9
Probabilities in Quantum Mechanics

9.1 Introduction

Much of the discussion to this point has proceeded as if the world conforms to classical physics. It does not. The question naturally arises, therefore, as to what, if any, of the foregoing is applicable to a world such as ours.

On the face of it, probabilities play a very different role in quantum theory than they do in classical physics. Unlike classical physics, quantum theory as typically formulated invokes an explicitly probabilistic postulate, the Born rule, which tells us how to use a quantum state description of a system to calculate probabilities of outcomes of experiments performed on the system.

Now, of course, one might take these probabilities as playing much the same role as they do in classical physics. Einstein himself suggested that quantum state ascriptions might be thought of as analogous to the probability distributions invoked in classical statistical mechanics (Einstein 1936: 340; 1954: 316–17). On this view, a complete microphysical description of a physical system would make no mention of quantum states; these would be associated with preparation procedures and, like the probability distributions of classical statistical mechanics, reflect less than complete knowledge of the state of a system. This motivates a project of attempting to construct a theory on which quantum states play that role. But that project, though well-motivated, does not succeed; there are very good reasons to believe that quantum states represent something in the furniture of the world (see Myrvold 2020b for an argument to that effect).

Quantum states, thus, are not like classical statistical mechanical distributions. Furthermore, it is frequently said that the fundamental revision in world-view that is required by acceptance of quantum theory is a renunciation of determinism. If this is right, it might be thought that the need for a notion that plays a role akin to that of objective chance and which is compatible with deterministic physics evaporates.

Things are not so simple! For one thing, there exists a theory, alternatively known as the de Broglie–Bohm pilot wave theory or Bohmian mechanics,

Beyond Chance and Credence. Wayne C. Myrvold, Oxford University Press (2021).

DOI: 10.1093/oso/9780198865094.003.0009

that is completely deterministic and (provided that probabilistic statements are attached to it in an appropriate way) reproduces all the probabilistic predictions of ordinary non-relativistic quantum mechanics. In order for the theory to function, the probabilities invoked in it cannot be mere credences, as, unlike credences but like objective chances, they can be imperfectly known and hypotheses about their values can be subjected to empirical test. If there is no way to make sense of this in the context of deterministic physics, the de Broglie–Bohm pilot wave theory is nonsense. This, of course, will not be my conclusion; I will argue in §9.4 that the probabilities in a theory of this sort ought to be thought of as epistemic chances.

Another reason that things are not so simple is that quantum mechanics, as it is invoked in quantum statistical mechanics, *is* a deterministic theory. Quantum theory, in its standard textbook presentation, has the peculiarity that it has two rules for evolving quantum states. In ordinary circumstances, the state is to be evolved via a deterministic law of evolution, expressed by the Schrödinger equation. When an experiment is done, the state is to be replaced by a new one, corresponding to the experimental outcome, and it is there, and only there, that probabilities enter, in the standard formulation. Quantum statistical mechanics—including, crucially, studies of equilibration—invokes only deterministic, Schrödinger evolution. There can, therefore, be no hope of deriving a monotonic approach to equilibrium from quantum mechanics alone, for the same reasons as in classical physics: reversibility and Poincaré recurrence. As in classical physics, there exist states that will evolve away from, rather than towards, equilibrium. We should not regard it as impossible that the systems we study in fact, possess those states, but merely improbable. And this means we will have to make sense of probability talk of that sort.

There is no consensus as to how (or, indeed, whether) to understand quantum theory as a description of the world. This gives rise to the literature on what is often called the "interpretation of quantum mechanics" (a potentially misleading phrase, because it might suggest that we have a bare uninterpreted formal theory, which we are to endow with a physical interpretation, and because different avenues of approach to the problem may involve distinct physical theories). Avenues of approach to this issue may be broadly classified by their approach to the so-called *measurement problem*. Probabilities play different roles on each of these classes of approach. I will argue, however, the considerations of the sort invoked in previous chapters, namely, evolution of reasonable credences via deterministic physical laws, have a role to play on each class of approach.

In this chapter, we will begin with some remarks about the process of equilibration, as it is studied in quantum statistical mechanics, and then dive into the measurement problem and the role of probabilities in various approaches to it.

For those unfamiliar with quantum theory and the philosophical issues associated with it, I present, in an appendix to this chapter, a brief overview of the formalism of quantum theory, and some of the philosophical issues. This will not be as self-contained as the introductions to probability, thermodynamics, and statistical mechanics. As a supplement, I suggest a pair of *Stanford Encyclopedia of Philosophy* entries: the entry "Quantum Mechanics" (Ismael 2015) for the basic formalism, and the entry "Philosophical Issues in Quantum Theory" (Myrvold 2018b), for an introduction to the basic philosophical issues, and pointers to further references.

9.2 Quantum equilibration

There are two approaches, at first glance strikingly different, towards the study of equilibration. On one approach, one considers an isolated system, but focusses attention on a limited set of dynamical variables of the system, typically thought of as its macrovariables. The other considers a nonisolated system, in interaction with its environment, and tracks the evolution of the state of the system. The two approaches are not as different from each other as might seem at first glance. In each case, we are investigating the evolution of a limited set of degrees of freedom of a larger system and disregarding the rest. The larger system is itself treated as isolated, and hence undergoing Hamiltonian evolution.

Suppose we have a system that consists of one subsystem S, the system of interest, and its environment, E. Suppose the joint system SE starts out in a state in which both S and E are in pure states—i.e. the initial state is a product state. If there are interactions between S and E, the joint system will typically evolve into a state in which the two are entangled. The reduced state of S will then be a mixed state—what is called an "improper mixture," to distinguish it from a proper mixture, which represents a situation in which the system has some pure state, and we are uncertain which state it is. But, as long as attention is focussed on the system S, and not on joint properties of the system and environment, then the improper mixture is operationally and observationally indistinguishable from a proper mixture. The upshot is that, if thermal states are to be represented by mixed states, then these need

not represent any sort of subjective uncertainty about the quantum state of the whole joint system *SE*.

Suppose we have an isolated quantum system that begins in some pure state, and evolves undisturbed. We focus our attention on some limited set of observables of the system, perhaps those that pertain to some subsystem. If the system has a discrete set of energy levels, then the quantum recurrence theorem (see Appendix to this chapter) applies, and the quantum state of the system, and hence also any subsystem, will be uniformly quasi-periodic. That is, after some time τ, which doesn't depend on the initial state, the system will return to a close approximation to its initial state.

In spite of this, it is possible for the state of a subsystem to approach some sort of equilibrium state, and remain there for an enormously long time, much longer than the time period of interest in which we will be interacting with it. Moreover, it is possible—this is the content of the theorem of Linden et al. (2009)—to show that for certain sorts of initial states of the combined system (namely, those that are superpositions of energy states to which a large number of energy levels contribute significantly, where what counts as "large" is: large compared to the Hilbert-space dimension of the subsystem), the state of the subsystem equilibrates, in the sense that the reduced state of *S* approaches some quasi-stable state and subsequently spends most of its time near it, with only rare fluctuations away from it.

A result such as this deals with long-term average behaviour of all initial states satisfying the specified condition. That tells us quite a bit, but, by itself, it doesn't tell us *anything at all* about what to expect in the immediate future. Some initial non-equilibrium states will head further away from equilibrium before heading back towards it; others will exhibit the sort of behaviour that experience leads us to expect. To apply equilibration considerations to any actual systems, we need to be reasonably confident that the state we have prepared is not one of those odd states that initially move further away from equilibrium before heading towards equilibrium.

Thus, results such as these, if they are to guide our expectations about near-term behaviour of the systems, require supplementation with considerations about the sorts of pure states that we, or nature, will be able to prepare. We don't want to *completely* exclude states that exhibit bizarre behaviour; what is wanted is that they be deemed improbable. We require, it seems, some sort of epistemic probability distribution over the quantum states of a system at a given time. The prospects for prescribing unique credences over possible initial states are no better in the quantum realm than in the classical. What

we can, however, hope for is that differences between reasonable credences will wash out.

Similar considerations apply to studies of the process called *decoherence*. These concern systems that interact with a large, noisy environment, and the tendency for the reduced density operator for such a system to approach a mixture of quasi-classical states. The dynamics of the joint system are unitary and reversible. For that reason, they cannot lead to decoherence for arbitrary initial states of the composite consisting of the system under study and its environment. Models of decoherence typically start with a state in which there are no correlations between system and environment, and evolve it forward. During the process, entanglement builds up between the system and its environment, but it is expected that the correlations that emerge are largely irrelevant to the subsequent evolution of the system *S*. As in the case of equilibration, states in which the process of decoherence is reversed ought not to be regarded as *impossible*, but only unlikely to be realized.

The phenomenon of decoherence raises the same questions about irreversibility raised by equilibration. The unitary evolution involved is invariant under temporal reversal. Thus, it can't be the case that arbitrary initial states of system + environment induce decoherence of the system. What makes the demonstrations of decoherence work is that one starts with a state in which the system and environment are unentangled. Forward evolution entangles the system with its environment, and the originally pure state of the system of interest becomes a mixed one. If one were to run the evolution backwards in time, one would get the same result in the backwards direction. But typically, models of decoherence are meant to be models of situations in which a quantum system has been subject to some preparation procedure which effectively screens off any correlation it might have with its environment.

That preparations of this sort are possible is a widespread assumption of scientific experimentation. It is worth asking what the grounds are for an assumption of this sort. After the preparation, it remains true that the system and the environment share a common past. And, if the dynamics are deterministic and invertible, it will simply not be true that there are *no* variables of system and environment that are correlated with each other. What happens, instead, is that these correlations become buried so deeply in the states of the system that they become essentially irrelevant to future evolution of macrovariables.

The application to actual systems of results about equilibration or decoherence in quantum mechanics, even if they concern states that are initially pure, will proceed much as the sorts of classical results that we have been considering. We begin with a class $\mathcal{C}$ of reasonable credences over the initial state of the system, which contain some uncertainty about the initial state, due to limitations of precision of state preparation. The goal should be to track the evolution of these epistemic mixtures, and show that the dynamics takes all mixtures in $\mathcal{C}$ to states that yield effectively the same probabilities for outcomes of feasible experiments.

9.3 The measurement problem

9.3.1 The measurement problem formulated

If quantum theory is meant to be (in principle) a universal theory, it should be applicable, in principle, to all physical systems, including systems as large and complicated as our experimental apparatus. Consider, now, a schematized experiment. Suppose we have a quantum system that can be prepared in at least two distinguishable states, $|0\rangle_S$ and $|1\rangle_S$. Let $|R\rangle_A$ be a ready state of the apparatus, that is, a state in which the apparatus is ready to make a measurement.

If the apparatus is working properly, and if the measurement is a minimally disturbing one, the coupling of the system S with the apparatus A should result in an evolution that predictably yields results of the form

$$\begin{aligned} |0\rangle_S|R\rangle_A &\Rightarrow |0\rangle_S|\text{“0”}\rangle_A \\ |1\rangle_S|R\rangle_A &\Rightarrow |1\rangle_S|\text{“1”}\rangle_A \end{aligned} \tag{9.1}$$

where $|\text{“0”}\rangle_A$ and $|\text{“1”}\rangle_A$ are apparatus states indicating results 0 and 1, respectively. Now suppose that the system S is prepared in a superposition of the states $|0\rangle_S$ and $|1\rangle_S$.

$$|\psi(0)\rangle_S = a|0\rangle_S + b|1\rangle_S, \tag{9.2}$$

where a and b are both non-zero. If the evolution that leads from the pre-experimental state to the post-experimental state is linear Schrödinger evolution, then we will have

$$|\psi(0)\rangle_S|R\rangle_A \;\Rightarrow\; a|0\rangle_S|\text{“0”}\rangle_A + b|1\rangle_S|\text{“1”}\rangle_A. \tag{9.3}$$

This is not an eigenstate of the instrument reading variable, but is, rather, a state in which system and apparatus are entangled with each other. The eigenstate-eigenvalue link, applied to a state like this, does not yield a definite result for the instrument reading. The problem of what to make of this is called the "measurement problem."

9.3.2 Approaches to the measurement problem

If quantum state evolution proceeds via the Schrödinger equation or some other linear equation, then typical experiments will lead to quantum states that are superpositions of terms corresponding to distinct experimental outcomes. It is sometimes said that this is in conflict with our experience, according to which experimental outcome variables, such as pointer readings, always have definite values. This is a misleading way of putting the issue, as it is not immediately clear how to interpret states of this sort as physical states of a system that includes experimental apparatus, and, if we can't say what it would be like to observe the apparatus to be in such a state, it makes no sense to say that we never observe it to be in a state like that!

Nonetheless, we are faced with an interpretational problem. If we take the quantum state to be a complete description of the system, then the state is, contrary to what we would antecedently expect, not a state corresponding to a unique, definite outcome. This is what led J. S. Bell to remark, "Either the wavefunction, as given by the Schrödinger equation, is not everything, or it is not right" (Bell 1987a: 41; 1987b; 2004: 201). This gives us a (*prima facie*) tidy way of classifying approaches to the measurement problem:

I. There are approaches that involve a denial that a quantum wave function (or any other way of representing a quantum state), yields a complete description of a physical system.
II. There are approaches that involve modification of the dynamics to produce a collapse of the wave function in appropriate circumstances.
III. There are approaches that reject both horns of Bell's dilemma, and hold that quantum states undergo unitary evolution at all times and that a quantum state-description is, in principle, complete.

We include in the first category approaches that deny that a quantum state should be thought of as representing anything in reality at all. These include variants of the Copenhagen interpretation, as well as pragmatic and other anti-realist approaches. Also in the first category are approaches that seek a completion of the quantum state description. These include hidden-variables approaches and modal interpretations. The second category of interpretation motivates a research programme of finding suitable indeterministic modifications of the quantum dynamics. Theories that incorporate such modifications are called *dynamical collapse theories*. Approaches that reject both horns of Bell's dilemma are typified by Everettian, or "many-worlds" interpretations.

As already mentioned, I don't think that the project of constructing an adequate theory on which quantum states do not represent anything in physical reality can succeed. Among those who deny the reality of quantum states, the most prominent is the school that goes by the name of *QBism* (for Quantum Bayesianism), who claim that a quantum state is nothing other than the credence of some agents. My own view is that proponents of this view don't sufficiently appreciate the impact of theorems such as those of Pusey et al. (2012) and Barrett et al. (2014) for a view of that sort. See Myrvold (2020c) for defense of this claim.

9.4 Probabilities in de Broglie–Bohm pilot-wave theories

The de Broglie–Bohm pilot-wave theory, also known as Bohmian Mechanics, is a deterministic theory on which experiments have definite outcomes. This is achieved by taking the rhetoric of "wave-particle duality" seriously; on this theory, the quantum state, represented by a wave function, is physically real, but so are particles, whose motion is guided by the wave-function. The number of physicists actively working on the theory has never been large, but even within this group, there is a divide on how to think of the probabilities in the theory. The divide mirrors a divide on how to think of probabilities in statistical mechanics. One group—which we may call the *Bohmian Mechanics* group—includes Shelly Goldstein, Detlef Dürr, Nino Zanghì, and their collaborators. They take it as a fundamental postulate of any physical theory that there be some measure over the possible ways that the world can be that can be used to judge typicality of properties—a property P is typical just in case the set of states having P is overwhelmingly larger, on the given typicality measure, than the set of states that lack it. The

other group, which includes Antony Valentini and his collaborators, seeks to understand the probabilities standardly invoked in the theory as the result of a process of relaxation to equilibrium, thought to have taken place in the early universe.

More details on this, below. First, a brief presentation of the theory.

As the quantum state of an n-body system evolves according to the Schrödinger equation, the probability density over positions,

$$\rho(q,t) = |\Psi(q,t)|^2 \tag{9.4}$$

evolves also. (Here we are using q as a variable for a point in $3n$-dimensional configuration space.) Its evolution satisfies a continuity equation of the same form as the density of a conserved fluid flowing around configuration space,

$$\frac{\partial \rho}{\partial t} + \sum_{k=1}^{3n} \frac{\partial}{\partial q_k} j_k = 0, \tag{9.5}$$

where $\mathbf{j}$ is the probability current density on phase space,

$$j_k(q,t) = -\frac{i\hbar}{2m_k}\left(\psi(q,t)\frac{\partial}{\partial q_k}\psi^*(q,t) - \psi^*(q,t)\frac{\partial}{\partial q_k}\psi(q,t)\right). \tag{9.6}$$

The probability density thus acts like a fluid with a velocity field $\mathbf{v}(q,t)$, given by,

$$v_k(q,t) = j_k(q,t)/\rho(q,t). \tag{9.7}$$

The de Broglie-Bohm pilot wave theory posits that such a system actually does consist of n particles, obeying a non-Newtonian law of motion, which requires that the instantaneous velocities of the particles always satisfy (9.7).

The guidance condition (9.7), together with the Schrödinger equation for the system's wave function, defines a flow on the configuration space of the n-particle system. This permits us to evolve probability distributions over configurations of the particles. Let $\omega(q,t)$ be a density function on configuration space. The de Broglie-Bohm law of motion for the particles, (9.7), has the pleasing feature that, *if* for some time t_0,

$$\omega(q,t_0) = |\Psi(q,t_0)|^2, \tag{9.8}$$

then the same holds for all other times: for all times t,

$$\omega(q,t) = |\Psi(q,t)|^2 , \tag{9.9}$$

Thus, evolution of the density $\rho = |\Psi|^2$ via the guidance condition meshes with evolution of the wave function via the Schrödinger equation. This property of the probability density $\rho(q,t)$ is called *equivariance*. It is a cousin of invariance of probability distributions.[1]

This raises the question: what sense can we give to talk of probability in this context? We have a deterministic theory, and, recall, it was a commitment to determinism that led Bernoulli, Laplace, et al. to hold that probability was entirely an epistemic matter. But, if the probabilities involved are to be thought of as credences, reflecting our ignorance of the exact positions of the particles, then we need to ask: Why should these credences have anything to do with the standard, Born-rule probability distribution, represented by $|\Psi(q,t)|^2$?

Insight into this matter is found in an article by Detlef Dürr, Shelly Goldstein, and Nino Zanghì (Dürr et al. 1992), whom will we will refer to as DGZ.

They first explain what it might mean, in the absence of the collapse postulate, to talk about the wave function of a subsystem of the Universe (Recall that the collapse postulate is introduced to deal with the fact that, in its absence, at the end of an experiment, the system is not in an eigenstate of the observable "measured", but, rather, is entangled with the apparatus.). This has to do with the *effective wave function* of a system.

Suppose that we have partitioned the universe into a system of interest, S, and everything else, which we will call the environment, E. Let x be a variable that ranges over the possible configurations of S's particles, and let X be its actual (perhaps unknown) configuration. Let y be a variable that ranges over the possible configurations of the environment E, and let Y be its actual configuration.

[1] In the case of classical evolution, for a time-independent Hamiltonian, we had a time-independent flow on phase space, and hence it was possible to have an invariant probability distribution. For the de Broglie–Bohm dynamics, except in the special case of a stationary wave function, the velocity field on configuration space changes with time, which gets in the way of there being an invariant probability distribution. But a probability distribution *can* keep up with the changing flow.

Define the *conditional wave function* of S by

$$\psi_Y(x, t) = \Psi(x, Y, t). \tag{9.10}$$

Suppose that the system S is isolated from the environment (or at least effectively so), a necessary condition for S to be in a pure state more than momentarily. In that case, the Hamiltonian is a sum of a Hamiltonian for S and a Hamiltonian for E.

$$\hat{H} = \hat{H}_S + \hat{H}_E. \tag{9.11}$$

Suppose, also, that the universal wave function Ψ has evolved to a form

$$\Psi(x, y, t) = \psi(x, t)\Phi(y, t) + \Psi^{\perp}(x, y, t), \tag{9.12}$$

where the sets of values of y for which $\Phi(y, t)$ and $\Psi^{\perp}(x, y, t)$ are non-zero are macroscopically disjoint, meaning that there is some macroscopic function of y—say, a pointer position—whose value distinguishes the supports of $\Phi(y, t)$ and $\Psi^{\perp}(x, y, t)$. This is the sort of state that would be the result of a measurement whose result is recorded with the pointer position. Suppose also that Y, the actual value of y, is in the support of Φ.

Then the conditional wave function takes the form

$$\psi_Y(x, t) = \psi(x, t)\Phi(Y, t). \tag{9.13}$$

Moreover, as a consequence of the absence of interaction terms between S and E in the Hamiltonian, the conditional wave function obeys the Schrödinger equation with H_S as its Hamiltonian, and the system has, in effect, its own wave function $\psi(x, t)$. Under such conditions, DGZ call ψ the *effective wave function* of the system S.

DGZ then demonstrate that, if the initial configuration Q is chosen at random according to the standard quantum probability measure, then the conditional distribution of X, conditional on the *actual* value of Y, is correctly given by the conditional wave function $\psi(x)$. Moreover, if one were to perform multiple preparations on a large number of distinct subsystems, and select those with the same effective wave functions $\psi(x)$, and then perform the same experiment on each system, with high probability the frequencies of experimental outcomes would closely match the Born-rule probabilities calculated from ψ.

What DGZ have proven, then, is, in their own words,

> that for *every* initial ψ, this agreement with the predictions of the quantum formalism is obtained for *typical*—i.e., for the overwhelming majority of—choices of initial q. And the sense of typicality here is with respect to the only mathematically natural—because equivariant—candidate at hand, namely, quantum equilibrium.
>
> Thus, on the universal level, the physical significance of quantum equilibrium is as a measure of typicality, and the ultimate justification of the quantum equilibrium hypothesis is, as we shall show, in terms of the statistical behavior arising from a typical initial configuration.
>
> (Dürr et al. 1992: 859)

All of this is a valuable, and, indeed, indispensable contribution towards understanding the role of probabilities in Bohmian mechanics. It still leaves us, however, with the same question as was raised in previous chapters, about the standard equilibrium measures invoked in classical statistical mechanics—what is the status of this typicality measure? In what sense is it a "natural" measure?

Here's what DGZ say about this (p. 868).

> we regard the quantum equilibrium distribution **P**, at least for the time being, solely as a mathematical device, facilitating the extraction of *empirical* statistical regularities from Bohmian mechanics …and otherwise *devoid of physical significance*. (However, as a *consequence* of our analysis, the reader who so wishes can safely also regard **P** as providing a measure of *subjective* probability for the initial configuration Q. After all, **P** could in fact be *somebody's* subjective probability for Q).

We *can* regard the quantum measure as providing a measure of subjective probability. But *must* we, on pain of unreasonableness?

DGZ say that the standard quantum measure is mathematically natural because equivariant. In a similar vein, one could say (and some do; see Lebowitz 1999a: 520; 1999b: S348; 2001: 53) that the equilibrium measures of standard statistical mechanics are mathematically natural because invariant. This still leaves the question of the significance of equivariance and invariance, in their respective contexts. The answer we have advanced, in

the previous chapter, for the significance of invariance is that the standard equilibrium measures are singled out as appropriate *for systems that have undergone a process of equilibration*, where equilibration is understood in terms of convergence of measures—a system that has been evolving undisturbed since a given time t_0 has equilibrated by time t if its dynamics are such that all reasonable credence functions about conditions at t_0 yield, when evolved forward via the system's dynamics, effectively the same probabilities about all macroscopically measurable properties. Any attractor measure of this sort would have to be an invariant measure.

This suggests that we take probabilities in the de Broglie-Bohm theory in the same way, as equilibrium probabilities, the result of a process of "washing-out" of initial conditions. And, indeed, this is precisely what has been suggested by Antony Valentini (1991a; 1991b). He shows that an analog of what Tolman (1938) called the "Generalized *H*-theorem" (which, as you may recall, we met in §8.5.3) holds in Bohmian mechanics. He calls this the "subquantum *H*-theorem." As in the statistical mechanical case, the proof proceeds by first defining a coarse-graining operation on probability distributions, then demonstrating that the de Broglie-Bohm dynamics lead to an increase of an appropriately modified entropy.

Since that time, there has been an abundance of theoretical work and computer simulations demonstrating the effectiveness of Bohmian dynamics in bringing about convergence to the equivariant distribution (with exceptions; one can also find systems for which convergence does not occur). See Efthymiopoulos et al. (2017) for an overview.

These results, of course, raise the question of the status of the input measures. You can probably guess what my answer is—we should take them to be reasonable credences, and take the probabilities in de Broglie–Bohm theory to be epistemic chances, in the sense of Chapter 5.

As Dürr, Goldstein, Zanghì, and Valentini all stress, there is a deep analogy between equivariance in Bohmian mechanics and equilibrium in statistical mechanics. One common feature is the value of nonequilibrium as a resource. As Valentini (2002a; 2002b; 2002c) has shown, on the de Broglie–Bohm theory, if one had knowledge of the configuration of a system more precise than the quantum equilibrium probabilities allow, this informational nonequilibrium would permit one to perform a variety of tasks that are provably impossible within standard quantum mechanics.

9.5 Dynamical collapse theories: probabilities as objective chances

The collapse postulate, as formulated by von Neumann and Dirac, is a bit mysterious, as it invokes distinct laws of evolution depending on whether a measurement is occurring. Measurements, however, are just physical interactions that have a certain purpose, and it is not at all clear how it could be that physical systems obey fundamentally different laws depending on whether they are involved in a measurement.

It turns out, however, that it is possible to formulate a unified dynamics for microscopic and macroscopic systems (as Ghirardi et al. 1986 put it), that with high probability, closely mimics the ordinary Schrödinger evolution for an isolated system consisting of a small or not-too-large number of particles, and nevertheless, with high probability, suppresses superpositions of macroscopically distinct states. The best known of these is the GRW theory (Ghirardi et al. 1986), which the authors themselves refer to as *Quantum Mechanics with Spontaneous Localization* (QMSL). This theory, however, does not respect the symmetrization and antisymmetrization requirements on systems of identical particles. There is, however, a collapse theory that does; this is the *Continuous Spontaneous Localization* theory (CSL) (Pearle 1989; Ghirardi et al. 1990).

Among non-Everettian theories, that is, theories in which experiments have unique outcomes, dynamical collapse theories have the advantage over hidden-variables theories that they need not involve any action-at-a-distance. Though probabilities of spatially separated events are not independent of each other, it is probabilities of random experimental outcomes that are correlated, and interventions on a system do not affect the probability of any events at a distance. Nor do such theories require any distinguished relation of distant simultaneity; they are compatible with relativistic causal structure (see Myrvold 2016 for further discussion of this point, and Myrvold 2018a; 2019b for a discussion of ontology for such theories). There, are, in fact, relativistic versions of these theories. There is a relativistic version of the GRW theory, which is restricted to a finite number of non-interacting particles (Dove and Squires 1996; Dove 1996; Tumulka 2006). Also, in different ways Bedingham (2011a,b) and Pearle (2015) have extended the CSL model to relativistic quantum field theories, without the restriction to non-interacting theories.

There is no trouble in interpreting the probabilities in these theories; they are objective chances. The laws of motion of these theories are

indeterministic. Given a quantum state at a time t_0, the theory does not prescribe a unique state at a future time t_1. What we obtain from the theory is a probability distribution over future states.

What role will these probabilities play in statistical mechanics? As mentioned above, equilibration results in quantum statistical mechanics are invariably couched in terms of unitary, deterministic evolution. The reversibility argument, therefore, applies in full force. It cannot be proven that we will observe relaxation to equilibrium in the near term for *all* initial states; as in the classical case, if there are initial states that evolve towards equilibrium, there are also states that evolve away from it. We ought not to regard such states as impossible, but merely improbable. As in the classical case, one can hope to prove that, given reasonable credence about states at a given time, one ought to expect equilibration in the near term. Stochastic evolution will place some constraints on credences about states. Even if your credences about the initial state of a system were concentrated on a subset of initial conditions that lead to antithermodynamic behaviour, knowing that a stochastic collapse mechanism is in place will fuzz them out.

Can such lawlike fuzzification completely replace epistemic considerations, rendering them entirely superfluous, as suggested by Albert (1994; 2000: ch. 7)? To answer this question requires investigation—and, as far as I know, this has never been done—whether it can be proven, for a suitable class of Hamiltonians that we believe contains realistic ones, that the sort of stochastic dynamics provided by the GRW or CSL theories yields, for macroscopic systems, an approach to standard equilibrium states for all initial states.

If a result of that sort is obtainable, it will not (contrary to what one might expect) take the form of relaxation towards a stationary equilibrium state. The reason for this is that, on theories of that sort, the standard equilibrium states are *not* stationary states. These theories involve (very slight) violation of conservation of energy, which would result in a gradual warming of an isolated body (too slight to be observed by currently feasible experiments). What one could hope to demonstrate is convergence towards, not a stationary distribution, but one that is equivariant under the stochastic dynamics.

There is a worry about this sort of project, mentioned by Albert (2000: 156–9) as having been raised by Larry Sklar and Philip Pearle. This sort of worry has to do with equilibration, or lack thereof, in systems for which we should expect a good approximation to unitary evolution. One such case would be a gas consisting of around 10^5 molecules. Would it have a

tendency to spread out, if originally confined to a small region of the available volume? If so, then this tendency cannot be attributed to the stochasticity of GRW dynamics, as, for a system that small, the theory predicts a close approximation to unitary evolution. Another case has to do with spin echo experiments (recall these from §8.5.1), in which a tendency to approach a state that looks macroscopically random can be observed, and, yet, demonstrably, the evolution has been unitary, or close to it; otherwise, it would not be possible to restore the initial state by the reversing pulse. In both cases, there exist states that don't tend to equilibrate, but these are sensitive to small perturbations, and it would be extraordinarily difficult to reliably prepare them, and we don't expect them to occur in nature. There is, it seems, to be an ineliminable role for epistemic uncertainty about initial conditions to play.

9.6 Everettian interpretations: no probabilities, but a working substitute (perhaps)

Everettian approaches reject both horns of Bell's dilemma. On such an approach, the quantum state obtained by taking the state at some time and applying deterministic, unitary evolution, is taken to be correct, and a quantum state is taken to be capable of representing all of physical reality. When an experiment or some other event occurs that, on the usual way of thinking about things, has a plurality of possible outcomes, whose respective probabilities are calculated from the quantum state by the Born rule, on Everettian interpretations, all eventualities are realized. Provided that the outcomes get recorded in macroscopic variables that are subject to decoherence, any opportunity to demonstrate this plurality of outcomes via an experiment that exhibits interference between these terms will quickly be lost, giving rise to what is, in effect, a branching of worlds. You, the experimenter, will have successors that share your memories and differ in the outcomes they perceive, and none of these is privileged as being the unique future you.

This is a deterministic theory. The evolution at all times is the deterministic, unitary evolution, and the dynamics, together with the state at some time, uniquely determines what will happen after the experiment—there will be branching. Nor is there any room for asking, prior to the experiment, "Which outcome will I perceive?"—the answer is that multiple versions of you will perceive all of the outcomes with nonvanishing Born-rule probability.

In anything like any of the usual senses of "probability" it makes no sense, in the context of a branching theory like this, to talk of a probability that this or that experimental result will be the one that occurs (or the one that is perceived) as there is no such thing as *the* experimental result that will occur. They all will occur, on various branches. It is a presupposition of ordinary probability talk that we assign probabilities to the elements of a set of mutually exclusive alternatives, one, and only one, of which is realized. This presupposition seems not to get a grip in the context of a branching theory.

Is this a problem for Everettian approaches?

One attitude that might be taken is that, beyond epistemic uncertainty about the quantum state, there is no need for any talk of probability. It *is* a deterministic theory, after all!

There are two closely related problems associated with adopting such an attitude. One is how one would ever get to know what the quantum state of a system is in the first place. On the usual account, if one wants to know what quantum state to associate with a given preparation procedure, one does a statistical analysis of experiments involving multiple iterations of the preparation procedure. This proceeds as outlined in §2.4. This requires that one treat the Born rule probabilities as chances, about which one can gain information; in particular, it requires that one have conditional credences, conditional on hypotheses about what the quantum state is, that satisfy the Principal Principle with the Born rule probabilities taken as chances.

Another problem has to do with how one might come to accept Everettian quantum mechanics in the first place. A great deal of the evidence that quantum mechanics is getting something right is statistical in nature. This is evidence that quantum mechanics is getting the *chances* of outcomes of experiments right. If this is jettisoned as nonsensical in an Everettian context, one runs the risk of rendering the theory empirically self-undermining.

The problem of either making sense of probabilities of outcomes of experiments or else finding something else that does the same job, in such a way that the statistical evidence counts in favour of quantum mechanics, even when interpreted in an Everettian vein, is known as the *Everettian evidential problem*.[2]

[2] The term "evidential problem" stems, I believe, from Myrvold (2005), though this was not the first occasion on which the problem was raised. See Wallace (2006; 2012: ch. 6) and Greaves and Myrvold (2010) for somewhat different approaches to addressing the problem.

A significant step towards dealing with the problem of probability in an Everettian context was taken by David Deutsch (1999), who introduced the use of decision theory into this context. The logic of Deutsch's argument has been clarified by David Wallace (2003; 2007; see Wallace 2012: ch. 5 for his fullest statement of the argument). The question addressed is: How would an agent who accepted Everettian quantum mechanics and knew the quantum state make decisions? Suppose, for example, that an experiment is about to be done, and goods are to be distributed on the various post-experimental branches in a manner dependent on the experimental outcome. The agent is to indicate preferences between various options of post-branching distributions of goods. (Barry Loewer has dubbed such scenarios *brambles*, as the branching analog of gambles.) For example, given a certain quantum state preparation of a spin-1/2 particle, to be followed by a spin "measurement," you might be asked whether you prefer a scenario in which your successors who see a "+" get a reward and those who see "−" get nothing to a scenario in which the rewards are reversed. The Deutsch–Wallace strategy is to argue that, on the basis of a set of reasonable constraints on the agent's preferences between brambles, an agent's choices between actions will maximize a weighted average of utilities across branches, using the Born-rule weights. That is, such an agent's preferences will match those of someone who thought of the experiments in the usual way, as having unique outcomes with the Born-rule weights as chances.

The argument can be extended to permit an agent who is uncertain about what quantum state to associate with a given preparation to update her credences upon receiving statistical information (see Greaves 2007). The idea is to bring in considerations of accuracy as an epistemic utility, and apply Everettian decision theory to the choice of a credence-updating strategy.

Suppose that one accepts the conclusion of the Deutsch–Wallace argument. This still doesn't solve the Everettian evidential problem, as the argument applies only to an agent who already accepts Everettian quantum mechanics. But it does suggest a strategy for dealing with that problem, outlined by Greaves (2007) and developed more fully by Greaves and Myrvold (2010). One can consider an agent who thinks of experiments as branching events, without commitment to Everettian quantum mechanics or any other theory of how the branching proceeds. The idea is to develop a decision theory for such an agent, by imposing reasonable constraints on her preferences between brambles. Further constraints lead to a representation of the agent's credences as credences about *objective branch weights*. Under the conditions of the representation, hypotheses about the values of these objective branch

weights can be tested empirically, in much the same way that hypotheses about chances are, with the result that these branch weights can be estimated, independently of any theory about them. If, then, a theory, such as Everettian quantum mechanics, is formulated, which furnishes a prediction about these weights, one can compare the branch weights predicted by the theory to those estimated empirically and thereby confirm or disconfirm the theory.

Another approach to probability is due to Sebens and Carroll (2018). Their approach focuses on a sort of uncertainty that exists in an Everettian context, even if the agent knows the quantum state and how it evolves. After an experiment has been performed and the result recorded, before becoming aware of the result an agent might have a sort of self-locating uncertainty—though she knows that she has counterparts on branches corresponding to each result, she doesn't yet know which sort of branch she is on. This sort of uncertainty has been called *Vaidman uncertainty*, after the discussion in Vaidman (1998). Sebens and Carroll impose conditions on an agent's post-branching credences that have the effect that her credence in a given result should match the Born-rule weight of branches on which that result obtains.

What drives arguments such as the Deutsch–Wallace argument and the Sebens–Carroll argument is the fact that, in an Everettian context, there is nothing except the quantum state that could be as a basis for formulating a decision rule or a rule for assigning credences to branches. If all one has is the quantum state, and, if one is to respect the condition that zero probability be assigned to subspaces orthogonal to the quantum state, then Born-rule probabilities are pretty much the only option. One may have reservations about the precise form of these arguments, but it is not a simple matter to formulate a coherent alternative to taking Born-rule weights to play the role of probabilities.

Let us assume, for the moment the arguments or a suitable modification of them succeed. What, then, will we say about the role of probabilities, or their surrogates, in the theory? Do they eliminate the need for the sorts of concerns that have been the subject of this book, having to do with epistemic limitations on our knowledge of the physical state?

First of all, decoherence is required to get the theory off the ground, and to obtain a branching structure in the first place. As we have already observed, on reversibility grounds it cannot be the case that this occurs for arbitrary initial states. As with equilibration, we should regard the realization of states that fail to decohere, or in which decoherence appears and then is reversed, as not impossible, but only very improbable.

In addition, since equilibration results in the literature typically involve only deterministic, unitary evolution, and the only role that the Born-rule probabilities play in them is in shaping one's expectations about what will be observed when one performs an experiment on the system, the Everettian account of them will be pretty much the same as everyone else's, and, as argued above, will rely on considerations about limitations of control by us, or, indeed, of any physical process, over what states can be reliably produced.

Thus, it seems, on this approach, as on the other major approaches to the measurement problem, there will be a role for something akin to epistemic chances to play.

9.7 Can classical statistical mechanics stand on its own two feet?

Classical statistical mechanics is applicable when quantum effects can be ignored. A sufficient condition for this is a positive Wigner function that evolves according to the Liouville equation (see §9.9.4). In such a case, there is a built-in restriction on the density function—it cannot be one that violates the uncertainty relations. We have argued that equilibration should be regarded as a process of convergence of probability distributions in an appropriately restricted class. Wigner functions have built-in limitations on how sharply focussed they can be on small regions of phase space.

Might it be the case that quantum uncertainty is all that is needed for classical statistical mechanics? Perhaps. This has, indeed, been suggested by David Wallace (2016a), who regards the status of the standard probabilistic posits in classical statistical mechanics as otherwise a bit mysterious. A case can be made that, in a variety of cases (including, crucially, the paradigm case of a dilute gas), the limitations on phase-space density functions imposed by quantum mechanics suffice to guarantee equilibration.

Quantum mechanics can lend probabilities to classical statistical mechanics. But *need* it do so? Is classical statistical mechanics, regarded as a discipline in its own right, intrinsically conceptually incoherent?

Here's another way to ask the question. Classical statistical mechanics, as developed in the latter half of the nineteenth century, achieved some remarkable successes, which made the domains in which it failed so striking. The domains in which classical statistical mechanics failed included the problem of specific heats (identified by Kelvin at the turn of the century

as one of two nineteenth-century clouds over the dynamical theory of heat and light), and the problem of thermal equilibrium of matter with radiation (which, of course, led Planck to the quantum postulate). The areas in which the classical theory broke down were clues that classical physics wasn't quite right. But should nineteenth-century physicists have taken the *successes* of classical statistical mechanics as indications that classical physics is not quite right?

I have argued that the answer is no. All that is needed for classical equilibration results to get off the ground is some uncertainty, expressed as a limitation on the range of density functions over initial conditions that could represent the credences of a reasonable agent. From the standpoint of statistical physics, the grounds for this uncertainty are of no consequence. The case of molecules in a gas interacting with each other, yielding density functions that evolve as if they were probability density functions over the states of a collection of hard spheres bouncing off each other (though in reality the gas is nothing of the sort) goes over smoothly to cases of roulette wheels spinning or billiard balls being shaken, where our uncertainty may be much greater than the absolute minimum required by quantum mechanics. Though the world is, indeed, quantum, there remains a place for considerations of epistemic limitations in statistical physics.

9.8 Conclusion

Much of the philosophical literature on statistical mechanics treats of classical statistical mechanics. This is in sharp contrast to the contemporary *scientific* literature on statistical mechanics. Does it make a difference?

When thinking about the foundations of statistical mechanics, we should always bear in mind that classical physics is not a fundamental theory (and, indeed, that we are not, and never have been, in possession of any theory that is a serious candidate for a complete and fundamental theory). Insofar as our interest in classical statistical mechanics is not merely historical, we should make sure that any conclusions we draw from consideration of classical statistical mechanics carry over to the quantum domain.

Recall that the probabilistic turn in statistical mechanics came about as a result of the reversibility argument. Insofar as quantum statistical mechanics deals with unitary evolution, it is subject to reversibility considerations. The founders of statistical mechanics concluded that the lesson of the reversibility argument is that the sorts of physical states that resist equilibration are

ones that cannot be produced with equal facility (as Bernoulli would put it) as those that equilibrate. Much of this book has been devoted to making sense of that. Our proposal has been to consider the effect of dynamical evolution on the sorts of credences an agent like us could have about the physical states produced by the processes we study. In this chapter, we have seen that considerations of this sort are needed in the quantum realm, as well.

9.9 Appendix: a brief introduction to the basics of quantum theory

9.9.1 Quantum states and classical states

In classical mechanics, a maximally specific state-description picks out a point in the system's phase space, which, in turn, yields a definite value for all dynamical variables of the system, which are represented by functions on phase space. A probability distribution that assigns any probability other than one or zero to some physical proposition about the value of a dynamical variable is an incomplete specification of the state of the system. In quantum mechanics, things are different. There are no quantum states that assign definite values to all physical quantities, and probabilities are built into the standard formulation of the theory.

Construction of a quantum theory of some physical system proceeds by first associating the dynamical degrees of freedom with *operators*. These form an algebra; that is, there are well-defined notions of addition and multiplication. Multiplication of these operators is unlike multiplication of numbers, in that order of multiplication can make a difference. That is, it will not always be the case that $\hat{A}\hat{B} = \hat{B}\hat{A}$. When the result of multiplying two operators doesn't depend on their order, the operators are said to *commute*. Under standard assumptions about the algebra of operators, we will be able to represent them as operators on an appropriately constructed Hilbert space (see Ismael 2015 for definition of this, if you're not familiar).

A state can be characterized by an assignment of expectation values to physical quantities ("observables"). These are required to be *linear*: that is, if ρ is the function assigning expectation values to observables, for any observables A, B, and any real numbers α, β,

$$\rho(\alpha A + \beta B) = \alpha\,\rho(A) + \beta\,\rho(B). \tag{9.14}$$

A complete set of such expectation values is equivalent to a specification of probabilities for outcomes of all experiments that could be performed on the system.

Two physical quantities are said to be *compatible* if there is a single experiment that yields values for them both; these are associated with operators that commute.

A *pure* state, that is, a maximally specific assignment of expectation values, may be represented in a number of physically equivalent ways, for instance by a family of parallel vectors in the Hilbert space (vectors that are non-zero multiples of each other represent the same state), or a projection operator onto a one-dimensional subspace. In addition to pure states, one can also consider non-pure states, called *mixed*. For example, an experimenter might flip a coin, and subject the system to one or the other of a pair of state-preparations, depending on the outcome of the toss. This procedure yields well-defined probabilities for the outcome of any experiment that can be performed on the systems subjected to this procedure, and so counts as a state-preparation in its own right. One way of representing a state is via a *density operator*. These include both pure states (represented by projection operators) and non-pure states.

For physical quantities that can take on a continuous range of values, such as the position or momentum of a particle, we can represent the state via a function on the space of possible values of the quantity. Thus, for a system consisting of n spinless particles, its state can be represented as a function on its $3n$-dimensional configuration space, or, equivalently, as a function on its $3n$-dimensional momentum space. These are often called "wave functions." For a system of particles with spin, we have to include also a specification of spin states. The spin-state of a finite number of particles that are not all spinless can be represented via a vector in a finite-dimensional Hilbert space. The wave function for such a system will assign a vector in that space to each point in the configuration-space or momentum-space of the system.

If a pure state assigns a definite value to a physical quantity A, a vector that represents the state will be an eigenvector of the operator $\hat{A}$. This gives rise to what has been called the *eigenstate-eigenvalue link*, that is, the interpretative principle that, if a system is assigned a state vector that is an eigenvector of some operator representing a physical quantity, then the corresponding dynamical quantity has the corresponding value, and this can be regarded as a property of the physical system.

The noncontroversial core of quantum theory consists of rules for identifying, for any given system, appropriate operators to represent its dynamical

quantities, and an appropriate Hilbert space for these operators to act on. In addition, there are prescriptions for associating quantum states with specified preparation procedures, and for evolving the state of system when it is acted upon by specified external fields or subjected to various manipulations. From the quantum state one can calculate probabilities of outcomes of experiments.

9.9.2 Quantum state evolution

9.9.2.1 The Schrödinger equation

The equation of motion obeyed by a quantum state vector is the Schrödinger equation. It is constructed by first forming the operator $\hat{H}$ corresponding to the total Hamiltonian of the system, which represents the total energy of the system. The rate of change of a state vector is proportional to the result of operating on the vector with the Hamiltonian operator $\hat{H}$.

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H}|\psi(t)\rangle. \tag{9.15}$$

There is an operator that takes a state at time t_0 into a state at time t; it is given by

$$\hat{U}(t; t_0) = exp(-i\hat{H}(t - t_0)/\hbar) \tag{9.16}$$

This operator is a linear operator that implements a one-to-one mapping of the Hilbert space to itself that preserves the inner product of any two vectors. Operators with these properties are called *unitary operators*, and, for this reason, evolution according to the Schrödinger equation is called unitary evolution.

For our purposes, the most important features of this equation is that it is deterministic and linear. The state vector at any time, together with the equation, uniquely determine the state vector at any other time. Linearity means that, if two vectors $|\psi_1(t_0)\rangle$ and $|\psi_2(t_0)\rangle$ evolve into vectors $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$, respectively, then, if the state at time t_0 is a linear combination of these two, the state at any time t will be the corresponding linear combination of $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$.

$$a|\psi_1(t_0)\rangle + b|\psi_2(t_0)\rangle \;\Rightarrow\; a|\psi_1(t)\rangle + b|\psi_2(t)\rangle. \tag{9.17}$$

9.9.2.2 Time reversal in quantum mechanics

In classical mechanics, the class of dynamically possible motions is invariant under time reversal provided that the Hamiltonian is invariant under an operation that leaves position coordinates unchanged and reverses the sign of their conjugate momenta.

In the quantum realm, there is a result known as *Wigner's theorem*, to the effect that any symmetry operation in quantum mechanics must be implementable by an operator that is either unitary or anti-unitary. A unitary operator is one that preserves the inner product of any two vectors in Hilbert space, and is linear. An anti-unitary operator $\hat{V}$ is a mapping of the Hilbert space that maps the inner product of any two vectors to its complex conjugate, and is antilinear:

$$\hat{V}(\alpha|\psi\rangle + \beta|\phi\rangle) = \alpha^* \hat{V}|\psi\rangle + \beta^* \hat{V}|\phi\rangle. \tag{9.18}$$

See Weinberg (1995: ch. 2, appendix A) for an exposition of Wigner's theorem.

If we have a quantum version of a system whose state is characterized by spatial coordinates $\{q_i\}$ and their conjugate momenta $\{p_i\}$, we demand that the time-reversal operation $|\psi\rangle \Rightarrow |\psi\rangle^T$ be such that the expectation values of the operators $\{\hat{q}_i\}$ corresponding to the spatial coordinates be unchanged, and that the expectation values of the operators $\{\hat{p}_i\}$ change sign. If spin is involved, we demand that it change sign, too. It can be shown that these conditions are implemented by an anti-unitary operator that has the effect of complex-conjugating wave functions.

$$\psi^T(q) = \psi^*(q). \tag{9.19}$$

Under this operation, if the Hamiltonian operator is invariant under temporal inversion, then the dynamical laws are also invariant under that operation.

9.9.2.3 The Quantum recurrence theorem

For a quantum system with a discrete set of energy levels, there is a recurrence theorem, just as there is in classical mechanics (Bocchieri and Loinger 1957). But the quantum version is simpler. For one thing, it applies to *all* initial states, whereas the classical version applies to all but a set of measure zero. Second, though, in the classical case, the recurrence time can vary wildly depending on the initial point (so that, though you know that the

system it will return to any neighbourhood of its initial state, you don't know exactly when), the quantum version has a uniform recurrence time.

> If a system's Hamiltonian has a discrete spectrum, then, for any $\varepsilon > 0$ and any τ, there exists a time $T > \tau$ such that, for any initial state $|\psi\rangle$,
>
> $$\| |\psi(t_0 + T)\rangle - |\psi(t_0)\rangle \| < \varepsilon.$$

9.9.2.4 The collapse postulate

Textbook formulations of quantum mechanics usually include an additional postulate about how to assign a state vector after an experiment. In his influential formulation of quantum theory, von Neumann distinguished between two types of processes: Process 1, which occurs upon performance of an experiment, and Process 2, the unitary evolution that takes place as long as no experiment is made (see von Neumann 1932; 1955: §V.1). In Dirac's formulation, the postulate is

> When we measure a real dynamical variable ξ, the disturbance involved in the act of measurement causes a jump in the state of the dynamical system. From physical continuity, if we make a second measurement of the same dynamical variable ξ immediately after the first, the result of the second measurement must be the same as that of the first. Thus after the first measurement has been made, there is no indeterminacy in the result of the second. Hence, after the first measurement has been made, the system is in an eigenstate of the dynamical variable ξ the eigenvalue it belongs to being equal to the result of the first measurement. This conclusion must still hold if the second measurement is not actually made. In this way we see that a measurement always causes the system to jump into an eigenstate of the dynamical variable that is being measured, the eigenvalue this eigenstate belongs to being equal to the result of the measurement. (Dirac 1935: 36)

Dirac's "jump" has come to be known as *state vector collapse* or *wave-function collapse*, and the postulation of a jump of this sort is called the *collapse postulate*, or *projection postulate*.

If the quantum state vector is thought to represent only a state of belief or knowledge about a physical system, and not the physical state of the system, then one could regard an abrupt shift in the state vector upon measurement as a shift corresponding to incorporating the result of the measurement into one's belief state. Neither von Neumann nor Dirac, however, seem to have

thought of it this way. Note that neither expresses the postulate in terms of "observation"; they speak, instead, of "measurement," treated as a physical process, and there is no suggestion that a conscious observer must become aware of the result of the measurement in order for the collapse postulate to apply. A formulation of a version of the collapse postulate according to which a measurement is not completed until the result is observed is found in London and Bauer (1939). They deny, however, that it represents a mysterious kind of interaction between the observer and the quantum system; for them, the replacement of the pre-observation state vector with a new one is a matter of the observer acquiring new information. These two interpretations of the collapse postulate, as either a real change of the physical state of the system, or as a mere updating of information on the part of an observer, have persisted in the literature.

If state vector collapse is to be regarded as a physical process, this raises the question of what physically distinguishes interventions that are to count as "measurements," capable of inducing an abrupt jump in the state of the system, from other interventions, which induce only continuous, unitary evolution. As John S. Bell (1990) has argued, "measurement" is not an appropriate concept to appear in the formulation of any physical theory that might be taken to be fundamental. If, however, one dispenses with the postulate, this gives rise to the so-called "measurement problem."

9.9.3 Quantum statistical mechanics

Given the structural difference between classical and quantum mechanics, it is remarkable how much of the formalism of classical statistical mechanics goes over into quantum statistical mechanics. Much of the transition consists simply of the replacement of a density function ρ representing a probability distribution on a classical phase space with a density operator $\hat{\rho}$. It might, in fact, be possible to write a statistical mechanical textbook in which almost every formula has a dual meaning, and can be read as classical or quantum, and come out correct on either reading. We have already seen instances of formulas of this sort, in (7.18) and (7.19).

The evolution of a quantum density operator for a system with Hamiltonian operator $\hat{H}$, is given by

$$\frac{d}{dt}\hat{\rho} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}], \tag{9.20}$$

where $[\hat{H}, \hat{\rho}]$ is the commutator of $\hat{H}$ and $\hat{\rho}$,

$$[\hat{H}, \hat{\rho}] = \hat{H}\hat{\rho} - \hat{\rho}\hat{H}. \tag{9.21}$$

As in classical physics, a density operator that is a function of constants of the motion will be stationary in time. In particular, for a time-independent Hamiltonian, any density operator that is a function only of $\hat{H}$ will be stationary.

Comparison of (9.20) with the Liouville equation is facilitated by writing the latter in terms of the Poisson bracket. The Poisson bracket of any two phase-space functions f, g is defined as

$$\{f, g\} = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right). \tag{9.22}$$

With this in hand, Liouville's equation (7.3) can be written as

$$\frac{\partial \rho}{\partial t} = -\{\rho, H\}. \tag{9.23}$$

There are quantum versions of the microcanonical and canonical distributions. For the microcanonical distribution, consider a narrow band of energy $[E, E + \delta]$, and consider the set of energy eigenstates with energy in that band. If the system has a discrete set of energy levels, and each energy level has at most a finite degree of degeneracy, this will be a finite set of energy eigenstates, and one can define a mixed state that is an equally weighted mixture of all of these.[3]

The quantum version of the canonical distribution is

$$\hat{\rho} = Z^{-1} e^{-\beta \hat{H}}, \tag{9.24}$$

[3] The fact that we are here dealing with a finite set of eigenstates, each of which gets an equal weighting, has encouraged the idea that invocation of a quantum-mechanical microcanonical distribution is a mere application of a Principle of Indifference. The fact that we are dealing with a finite set means that complications that arise in attempting to apply the Principle to infinite sets are bypassed. But the fundamental issue remains: without a choice of a set of states to regard as equiprobable, the Principle offers us no advice. In this case, we must ask: Why energy eigenstates, rather than some other set of states? The fact that these are stationary states makes this a candidate for an equilibrium measure, but, if we invoke considerations of this sort, that is, dynamical considerations, we have left the realm of a pure application of Indifference, and are well on the way to asking the question that is relevant: Is this the sort of state that a process of equilibration will lead to?

where, in order that $\hat{\rho}$ have trace one, we set

$$Z = \mathrm{Tr}[e^{-\beta \hat{H}}]. \tag{9.25}$$

The analog of the Gibbs entropy is the von Neumann entropy,

$$S_N[\hat{\rho}] = -k\mathrm{Tr}[\hat{\rho} \log \hat{\rho}]. \tag{9.26}$$

For an isolated system, it is a constant of the motion.

As in the classical case, we can define a coarse-grained density operator $\hat{\bar{\rho}}$ that is a function only of some set of macrovariables of interest, and a corresponding coarse-grained entropy

$$\overline{S}_N[\hat{\rho}] = S_N[\hat{\bar{\rho}}]. \tag{9.27}$$

9.9.4 From quantum to classical

Classical and quantum statistical mechanics are usually dealt with as separate subjects, and we may well ask what the relation is between the two.

There is, I'm afraid, a picture of the relation that seems to be in the back of may people's minds, when thinking about a gas, treated classically. It is often, I think, implicitly assumed that, even if we take quantum indeterminacy into account, we can think of the molecules as effectively classical, as they are sufficiently massive that they can be treated as if they were more-or-less well-localized objects. Quantum mechanics may be needed to account for their internal structure, but the familiar picture of billiard balls bouncing off of each other will not be too far off.

It turns out that this is completely wrong.[4]

Consider a quantum system (which may have some internal structure, in which case we're talking about the centre of mass wavepacket), of total mass m, which is moving freely, i.e. subject to no external influences. Let $\Delta x(t)$ be the wave-packet spread in x (or any one of the three spatial dimensions) at time t, and let $\Delta p_x(t)$ be the spread in momentum. Take t_0 to be the time at which $\Delta x(t)$ has its smallest value. Then the standard quantum mechanical evolution of the wave packet gives us

[4] It was David Wallace who first drew my attention to this. See Wallace (2016a: §8) for a similar calculation.

$$\Delta x(t)^2 = \Delta x(t_0)^2 + \left(\frac{\Delta p_x}{m}\right)^2 (t - t_0)^2. \tag{9.28}$$

By the uncertainty relation for position and momentum, Δp_x (which, since we're dealing with force-free evolution, is a constant of the motion) must satisfy,

$$\Delta p_x \geq \frac{\hbar}{2\,\Delta x(t_0)}, \tag{9.29}$$

and so we have

$$\Delta x(t)^2 \geq \Delta x(t_0)^2 + \left(\frac{\hbar}{2m\Delta x(t_0)}\right)^2 (t - t_0)^2. \tag{9.30}$$

It follows from this that,[5] no matter what the minimum spread $\Delta x(t_0)$ is, for all t,

$$\Delta x(t)^2 \;\geq\; \left(\frac{\hbar}{m}\right) |t - t_0|\,. \tag{9.31}$$

This means that the longest period of time during which a freely moving wavepacket can have a spread less than a certain amount d is

$$\Delta t = \frac{2m\,d^2}{\hbar}. \tag{9.32}$$

Now let's put some numbers in. Most of our atmosphere consists of nitrogen molecules, N_2, which have mass of about 28 atomic mass units, or about 5×10^{-26} kilograms. The mean time between collisions, τ, for nitrogen molecules, at room temperature and atmospheric pressure, is about 6×10^{-10} seconds (see Reif 1968, §12.2 for the calculation). We plug that into equation (9.31), using a value of about 7×10^{-34} m^{2}kg/s for $\hbar$. This gives us the result that, after a time equal to the mean time between collisions, the wavepacket has to spread to at least the size

$$\sqrt{\frac{\hbar\,\tau}{2m}} \approx 2 \times 10^{-9} \text{ metres}. \tag{9.33}$$

[5] Hint:

$$\left(\Delta x(t_0) - \frac{\hbar}{2m\Delta x(t_0)}\,|t - t_0|\right)^2 \geq 0.$$

Consider the fact that molecular dimensions are on the order of a few multiples of the Bohr radius, which is about 5×10^{-11} metres. The wave-spreading that a molecule undergoes in the intervals between collisions is two orders of magnitude larger than the dimensions of the molecules. In the time required for several collisions, we have nothing at all like well-localized wavepackets.

What gives, then? If gases are actually nothing at all like a collection of classical billiard balls colliding, what's the point of all the calculations made in the framework of the kinetic theory of gases, and why do they yield anything at all like accurate results for the macroscopic properties of gases?

The explanation has to do with the nature of the limiting relation between quantum and classical mechanics. Textbooks often give the impression that, for large, massive systems, quantum mechanics yields an approximation to the classical picture of the world. In fact, it yields nothing of the sort, a fact that is vividly illustrated by Schrödinger cat-style experiments. For any experiment like that, a bare quantum-mechanical treatment (i.e., without collapse, and without extra structure of the sort invoked by hidden-variables theories) yields a state that involves superpositions of macroscopically distinct states of affairs.

What we get, from quantum mechanics, in an appropriate limiting regime, is *probability distributions* on classical phase space that evolve approximately in the same way that classical probability distributions do. That is, the classical limit of a quantum state is a classical probability distribution. The pioneers of statistical mechanics who found themselves studying the behaviour of probability distributions on phase space were, unknowingly, probing a deeper level of physical reality, and constructing a more physically realistic picture, than they would have had they confined themselves to state-descriptions in terms of classical microstates.

And this means that all that effort expended in the name of the kinetic theory of gases, to the extent that it attempted to track, not the evolution of the microstate of the gas, but rather, the evolution of a probability distribution over microstates, was not wasted, and could, indeed, yield results that approximate the quantum results.

One tool that is useful in connection with this is the *Wigner function*, or *Wigner quasi-distribution*.[6] Quantum states, notoriously, do not yield a joint

[6] See Case (2008) for an accessible introduction to the Wigner function, and Hillery et al. (1984) and Lee (1995) for more in-depth overviews of the Wigner function and other phase-space quasi-distributions.

probability distribution over all observables; what they give us instead is, for each set of mutually compatible observables (represented by commuting operators), a probability distribution over those observables. In light of that, it is perhaps surprising that, for any quantum state of an n-body system, it is possible to define a real function $W(\mathbf{x}_1, \ldots, \mathbf{x}_n; \mathbf{p}_1, \ldots, \mathbf{p}_n)$ over classical phase space that returns a density function for that state's probability distribution over position, when integrated over all momenta, and a density function for the state's probability distribution over momenta, when integrated over all positions (the Wigner function is just one way to do it, but it's the best-known).

Given a quantum state, which can be represented by a position-space wave-function $\psi(\mathbf{x})$, or by its Fourier transform, the corresponding momentum-space wave-function $\phi(\mathbf{p})$, the Wigner function is defined by,[7]

$$W(\mathbf{x}, \mathbf{p}, t) = \frac{1}{h^3} \int e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\, \psi(\mathbf{x} + \mathbf{y}/2, t)\, \psi^*(\mathbf{x} - \mathbf{y}/2, t)\, d^3\mathbf{y} \qquad (9.34)$$

$$= \frac{1}{h^3} \int e^{i\mathbf{x}\cdot\mathbf{u}/\hbar}\, \phi(\mathbf{p} + \mathbf{u}/2, t)\, \phi^*(\mathbf{p} - \mathbf{u}/2, t)\, d^3\mathbf{u}. \qquad (9.35)$$

It is easy to verify that $W(\mathbf{x}, \mathbf{p}, t)$ is real-valued, and that, as advertised,

$$\int W(\mathbf{x}, \mathbf{p}, t)\, d^3\mathbf{p} = |\psi(\mathbf{x}, t|^2\,; \qquad (9.36)$$

$$\int W(\mathbf{x}, \mathbf{p}, t)\, d^3\mathbf{x} = |\phi(\mathbf{p}, t|^2\,. \qquad (9.37)$$

It does not, however, serve, in general, as a density function for a probability distribution on phase space because it can take on negative values. In special cases—such as a Gaussian wave-function, for example—it *is* positive everywhere.

We can also define a Wigner function for a mixed state, represented by a density operator $\hat{\rho}$.

These functions were introduced into the literature by E. P. Wigner (1932), in an article appropriately entitled "On the Quantum Correction to Thermodynamic Equilibrium," accompanied by a footnote that reads, "This expression was found by L. Szilard and the present author some years ago for another purpose." We should probably call it the *Wigner–Szilard function*, but, in conformity with prevailing usage and with Stigler's Law of Eponymy, we will stick with "Wigner function."

[7] I'm writing this down for a single particle; the extension to the $6n$-dimensional phase space of an n-particle system is straightforward.

$$W(\mathbf{x}, \mathbf{p}, t) = \frac{1}{h^3} \int e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar} \langle \mathbf{x} + \mathbf{y}/2, t| \hat{\rho} |\mathbf{x} - \mathbf{y}/2, t\rangle \, d^3\mathbf{y} \tag{9.38}$$

$$= \frac{1}{h^3} \int e^{i\mathbf{x}\cdot\mathbf{u}/\hbar} \langle \mathbf{p} + \mathbf{u}/2, t| \hat{\rho} |\mathbf{p} - \mathbf{u}/2, t\rangle \, d^3\mathbf{u}. \tag{9.39}$$

As the quantum state evolves in time, so will the corresponding Wigner distribution. It is interesting to compare its evolution to the evolution of a classical density function. Recall that this satisfies the *Liouville equation*, which was eq. (7.3).

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{N} \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0. \tag{9.40}$$

For a single particle subject to a potential $U(\mathbf{x})$, the Hamiltonian is

$$H(\mathbf{x}, \mathbf{p}) = \frac{p^2}{2m} + U(\mathbf{x}), \tag{9.41}$$

and so a classical density function $\rho(\mathbf{x}, \mathbf{p}, t)$ on its phase evolves according to

$$\frac{\partial \rho}{\partial t} = -\sum_{i=1}^{3} \left(\frac{p_i}{m} \frac{\partial \rho}{\partial x_i} - \frac{\partial U}{\partial x_i} \frac{\partial \rho}{\partial p_i} \right). \tag{9.42}$$

Now consider a quantum system, with a Hamiltonian operator $\hat{H}$ having the same form as (9.41). The evolution of the Wigner function W for such a system satisfies

$$\frac{\partial W}{\partial t} = -\sum_{i=1}^{3} \left(\frac{p_i}{m} \frac{\partial W}{\partial x_i} - \frac{\partial U}{\partial x_i} \frac{\partial W}{\partial p_i} \right.$$
$$\left. - \sum_{n=1}^{\infty} \left(-\frac{\hbar}{2} \right)^{2n} \frac{1}{(2n+1)!} \left(\frac{\partial^{2n+1} U}{\partial x_i^{2n+1}} \right) \left(\frac{\partial^{2n+1} W}{\partial p_i^{2n+1}} \right) \right). \tag{9.43}$$

Notice that the first two terms of the right-hand-side of (9.43) are the same as in the Liouville equation. The remaining terms involve increasing powers of $\hbar$ and increasing derivatives of U and W—odd powers only; the even derivatives play no role. In the special case in which third-order and higher derivatives of $U(\mathbf{x})$ are all zero (which will be true for a free particle or a particle in a harmonic oscillator potential), then these terms all vanish

and the evolution equation for the Wigner function is just the Liouville equation.

If the remaining terms of (9.43) are negligible compared to the first two, then the evolution of the Wigner function will approximately satisfy the Liouville equation. This will be the case when the Wigner function and/or the potential is relatively smooth, where "relatively smooth" is to be cashed out as meaning that the higher derivatives are such that all but the classical terms of (9.43) are of negligible size. If, in addition the Wigner function is positive, then we will have recovered from the quantum-theory a classical-like object—a function that behaves like a classical phase-space density function.

Thus, in some special cases, a Wigner function acts like a classical phase-space density function. But not every classical phase-space density function can be obtained this way. For one thing, since the Wigner function has to yield quantum probabilities for position and momentum as marginals, no Wigner function can violate the Heisenberg uncertainty relation,

$$V(X)\,V(P) \;\geq\; \frac{\hbar^2}{4}. \tag{9.44}$$

It can also be shown that there is a bound on the value of Wigner functions. For a Wigner function on the $6n$-dimensional space of n particles,

$$|W(q,p)| \leq \left(\frac{2}{h}\right)^{3n}. \tag{9.45}$$

This is not a restriction that arbitrary density functions are obliged to respect.

Studies of the classical, or quasi-classical, limit of quantum mechanics often emphasize the role of environmentally induced decoherence. Though the quantum state of an isolated system evolves unitarily, and hence a pure state remains pure, a system that interacts with the outside world becomes entangled with it, and its reduced state—that is, the restriction to the system of the global state of the system plus its environment—will become a mixed state. For the right sort of interactions with a large, complex environment, typical initial states lead to states that satisfy some criterion of classicality. One such criterion is positivity of the Wigner function.

Beyond Chance and Credence

A Theory of Hybrid Probabilities

WAYNE C. MYRVOLD

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2021

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2020950770

ISBN 978–0–19–886509–4

DOI: 10.1093/oso/9780198865094.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY