (2005). See also Wallace (2012 p.99-100).

[5] See Kiefer (1996) for an overview. For realistic cases, involving macroscopic numbers of particles, decoherent history spaces are quasiclassical domains, in Gell-Mann and Hartle's sense, depending on the Hamiltonian, degrees of freedom involved, and initial quantum state.

[6] Stoica (2022) also proposes a synchronic rule, but involving an infinite ensemble.

probability. it is similar to the Graham rule in treating 'non-zero amplitude' as a sufficient as well as a necessary criterion for the existence of branches: a branch is a certain sort of state, and it exists if and only if it has non-zero amplitude (whereas on the Gibbs approach, the amplitude is used to define branches). For all its intuitive appeal, it fails as a synchronic probability rule, even fixing the state once and for all. Simple variants of the rule are also considered and are similarly shown to fail. In effect, they fall victim to the reference-class problem – the problem of reconciling relative frequencies for an event when referred to different but equally permissible ensembles. The Gibbs rule passes this test. It is also consistent with the Born rule. In view of the universality of quantum mechanics, it then seems that finite frequentism is able to explain probability in the objective or physical sense wherever it occurs in the physical sciences.[7]

Or so I claim, embracing realism about decohering microstates, when the ensembles involved are actual as well as finite. Different ensembles are different ways of parcelling up the same total state, with respect to the same decoherence basis, into finitely-many parts. Might antirealists make use of this method as well? For them, like Gibbs, the ensembles will be imaginary, but still finite: that may be viewed as an improvement on hypothetical frequentism, where the ensembles are usually infinite. Perhaps. But as we shall see, permissible ensembles, for any event *E*, will in general include Schrödinger-cat states for *E* – states that are superpositions of states in which *E* occurs, and states in which *E* does not occur. Because included in ensembles, they have non-zero relative frequency, and so non-zero probability. That may well prove unacceptable to antirealists, even as imaginary states of affairs, since they have a chance of being actual.

Pilot-wave theory (also known as Bohmian mechanics) adds additional 'hidden' degrees of freedom, but otherwise preserves the Schrödinger equation with the same initial state: it therefore must countenance the same ensembles of decohering microstates. It does both better and worse than antirealism: better, in that those microstates have some kind of reality, if only as guiding ('piloting') the values of the hidden-variables; worse, in that on introducing those additional variables, the statistical patterns we are considering become only a part of a larger physical picture, that may tell a quite different story about probability.[8]

A last worry, now, for realism and the Everett interpretation. Might the new rule, in avoiding the usual pitfalls of finite frequentism, be too good to be true? Perhaps it isn't a form of frequentism at all, but a new and sui generis theory of physical probability. The epistemology, in particular, seems quite different from anything in von Mises' writings, as the ensembles are in principle unobservable.

I refuse to be so flattered. The origins of the new rule lie squarely in the ideas of Boltzmann and Gibbs, and they were surely frequentists. They employed imaginary, infinite ensembles, so for them too the ensembles were unobservable in principle. There are other weighty tomes in the literature on frequentism that appealed to ensembles of possibles,

---

[7] Whether the notion of probability is required in quantum mechanics at the microscopic level, absent decoherence, is an open question. If so, it may yet be defined in frequentist terms, but now requiring variation in basis (see the comments preceding Eq.(7)).

[8] For example, a non-equilibrium story, as introduced in Valentini (1991), the reason too for rejecting the 'Everett in denial' criticism of pilot-wave theory (Valentini 2010).

again unobservable in principle, for example von Kries (1886).[9] But that is only to say that the challenge on epistemology is common to them all. I shall come back to this question, albeit briefly, at the end.

2. Microstate counting in classical statistical mechanics

The key idea in Boltzmann (1877) was to partition the accessible phase space for a gas into a finite number of cells, all of the same size, and to consider the relative numbers of such cells consistent with one macrostate (in which the gas has a certain pressure and volume, for given total energy), in comparison with another. The macrostate with the greatest number, relative to all the others, is the equilibrium state, also, according to frequentism, the most probable state. In this way he was able to compute the equilibrium entropy and obtain the equation of state.

The concept of equiprobability and ensemble probability fit together seamlessly. If the probability of E is the relative frequency of E in an ensemble, the ensemble elements all have the same probability. Conversely, if all the elements in an ensemble have the same probability, then the relative frequency of E in the ensemble is its probability.

The size of a cell γ is given by their volumes in phase space, as defined by Liouville measure. In terms of rectilinear coordinates $\langle p_1 \ldots p_{3N}; x_1 \ldots x_{3N}\rangle$ for the positions and momenta of N particles in 3-dimensional space, it is:

$$\sigma(\gamma) = \int_{\gamma} dp_1 \ldots dp_{3N} dx_1 \ldots dx_{3N}.$$

To ensure that the choice of partitioning into cells $\{\gamma\}$ be irrelevant, Boltzmann took the infinite limit $\sigma(\gamma) \to 0$. In this way he avoided any dependence of the probability on the sizes and shapes of the cells – he solved the reference-class problem. The entropy of a macrotate was proportional to the logarithm of its probability – what Einstein, in 1905, called 'Boltzmann's principle'. Irreversible behaviour in thermodynamics, Boltzmann reasoned, is to be understood as involving an increase in total entropy -- from less probable to more probable states, from smaller to larger volumes of phase space explored by the gas, to macrostates that include or are realised by the greatest number of microstates.

These ideas proved immensely important to the subsequent history of discovery of quantum mechanics, first in the work of Planck, in his attempt to ground the new black-body radiation law on the Boltzmann entropy, as defined by ratios in numbers of microstates, and second in Einstein's 1905 calculations, similar to Planck's, but based on the Wien black-body law, that only held in the high-frequency, low temperature regime. That likewise used the Boltzmann entropy, as defined by ratios in numbers of microstates. But here the situation was greatly complicated by the underlying indistinguishability of light

[9] Everett's branches have also been interpreted in terms of possible worlds in Wilson (2020) (although embracing a form of modal realism), and in terms of perspectives on a single world (Barrett 2011).

quanta, that only made a difference to the microstate counting rule away from the Wien limit, linking Einstein's and Planck's discoveries.[10]

It may be for that reason that philosophers of probability have ignored this frequentist background to the discovery of quantum mechanics,[11] but Einstein highlighted it, in one of the few places where he returned to this history:

> On the basis of the kinetic theory of gases Boltzmann had discovered that, aside from a constant factor, entropy is equivalent to the logarithm of the "probability" of the [macro]state under consideration. Through this insight he recognized the nature of the course of events which, in the sense of thermodynamics, are "irreversible". Seen from the molecular-mechanical point of view, however, all courses of events are reversible. If one calls a molecular-theoretically defined state a microscopically described one, or, more briefly, microstate, then an immensely large number of states belong to a macroscopic condition. [This number] is then a measure of the probability of a chosen macrostate. This idea appears to be of outstanding importance also because of the fact that its usefulness is not limited to microscopic description on the basis of [classical] mechanics. (Einstein 1949 p.43).

What is more, the idea led to the discovery of quantum mechanics precisely by *not* taking the $\sigma(\gamma) \to 0$ limit (this was Planck's daring step in 1900, departing from Boltzmann's method). It was finite frequentism indeed, albeit not of the kind we shall consider.[12]

Boltzmann's ideas are today sill familiar, but mainly for what they say about entropy. Except for special cases, like equilibrium states for finite-dimensional Hilbert spaces, the link with frequentism has been lost. Classically, microstate-counting is seen as little more than a calculational device; conceptually, since dependent on a measure λ (for a continuous phase space Γ), it seems more straightforward to identify thermodynamic probability outright with the volume of the accessible phase-space. Here the very term 'thermodynamic probability' is suspect; as argued by Albert (2000), Liouville measure enters in two quite distinct ways, the first in defining the Boltzmann entropy of a given macrostate (logarithm of accessible phase-space volume) and the second in defining a probability distribution over microstates. The sense in which the Boltzmannian entropy of a closed system will only *probably* increase, is made out with respect to the second. Increasingly, this has been replaced by the concept of 'typicality', especially in the adjacent literature on probability in pilot-wave theory[13]. Everett too spoke of 'typical' branches and observers, rather than probability.[14]

In contrast, in physics, Gibbs' definition of entropy is more common, and defined not in terms of the properties of an individual gas, but in terms of a probability distribution on phase space. But like Boltzmann, Gibbs also proposed a frequentist account of these

---

[10] See Saunders (2020) for a recent history. I remarked on it further in my (2021), but there erred in the suggestion (p.16-17) that for a completely degenerate density matrix, the decoherence condition is automatically satisfied whatever the fine-graining.

[11] Including some who know it well: Myrvold (2021), for example, in a chapter devoted to frequentism, does not mention it.

[12] As involving the dimensionality of certain finite-dimensional Hilbert spaces, and a completely degenerate initial state (density matrix).

[13] Beginning with Dürr et al (1992); see Goldstein (2001), Myrvold (2016).

[14] Barrett (2016), and, for a defence of Everett's treatment of probability, Lazarovici (2023).

probabilities -- it is he who coined the term ‘statistical mechanics’ – but now for a continuum infinity of phase-space points, a continuous infinity of gases, making up one or another ensemble. He defined various equilibrium ensembles (microcanonical, canonical, and grand canonical), depending on external constraints (on the total energy, temperature, and particle number). He was explicit on the status of members of the ensemble: they were ‘creatures of the imagination’ (Gibbs 1902 p.188); so this was hypothetical (imaginary, uncountably infinite) frequentism.

If we follow Boltzmann’s procedure, and coarse-grain phase space into equiprobable cells, there is now a clear alternative to the equivolume rule. Cells should be equiprobable with respect to a suitable Gibbs ensemble, in the following sense: they should contain equal numbers of gases in the ensemble (equal numbers of representative points of gases in the ensemble). For this Gibbs introduced a new function on phase space, the *density in phase*, $D\colon \Gamma \to \mathbb{R}$, where $D(q)d\sigma$ is the number of imaginary gases in the ensemble in the infinitesimal neighbourhood $d\sigma$ of the point q. Equiprobability of γ then requires equality in the quantities:

$$N(\gamma) := \int_{\gamma} D(q)d\sigma \tag{1}$$

rather than $\lambda(\gamma)$ – or more precisely, since Gibbs thought the numbers involved may all be infinite, equality in the ratios $N(\gamma)/N$, where N is the integral (1) extended over all of phase space Γ. That is, cells γ are chosen so that these ratios are all equal, and thereby equiprobable.

Here, in his own words, is how Gibbs made the connection with probability:

> Now, if the value of D is infinite, we cannot speak of any definite number of systems, within any finite limits, since all such numbers are infinite. But the ratios of these infinite numbers may be perfectly definite. If we write N for the total number of systems, and set
>
> $$P = \frac{D}{N},$$
>
> P may remain finite, when N and D become infinite. The integral
>
> $$\int Pd\sigma$$
>
> taken within any given limits, will evidently express the ratio of the number of systems falling within those limits to the whole number of systems. This is the same thing as the *probability* that an unspecified system of the ensemble (i.e. one of which we only know that it belongs to the ensemble) will lie within the given limits. The product $Pd\sigma$ expresses the probability that an unspecified system of the ensemble will be found in the element of extension-in-phase $d\sigma$. We shall call P the *coefficient of probability* of the phase considered. (Gibbs 1902 p.16).

The two procedures yield the same results for the equilibrium states of isolated gases of non-interacting particles, like black-body radiation (when D is essentially uniform), but they differ for non-ideal material gases, and for non-equilibrium ensembles. Which should we use in extending the method to the quantum case? Gibbs’ approach to statistical mechanics has proved much more fruitful in physics than Boltzmann’s, so one might well incline to that; however, as a method for defining finite ensembles, it presupposes the notion of a continuously-infinite background ensemble, with its associated density-in-phase. What this might mean in quantum mechanics will be clearer shortly.

3. Decoherent history space

A decoherent history space is a structure $\langle \mathcal{M}, \lambda, |\psi\rangle, \widehat{H}, \mathcal{H}, \{t_k\}\rangle$, where $\mathcal{M}$ is a parameter space, the analogue of phase space (in the simplest case defined in terms of the spectra of some value of commuting self-adjoint operators), λ is an additive, dimensional measure on $\mathcal{M}$ (the analogue of Liouville measure), $\widehat{H}$ is the Hamiltonian, containing the dynamics, and $|\psi\rangle \in \mathcal{H}$ is the total ('universal') state in the Hilbert space $\mathcal{H}$. We further assume that $\mathcal{M}$ includes the spectrum of at least one continuous operator (like position or momentum). This is physically reasonable; all events take place in space, and the spectrum of the position operator is continuous.[15] We further assume $\mathcal{M}$ is bounded, $\lambda(\mathcal{M}) < \infty$.[16]

More is needed to define quantum histories, notably a discretisation of the time, $t_k\}, k = 1, \ldots, n$, from which n-step histories can be defined. But we will not need these further tools, and we can hide reference to the Hilbert space and dynamics as well; the only structure we shall need is $\langle \mathcal{M}, \lambda, |\psi\rangle\rangle$. Formally, this is similar to a classical probability space $\langle \mathcal{M}, \sigma, \mu\rangle$, where $\mathcal{M}$ (the event space) is a continuous bounded space with additive measure σ (this could be phase space and Liouville measure respectively, so we use the same symbol), equipped with an additive probability function μ.

As before we call the partitions cells, and as before, for the continuous part of $\mathcal{M}$, we assume the cells are connected. In comparison to the classical case, the main novelty is the mapping from $\gamma \subseteq \mathcal{M}$ to vectors $P_\gamma|\psi\rangle \in \mathcal{H}$, where projectors P_γ come by construction of $\mathcal{M}$ (via Stone's theorem) from some commuting family of self-adjoint operators. In this way algebraic operations involving projectors mirror set-theoretic operations on $\mathcal{M}$, in that for disjoint γ, γ', projections onto unions go over to sums of projectors:

$$\gamma \cap \gamma' = \emptyset \Rightarrow P_{\gamma \cup \gamma'} = P_\gamma + P_{\gamma'} \tag{2}$$

and projections onto intersections of γ, γ', now without restriction, go over to products:

$$P_{\gamma \cap \gamma'} = P_\gamma P_{\gamma'} = P_{\gamma'} P_\gamma \, .$$

It follows that if γ, γ' are disjoint, the states $P_\gamma|\psi\rangle$, $P_{\gamma'}|\psi\rangle$ are (exactly) orthogonal. For a classical probability space $\langle \mathcal{M}, \sigma, \mu\rangle$, in place of the map $\gamma \to P_\gamma$ from regions of $\mathcal{M}$ to projection operators and states, we have the map $\gamma \to \mu(\gamma)$ from events to probabilities. Eq.(2) is then additivity of μ, and has to be assumed separately.

Consider now some fine-grained partitioning $\gamma_1, \gamma_2, \ldots.$, of $\mathcal{M}$, where $\gamma_j \cap \gamma_k = \emptyset$ for $j \neq k$, and $\bigcup_k \gamma_k = \mathcal{M}$ (were $\mathcal{M}$ a discrete space, these could be eigenvalues of the associated operators). Since $P_{\mathcal{M}}$ is the identity, it follows, at any time t

[15] As Bell (1987) argued in defence of pilot-wave theory, all experiments ultimately involve position measurements (although the claim has been challenged, for example, in Albert (1992 Ch.5)).

[16] This could be relaxed to require only that the support of $|\psi\rangle$ on $\mathcal{M}$ is bounded, but it anyway only plays a role in the Boltzmann case (assuming the state is normalizable).

$$|\psi\rangle = P_{\cup_k \gamma_k}|\psi\rangle = \sum_k P_{\gamma_k}|\psi\rangle\,.$$

In this way the total state at t is represented as a superposition of at most countably-many states $P_{\gamma_k}|\psi\rangle$, complete with relative phases. The state $|\psi\rangle$ we are interested in is that following a measurement process, under the Schrödinger equation, where states $P_{\gamma_k}|\psi\rangle$ may differ macroscopically (in particular, as to experimental outcome at t). They are not yet *decohering* states, however. For decoherence, except in contrived examples, some coarse-graining of the γ_k's is needed. In this way we arrive at a partitioning of $\mathcal{M}$ into cells $\alpha_1, \alpha_2, \ldots.$, again assumed connected, to obtain decohering microstates $P_{\alpha_k}|\psi\rangle$. These are the objects of our counting rules.

The further consideration we need, stemming from the underlying physics, is that decohering microstates may provide considerable microscopic detail at molecular and event atomic scales, consistent with decoherence, and for Avogadro numbers of particles. The development of superpositions of states like these over time will be rapid and at ordinary temperatures, quite uncontrollable.[17] Quantum jumps, in other words, are everywhere. There are of course limits, as Schrödinger, in this Journal, famously argued,[18] but for our purposes we may suppose the ensembles of microstates may be chosen as very large indeed. That said, small ensembles may also be chosen, for coarse-graining always preserves decoherence: relative frequencies as defined in small ensembles had better make sense, just as in large ones, so long as the ensembles conform to the same definite rule. But which rule?

4. The new microstate-counting rules

In light of §2, there are the two obvious candidates: one, following Boltzmann, based on the idea of equivolume cells at an instant in time, that are accessible to the individual gas; and the other, following Gibbs, based on the idea of an ensemble, and cells of equal fractions of the ensemble, at a given time, as defined by the density in phase.

It might be thought that there is a third option: why not count branches *simpliciter*, without appeal to a partitioning of $\mathcal{M}$ at all? But the idea is confused. At best it amounts to 'count non-zero eigenstates' of some self-adjoint operator. But as noted, even were $\mathcal{M}$ a discrete manifold, with the $P_{\gamma_k}|\psi\rangle$'s eigenstates of operators, they will not in general decohere. As Wallace puts it:

> [T]here is no sense, in which these phenomena lead to a naturally discrete branching process: as we have seen in studying quantum chaos, while a branching structure can be discerned in such systems, it has no natural "grain". To be sure, by choosing a certain discretisation of (configuration-) space and time, a discrete branching structure will emerge,

[17] For quantitative estimates, see Tegmark (1993) and, for a case of special interest, Tegmark (2000). The reason localised states decohere is because entanglement with the environment is mostly produced by local (electromagnetic) couplings, the crucial insight of Zurek (1981).

[18] Schrödinger (1952): 'quantum jumps' cannot be occurring at the level of *coherent* interactions between molecules, involving individual electrons, for they are needed to explain the physics of ordinary matter.

> but a finer or coarser choice would also give branching. And there is no "finest" choice of branching structure: as we fine-grain our decoherent history space, we will eventually reach a point where interference between branches ceases to be negligible, but there is no precise point where this occurs. (Wallace 2012 p.99-100).

Counting numbers of eigenstates, with $\mathcal{M}$ a purely discrete manifold, is not to count decohering states. In any case, in realistic cases, $\mathcal{M}$ involves a continuous space, so there is no 'natural grain'.

We have the two methods as stated. They each promise a reductive account of probability insofar as the reductive base is defined independent of probabilistic concepts. For the first method, where the cells are fixed in size, we suppose the idea 'accessible cell' is made out in terms of *cell of non-zero amplitude*. This seems the most natural requirement, once the cells are fixed, and it is the one used in Graham's branch-counting rule, where the cells were macrostates defined as experimental outcomes. Once states are fixed by specifying cells, it only remains to determine if they exist or not – on whether or not they have non-zero amplitude. Since the cells all have the same finite size, and given that $\mathcal{M}$ is bounded, it follows that the number of decohering microstates is strictly finite.

For the second method, we need the analogue of Gibbs' ensemble and its associated density in phase. To see what this should be, observe that we are replacing Gibbs' infinite collection of classical (pointlike) microstates, each the state of an imaginary gas, complete with the density in phase D, comprising the ensemble, by an ensemble of actual (non-pointlike) quantum microstates, each the state of an experimental apparatus, complete with their amplitudes and relative phases, comprising -- superposing to give -- the total state. The latter, as a superposition of decohering states, is the analogue of Gibbs' ensemble. We suppose further that the state is normalizable, i.e. that it has finite amplitude.[19] Rather than dividing up the classical ensemble into microstates with equal fractions of the ensemble, as defined by the density in phase, *we divide the quantum state into microstates of equal amplitude* -- that is, into branches of equal amplitude. Equivalently, if an additive quantity is wanted (corresponding to a density which can be integrated, as in (1)), so that the fractions add, we divide the quantum state into microstates of equal *squared* amplitude. But of course the two methods are the same: amplitudes are non-negative real numbers, so equality in amplitude and in squared amplitude are interchangeable.

Gibbs' density in phase as a function on phase space is thus identified with the squared norm of the quantum state as a function on $\mathcal{M}$. Whereas, for an infinitesimal neighbourhood $d\sigma$ of a point q in phase space, $D(q)d\sigma$ is *the number of systems* in the neighbourhood $d\sigma$ of q, we now have $|\langle q|\psi\rangle|^2 d\lambda$ as *the amount of wave-function* in the neighbourhood $d\lambda$ of $q \in \mathcal{M}$.

Branches thus defined had better correspond to cells of non-zero measure σ. This follows more or less automatically, as if $\lambda(\alpha_k) = 0$ then $P_{\alpha_k}|\psi\rangle = 0$ for any reasonable volume measure on $\mathcal{M}$. But it gives pleasing symmetry to the two branch-counting rules to make

---

[19] Of course it must, if a vector in Hilbert space, but unnormalizable states arise naturally in certain contexts (for example, as solutions to the Wheeler-DeWitt equation), and as goes probability the requirement of normalizability may well be dropped, for the reasons pointed out by Gibbs: like phase, only *ratios* in amplitudes have any physical meaning.

this explicit: the one, following Boltzmann, counts cells of equal volume that have non-zero amplitude, whilst the other, following Gibbs, counts cells of equal amplitude that have non-zero volume.

With that it is clearer that *both* rules are in question as reductive accounts of probability. For of course 'amplitude' in ordinary quantum mechanics *does* have a connection with probability, through the Born rule -- the only place in quantum mechanics where probability enters the picture. 'Non-zero amplitude' means 'non-zero probability', and 'equiamplitude' means 'equiprobable'.

Yet that is not quite right, as on the Boltzmann rule, equivolume cells with non-zero amplitude are equiprobable, not just of non-zero probability. But the Gibbs rule looks more vulnerable: the quantity

$$\frac{|\langle q|\psi\rangle|^2}{\int_{\mathcal{M}} |\langle q|\psi\rangle|^2 d\lambda} \tag{3}$$

is just the Born rule probability density, and corresponds to Gibbs' index of probability:

$$\frac{D(q)}{\int_{\Gamma} \ D(q)d\sigma}. \tag{4}$$

Multiplied by the volume of an infinitesimal neighbourhood $d\sigma$ about q, this is, as Gibbs said, 'the *probability* that an unspecified system of the ensemble will be found in $d\sigma$ of q'. Yet recall *why* he said that: it is because this expression 'will evidently express the ratio of the number of systems falling within those limits to the whole number of systems', and immediately following, 'this is the same thing as the *probability*….'. The interpretation of the expression (4) in terms of a probability density, in other words, is a *consequence* of frequentism. Likewise whatever right the expression (3) has to be called a probability density derives from frequentism, and specifically, only if we can identify 'fraction of wavefunction' with 'ratio in numbers of microstates' – but that follows for our ensembles by construction. We are *defining* microstates so that they have the same fraction of wavefunction. As for the notion of amplitude itself, it is clearly intelligible in contexts that have nothing to do with probability. Like phase, it is a new physical primitive of quantum theory.

There is another concern about the legitimacy of the reductive base, attaching now to decoherence itself. Does decoherence theory depend, in its formulation or derivation from the Schrödinger equation, on probabilistic reasoning or justification? It is usually *interpreted* in terms of probability, and its practitioners, when pressed as to the meaning of given expressions, may well appeal to the Born rule. But the appeal can seem forced, not least for the decoherence condition itself. The central concept 'interference' was never happily explained in terms of 'probability waves', whatever they might be; but it is immediately intelligible in the cancelling or reinforcing of physical quantities that take positive and negative values, as in any wave theory involving amplitude and phase. Interference is distinct from an interference *effect*, whereby interference is made visible, and measurement, arguably, is involved; then, it may be granted, in ordinary quantum mechanics, we make explicit or implicit use of the Born rule; but (at least for realists about states) interference takes place at the locus of the screen even when the screen is removed, and no observation is made.

To say more on this topic would take us in two different directions, one technical, involving more decoherence theory, and one philosophical, and the need in the physical sciences, if it is a need, for operational justifications of approximations, model-building, or representation. I forego both expeditions, noting that the same issues arise for the decision-theoretic argument of Wallace (2010, 2012), which first attracted this kind of criticism.[20] They even arise for Deutsch (1999), and Everett (1957), although both were innocent of decoherence theory; as Everett admitted, his abstract model of a measurement should be backed up with a concrete model, shown to behave appropriately, on the basis of the Schrodinger equation alone; and even when restricted to a mechanical model of a measuring apparatus, that will involve approximations,[21] so may be subject to a similar critique.

We conclude that there is a prima facie case that 'amplitude' and 'branch state' are intelligible physical notions that need have nothing to do with probability. It follows that *both* branch-counting rules, if successful, provide reductive theories of probability. However, they cannot both be correct (except, perhaps, for equilibrium cases), and perhaps neither is satisfactory.

5. Consistency

Because of the decoherence condition, and the assumption that $\mathcal{M}$ is bounded, on either rule we are limited to finite partitionings of $\mathcal{M}$ at any time, and hence to probabilities as rational numbers. But there is no conflict with Kolmogorov's axioms. On the contrary, probabilities defined as relative frequencies in a finite ensemble are automatically additive, normalised, and contained in the interval $[0,1]$. Conditional probabilities, defined as ratios of rational numbers, remain rational. As one of the many attractions of frequentism, Kolmogorov's axioms are derived and not assumed.[22]

However, having defined the kind of ensemble (equiamplitude or equivolume), there are still infinitely many precisifications available, as we are assuming $\mathcal{M}$ varies continuously. It is the reference-class problem, familiar in any version of finite frequentism: to which ensemble, exactly, should a macrostate be referred? Our strategy is similar to von Mises': *any* admissible ensemble should be acceptable, and yield, for any macrostate, a perfectly meaningful relative frequency, save only that some may be more informative than others (we shall see how this works in a moment) – so long as they are all consistent with one other; so long as they may all be jointly true. This avoids the deficiencies of von Mises-style pluralism.

[20] For example, by Kent (2010), Dawid and Thèbault (2015); for replies see Saunders (2022), Franklin (2024).

[21] As in Everett's sketch of how to derive classical motions for wave-packets from the Schrödinger equation (Everett 1973 p.86-90).

[22] Admittedly *countable* additivity, Kolmogorov's main innovation, requires more, but we make no use of that here. Note also that Kolmogorov took conditional probability as a derived notion, as do we, not as an axiom.

To make this precise, we need to specify the rule, as the details vary. But not by much, so it will be enough to work with a single rule; the arguments and the consistency conditions we end up with are the same. We use the equivolume rule.

Consider then a partitioning $\{\alpha_k\}$ of $\mathcal{M}$ into cells of size τ, $k = 1, \dots, s$, so that $\lambda(\alpha_k) = \tau$, and $\lambda(\mathcal{M}) = s\tau$. Let there be $n \leq s$ cells with non-zero amplitude (so n microstates in all). Suppose that a macrostate in $\mathcal{M}$ represents an experimental outcome – say the configuration $\beta_\uparrow$ of a Stern-Gerlach apparatus, on obtaining the outcome 'spin-up' – and suppose further that $\beta_\uparrow$ is exactly partitioned by $\{\alpha_k\}$ -- that is, for some index set *K*, we have $\beta_\uparrow = \bigcup_{k\in K} \alpha_k$, where *K* is of some finite cardinality bounded by s. Then for any state $|\psi\rangle$ there will be an integral number $n_\uparrow$ of microstates in the macrostate $\beta_\uparrow$, where $n_\uparrow \leq n$, out n in total. The probability of 'spin-up' is then (exactly) $n_\uparrow/n$.

Our first consistency condition is that (for the same state $|\psi\rangle$ and decoherent history space) *every* admissible partitioning that exactly partitions $\beta_\uparrow$ yields the same ratio. If on a new exact partitioning of $\beta_\uparrow$ there are $n_\uparrow{}'$ spin-up branches out of n' in total, we require $n_\uparrow/n = n_\uparrow{}'/n'$.

The question now arises: What if $\beta_\uparrow$ is *not* exactly partitioned by the $\{\alpha_k\}$? Indeed, unless $\lambda(\beta_\uparrow)/\tau$ is an integer, there can be no index set *K* with the required properties. One might hope to vary τ to obtain an exact partitioning of $\beta_\uparrow$, but that trick only works once; since $\mathcal{M}$ is exactly partitioned, this is only possible if $\lambda(\beta_\uparrow)/\lambda(\mathcal{M})$ is rational.

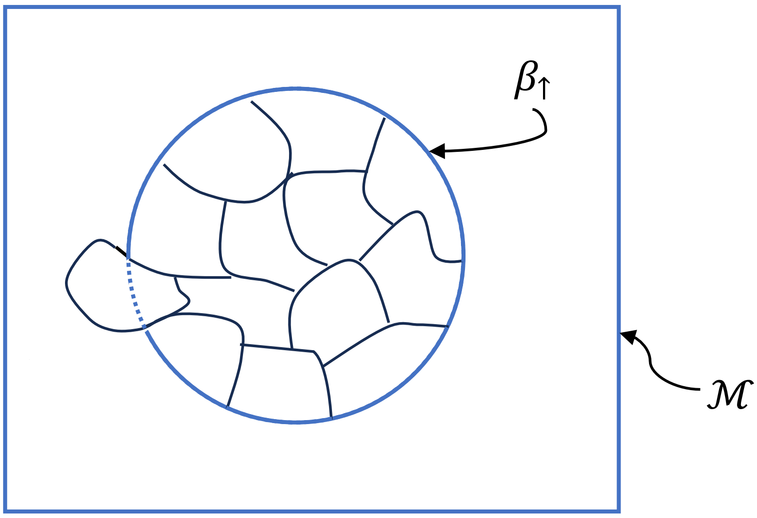

Fig 1

A bounded 2-dimensional continuous state space $\mathcal{M}$ and macrostate $\beta_\uparrow$. On any equi-area partitioning of the state space, at least one cell will straddle the boundary of $\beta_\uparrow$ (will be indefinite on $\beta_\uparrow$), unless the ratio of the areas of $\mathcal{M}$, $\beta_\uparrow$ is rational. (Only cells partitioning $\beta_\uparrow$ are shown.)

For a simple illustration, see Fig.1. Here $\mathcal{M}$ is a connected, bounded region in 2-dimensional Euclidean space, partitioned into n cells of arbitrary shapes, but each with exactly the same area τ. Let the boundary of the macrostate $\beta_\uparrow$ be a circle of radius r. Unless for some positive integer $m \leq n$:

$$r = \sqrt{\frac{m\tau}{\pi}}$$

it follows $\beta_{\uparrow}$ cannot be exactly partitioned, whatever the shapes of the cells – some cells will neither be contained $\beta_{\uparrow}$, nor in its complement. We shall say they are cells *indefinite on* $\beta_{\uparrow}$. If of non-zero amplitude, the associated branches are likewise indefinite as to $\beta_{\uparrow}$. But equally obviously, on an optimal partitioning for given $\beta_{\uparrow}$ and τ (reducing the number of indefinite cells to a minimum), at most *one* cell need be indefinite.

I take it that macrostates in $\mathcal{M}$ are in general defined independent of any partitioning.[23] Is the accompanying inevitability of indefinite cells, and indefinite branches, a difficulty? But indefiniteness per se is hardly novel, for it arises in a classical probability space $\langle \mathcal{M}, \sigma, \mu \rangle$. If we understand events in terms of sentences, a union of disjoint cells goes over to a disjunction of contradictory sentences, and intersections go over to conjunctions. Any sufficiently coarse-grained description will fail to specify some macroscopic property (will be indefinite on that property) – meaning, it will be disjunctive in this sense. But further, classically, the disjuncts will be assigned definite probabilities under μ , and it will follow that the sum of the probability of the disjuncts equals the probability of the disjunction, as follows from the additivity of μ at that finer-grained level.

In the quantum case there is no fixed, background probability measure μ, and we have only families of finite ensembles, families of relative-frequency probability distributions. But we require the same condition: the probability assigned to a cell α indefinite on $\beta_{\uparrow}$, referred to any admissible ensemble, where $\alpha = \gamma \cup \gamma'$ and γ is contained in $\beta_{\uparrow}$, whilst γ' is disjoint, is the sum of the probabilities of γ, γ', referred to any other admissible ensembles. We can also speak of the probabilities of the associated branches: the probability of the branch

$$P_{\alpha}|\psi\rangle = P_{\gamma}|\psi\rangle + P_{\gamma'}|\psi\rangle,$$

a Schrödinger-cat state, a superposition of a 'spin-up' state and a 'not spin-up' state, referred to one admissible ensemble, must equal the sum of the probabilities of the branches $P_{\gamma}|\psi\rangle$, $P_{\gamma'}|\psi\rangle$, as referred to any other admissible ensemble.

On reflection, it is clear that the same should apply when α is indefinite with respect to any region in $\mathcal{M}$, whether or not a macrostate – and, we belatedly realise, the new condition is a special case of the consistency condition already stated, likewise allowing cells of any size. Notice further that were there a unique 'finest' partitioning of $\mathcal{M}$, from which all others could be constructed by sums, the consistency condition would follow automatically from additivity of probabilities at that finest level; and that additivity, in turn, follow from frequentism, defined at that finest level. The consistency condition adds to frequentism only if there is no such finest partitioning.[24]

The key point is that there is nothing *incorrect* or *approximate* in the probability assigned to a cell α indefinite with respect to a macrostate. If $P_{\alpha}|\psi\rangle$ is itself a member of an admissible ensemble of n microstates, its probability is $1/n$, end of story, Schrödinger-cat

[23] Might macrostates not be definable in $\mathcal{M}$ at all? If so, we have the wrong decoherent history space (as specific to degrees of freedom, Hamiltonian, and state).

[24] Khawaja (2024) proposes to define such a finest partitioning on the basis of a supervaluationist approach to precisification, using the non-zero amplitude definition. However, that would seem to require consistency in our sense, and fail for the same reasons as the Boltzmann rule fails (as we shall see in the next section).

state or no. The difficulty is rather that such a microstate in an admissible ensemble, indefinite on $\beta_\uparrow$, counts neither positively nor negatively to its probability in that ensemble. But there is an obvious way of dealing with this. We can still define an *interval* for the probability of $\beta_\uparrow$: a lower bound, on counting branches in $\beta_\uparrow$, discarding all the indefinite cases (branches $P_\alpha|\psi\rangle$ that are not eigenstates of $P_{\beta_\uparrow}$); and an upper bound, on counting branches in $\beta_\uparrow$, including all the indefinite cases. So, if there are n branches in all, of which m are spin-up and r are indefinite on $\beta_\uparrow$, the probability of spin-up lies in $\left[\frac{m}{n}, \frac{m+r}{n}\right] \subset \mathbb{R}$. Call probabilities of this kind *interval probabilities*, as not rational numbers, but intervals of real numbers bounded by rational numbers.[25]

The further natural requirement is that all these intervals probabilities be consistent with one another. For this it is sufficient if they may all be true, and that in turn is assured if they have non-zero intersection. This is our second consistency condition: for any macrostate $\beta_\uparrow$, all interval probabilities, obtained from any admissible partitioning, must have non-zero intersection. If $\beta_\uparrow$ has an exact probability, that should be contained in that intersection too.

The two consistency conditions just adduced are similarly motivated and apply equally in the case of the equiamplitude rule. If we write exact probabilities $p \in \mathbb{Z}$ as $[p, p]$, and so treat all probabilities in a uniform notation,[26] the conditions can be reduced to the single requirement: the intersection of probabilities for any macrostate, defined in any admissible ensemble, is non-empty. We go further, as before, and require the same to hold for cells in $\mathcal{M}$ of any size, whether or not they are macrostates, consistent with decoherence.

In summary, we require, for a given decoherent history space:

C1 The probability of a cell in $\mathcal{M}$ in any exact partitioning is the same, for every admissible ensemble.

C2 The intersection of probabilities of any cell in $\mathcal{M}$ is non-empty, for every admissible ensemble.

As noted, *C1* is a special case of *C2*, but it will help in the sequel to treat them separately. We require them to be satisfied for any choice of $|\psi\rangle$, consistent with a given dynamics and decoherent history space, so they are far from trivial.

6. The equivolume rule

According to this rule, an admissible partition is any division of $\mathcal{M}$ into cells of the same size τ under σ, and the microstates of the associated ensemble are in 1:1 correspondence with cells of non-zero amplitude.

Condition *C1* now admits obvious counter-examples, even for fixed τ. Let $\mathcal{M}$ be partitioned into three macrostates $\beta_k, k = 1,2,3$, all with the same volume τ, and suppose

---

[25] Why not intervals of rational numbers? No reason, save as a matter of technical convenience, as we will later want to relate probability intervals to Born rule quantities (that in general are not rational numbers).

[26] Addition and multiplication are given in the obvious way ($[p, q] + [p', q'] = [p + p', q + q']$, $[p, q] \times [p', q'] = [p \times p', q \times q']$).

the corresponding branches $P_{\beta_k}|\psi\rangle$ all have non-zero amplitude. It follows that the probability of β_1 is 1/3. Keeping β_1 fixed, we vary the shapes of β_2' and β_3' at constant volume to obtain a new partitioning of $\mathcal{M}$, with equivolume macrostates β_1, β_2', β_3' (so $\beta_2 \cup \beta_3 = \beta_2' \cup \beta_3'$). Evidently at least one of $P_{\beta_2'}|\psi\rangle$ and $P_{\beta_3'}|\psi\rangle$ must be non-zero, since

$$P_{\beta_2'}|\psi\rangle + P_{\beta_3'}|\psi\rangle = P_{\beta_2'\cup\beta_3'}|\psi\rangle = P_{\beta_2\cup\beta_3}|\psi\rangle = P_{\beta_2}|\psi\rangle + P_{\beta_3}|\psi\rangle \neq 0;$$

but there is no reason at all why both should be non-zero. For example, $|\psi\rangle$ might vanish outside of $\beta_1 \cup \beta_2'$, with the region in β_2' where it is non-zero straddling the boundary of β_2 and β_3. In that case, β_1 is exactly partitioned in a new ensemble containing just two branches, $P_{\beta_1}|\psi\rangle$ and $P_{\beta_2',}|\psi\rangle$, so has relative frequency 1/2, not 1/3. Contradiction.[27]

Since exact probabilities for different ensembles are contradictory, interval probabilities are too. For any exact probability $p \in \mathbb{Z}$ is contained in a interval probability $\left[\frac{m}{n}, \frac{m+1}{n}\right], m < n$, for some partitioning of $\mathcal{M}$ into n equivolume cells of non-zero amplitude. If there are two exact probabilities for the same macrostate, differing by more than $2/n$, there will be two non-intersecting interval probabilities as well. Hence *C2* will be violated too.

The rule delivers inconsistent probabilities for the same macrostate, referred to different ensembles. Here is an alternative: modify the rule so that it applies only to the region of $\mathcal{M}$ on which $|\psi\rangle$ (as a wave-function in $L^2(d\lambda, \mathcal{M})$) is non-vanishing.[28] That would seem to exclude the kind of counterexample just given. The proposal is that the support of $|\psi\rangle$ on $\mathcal{M}$, the set $\mathcal{D}_{|\psi\rangle} = \{q \in \mathcal{M}; \langle q|\psi\rangle \neq 0\}$, is to be divided into equivolume cells, whereas before it was $\mathcal{M}$ that was so divided, with everything otherwise proceeding the same. Any cell wholly contained in the support of $|\psi\rangle$ on $\mathcal{M}$, of whatever size or shape, will have non-zero amplitude, so branch-counting reduces to cell-counting in $\mathcal{D}_{|\psi\rangle}$. As such, it will be fully determined by the additive volume measure σ; plausibly, consistency is then ensured.

The suggestion is ingenious but fails on what looks like a technicality. The set of points $\mathcal{D}_{|\psi\rangle} \in \mathcal{M}$ will not in general be a connected sub-manifold of $\mathcal{M}$; in 2-dimensions, excising points $\{q \in \mathcal{M}; \langle q|\psi\rangle = 0\}$ can leave as complicated a pattern of disconnected components imaginable. In general, no exact partitioning of even two components will be possible, for the same reason as in Fig.1. What then is the status of a cell α borderline, in this sense, on a component of $\mathcal{D}_{|\psi\rangle}$? Cells like this threaten to be ubiquitous. Since $\alpha \not\subset \mathcal{D}_{|\psi\rangle}$, by the amended rule, $P_\alpha|\psi\rangle$ does not count as a microstate, yet $\alpha \cap \mathcal{D}_{|\psi\rangle} \neq \emptyset$, so $P_\alpha|\psi\rangle \neq 0$, so it is not nothing, either. Before we considered cells of non-zero amplitude indefinite on a macrostate; that did not impugn their status as microstates, as we took pains to explain. The new sort of indefiniteness is more akin to indefiniteness as to *existence*. For borderline α, is $P_\alpha|\psi\rangle$ a superposition of existence and non-existence of α?

Maybe metaphysical sense can be made of the idea, maybe not; either way, there is the question of how cells borderline in this sense are to contribute – or not contribute – to relative frequencies of macrostates. One strategy is to include them all, but that can immediately be dismissed, for it simply returns us to the previous rule – which, as we have seen, is inconsistent. But to exclude all borderline cells also leads to inconsistency. In

[27] The argument is in effect a synchronic version of the one mounted against Graham's branch-counting rule, concerning diachronic consistency; see Saunders (1998 p.388-9), Wallace (2012 p.120).

[28] My thanks to an anonymous referee for this suggestion.

illustration, see Fig.2, where $\mathcal{D}_{|\psi\rangle}$ is the elliptical region, and the rectangle in its interior is the macrostate β. Two equivolume partitionings are considered, yielding contradictory relative frequencies for β. Nor does it help to use a mix of the two, and define a new kind of interval probability, with one bound obtained by not counting any borderline cells, the other by including them. Doing this, it is not even clear which bound is the greater (borderline cells may add to or subtract from the denominator, as well as numerator). Counter-examples to $C2$ are then not hard to construct.

Might we insist instead that $\langle q|\psi\rangle$ be non-vanishing everywhere on $\mathcal{M}$, with the possible exception of sets of zero-volume under λ? With that all these difficulties disappear. Take the original state $|\psi\rangle$, and add to it any analytic L^2-function on $\mathcal{M}$, with a sufficiently tiny amplitude: the resulting state can be as close as you like to the original (in the norm topology), yet have the desired effect. But it is a pointless exercise. Combined with the non-zero amplitude requirement, it ensures that the state be entirely irrelevant to the probabilities, for they will then be determined by β , $\mathcal{M}$ and σ alone. Why suppose that?

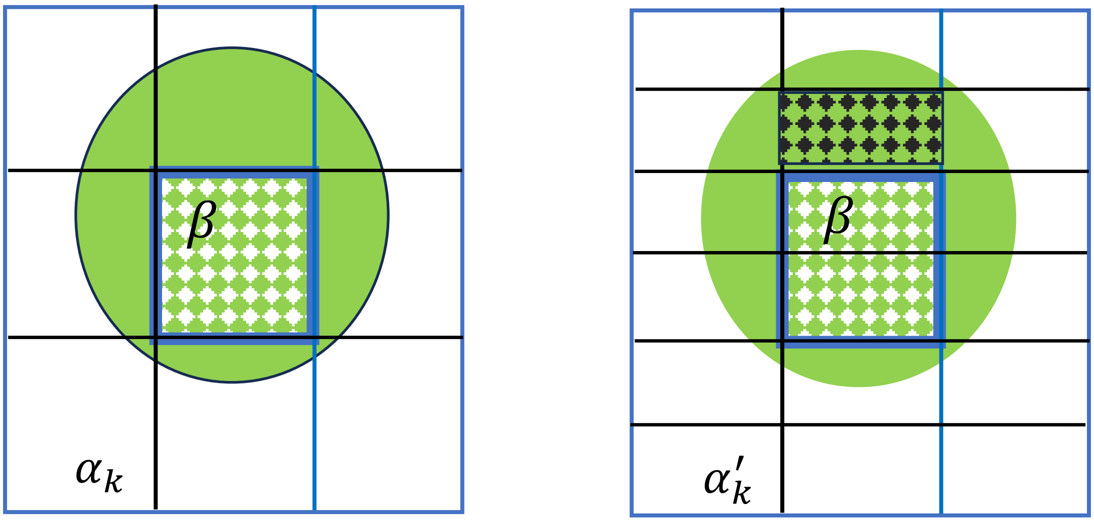

Fig 2a

Fig 2b

Fig 2: Macrostate β (outlined in blue) is exactly partitioned by $\{\alpha_k\}$, $\{\alpha'_k\}$, stippled in white, and is contained in the support of $|\psi\rangle$ on $\mathcal{M}$, the elliptical region shaded green. In Fig 2a no cell in $\{\alpha_k\}$ disjoint from β is contained in the ellipse, so the probability of β in $\{\alpha_k\}$ is 1, whereas in Fig 2b there is such a cell (stippled black), with 2 cells in β, so the probability of β in $\{\alpha'_k\}$ is 2/3, contradiction.

It appears that the only hope of the equivolume rule is to restrict all relative frequencies to a single, unique, fine-graining of $\mathcal{M}$, a unique ensemble, determined we know not how. But absent such a prescription, we do not have a rule for the probabilities at all.

7. The equiamplitude rule

We choose instead partitionings $\{\alpha_k\}$ in such a way that decohering microstates $P_{\alpha_k}|\psi\rangle$ all have equal amplitude, denote ω. Only ensembles of microstates like these are admissible. Assuming $|\psi\rangle$ has finite amplitude, $\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle} < \infty$, it follows that ω is restricted to the values $\omega(n) = \||\psi\rangle\|/\sqrt{n}$, where n is an integer (the number of branches). Here as before squares of amplitudes enter the picture, and with them square roots, traceable to the fact that Hilbert space is an inner-product space, equivalently, an L^2space.

Partitionings like this can always be found, as we require $\mathcal{M}$ to include a continuous variable. For a simple example, consider again the case where $\mathcal{M}$ is a connected region in $\mathbb{R}^2$, equipped with Lebesgue measure. For any n, and any function $f(x,y) \in L^2(\mathcal{M}, dxdy)$, we may always partition $\mathcal{M}$ into connected regions $\alpha_k, k = 1, \dots, n$, where $\bigcup_k \alpha_k = \mathcal{M}$, such that

$$\int_{\alpha_1} |f|^2 dxdy = \int_{\alpha_2} |f|^2 dxdy = \cdots = \int_{\alpha_n} |f|^2 dxdy.$$

It follows that the functions $\chi_{\alpha_k}(x,y)f(x,y), k = 1, \dots, n$, where χ_{α_k} is the characteristic function of α_k, are all orthogonal and have equal amplitude, where

$$f(x,y) = \sum_{k=1}^{n} \chi_{\alpha_k}(x,y)f(x,y)$$

is an expansion into n equiamplitude states.

As with the equivolume rule, the larger the number of branches, the smaller the cells α_k, so that eventually the decoherence condition will not be met. It follows that the number of equiamplitude branches definable at any time is strictly finite (this is true whether or not $|\psi\rangle$ is normalizable). Again, we expect there to remain a great deal of arbitrariness in the choice of partitioning, even once $\omega(n)$ is fixed, as the above example makes clear.

The consistency conditions remain as before. Consider first *C1*, and our previous counter-example, a partitioning of $\mathcal{M}$ into three cells $\beta_k, k = 1,2,3$, save now they have equal amplitude rather than equal volume. It follows that $\left|P_{\beta_k}|\psi\rangle\right| = \omega(3)$. As before, β_1 then has (exact) probability 1/3. Let $\{\beta'_k\}$ be a new equiamplitude partitioning into n' branches of amplitude $\omega(n')$, of which $P_{\beta_1}|\psi\rangle$ is one, so $\beta_1 = \beta'_1$, as before. If follows $P_{\beta_1}|\psi\rangle = P_{\beta'_1}|\psi\rangle$, as before, so their amplitudes are equal too; so $\omega(n') = \omega(3)$, and $n' = 3$. Therefore the relative frequency of β_1 relative to the new ensemble must again be 1/3. The counter-example fails.

To prove *C1* in full generality, suppose that for given $|\psi\rangle$ a macrostate $\beta_\uparrow$ is exactly partitioned by equiamplitude cells $\{\alpha_i\}$, n in total. It follows that for some finite index set K:

$$P_{\beta_\uparrow} = \sum_{i \in K} P_{\alpha_i}. \tag{5}$$

Let the cardinality of K be $n_\uparrow$. It follows that the relative frequency of 'spin up' is $n_\uparrow/n$. Since by assumption $\sum_{j=1}^{n} P_{\alpha_j}|\psi\rangle = |\psi\rangle$, from the definition of $\omega(n)$ it follows:

$$\langle\psi|\psi\rangle = \langle\psi|\left(\sum_{i=1}^{n} P_{\alpha_i}\right)\left(\sum_{j=1}^{n} P_{\alpha_j}\right)|\psi\rangle = n\omega(n)^2.$$

Further, from (5)

$$\langle\psi|P_{\beta_\uparrow}|\psi\rangle = \langle\psi|\left(\sum_{i=1}^{n} P_{\alpha_i}\right)\left(\sum_{k\in K} P_{\alpha_i}\right)\left(\sum_{j=1}^{n} P_{\alpha_j}\right)|\psi\rangle = n_\uparrow\omega(n)^2.$$

Taking the ratio:

$$\frac{\langle\psi|P_{\beta_\uparrow}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{n_\uparrow}{n}. \tag{6}$$

The LHS is the Born rule quantity for $\beta_\uparrow$. The agreement is exact; no approximation is involved. The same agreement with the Born rule follows for any other exact partitioning $\{\alpha_i{}'\}$ of $\beta_\uparrow$ with index set $K'$and n' branches (the argument is identical). Since the LHS of (6) is the same in each case, the RHS is too; therefore $n'_\uparrow/n' = n_\uparrow/n$, and *C1* is proved in full generality.

In the case of a macrostate $\beta_\uparrow$ defined independent of any partitioning, a similar argument shows that the Born rule quantity is contained within the limits of any (hence every) interval probability for $\beta_\uparrow$; therefore the intersection of all interval probabilities for $\beta_\uparrow$ must be non-empty, and condition *C2* is satisfied as well. We forego the proof.

There remains the question of whether the probabilities thus defined vary continuously with the state. A first point is that unlike for the equivolume rule, the partitioning cannot be held fixed whilst the state is varied. The α_k's must be varied (and the volumes $\lambda(\alpha_k)$ will vary) as the state is changed, to maintain the equiamplitude condition, even maintaining the same total amplitude and the same $\omega(n)$. Each of the n orthogonal branches $P_{\alpha_k}|\psi\rangle$ will vary continuously, each preserving its amplitude. The probabilities, the relative frequencies, will be unchanged, but what the probabilities are *about* is continuously changed. (This is easy to see if we allow the basis to change as well, whereupon we can work with finite-dimensional Hilbert spaces. Consider the simplest case of the state of a spin- ½ particle, and choose a basis, at each time, in which it is a superposition of orthogonal equiamplitude spin states. As the state is rotated, the later rotate as well, the basis and projectors on the basis continuously changing in time, each with constant relative frequency 1/2.)

The situation for probabilities of a macrostate defined independent of any partitioning is quite different. In general, for n branches, there is no exact partitioning, and we have only interval probabilities. But the optimal interval probability, as we saw, is of width 1/ n. Suppose, in the state $|\phi\rangle$, this is $\left[\frac{n_\uparrow}{n}, \frac{n_\uparrow+1}{n}\right]$, and consider an infinite sequence of states $|\phi_k\rangle, k = 1,2, \ldots$ and a corresponding infinite sequence of optimal interval probabilities $\left[\frac{n_\uparrow^k}{n}, \frac{n_\uparrow^k+1}{n}\right]$. Then if $|\phi_k\rangle \to |\phi\rangle$ in the norm topology, it follows $n_\uparrow^k \to n_\uparrow$, meaning there is an integer m such that if $k > m$, then $n_\uparrow^k \in \{n_\uparrow + 1, n_\uparrow, n_\uparrow - 1\}$. In the limit $n_\uparrow^\infty = n_\uparrow$.

To prove this, let the Born rule quantities be p_k, p, respectively:

$$p_k = \frac{\langle\phi_k|P_{\beta_\uparrow}|\phi_k\rangle}{\langle\phi_k|\phi_k\rangle}\ ; \qquad p = \langle\phi|P_{\beta_\uparrow}|\phi\rangle.$$

From before:

$$p_k \in \left[\frac{n_\uparrow^k}{n}, \frac{n_\uparrow^k + 1}{n}\right], \qquad p \in \left[\frac{n_\uparrow}{n}, \frac{n_\uparrow + 1}{n}\right]. \tag{7}$$

Since $|\phi_k\rangle \to |\phi\rangle$ as $k \to \infty$, $p_k \to p$, because the Born rule is continuous in the norm topology. If now p is not an integral multiple of $1/n$, then there exists $\epsilon \in \mathbb{R}$, $1/n > |\epsilon| > 0$ such that $p = n_{\uparrow}/n + \epsilon$. Since for any $|\epsilon| > 0$, there exists an integer m such that $|p - p_k| < |\epsilon|$ for all $k > m$, from Eq.(7) it follows $n_{\uparrow}^k = n_{\uparrow}$ for all $k > m$, and the conclusion follows. If instead p is an integral multiple, $p = n_{\uparrow}/n$, there exists integer m such that $|p - p_k| < 1/n$ for all $k > m$, when from Eq.(7), again the conclusion follows.

8. Epistemology

The equiamplitude rule gives consistent probabilities of macrostates as relative frequencies for every admissible partitioning, consistent too with the Born rule quantities, and varying continuously with the state. It is a reductive analysis, defined in terms of primitives that are widely used, in contexts that have nothing to do with probability. The ensembles are actual as well as finite, so unlike the decision-theory approach to probability, it is certainly about the world, or worlds. It is inspired by actual scientific practise and theory – not just quantum mechanics, but the ideas of Boltzmann and Gibbs in classical statistical mechanics. It is independent of metaphysics – for example, of personal identity, or self-locating uncertainty – and may even apply to other approaches to quantum foundations, for example pilot-wave theory.[29]

It appears, in short, to have all the virtues of finite frequentism, as advertised in Hájek [1996 p.72):

> Any aspiring frequentist with serious empiricist scruples should not give up on finite frequentism lightly. The move to hypothetical, frequentism say, comes at a considerable metaphysical price, one that an empiricist should be unwilling to pay. Finite frequentism is really the only version that upholds the anti-metaphysical, scientific inclination that might make frequentists of us in the first place. In any case, at first blush, it is an attractive theory. It is a reductive analysis, whose primitives are well understood; it apparently makes the epistemology of probability straightforward; unlike the classical and logical theories of probability, it appears to be about the world; and it seems to be inspired by actual scientific practice.

To those we might add: it also explains Kolmogorov's axioms. But one virtue is not entirely obvious. Is it true that the resulting epistemology is straightforward? Is it straightforward, even, in the classical case?

It is certainly simple for certain kinds of empiricists: those who, correctly insisting that only observed statistics provides an empirical basis for inferences about probabilities, identify those probabilities with those self-same statistics. Realists will insist that *all* extant statistics, whether or not observed, determine the probabilities – so in a single world, the statistics of ensembles spread out over all time, or all space, or both. It is to *those* relative

[29] Typicality, as introduced in Dürr et al (1992), evoked (what looks like hypothetical) frequentism; Hubert (2021) suggests that typicality is always at bottom a form of frequentism, so presumably also the typicality arguments attributed to Everett. They may better be grounded on our finite frequentism. However, the challenge posed by Valentini (1991, 2010) to pilot-wave theory remains the same.

frequencies that we wish to infer, on the basis of our local, observed collection of data. But as soon as a gap is opened between the inferred ensemble, and the one used as evidence for the inference, the epistemology of probability seems far from straightforward.

Are things any different in the quantum case? True, we face the apparently unnerving scenario that only a single microstate is available, to any one observer, at any one time, for any single trial; it seems that only a single member of a relevant ensemble is ever selected. In a single world, where the ensemble is spread out over space and time, we assume we make multiple selections from that one ensemble, at different places and times. But this comes at a certain price. The chance set-ups will not be *precisely* the same, on each trial; the ensemble is cobbled together out of several trials, only more or less physically similar, and of course at most one outcome can be observed for each trial. It is the cobbled-together ensemble that is supposed to define the physical probability (in terms of statistics), one member for each trial; of which a sub-ensemble is actually observed, and similarly involves trials of chance set-ups that are not exactly the same. Both target ensemble and observed ensemble involve these difficulties.

In contrast, in the quantum case, the ensemble arises for each trial, and the chance set-up for each member of the ensemble is *identically* the same, differing, at most, in outcome. There is no need to cobble together an ensemble out of sufficiently similar chance trials to define the physical probabilities. But to produce an observed ensemble, for a single observer or epistemic community, repetitions of trials are needed, involving chance set-ups that again, are not precisely the same – and we are back to a cobbled-together ensemble, as comprised by observed trials. But at least this complication afflicts only the evidence, not the probabilities themselves.

In brief: the epistemology of probability encounters much the same difficulties, whether in one-world or in many-world frequentism, and is not at all straightforward on either account of probability; and we ruefully note Hájek's careful wording, that finite frequentism *apparently* makes the epistemology of probability straightforward.

Going the other way, suppose the ensemble produced by a measurement process is known (on the basis, say, of quantum calculations, supposing Everettian quantum mechanics is true). What credences should we form? There is of course the question of what credence even *means*, in face of branching, when all the relevant details are known in advance, but the answer to this question has been well-rehearsed in the decision-theory literature: regardless of one's views on personal identity, or self-locating uncertainty, or language use, agents confronted with different branching scenarios must still choose among actions, even if only to curl up in a ball and hide. These actions, even no action, involve the allocation of finite present resources -- and therein lies the meaning of credence. In the one-world case, uncertainty is needed about future scenarios with respect to existence; for if lacking, why consider scenarios that we know will not arise? In the many-world case, in contrast, it is precisely *because* we are certain of the existence of those future scenarios, that we must take them into calculation – because we *know* they will all arise.

The question is then how, exactly, that calculation goes. It is clear where we want to end up: with Lewis' 'Principal Principle' (roughly, that if agents know the probabilities for various scenarios, then their credences in those scenarios should be weighted accordingly). Of course we could simply *use* the Principal Principle, as we do in conventional quantum mechanics, but now that we have an independent account of physical probability, we should aim higher: *why* ought an agent's credence in an experimental outcome, following a quantum measurement, match the relative frequency of that outcome, in the ensemble produced by that experiment?

I offer two answers. First, because it is prima facie reasonable. What better reason to give one eventuality more weight in our deliberations than another, when all else being equal, it is much more frequent than the other? And all else surely *is* equal: the agent, prior to measurement, bears exactly the same physical relations -- in space, time, *and* amplitude – to every element in the ensemble produced following the measurement.[30] That does not mean we should be careless of rare eventualities with very large negative utilities – again, just because, for an Everettian, those eventualities, though rare, are still there. To those branches we funnel compensation, siphoned off from those that are not so afflicted, otherwise known as taking out an insurance policy.

My second answer is that it is reasonable that credence match relative frequency insofar as the axioms of the decision-theory approach are reasonable. For it follows that agents who are reasonable in this latter sense will conform their credences to ratios in the Born rule quantities, and hence to relative frequencies, for the two agree.

However, those axioms are not entirely self-evident. Again, one may hope to do better. For at the core of the decision-theory derivations are the symmetries of equiamplitude states, to be realised at the level of observably distinct macrostates. Obtaining equiamplitude macrostates, for a given initial state, requires careful engineering. Several of the axioms used by Deutsch and Wallace are directed to underwriting an agent's indifference to this kind of fine-tuning. But on the present account, equiamplitude states, at the microscopic level, are already there for the taking, along with their symmetries, with no contrived experiments needed to reveal them. Maybe they can be exploited more directly.[31]

Acknowledgments

My thanks to two anonymous referees for very helpful suggestions and criticism.

References

Albert, D. [1992], *Quantum Mechanics and Experience*, Harvard University Press, Cambridge.

-- [2000], *Time and Chance*, Harvard University Press, Cambridge.

Barrett, J. [2011] 'Everett's pure wave mechanics and the notion of worlds', *European*

[30] Differences in phase, properly speaking, reduce to phase relations among the branches, not between the agent prior to branching and the individual branches (because of the irrelevance of the overall phase).

[31] I make some brief suggestions on how this may go in Saunders (2021). Short (2023), viewed as a derivation of an agent's credence function rather than physical probability, is a further step in this direction (with his 'invariance' condition following from frequentism).

Journal of Philosophy of Science **1**: 277-302.

-- [2016], 'Typicality in pure wave mechanics', *Fluctuation and Noise Letters* **15**, 1640009.

Bell, J. [1987], *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, Cambridge.

Boltzmann, L. [1877], 'Über die Beziehung zwischen den zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respective den Sätzen über das Wärmegleichgewicht', *Kaiserliche Akademie der Wissenschaften (Wien), Proceedings*, **76**, 373-435. Translated by K. Sharp and F. Matschinky in *Entropy* **17**, 1971-2009 (2015).

Brown, H., and G. Porath [2020] 'Everettian probabilities, the Deutsch-Wallace theorem, and the principal principle', in M. Hemmo and O. Shankar, *Quantum, Probability, Logic: the work and influence of Itamar Pitowsky*, Springer, Berlin.

Dawid, R., and K. Thébault (2015), 'Many worlds: decoherent or incoherent?', *Synthese* **192**, 1559-80.

Deutsch, D. [1999], 'Quantum theory of probability and decisions', *Proceedings of the Royal Society of London* **A455**, 3129–37.

-- [2016], 'The logic of experimental tests, particularly of Everettian quantum theory', *Studies in History and Philosophy of Modern Physics* **55**, 24-33.

Dürr, D, S. Goldstein, and N. Zanghi [1992], 'Quantum equilibrium and the origin of absolute uncertainty', *Journal of Statistical Physics* **67**, 843-907.

Einstein, A. [1949], 'Autobiographical notes', in *Albert Einstein, Philosopher-Scientist, Library of Living Philosophers*, P. Schilpp (ed.), Open Court, Chicago.

Everett III, H. [1957], '"Relative state"' formulation of quantum mechanics', *Reviews of Modern Physics* **29**, 454–62.

-- [1973], 'Theory of the universal wavefunction', in DeWitt, B and N. Graham (eds) *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton University Press, Princeton pp.3–140.

Franklin, A. [2024], 'Incoherent? No, just decoherent: how quantum many worlds emerge', *Philosophy of Science*, forthcoming.

Gell-Mann, M., and J. Hartle [1990] 'Quantum mechanics in the light of quantum cosmology', in *Complexity, Entropy, and the Physics of Information,* W.H. Zurek (ed.), Addison-Wesley, Reading.

Gibbs, J. [1902], *Elementary Principles in Statistical Mechanics*, Yale Centenary Publications, New Haven.

Goldstein, S. [2001], 'Boltzmann's approach to statistical mechanics', in *Chance in Physics: Foundations and Perspectives*, J. Bricmont, D. Durr, M. Galavotti, G. Ghirardi, F. Petruccione, and N. Zanghi, Lecture Notes in Physics 574, Springer-Verlag, Berlin (2001)

Graham, N. [1973], 'The measurement of relative frequency', in DeWitt, B and N. Graham (eds) *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton University Press, Princeton, pp.229-53, (1973).

Hájek, A. [1996], ‘‘Mises redux’-- redux: fifteen arguments against finite frequentism’, *Erkenntnis* **45**, 209-227.

Hubert, M. [2021], ‘Reviving frequentism’ *Synthese* **199**, 5255–5584.

Joos, E., and H. D. Zeh [1985], ‘The emergence of classical properties through interaction with the environment’, *Zeitschrift für Physik B Condensed Matter* **59**, 223–243.

Khawaja, J. (2024), ‘Conquering Mt Everett’, *British Journal for the Philosophy of Science*, forthcoming.

Kent, A. [2010], ‘One world versus many: The inadequacy of Everettian accounts of evolution, probability, and scientific confirmation’, in Saunders et al (2010 pp.307-54).

Kiefer, C. [1996], ‘Consistent histories and decoherence’, in D. Giulini, E. Joose, C. Kiefer, J. Kupsch, I.-O Stamatescu, and H.D. Zeh, *Decoherence and the Appearance of a Classical World in Quantum Theory*, Springer, Berlin (1996).

La Caze, A. [2016], ‘Frequentism’, in A. Hájek and C. Hitchcock, (eds.), *The Oxford Handbook of Probability and Philosophy*, pp.341-59, Oxford University Press, Oxford.

Lazarovici, D. [2023], ‘How Everett solved the probability problem in Everettian quantum mechanics’, *Quantum Reports* **5**, 407-17.

Myrvold, W [2016], ‘Probabilities in statistical mechanics’, in A. Hájek and C. Hitchcock, (eds.), *The Oxford Handbook of Probability and Philosophy*, pp.341-59, Oxford University Press, Oxford.

-- [2021], *Beyond Chance and Credence: a theory of hybrid probabilities*, Oxford University Press, Oxford.

Rowbottom, D. [2015], *Probability*, UK Polity Press, Cambridge.

Saunders, S. [1993], ‘Decoherence, relative states, and evolutionary adaptation', *Foundations of Physics*, **23**, 1553-1585.

-- [1998], ‘Time, quantum mechanics, and probability’, *Synthese* **114**, 373-404.

-- 2005 ‘What is probability?’, in *Quo Vadis Quantum Mechanics*, A. Elitzur, S. Dolev, and N. Kolenda, (eds.), Springer, Berlin.

-- [2020], ‘The concept ‘indistinguishable’, *Studies in History and Philosophy of Modern Physics* ***71***, 37–59.

-- [2021], ‘Branch-counting in the Everett interpretation of quantum mechanics’, *Proceedings of the Royal Society A* **477**, 20210600.

-- [2022], ‘The Everett interpretation: probability’, in E. Knox and A. Wilson, *The Routledge Companion to Philosophy of Physics,* Routledge, Abingdon.

Saunders, S., J. Barrett, A. Kent, and D. Wallace [2010], *Many Worlds? Everett, quantum theory, and reality*, Oxford University Press, Oxford.

Schrödinger, E. [1952], ‘Are there quantum jumps?’, *British Journal for the Philosophy of Science* **3**, 109-123, 233-242.

Short, A. [2023], ‘Probability in many-worlds theories’, *Quantum* **7**, 971.

Stoica, C. [2022], ‘Born rule: quantum probability as classical probability’, available online at https://arxiv.org/pdf/2209.08621.pdf

Tegmark, M. [1993], ‘Apparent wave function collapse caused by scattering’, *Foundations of Physics Letters* **6**, 571-590.

-- [2000], ‘The importance of quantum decoherence in brain processes’, *Physical Review* **E 61**, 4194–206.

Valentini, A. [1991], ‘Signal-locality, uncertainty, and the sub-quantum H-theorem’*,* *Physics Letters* **A156**, 5-11.

-- [2010], ‘De Broglie-Bohm pilot-wave theory: many worlds in denial?’, in Saunders et al (2010), pp.477-509.

von Kries, J. [1886], *Die Principien der Wahrscheinlichkeitsrechnung.* J. C. B. Mohr, Frankfurt.

Wallace, D. [2010], ‘How to prove the Born rule’, in Saunders et al (2010 pp.227-263).

-- [2012], *The Emergent Multiverse: Quantum theory according to the Everett Interpretation*, Oxford University Press, Oxford.

Wilson, A. [2020], *The Nature of Contingency: quantum physics as modal realism*, Oxford University Press, Oxford.

Zurek, W. [1981], ‘Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse?’, *Physical Review D* **24**, 1516-25.