Phil Physics: Week 3

Since we went a bit slow last week, there is some overlap with the end of last week's notes and the beginning of this week's notes.

Introduction

What's the new thing in quantum mechanics? What's the philosophical take-away? As we discussed previously, different people give different answers. Some say: many worlds. Others day: consciousness cannot be reduced to matter. Yet others say: nonlocal cause and effect. Yet others say: failure of classical logic.

In order to make progress on that question, it can help to look at a similar, but more precise question: what's special and different about the mathematical models that quantum mechanics provides? Here are some options:

superposition	non-commutativity	entanglement	dynamics

For reasons of mathematical exposition, we'll look at superposition and entanglement before looking at non-commutativity and dynamics.

[Added after lecture: the distinction between **synchronic** and **diachronic**. The former has to do with how things are at a time, and the latter has to do with how things change over time.]

Superposition

As you know, vectors can be added. The math is straightforward. Pictorially, the sum of two vectors is the vector on the diagonal of the parallelogram formed from the original two vectors. In terms of coordinate representations of vectors, the sum can be taken "pointwise". That is,

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 \\ b_1 + b_2 \end{pmatrix}$$

We will call the sum of vectors $\vec{a}$ and $\vec{b}$ their **superposition**. The metaphorical language here comes from a different application of vector addition which we'll see below.

Another mathematical operation we can perform on vectors is scaling, i.e. we can multiply a vector by a real number. Pictorially speaking, to multiply v by 2 is to stretch it to twice its original length (although mind you, we don't have a definition yet of the length of the vector v). To multiply v by 3 is to stretch it to three times its original length. On the other side, to multiply v by -1 is to reverse its direction, and to mutiply v by 0 is to squish it down to a vector of zero length.

Now, as to the origin of the word **superposition** that comes from studying wave phenomena in physics. Imagine that you're on the beach in Maui, and two waves are approaching the shore from slightly different directions. When these waves come together, what happens? They don't collide like solid objects and repel each other. Instead, they start to weave themselves together. At some points, the peaks of the original waves meet to form a higher wavecrest, and at other points, the troughs of the two waves meet to form a lower depression. Of course, there can also be points where the two waves **interfere** with each other, or cancel each other out.

What is the mathematical representation of waves? If we think of the sea floor as represented by the plane $\mathbb{R}^2$ of real numbers, then a wave can be represented by a function $\psi : \mathbb{R}^2 \to \mathbb{R}$. Here we stipulate that $\psi(q) = 0$ represents sea level, i.e. the height of the water when it is at rest. Now, if two waves ψ_1 and ψ_2 are coming in to shore, then we define the superposition wave $\psi_1 + \psi_2$ by:

$$(\psi_1 + \psi_2)(q) = \psi_1(q) + \psi_2(q), \qquad (q \in \mathbb{R}^2).$$

(Obviously the example here is imperfect, because it predicts that if $\psi_1(q) = \psi_2(q) = -0.75$, then the superposition wave would be below the ocean's floor at point q.) The point is simply that the mathematical representations of waves — call them **wave-functions** — form a vector space. In the present example, the constructed vector space would be much larger than the space

[Figure to be supplied in lecture]

of directions in $\mathbb{R}^2$. Nonetheless, the idea that material things (such as electrons) display wave-like behavior is the reason why we represent their states by vectors in a vector space.

Let's look now at the simplest possible waves. Consider a string stretched between two points. In fact, let's idealize to the point where the string has just two locations: the left side, 0, and the right side, 1. The state of the string can then be represented by an assignment ψ of real numbers to those two points. What's more, such states can be scaled, and any two such states can be superposed. In other words, the states of the string form a vector space. If we wanted to be fancy, we could call this vector space $l_2(\mathbb{Z}_2)$, i.e. it's the space of functions from $\mathbb{Z}_2 = \{0, 1\}$ to the real numbers.

So now here's the idea behind using a vector space to represent "spin": the spin state of an electron is like a string with two locations, subject to the condition that the values at the two locations sum, after being squared, to 1. In other words, we require that an electron's spin wave-function ψ has the feature that $(\psi(0))^2 + (\psi(1))^2 = 1$.

To get the feel for superposition, it might help to look at another kind of experiment: the famous two-slit interference experiment. Suppose that there's a stream of particles directed toward a screen with two slits, and behind the screen there is another detector screen. Suppose also that there are little doors on the slits that we can open and close.

In the first experiment, we close the bottom door so that the stream only goes through the top door, and we see a pattern of detections on the back screen like this:

That's not surprising: we expect that the particles emerge from the slit with fairly random momentum. What's surprising is what happens when we open the second door. If the source were producing discrete particles, then the prediction of classical physics would be two lumps on the back screen, like this:

In contrast, if the source were producing waves, then classical physics would predict that the waves coming out of the two slits would interfere with each other, producing an interference pattern on the back screen.

Quantum mechanics also predicts the interference pattern, and the explanation goes like this: if only the top slit is open, then it prepares particles in the state $|z+\rangle$. If only the bottom slit is open, then it prepares particles in the state $|z-\rangle$. However, if both slits are open, then the state is $\frac{1}{\sqrt{2}}(|z+\rangle + |z-\rangle)$. This latter state is *not* a state in which the particle definitely goes through the top or bottom slit. Instead, it's more like a wave that goes through *both* the top and bottom slits, and then interferes with itself on the other side.

Now back to our Stern-Gerlach magnets. We have already represented the spin state of an electron by a vector in $\mathbb{R}^2$. Mathematically, these vectors can be superposed, i.e. added together. But what does that mean physically? Suppose we take the state $|z-\rangle$ where the electron has the property of "down" for S_z, and the state $|z+\rangle$ where the electron has the property of "up" for S_z, and then we add them together. Does the resulting vector define a physical state, and what is that state like?

Since we have

$$|z-\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |z+\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

it follows that

$$|x+\rangle = \frac{1}{\sqrt{2}}(|z-\rangle + |z+\rangle), \qquad |x-\rangle = \frac{1}{\sqrt{2}}(|z+\rangle - |z-\rangle).$$

Hence, $|x+\rangle$ is a superposition of $|z-\rangle$ and $|z+\rangle$, and $|x-\rangle$ is a different superposition of $|z-\rangle$ and $|z+\rangle$.

That is curious for several different reasons. First, what in the world does S_z have to do with S_x? Aren't these supposed to be independent axes? How could summing a state with one apple and a state with two apples yield a state with one orange? Second, how can summing together states where S_z is sharp give rise to states where S_z is fuzzy? That's especially puzzling because electrons can't remain ambivalent about which way they'll go through a S_z magnet: they have to go up or down.

Summary of the philosophical issues: superposition of quantum states is a whole new kind of thing that we've never quite seen before. In one sense, understanding superposition is *the* problem of interpreting quantum mechanics.

Summary of the mathematical issues: quantum states are represented by vectors, i.e. elements of a vector space. A (real) **vector space** H is a set with a special element $0 \in H$, an operation $+$ that sums vectors, and a scalar multiplication operation. (Here our scalars are real numbers. Later they will be complex numbers.) These operations satisfy the axioms:

$$\begin{aligned}
u + 0 &= u \\
u + v &= v + u \\
u + (v + w) &= (u + v) + w \\
a(u + v) &= au + av \\
(ab)v &= a(bv)
\end{aligned}$$

At this point, we could define the notions of a **basis** for H, a **subspace** of H, and the **dimension** of H. (Not surprisingly, $\mathbb{R}^2$ comes out as having dimension 2.) But we won't need those notions for a while yet.

Another example

Suppose that there are two boxes, left and right, and a marble that can be in either of the two boxes. We use the vector $|L\rangle$ to represent the state where the marble is in the left box, and $|R\rangle$ to represent the state where the marble is in the right box. We use $|0\rangle$ to represent the state where the marble is sitting still, and $|1\rangle$ to represent the state where it is moving from one box to the other.

According to QM, the relation of momentum (velocity) to position is represented by the following equations:

$$\begin{aligned}
|0\rangle &= \tfrac{1}{\sqrt{2}}\left(|L\rangle + |R\rangle\right), \\
|1\rangle &= \tfrac{1}{\sqrt{2}}\left(|L\rangle - |R\rangle\right).
\end{aligned}$$

In other words, "sitting still" is a superposition of being in the left and the right boxes, and "moving" is a different superposition of being in the left and the right boxes. Notice that the relation here is directly analogous to the relation between S_z and S_x.

How are we to interpret superposition? What does it *mean* to say that the state is a superposition of $|L\rangle$ and $|R\rangle$? It's tempting to think that it can be interpreted probabilistically, i.e. that it means that there's a fifty percent chance that the marble is in the left box and a fifty percent chance that it's in the right box. But if you say that, then you're subject to a "Danish book" argument (a name I made up, in honor of the so-called Copenhagen interpretation QM):

1. Suppose that the state is $|0\rangle$, which is a superposition of $|L\rangle$ and $|R\rangle$.

2. There's a fifty percent chance that the marble is in the left box.

3. If the marble is in the left box, then its state is $|L\rangle$.

4. But $|L\rangle$ is a superposition of $|0\rangle$ and $|1\rangle$.

5. Hence, there is a fifty percent chance that the state is $|1\rangle$, which contradicts our assumption that the state is $|0\rangle$.

Inner product and length

Last week we talked about the **Born rule**, which is a rule for calculating the probability of a measurement outcome, conditional on the given quantum state (i.e. state vector). We can use the notation:

$$\text{Prob}(A = a|w) \;=\; |\langle v_a, w\rangle|^2.$$

The thing on the right-hand side is pure math; the thing on the left-hand side is our interpretation of that math. The equation "$A = a$" doesn't really make sense as an equation; instead, it is shorthand for the statement that "the quantity A takes value a." Here we are assuming that v_a is the quanum state in which A has value a. The official terminology is that v_a is an **eigenstate** of A with **eigenvalue** a. (For the technically inclined: the quantity A will be represented by a linear operator on the state space. But you don't need to know that yet!)

I also mentioned last week that $\langle v_a, w\rangle$ is something like the angle between the vectors v_a and w. Its real name is the **inner product**. To make that idea more precise: given vectors in coordinate representation, we define their inner product as follows:

$$\left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\rangle \;=\; a_1 b_1 + a_2 b_2.$$

Exercise. Show that the inner product, as defined above, is linear in both arguments. For the first argument, you'll show that $\langle u_1 + u_2, v\rangle = \langle u_1, v\rangle + \langle u_2, v\rangle$, and $\langle ru, v\rangle = r\langle u, v\rangle$.

Exercise. Show that $\langle v, v\rangle \geq 0$, and that $\langle v, v\rangle = 0$ only if $v = 0$.

We can use the properties derived in the previous two exercises as the official definition of a (positive definite) inner product on a vector space H: it's a function from pairs of elements of H to real numbers that is linear in both arguments, and such that $\langle v, v\rangle \geq 0$, with $\langle v, v\rangle = 0$ only in the case that $v = 0$.

Definition. Given a vector space H with an inner-product $\langle -, -\rangle$, we define $\|v\| = \langle v, v\rangle^{1/2}$. We call $\|v\|$ the **length** or **norm** of the vector v.

Convention. Quantum states will be represented by unit-length vectors.

For any innner product, we have the following result, called the **Cauchy-Schwartz inequality**:

$$|\langle u, v\rangle| \;\leq\; \|u\|\,\|v\|.$$

(The proof isn't difficult, but we won't go through it here.) Recalling the Born rule, this result shows that when u and v are unit vectors, then $0 \leq |\langle u, v\rangle|^2 \leq 1$, which permits its interpretation as the **transition probability** from u to v.

Definition. Vectors u and v are said to be **orthogonal** just in case $\langle u, v\rangle = 0$.

We have stipulated that quantities (aka observables), such as S_x and S_z, are represented by orthogonal pairs of vectors. In other words: a quantity A corresponds to an orthogonal decomposition of the vector space H, where distinct values for A correspond to distinct vectors in this decomposition. These vectors are the **eigenvectors** for A. Because of this stipulation, we have

$$\text{Prob}(A = a | A = a') = 0,$$

when $a \neq a'$. (In fact, quantities such as A will be represented by linear operators on H. But you don't need to know that yet.)

Mathematical summary: An **inner product** on H is a function that takes pairs of vectors and returns a real number. (Later, when we look at complex vector spaces, the inner product may return a complex number.)

Rotation of Stern-Gerlach magnets

We set things up so that the eigenvectors for S_x and S_z are skew to each other in the plane. That is, if φ is an eigenvector for S_x and ψ is an eigenvector for S_z, then

$$\langle \varphi, \psi \rangle \;=\; \pm \frac{1}{\sqrt{2}},$$

and hence

$$\mathrm{Prob}(S_z = \pm 1 | S_x = \pm 1) \;=\; \frac{1}{2}.$$

Tensor products

Let's turn back now to the experiment with the reflectors — where "down" outcomes of a S_z gate are channelled to a S_x gate, and the outcomes are then recombined and sent to another S_z gate. The results of this experiment (i.e. always down) are not explained by the simplified version of the Born+collapse rule that we gave above — at least not if we think of the S_x gate as a "measurement." How then can we think of what happens when the electron passes through that gate? For this, we'll need to adopt a more sophisticated formalism.

An electron is, in one sense, a complicated thing. It has more than one kind of property. It has its spin properties, but it also has location (and momentum) properties. So far, we have been idealizing away from those other properties. We have spoken as if the electron goes up or down, but we didn't explicitly encode that into its state vector.

So, an electrons state really has two parts: its spin part, and its position part. Let's use the math symbol "$\otimes$" to hold those two parts apart from each other. So, if an electron is up for σ_z, and is literally physically up, then we'll write its state as $z_1 \otimes u$. (For now, we can pretend like there are only three possible spatial locations: up u, down d, and middle m.)

Here then is another hypothesis about what happens when an electron, starting in position m, passes through a S_z gate: if the electron is in state $|z+\rangle$ then it goes to state u, and if the electron is in state $|z-\rangle$ then it goes to state d. More precisely:

$$\begin{aligned} m \otimes |z-\rangle &\longmapsto d \otimes |z-\rangle \\ m \otimes |z+\rangle &\longmapsto u \otimes |z+\rangle \end{aligned}$$

This assumes, of course, that the S_z gate doesn't disturb the spin part of the state, so long as it's z-up or z-down. (Do we have any evidence for thinking that's true?)

But if the state vector is something like $m \otimes \psi$, then what vector space are we talking about? We have assumed that the spin state ψ is in the space $\mathbb{R}^2$, but now we also have a second vector space for location, and the electron's quantum state is somehow a "product" of those two vectors.

In order for these **product states** to count as quanum states, we need to be able to superpose them. That is, we need to define linear combinations of two product states. We could have a very long discussion indeed about what justifies the formal rules for product states; that's an area of research in itself. For now, I'll just have to give you the rules of calculation as brute (unexplained) facts.

$$
\begin{aligned}
&(x_1 + x_2) \otimes y \;=\; x_1 \otimes y + x_2 \otimes y \\
&x \otimes (y_1 + y_2) \;=\; x \otimes y_1 + x \otimes y_2 \\
&r(x \otimes y) \;=\; rx \otimes y \;=\; x \otimes ry
\end{aligned}
$$

We'll also need some facts about how the tensor product relates to inner products. The basic fact is:

$$\langle x_1 \otimes y_1, x_2 \otimes y_2 \rangle \;=\; \langle x_1, x_2 \rangle \cdot \langle y_1, y_2 \rangle$$

This means, for example, that if x_1 is orthogonal to x_2, then $x_1 \otimes y_1$ is orthogonal to $x_2 \otimes y_2$.

Mathematical summary: given two vector spaces H and K, there is a vector space $H \otimes K$ that is generated by vectors of the form $x \otimes y$. However, not every vector ψ in $H \otimes K$ has the form $x \otimes y$. That is, there will be superpositions of product states, i.e. states of the form

$$x_1 \otimes y_1 + x_2 \otimes y_2,$$

that cannot be further simplified. Such states present a massive challenge for physical understanding; they are called **entangled states**. Here's the official definition:

Definition. A vector ψ in $H \otimes K$ is said to be a **product vector** if it has the form $\psi = u \otimes v$. If ψ is not a product vector, then it's said to be **entangled**.

Linear operators

In QM, linear operators play multiple roles — as representatives of dynamic transitions, and as representatives of quantities.

Definition. Let H be a vector space. A **linear operator** A on H is a function $A : H \to H$ such that

$$\begin{aligned} A(u+v) &= Au + Av, \\ A(ru) &= rAu, \end{aligned}$$

for all $u, v \in H$ and $r \in \mathbb{R}$.

Exercise. Convince yourself that if A and B are linear operators, then the composite function $B \circ A$ is also a linear operator. (We usually just write BA for the composite.)

Exercise. Convince yourself that if A and B are linear operators, then so if the operator $B + A$ whose action is defined by $(B + A)u = Bu + Au$.

As many of you know, one easy way to define linear operators is with matrices. First of all, suppose that we've equipped H with a system of coordinates so that each $\psi \in H$ is represented by a column vector of real numbers. (For simplicity, we continue to assume that H is isomorphic to $\mathbb{R}^2$.) Then for any four real numbers $a_{11}, a_{12}, a_{21}, a_{22}$, there is a linear operator given by

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}.$$

The dynamical changes of state will be represented by a particular kind of linear operator:

Definition. A linear operator U on H is said to be **unitary** just in case $\langle U\varphi, U\psi \rangle = \langle \varphi, \psi \rangle$, for all $\varphi, \psi \in H$.

Recall that we represented the quantity S_z as follows:

- Value $+1$ is associated with the vector $|z+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
- Value -1 is associated with the vector $|z-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

It's obvious that any vector in $H \cong \mathbb{R}^2$ is a linear combination of $|z+\rangle$ and $|z-\rangle$. Indeed, we have

$$\begin{pmatrix} a \\ b \end{pmatrix} \;=\; a|z+\rangle + b|z-\rangle.$$

In this case, we say that $|z+\rangle$ and $|z-\rangle$ form a **basis** for H. Indeed, they form an **orthonormal basis** because $|z+\rangle$ is orthogonal to $|z-\rangle$, and each vector is of unit length. We can now refine what we said before: quantities correspond to orthonormal bases of the state space.

Exercise. Show that $|x+\rangle$ and $|x-\rangle$ form an orthonormal basis for H.

Given an orthonormal basis x_1, x_2 of H it's easy to define linear operators: just choose where to send x_1 and x_2, and everything else will be defined automatically. For example, consider the linear operator S_z that sends $|z+\rangle$ to $|z+\rangle$, and $|z-\rangle$ to $-|z-\rangle$. We then have

$$S_z \begin{pmatrix} a \\ b \end{pmatrix} \;=\; S_z\left(a|z+\rangle + b|z-\rangle\right) \;=\; a|z+\rangle - b|z-\rangle \;=\; \begin{pmatrix} a \\ -b \end{pmatrix}.$$

In this case we say that $|z+\rangle$ is an **eigenvector** for S_z with **eigenvalue** $+1$. Similarly, $|z-\rangle$ is an eigenvector for S_z with eigenvalue -1.

Exercise. Show that S_z is unitary.

Since $|x+\rangle$ and $|x-\rangle$ also form an orthonormal basis, we can define S_x in an analogous way. That is, we set $S_x|x+\rangle = |x+\rangle$ and $S_x|x-\rangle = -|x-\rangle$, and then extend linearly to all vectors in H. The operator S_x is also unitary (since it sends one orthonormal basis to another orthonormal basis). Furthermore, since $|x+\rangle = \frac{1}{\sqrt{2}}(|z+\rangle + |z-\rangle)$ and $|x-\rangle = \frac{1}{\sqrt{2}}(|z+\rangle - |z-\rangle)$, it follows that

$$S_z|x+\rangle \;=\; |x-\rangle, \qquad S_z|x-\rangle = |x+\rangle.$$

Hence, in the current coordinate system

$$S_z \begin{pmatrix} a \\ b \end{pmatrix} \;=\; \begin{pmatrix} b \\ a \end{pmatrix}.$$

Exercise. Confirm that in the current corrdinate assignment, we have

$$S_z \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad S_x \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Modelling interference

We're finally ready to model the experiment with the two reflectors. For simplicity, we treat the down arrow after the first S_z gate as preparing the original state $m \otimes |z-\rangle$. (Later we will revisit the question of what the eraser E does to the spin-z up states.) We now adopt the following hypothesis (i.e. proposed dynamical law) for what happens when an electron passes through S_x gates.

- State $m \otimes |x+\rangle$ changes to $u \otimes |x+\rangle$.
- State $m \otimes |x+\rangle$ changes to $d \otimes |x-\rangle$.
- Linear extension: state $m \otimes (a|x+\rangle + b|x-\rangle) = a(m \otimes |x+\rangle) + b(m \otimes |x-\rangle)$ changes to state $a(u \otimes |x+\rangle) + b(d \otimes |x-\rangle)$.

This hypothesis gives the following particular result: since $|z-\rangle = \frac{1}{\sqrt{2}}(|x-\rangle - |x+\rangle)$, the initial state (just before S_x) can be rewritten as

$$\begin{aligned} m \otimes |z-\rangle &= m \otimes \left(\tfrac{1}{\sqrt{2}}(|x-\rangle - |x+\rangle) \right) \\ &= \tfrac{1}{\sqrt{2}} \left(m \otimes |x-\rangle - m \otimes |x+\rangle \right). \end{aligned}$$

By linearity, the S_x gate changes this state to the state

$$\tfrac{1}{\sqrt{2}} \left(d \otimes |x-\rangle - u \otimes |x+\rangle \right).$$

(This looks like an entangled state, and in fact it is!) The top reflector then serves as a unitary gate that changes u to m; the bottom reflector serves as a unitary gate that changes d to m. Hence, after the reflectors, the state is again

$$\tfrac{1}{\sqrt{2}} \left(m \otimes |x-\rangle - m \otimes |x+\rangle \right) \;=\; m \otimes |z-\rangle.$$

That predicts the result: at the final S_z gate, all electrons go into state $d \otimes |z-\rangle$.