Part one

Quantum chance

1

Quantum probability and the problem of interpretation

1.1 Quantum probability and quantum mechanics

1.1.1 The formalism of quantum probability theory

Discussions of quantum mechanics are often confused by a lack of clarity about what exactly constitutes 'quantum mechanics'. It is therefore useful to try at the start to isolate a consistent mathematical core of quantum mechanics, and consider anything that goes beyond this core to be 'interpretation'. For us, this core is quantum probability theory.

Quantum probability is a generalization of classical probability, and therefore I begin with a brief review of the latter. I assume that the reader has some familiarity with the ideas of probability theory. What follows is just to provide a quick review, and to establish some notation and terminology.[1]

In modern classical probability theory, probabilities are defined over algebras of events. The motivation is straightforward: we begin with a set of 'primitive', or 'simple', events (the 'sample space'), and form an algebra of events by taking all logical combinations of the simple events. For example, let us take the simple events to be the possible results of rolling a six-sided die one time, so that the sample space is the set $\{1, 2, 3, 4, 5, 6\}$. We then form an algebra of events from the sample space by taking all possible logical combinations of the simple events. Logical combinations include, for example, 'either 3 or 5' and 'not 3 and not 2'.

In classical probability theory, we represent logical combinations with the set-theoretic operations of intersection (which represents 'and'), union (which represents 'or'), and complement (which represents 'not'). Events are therefore given by sets whose elements are taken from the sample space. The event 'either 3 or 5' is represented by $\{3\} \cup \{5\}$, which is $\{3, 5\}$. The event 'not 3 and not 2' is represented by $\neg\{3\} \cap \neg\{2\}$, which is $\{1, 2, 4, 5, 6\} \cap \{1, 3, 4, 5, 6\}$, which is $\{1, 4, 5, 6\}$. (More precisely, we form an algebra of events from a

sample space by closing the sample space under complement and countable union. Doing so guarantees closure under countable intersection.)

We introduce probabilities to the picture with a probability measure, p, over the simple events. This measure is extended to the entire algebra of events by Kolmogorov's axioms:

Kolmogorov's axioms:

(1) $p(\emptyset) = 0$,
(2) $p(\neg F) = 1 - p(F)$,
(3) $p(F \cup F') = p(F) + p(F') - p(F \cap F')$

where F and F' are events in the algebra and $\emptyset$ is the empty set (the set with no elements). In our example, the probability measure over the simple events is given by $p(\{i\}) = 1/6$ for all i. Hence, for example, $p(\{3\} \cup \{5\}) = 1/3$ (by axioms (3) and (1)), and so on.

Finally, we define a conditional probability measure, the probability of some event *given* the occurrence of some other event. For example, we may want to know the probability that the die shows 6 given that it shows either 2 or 6. Conditional probabilities are defined by

$$p(F|F') = \frac{p(F \cap F')}{p(F')}. \tag{1.1}$$

For example, $p(\{6\}|\{2,6\}) = 1/2$.

To summarize, we may identify a classical probability theory with an ordered triple, $\langle \Omega, \mathscr{F}, p \rangle$, where Ω is the sample space, $\mathscr{F}$ is the algebra of events generated by Ω, and p is a Kolmogorovian probability measure.

Quantum probability theory also begins with an ordered triple, $\langle \mathscr{H}, \mathsf{L}_{\mathscr{H}}, \psi \rangle$. Here $\mathscr{H}$ is a Hilbert space, which is a (complete, complex) vector space with an inner product defined on it. (We also require that it have a countable basis.) Every (normalized) vector — or equivalently, every ray — in $\mathscr{H}$ corresponds to a simple event, so that $\mathscr{H}$ may be considered the sample space. We generate an algebra of events, $\mathsf{L}_{\mathscr{H}}$, from $\mathscr{H}$ as follows. Beginning with the rays, i.e., the one-dimensional subspaces of $\mathscr{H}$, close under the operations of span, intersection, and orthogonal complement. (The span of two subspaces, P and P', is the set of all vectors that can be written as a weighted sum of vectors from P and P'. For example, the span of two one-dimensional subspaces (rays) is the plane containing both of them. The intersection of two subspaces is the largest set of vectors contained in both of them. The orthogonal complement, or orthocomplement, of a subspace is the largest subspace entirely orthogonal (perpendicular) to it.) These operations correspond to the lattice-theoretic operations of join (denoted '$\vee$'), meet

(denoted '$\wedge$'), and orthocomplement (denoted '$^{\perp}$'), respectively, which leads some to interpret them as the quantum-mechanical representation of the logical operations of or, and, and not. (I will discuss this idea further in section 4.1.)

The algebra of quantum-mechanical events is denoted by $\mathsf{L}_{\mathscr{H}}$ because it forms a lattice, a partially ordered set for which the operations meet, join, and orthocomplement are defined between each pair of elements. The partial ordering is given by subspace inclusion. Alternatively, the algebra of quantum-mechanical events can be considered to be a partial Boolean algebra. I will discuss this alternative in chapter 4.

Finally, ψ is a vector in $\mathscr{H}$ with norm 1 (i.e., the inner product of ψ with itself is 1). It generates a probability measure, p^{ψ}, over the sample space $\mathscr{H}$ through the familiar rule:

$$p^{\psi}(\varphi) = |(\psi, \varphi)|^2, \tag{1.2}$$

where $(\cdot, \cdot)$ is the inner product. Or, using Dirac notation (which I will use from now on),

$$p^{\psi}(\varphi) = |\langle\psi|\varphi\rangle|^2.$$

Often I will speak of the elements of the sample space not as vectors, but as projections, or subspaces. Every vector can be represented, for present purposes, as the (one-dimensional) subspace that it spans. Also, I will often use the terms 'projection operator' and 'subspace' interchangeably (for there is a one-to-one correspondence between them), and I will use the same notation for both. I will even, at times, say things like 'the projection P is contained in the projection P'', meaning that the subspace onto which P projects is a subspace of the subspace onto which P' projects. None of this loose talk should cause confusion.

Now the story gets a bit more complex. It would be nice if Kolmogorov's axioms held in quantum probability (substituting the lattice-theoretic operations for the set-theoretic ones, of course). Axioms 1 and 2 do hold, but axiom 3 fails in general, though it holds when the events are orthogonal (more precisely, when the subspaces representing the events are orthogonal). That is, we have:

(1) $p(\mathbf{0}) = 0$,
(2) $p(P^{\perp}) = 1 - p(P)$,
(3) $p(P \vee P') = p(P) + p(P')$, when $P \perp P'$,

where $\mathbf{0}$ is the zero subspace (the zero element of the lattice).[2]

These axioms are somewhat unsatisfactory, because they do not, by themselves, tell us how to calculate the probability of events $P \vee P'$ for arbitrary P and P'. Some have argued that we should take this limitation to be a *lesson*: the probability of such events is just undefined. However, the point of my starting with quantum probability is to provide as neutral a basis for interpretation as possible. Hence we may just as well allow into the theory probabilities for such combinations of events, and then later, if we like, remove them.

Moreover, quantum probability as introduced thus far lacks generality in another sense. There exist probability measures on Hilbert spaces that are not representable by a vector through equation (1.2). To capture all of the probability measures over a Hilbert space, we need to represent them not by vectors, but by the so-called density operators,[3] which are (bounded, positive) operators on the Hilbert whose trace is 1. The trace of an operator, W, is given by

$$\mathrm{Tr}[W] = \sum_i \langle \varphi_i | W | \varphi_i \rangle,$$

where $\{|\varphi_i\rangle\}$ is any orthonormal basis for $\mathscr{H}$ — the value of $\mathrm{Tr}[W]$ is independent of the choice of the basis, $\{|\varphi_i\rangle\}$. Hence we alter the definition of a quantum probability theory, so that it consists of an ordered triple, $\langle \mathscr{H}, \mathsf{L}_{\mathscr{H}}, W \rangle$, where W is a density operator, and it generates a probability measure over all of $\mathsf{L}_{\mathscr{H}}$ by

$$p^W(P) = \mathrm{Tr}[WP]. \tag{1.3}$$

Notice that we have simply bypassed the method of classical probability theory: rather than extending a measure over the sample space to a measure over all events through axioms such as Kolmogorov's, we define the measure over $\mathsf{L}_{\mathscr{H}}$ directly, through (1.3). It becomes a matter for investigation what the properties of the measure generated by W are. As it happens, axioms (1)–(3) as stated for our original version of quantum probability theory hold here as well. (The differences are that now: (a) we have a theory that includes all probability measures over the sample space, and (b) we have a theory that tells how to calculate the probability of every event in $\mathsf{L}_{\mathscr{H}}$ directly. We could have gotten (b) without moving to the formalism of density operators, however.)

A *random variable* on a classical probability space is a map from simple events to real numbers. A probability measure over the events therefore induces a probability measure over the *range* of the random variables. In quantum probability theory, random variables are represented by self-adjoint

operators on $\mathscr{H}$. Such operators can be conceived as maps from *some* rays in $\mathscr{H}$ to real numbers. Recall that every self-adjoint operator, A, has eigenvectors, $|a_i\rangle$, each of which corresponds to some eigenvalue, a_i. Hence A can be conceived as a map from its eigenvectors to their corresponding eigenvalues.

The set of all eigenvectors of A corresponding to a given eigenvalue, a_i, forms a subspace of $\mathscr{H}$, and may be denoted $P^A_{a_i}$. The set $\{P^A_{a_i}\}$ for all eigenvalues, a_i, of A is a set of mutually orthogonal subspaces spanning $\mathscr{H}$, so that for any density operator, W, $p^W(\bigvee_i P^A_{a_i}) = 1$, and the linearity of the trace functional plus the orthogonality of the $P^A_{a_i}$ further guarantee that the usual sum rule for probabilities holds: $p^W(P^A_{a_i} \vee P^A_{a_j}) = p^W(P^A_{a_i}) + p^W(P^A_{a_j})$ for $i \neq j$. Hence p^W generates a probability measure over the set of all eigenvalues of A.

Finally, the conditional probability of P given P' in quantum probability is

$$p^W(P|P') = \frac{\mathrm{Tr}[P'WP'P]}{\mathrm{Tr}[WP']}. \tag{1.4}$$

This definition of conditional probabilities is sometimes called 'Lüders' rule'.[4] It is the only definition (given certain constraints) that meets the reasonable criterion that whenever P is contained in P', the conditional probability is given by[5]

$$p^W(P'|P) = \frac{p^W(P')}{p^W(P)}. \tag{1.5}$$

Quantum probability theory is a generalization of classical probability theory.[6] Therefore, not everything that is true in classical probability theories will be true of the more general quantum probability theories. We have already seen one example, in the failure of Kolmogorov's third axiom. Another important difference is that while joint probabilities (probabilities for arbitrary sets of events to be jointly occurrent) are always definable in a classical probability theory, in quantum probability it is not possible to define a joint probability measure for arbitrary sets of events (given some plausible assumptions about joint probabilities).[7] To put it differently: if you pick an arbitrary set of events from $\mathsf{L}_{\mathscr{H}}$, you are not guaranteed that there is any probability for this set of events to be jointly occurrent. We may put it yet another way: while joint probability distributions for pairs of random variables always exist in classical probability theory, they need not exist for pairs of operators (more precisely, for the sets of their eigenvalues) in quantum probability theory.

1.1.2 From quantum probability to quantum mechanics

Quantum probability theory is a consistent mathematical theory, but as yet has nothing to do with physics. I have given a few hints about how some relation might be made between quantum probability and physics — in particular, I noted that quantum probability measures can be interpreted as probability measures over all of the eigenvalues of each operator. However, that fact is not enough to generate a physical theory, even after we make the standard identification between operators and physical quantities, i.e., 'observables', so that the eigenvalues of an operator are the possible values of the corresponding observable.

The problem is that it is not at all clear how to get from quantum probability to a consistent and satisfactory physical theory. This problem will occupy us for parts of the next four chapters. To make the problem clear, I begin with a minimal extension of quantum probability, to arrive at the theory that I will call 'quantum mechanics'. Even this minimal extension has its difficulties, as I discuss in the next section. Nonetheless, by making the extension and exposing the problem, we will at least have a handle on the difficulties that face quantum mechanics.

The extension may be given first in the more familiar terms of vectors in Hilbert space. In these terms, the state of a quantum system is represented by a vector, $|\psi\rangle$. A system's state evolves in time according to Schrödinger's equation:

$$\frac{d|\psi(t)\rangle}{dt} = -i\hbar H|\psi(t)\rangle, \tag{1.6}$$

where H is the Hamiltonian operator for a given system. The state of a system at any time generates a probability measure over all possible values of each observable in the way already described.

More generally, the state of a physical system is given by a density operator on Hilbert space. The evolution of a density operator is easily derived from the Schrödinger equation. The result is that the state, $W(t)$, of a system evolves according to a unitary operator, $U(t)$:[8]

$$W(t) = U(t)W(0)U^{-1}(t), \tag{1.7}$$

where $U(t) = e^{-iHt}$. The probability measure over all possible values of each observable is given at each time, t, by $\mathrm{Tr}[W(t)P_a^A]$, where, recall, A represents some observable, and a is some eigenvalue of A. (I shall often not distinguish notationally between observables and operators.)

'Quantum mechanics' therefore makes two claims that go beyond quantum probability. First, it claims that the state of a system is given by a density

operator. Second, it claims that the state evolves according to the unitary operator $U(t) = e^{-iHt}$. These two claims may appear innocent enough, but as we shall see in the next section, they lead to a difficult problem. Solving, or avoiding, this problem is one of the central challenges facing interpreters of quantum mechanics.

1.2 Interpreting quantum mechanics

1.2.1 The 'measurement problem'

The problem that quantum mechanics faces — the 'measurement problem' — is that it sometimes assigns the *wrong* state to some systems.[9] (As we shall see, the name 'measurement problem' is misleading, because it suggests that the problem occurs only when one makes a measurement, whereas the problem is, in fact, generic.)

The problem is best described by way of illustration. Suppose that a quantum system begins in the state $|\alpha_1\rangle$, an eigenvector of A with eigenvalue a_1. Suppose we perform a measurement of A, as follows: the measuring device begins in a ready-to-measure state, $|M_0\rangle$, and, after the measurement, is perfectly correlated with the value of A possessed by the system. We may represent the measurement schematically by (assuming for simplicity that the interaction does not disturb the measured system)

$$\begin{array}{ccc} \text{initial state} & \begin{array}{c}\text{measurement}\\ \text{interaction}\end{array} & \text{final state} \\ |\alpha_1\rangle|M_0\rangle & \longrightarrow & |\alpha_1\rangle|M_1\rangle, \end{array}$$

where $|M_1\rangle$ is the state of the apparatus that indicates a value of a_1. (Juxtaposition of two vectors represents a tensor product. Readers less familiar with the tensor product formalism may read juxtaposition as 'and'.[10] For example, read '$|\alpha_1\rangle|M_0\rangle$' as 'the measured system is in the state $|\alpha_1\rangle$ *and* the apparatus is in the state $|M_0\rangle$'.) Similarly, if the quantum system begins in the state $|\alpha_2\rangle$, then the interaction would be

$$\begin{array}{ccc} \text{initial state} & \begin{array}{c}\text{measurement}\\ \text{interaction}\end{array} & \text{final state} \\ |\alpha_2\rangle|M_0\rangle & \longrightarrow & |\alpha_2\rangle|M_2\rangle, \end{array}$$

but now trouble is close at hand. The evolution of the pair of systems during these measurement-interactions must be described by some unitary operator,

$U(t)$, and $U(t)$ is always linear, which means that for any vectors $|\varphi_1\rangle$ and $|\varphi_2\rangle$,

$$U(t)[c_1|\varphi_1\rangle + c_2|\varphi_2\rangle] = c_1U(t)|\varphi_1\rangle + c_2U(t)|\varphi_2\rangle, \tag{1.8}$$

where c_1 and c_2 are any complex numbers. Applying (1.8) to the measurement-interactions described above, we get that if the quantum system begins in the state $c_1|\alpha_1\rangle + c_2|\alpha_2\rangle$, then the measurement-interaction yields

initial state	measurement interaction	final state
$(c_1\|\alpha_1\rangle + c_2\|\alpha_2\rangle)\|M_0\rangle$	$\longrightarrow$	$c_1\|\alpha_1\rangle\|M_1\rangle + c_2\|\alpha_2\rangle\|M_2\rangle.$

It is not at all clear what to say about the final state in this interaction. What is clear is that when we perform the experiment, we find the apparatus in either the state $|M_1\rangle$ or the state $|M_2\rangle$. Yet, the final state assigned by quantum mechanics is neither of these. Indeed, it is apparently not the sort of state that we ever witness — a 'superposition' of $|\alpha_1\rangle|M_1\rangle$ and $|\alpha_2\rangle|M_2\rangle$. It appears that the standard theory fails: the final state that it assigns to the system (or, the event that is occurrent with probability 1) is one that we never actually see when we perform the experiment. What we see is either $|M_1\rangle$ or $|M_2\rangle$, but quantum mechanics predicts something else entirely.

I have described the 'measurement problem' in the context of a measurement, but the problem is general. It seems likely that the sort of interaction that led quantum mechanics to attribute the 'wrong' state to the measuring apparatus could occur also in situations that we would not call 'measurements'. Indeed, quantum mechanics appears to face the very general problem of not adequately describing the world as we actually see it. The states that it attributes to macroscopic objects are not the states that we observe them to have. Quantum mechanics does a good job of describing the world *behind* the phenomena of our everyday experience, but it appears to fail miserably to describe our everyday experience itself.

1.2.2 Are the quantum probabilities epistemic?

The following line of thought might already have occurred to the reader: Why not suppose that the probabilities that the standard view prescribes are merely epistemic probabilities? That is, why not suppose that when the standard view says that the final state is $c_1|\alpha_1\rangle|M_1\rangle + c_2|\alpha_2\rangle|M_2\rangle$ all it *means* is that one or the other of $|\alpha\rangle|M_1\rangle$ and $|\alpha_2\rangle|M_2\rangle$ is occurrent, with probabilities $|c_1|^2$ and $|c_2|^2$, respectively? (The probabilities $|c_1|^2$ and $|c_2|^2$ are then 'epistemic' because one of the two events is *really* occurrent, but we

do not know which — or better, the theory simply does not tell us which.) If we can maintain this interpretation of the quantum probability measure, then the measurement problem apparently disappears.

Of course, the ignorance interpretation faces the problem of explaining what it means for the event given by (the one-dimensional subspace spanned by) $c_1|\alpha_1\rangle|M_1\rangle + c_2|\alpha_2\rangle|M_2\rangle$ to be occurrent with probability 1, but a little metaphysical creativity might produce such an explanation. The real challenge facing this view is that it is not at all clear that the quantum probabilities can be reasonably interpreted as epistemic probabilities. Indeed, the prevailing orthodoxy among physicists (or at least, what is reputed by philosophers of physics to be the prevailing orthodoxy among physicists) is that quantum probabilities are *not* merely epistemic.

Rarely does one find *arguments* for this orthodox view, but arguments do exist. In this section, I consider one argument against the epistemic interpretation of quantum probabilities, and how that argument might be answered.

Consider a quantum system in the state $|\psi\rangle$. (In the formalism of density-operators, the system is in the (pure) state $|\psi\rangle\langle\psi|$.) Consider two events, given by $|\chi\rangle$ and $|\xi\rangle$, where, for $c_1 \neq d_1$ and $|\varphi\rangle\perp|\psi\rangle$,

$$|\chi\rangle = c_1|\psi\rangle + c_2|\varphi\rangle \quad |c_1| > |c_2|,$$
$$|\xi\rangle = d_1|\psi\rangle + d_2|\varphi\rangle \quad |d_1| > |d_2|.$$

(When I say that an event is 'given by' a vector $|\chi\rangle$, I mean that it is represented by the subspace spanned by $|\chi\rangle$.) We therefore have that $p^\psi(|\chi\rangle\langle\chi|) = |c_1|^2$ and $p^\psi(|\xi\rangle\langle\xi|) = |d_1|^2$. Given these probabilities, if you were to adopt an ignorance interpretation of p^ψ then you should be willing to accept the following bets as fair.

Bet 1: If $|\chi\rangle\langle\chi|$ is the truly occurrent event, then you win $|c_2|^2$ dollars, and otherwise you lose $|c_1|^2$ dollars. (The expected value of this bet is $p^\psi(|\chi\rangle\langle\chi|)|c_2|^2 - [1 - p^\psi(|\chi\rangle\langle\chi|)] \times |c_1|^2 = 0$. Hence it is a fair bet.)

Bet 2: If $|\xi\rangle\langle\xi|$ is the truly occurrent event, then you win $|d_2|^2$ dollars, and otherwise you lose $|d_1|^2$ dollars. (The expected value of this bet is $p^\psi(|\xi\rangle\langle\xi|)|d_2|^2 - [1 - p^\psi(|\xi\rangle\langle\xi|)] \times |d_1|^2 = 0$. Hence it is a fair bet.)

For simplicity, assume that if any *other* event occurs, then no money changes hands. Then bets 1 and 2 together form a so-called 'Dutch Book'. That is, although you are committed to agreeing that they are both fair bets, you are guaranteed to lose money if you take both of them (when either $|\chi\rangle\langle\chi|$ or $|\xi\rangle\langle\xi|$ occurs). If $|\chi\rangle\langle\chi|$ occurs, then you get a total of $|c_2|^2 - |d_1|^2$ dollars. If $|\xi\rangle\langle\xi|$ occurs, then you get a total of $|d_2|^2 - |c_1|^2$ dollars. Both