Lecture Notes III: Introduction to Quantum Mechanics

Cory Schillaci

July 14, 2011

Quantum mechanics was one of two revolutionary paradigm shifts of the early 20$^{\text{th}}$ century[1]. Although many people will have heard that quantum mechanics is strange (i.e. unintuitive in terms of everyday experiences), relatively few know many details of how it works or the exact predictions. You will get to join that select few. Although we certainly don't have time to touch on all of the predictions of quantum mechanics, we will cover some of the most profound ideas and see how these appear again and again in contemporary physics research.

1 A simple non-classical experiment

One of the most simple experiments with an unquestionably quantum outcome is the Stern-Gerlach (SG) experiment. What we want to do is pass electrically neutral objects between two poles of a magnet then image them on a special film. In classical physics, the speed of rotation of the particles around one axis would determine their deflection in the field. Because this rotation speed could be anything, the classical outcome would be a line of particles with a continuous set of deflections. Stern and Gerlach performed this experiment with silver atoms and obtained a much different result. The silver atoms were detected at only two points on the film, not in a line! See Figure 1 for a schematic of this experiment.

Our goal now is to develop a theory which can explain this result. We will proceed by introducing some postulates that form the basis of quantum theory, motivated by the outcome of the SG experiment.

2 Describing systems in quantum mechanics

I'm going to make a leap based on knowing the right answer and suggest that we consider the following postulate:

[1]The other being Einstein's theories of relativity.

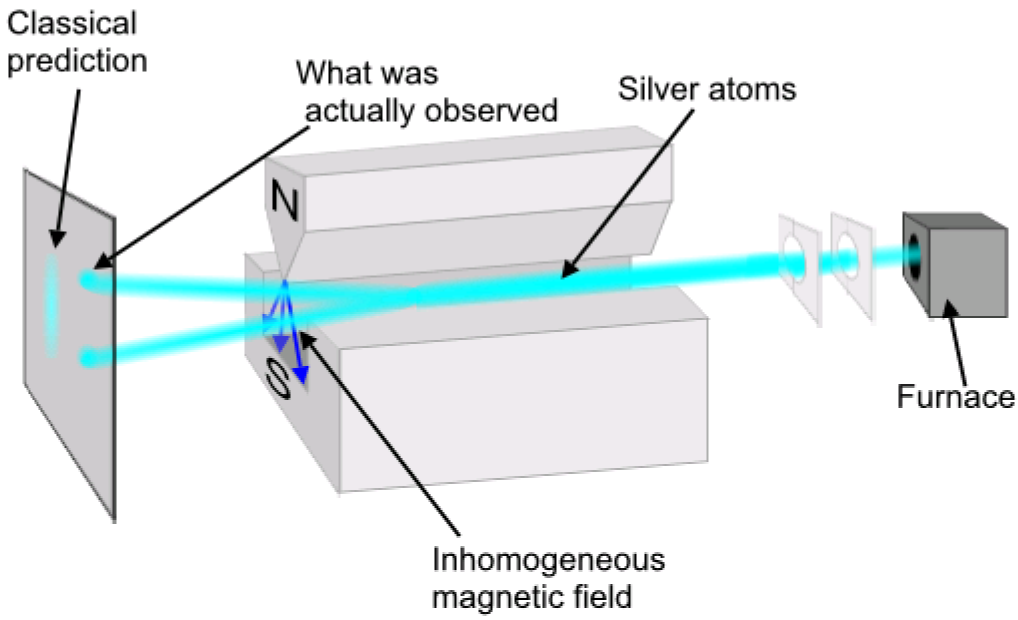

Figure 1: A schematic of the Stern-Gerlach experiment. From Wikimedia commons, originally by Theresa Knott.

Postulate I: A quantum mechanical system is described entirely by a unit vector, notated as $|\psi(t)\rangle$ and called a *state*, in a *Hilbert space*(a vector space with an inner product). Furthermore, every unit vector in this Hilbert space corresponds to a possible physical state of the system.

There's a lot of vocabulary here to unwrap. Formally, a vector space consists of a set V of vectors, $V = \{|\alpha\rangle, |\beta\rangle, |\gamma\rangle, \dots\}$ and a set of S of scalars $S = \{a, b, c, \dots\}$[2]. These sets are associated with operations of vector addition, scalar addition, and scalar multiplication which satisfy the following requirements:

i) The vector sum of any two vectors exists and is equal to another vector in the set. In other words, for any vectors $|\alpha\rangle$ and $|\beta\rangle$ in V there also exists a vector $|\gamma\rangle$ in V such that $|\alpha\rangle + |\beta\rangle = |\gamma\rangle$. This property is called *closure*.

ii) Addition of vectors is *commutative*: $|\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle$.

iii) Addition of vectors is *associative*: $|\alpha\rangle + (|\beta\rangle + |\gamma\rangle) = (|\alpha\rangle + |\beta\rangle) + |\gamma\rangle$.

iv) There is a *null vector* $|0\rangle$ (sometimes simply notated as 0) which is the additive identity. This means that for any $|\alpha\rangle$ in V, $|\alpha\rangle + |0\rangle = |\alpha\rangle$.

v) Scalar multiplication takes any scalar a from S and any vector $|\alpha\rangle$ in V and gives a vector $|\beta\rangle$ from V, $a|\alpha\rangle = |\beta\rangle$.

[2]These scalars should form what mathematicians call a *field*. This basically means they should be associated with the operations of addition, subtraction, multiplication, and division familiar from the real numbers

vi) Scalar multiplication is associative: $a(b\,|\alpha\rangle) = (ab)\,|\alpha\rangle = ab\,|\alpha\rangle$.

vii) Scalar multiplication is distributive over the vectors: $a\,(|\alpha\rangle + |\beta\rangle) = a\,|\alpha\rangle + a\,|\beta\rangle$.

viii) Scalar multiplication is distributive in the scalars: $(a+b)\,|\alpha\rangle = a\,|\alpha\rangle + b\,|\alpha\rangle$.

ix) The special cases $1\,|\alpha\rangle = |\alpha\rangle$ and $0\,|\alpha\rangle = |0\rangle$ hold, where 0 and 1 are the additive and multiplicative identities from S, respectively.

Our quantum mechanical vector space also has an inner product associated with it. The inner product between $|\alpha\rangle\,`\times'\,|\beta\rangle$ is notated $\langle\alpha|\beta\rangle = a$, where a is a scalar in S. An inner product must satisfy:

i) $|\alpha\rangle\,(|\beta\rangle + |\gamma\rangle) = \langle\alpha|\beta\rangle + \langle\alpha|\gamma\rangle$.

ii) $\langle\alpha|\,(a\,|\beta\rangle) = a\,\langle\alpha|\beta\rangle$.

iii) $\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*$.

iv) $\langle\alpha|\alpha\rangle >= 0$, where the equality holds if and only if $|\alpha\rangle = |0\rangle$.

That's a long list of rules. Fortunately, we are already familiar with a physically intuitive vector space from yesterday. For this vector space, V was collection of objects representing physical magnitudes and directions in two dimensions, basically arrows in the plane. The scalars S were just the real numbers.

Quick Question: Check that these arrows, combined with the real numbers, satisfy the requirements of a vector space with an inner product.

We had two ways of representing our "arrow" vector space. The first was to choose a set of orthonormal basis vectors, $\hat{x}$ and $\hat{y}$, which allowed any vector to be written as $\vec{V} = a\hat{x} + b\hat{y}$. Our other option was to specify a basis such as (x, y) and represent the vector as a row or column vector. Note that our new notation for inner products suggests that $|V\rangle$ is always a column vector and $\langle V|$ is always the row vector equal to its transpose. This ensures that the rules of matrix multiplication give the desired result.

For quantum mechanical vector spaces, the scalars are always the complex numbers. The set of vectors is the set of all possible physical states for the system. Generally, we choose a set of basis vectors which are in one-to-one correspondence with the possible outcomes of a certain physical measurement. Let's see how this works for the SG experiment.

In the Stern-Gerlach experiment there are exactly two outcomes. We choose the states which give each outcome to be basis vectors. Since the possible outcomes were that the silver atom went up or down, let's write the basis vectors as $|\uparrow\rangle$ and $|\downarrow\rangle$. By construction, if $|\psi\rangle = |\uparrow\rangle$, the SG apparatus will deflect the atom upwards (and similar for the down state).

Going back to Postulate 1 we are faced with a strange necessity. Since $|\psi\rangle$ can be any unit vector in the vector space containing $|\uparrow\rangle$ and $|\downarrow\rangle$, we can have states of the form

$$|\psi\rangle = \frac{a\,|\uparrow\rangle + b\,|\downarrow\rangle}{\sqrt{|a|^2 + |b|^2}} \tag{1}$$

for any a or b! We will explore what this means in the next section.

Quick Question: Why did I include the denominator of $\sqrt{|a|^2 + |b|^2}$ in (1)?

3 Extracting information from the state vector

One of the claims in Postulate 1 is that a state vector completely describes any system. This implies that there is no physical information that cannot be extracted from $|\psi\rangle$. This section will be an introduction to some of the predictions we can make from the state vector.

3.1 The death of determinism

Quantum mechanics radically changed the way we understand the universe. One of the paradigm shifts that quantum mechanics demanded was the abandonmnent of *determinism*. Determinism is the idea that, given sufficiently accurate information about the initial state of a system, an accurate prediction can in theory be made about the complete future behavior of every constituent[3].

All of this is a consequence of states such as the one in (1). If both a and b are nonzero, the system is neither in the state where we will see it go up nor the state where it will be deflected downwards. Classical physics had never come across any situation like this.

Before interpreting such a state, we must introduce the idea of measurement. The SG apparatus is an example of a measurement apparatus. When an atom reaches the film, the film measures the atom's position which is in one-to-one correspondence with the up and down states due to the magnet. The film always measures the atom in one or the other of the two spots. To understand this we introduce another postulate:

Postulate 2: Observable quantities are represented in quantum mechanics by Hermitian operators.

Let's work backwards through the first sentence this postulate, starting with the definition of an *operator*. An operator is simply a function which takes one vector and returns another. Because we can represent quantum states as column vectors,

[3]The discovery of chaos came later and also spelled trouble for determinism. A chaotic system requires exact knowledge of the initial conditions to make a long-term prediction with any accuracy at all.

square matrices are the most convenient way to represent linear operators[4] acting on the vectors in the Hilbert space.

Now that we know operators can always be represented by matrices, the requirement that they are Hermitian reduces to the requirement that their matrix representations are Hermitian as defined in Lecture II.

Defining the word *observable* is much harder. Essentially, an observable is anything we can measure directly from experiment. In the case of the SG experiment, we look at where the atom hit the photographic film and that tells us whether it was an $|\uparrow\rangle$ or a $|\downarrow\rangle$. Other observables include position, momentum, and energy. In practice, we sometimes know them when we see them and other times we work very hard to determine the relevant physical observables. Unfortunately, I am not aware of an unambiguous way to define observables.

Postulate 3: When the observable corresponding to the operator A is measured using some physical apparatus, a system in the state $|\psi\rangle$ will always be measured have a value of the observable corresponding to an eigenvalue of this operator.

So how do we know what matrix corresponds to any given observable? We need a matrix with the correct eigenvalues, but there may be more than one such matrix. The reason for this is that there is more than one choice of basis! We can, however, write it in terms of the states which give pure measurement outcomes like this:

$$A = \sum_{\lambda} \lambda \left|\lambda\right\rangle \left\langle\lambda\right| \tag{2}$$

The sum is over all possible outcomes of the measurement.

Quick Question: For the SG measurement, we can chose $\lambda = +1$ to indicate the up state and $\lambda = -1$ to represent the down state. Write (2) explicitly for this case. As soon as we pick a basis, we can write the kets as column vectors and operators as matrices. A convenient choice would be

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{3}$$

$$|\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4}$$

Quick Question: Writing bras and kets as row and column vectors, find S in this basis.

So far we have chosen a basis to work in and found the representation of the Hermitian operator S which corresponds to the Stern-Gerlach up/down observable. We need a prescription for predicting the outcome of an SG measurement using this formalism. From the beginning we have assumed that a state $|\psi\rangle = |\uparrow\rangle$ will definitely come out in

[4] It turns out that there is only one operator in QM that is not linear, and we won't have to worry about it.

the up position (and similar for the down state), so we have to recover that. Another reasonable thing to expect is that if I measure some observable for a system and then measure it yet again arbitrarily soon afterwards, I should get the same outcome. Our final postulate gives the prescription for doing this.

Postulate 4 When a state $|\psi\rangle$ is subject to a measurement of the observable A, the probability that it will be found to have the eigenvalue λ is given by $P(\lambda) = |\langle\lambda|\psi\rangle|^2$. Immediately after being measured, the system will be in the state $|\psi\rangle = |\lambda\rangle$ corresponding to the outcome λ of the measurement.

Let's say we are performing the SG measurement on an atom in the state $|\psi\rangle = (|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$. According to Postulate 3, we don't have a 100% chance of seeing the atom in any given spot on the film. Instead, if we do the measurement many times we will observe that half the time the atom is deflected upwards and half of the time it is deflected downwards.

What's new is that this probability is not due to any lack of knowledge about the system. According to Postulate 1, $|\psi\rangle$ is a complete description. This probabilistic description is fundamentally the only description which exists!

The last part of Postulate 3 tells us that measurement disturbs the state. If a system was in a superposition before the measurement, it *collapses* to an eigenstate of the observable during observation. Introductory treatments of quantum mechanics typically gloss over the details of how this collapse comes about. Often, students think that this collapse is something instantaneous. That's not correct! There is a framework called *decoherence* which fully treats the interaction of the system with the measuring device and explains why the instantaneous collapse picture is nearly correct. It's a bit complicated though, so I'll just leave you with the knowledge that there is a good understanding out there for now.