An informal tutorial introduction to Quantum Mechanics

Neeraj Kumar

July 25th, 2023

In this tutorial, I want to provide a basic introduction to the theory of quantum mechanics, in a mostly informal setting, but also not shying away from using equations. We will start with an odd experimental result that was observed in the 20th century and try to make sense of it by introducing and using our theory.

1 An odd experimental result

Suppose we have a physical system. Let's call our system S. Let's say that this system has two distinct physical properties and let's denote these properties by Z and X and both of these properties have two possible outcomes when measured : -1 and 1.

In the real world, Z and X are the spin in the z-direction, and x-direction respectively of an electron.

Now let's say you measure Z first, and get the value 1.

You then proceed to measure X and say you get the value -1.

If you measure Z again, what value would you expect to get ? I think it's fair to say that with our classical intution, you would expect to get 1 again. However, it turns out that if we do this experiment, we don't always get 1. In fact it turns out from experiment that there is a 50% chance of getting 1, and a 50% chance of getting -1.

What is happening ? It turns out that the laws governing our current experiment are quantum mechanical, and these results can be predicted and understood with the theory of quantum mechanics. L

We will focus in the next sections on how to represent our system and measurements on our system mathematically. Our discussion will be informal.

2 How do physicists explain this result ?

2.1 The state of the system

In quantum mechanics, with each physical system, a complex vector space is associated, and the state of the system at any time is a unit vector in this space. That maybe a lot of new words, if you haven't taken a course in Linear Algebra. But don't worry, we will run through this with an example right now.

A vector in Linear Algebra can be represented by a column matrix. In our case the elements of the vector are complex numbers. Since our system S is two dimensional, our column vectors will be two dimensional. A couple of examples of 2-dimensional column vectors include

$$A = \begin{bmatrix} 1+i \\ 1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

However,there is an additional constraint on our coulumn vectors, that they be unit vectors. Why ? This is because of the way we interpret the entries in our column vector. These entries are actually related to the probabilities of measuring a particular outcome if we measure our system. If our column vector is

$$C = \begin{bmatrix} a \\ b \end{bmatrix}$$

then the the probability of measuring outcome 1 if the property Z is measured is $(a * a^*) = |a|^2$ and the probabilty of measuring -1 if physical property Z is measured is $(b * b^*) = |b|^2$. (Recall that x^* is the complex conjugate of x, and $(x * x^*)$ is the square of the magnitude of x, which is denoted as $|x|$.)

An interesting follow up question that might come up in your mind is - what are the probabilities of outcomes 1 and -1 if X is measured. They turn out to be $|a + b|^2$ and $|a - b|^2$ respectively. We'll get to how we calculated these probabilities later, but for now we accept them. So acceptable state vectors for our system S are then of the kind

$$D = \begin{bmatrix} \frac{1+i}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

If the system is in the state D above, the probabilities of measuring 1 and -1 are $\frac{2}{3}$ and $\frac{1}{3}$ respectively. The state D above is probabilistic, we cannot be certain whether measurement will give result 1 or -1.

However consider states

$$F = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

F has probability 1 of measuring 1 for Z, and G has probability 1 of measuring -1 for Z.

I think by now, we have fair idea of how the current state of a system is represented. Let's move now to measurements.

2.2 Measurements

In quantum mechanics, each observable is associated with a hermitian operator. A hermitian operator can be represented as a matrix. Right now, it suffices to know that our observables Z and X can be associated with

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Why hermitian matrices ? To understand that, we first need to understand two new and related terms - eigenvectors and eigenvalues.

For any matrix, an eigenvector of that matrix is a coulumn vector that, when multiplied by the matrix gives a multiple of itself as the result. As an example, notice that

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} * \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 * \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} * \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -1 * \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

So Z has two eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Notice on the other hand that

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} * \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

So $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is NOT a eigenvector of Z.

The other related term is eigenvalue. This is the multiplier of the eigenvector in the result when the eigenvector is multiplied by the matrix . In this case, 1 and -1 are the eigenvalues of Z.

You should verify that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are the eigenvectors of X with eigenvalues 1 and -1 respectively.

Now let's come back to hermitian matrices, and our theory of measurement.

In our theory, if a physical quantity is associated with a hermitian matrix, then the results of our experiment will give eigenvalues of that matrix as outcomes. So Z can give as measurement outcome, one of two values 1 and -1 - as we had already mentioned before. In general the probability of getting any outcome can be calculated using the eigenvector(s) associated with that eigenvalue. To understand how to make that calculation, we first introduce the concept of the dual of a vector,inner product of vectors, outer product of a vectors and projectors associated with an eigenvalue.

But first let's introduce some new notation. We will now represent our vectors A, B etc as $|A\rangle$ and $|B\rangle$. This is called the a ket vector. There is a "dual" vector associated with ket, represented as bra vector $\langle A|$, $\langle B|$.
The dual of a ket vector $|\psi\rangle$ is the the bra vector $\langle\psi|$ which is a row vector with each entry being the complex conjugate of it's corresponding entry in $|\psi\rangle$. So the dual of $|\psi\rangle = \begin{bmatrix} a \\ b \end{bmatrix}$ is $\langle\psi| = [a^* \ b^*]$.

The inner product between two vectors $|P\rangle$ and $|Q\rangle$, denoted by $|P\rangle = \begin{bmatrix} a \\ b \end{bmatrix}$ and $|Q\rangle = \begin{bmatrix} c \\ d \end{bmatrix}$ is a scalar quantity denoted by $\langle P|Q\rangle$ and calculated as $\begin{bmatrix} a^* & b^* \end{bmatrix} * \begin{bmatrix} c \\ d \end{bmatrix} = (a^* * c) + (b^* * d)$

The outer product between P and Q is denoted as $|P\rangle\langle Q|$. In our case, it will be

$$\begin{bmatrix} a.c* & a.d^* \\ b.c* & b.d^* \end{bmatrix}$$

A projector associated with an eigenvalue m is the sum of outerproducts of all of it's associated eigenvectors. As an example, the projector associated with eigenvalue 1 of Z is

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} * \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

and the projector associated with the eigenvalue -1 of Z is

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} * \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

It can be proved that any hermitian matrix M can be written as the sum of eigenvalues multiplied by their projectors.

You can verify this for Z by noting that,

$$Z = (1 * \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}) + (-1 * \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Now with these tools in hand, we can talk about the probabilities of getting outcomes 1, or -1 when measuring Z - if our system is in some unknown state $|\psi\rangle$

If we are measuring an observable M with eigenvalues denoted by the set $\{m\} = [m1, m2]$, with corresponding projectors $\{P_m\} = [P_m^1, P_m^2]$, then the probability of measuring any particular outcome m is given by

$$P(m) = \langle\psi| P_m |\psi\rangle$$

Taking our example with Z, the probability of measuring 1 is

$$p(1) = \langle\psi| |1\rangle \langle 1| \langle\psi|$$

So if the state is $|\psi\rangle = \begin{bmatrix} a \\ b \end{bmatrix}$, then probability of measuring 1 is ,

$$\begin{bmatrix} a^* & b^* \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} = (a * a^*) = |a|^2$$

This is a result which we have mentioned before. Notice that it means, as already mentioned before that $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ will give the result 1 with probability 1, and the state $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ gives the result 1 with probability 0.

The theory also tells us the new state after measurement if the measurement outcome is the eigenvalue m, the new state is

$$\frac{P_m |\psi\rangle}{\sqrt{p(m)}}$$

In our case, if the outcome of our measurement is 1, we get the new state as

$$\frac{1}{|a|} * \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} = \frac{a}{|a|} * \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

We will learn later that the states like $\frac{a}{|a|}$ of magnitude 1 can be ignored, so the final state is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Similarly it can be shown using the above theory that if outcome is -1, then the new state is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

With this we have everything we need to explain our "unexpected" experimental results from the introduction.

As the last step in this, let's introuduce new terminology for something we will see quite often. We now denote the ket $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ by $|0\rangle$, and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ by $|1\rangle$

3 Explaining our result

Just as a quick reminder, we want to show that measuring Z first, then X, and then Z again does not neccessarily give the same output for the two Z measurements.

Let's say our qubit is in the initial state $|\psi\rangle = \begin{bmatrix} a \\ b \end{bmatrix}$. We don't know the outcome of this experiment when we measure Z. But we know that $p(1) = \langle\psi| * \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} * |\psi\rangle$ where $\langle\psi| = \begin{bmatrix} a^* & b^* \end{bmatrix}$. Cacluating $p(1)$ gives us $|a|^2$. Similarly $p(-1) = |b|^2$.

Let's assume our first experiment gives us 1 as the outcome. We can also calculate using that it will leave the system in the state $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Now, what happens when we do the X measurement on our system. Remember that $X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ It's two eigenvectors are $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with eigenvaule 1 and $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ with eigenvalue -1.

The two projectors for X are $(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} * \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix})$ corresponding to 1, and $(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} * \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix})$

The probability for getting the result -1 is then $\langle\psi| * \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} * |\psi\rangle$. Remember that after the first experiment, our state is $|\psi\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Using this value we can calculate $p(-1) = \frac{1}{2}$.

Let's assume we measure -1 though. What then is the new state of the sytem ? Remember that it is $\frac{P_m|\psi\rangle}{\sqrt{p(m)}}$. Putting in the values for $m = 1$, we get $\frac{\frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} * \begin{bmatrix} 1 \\ 0 \end{bmatrix}}{\frac{1}{\sqrt{2}}}$.

So the new state is $|\psi\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Now to the final stage of doing the Z measurement again.

The probablity of getting 1 for Z is now $p(1) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} * \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{2}$. What just happened ? We started with Z giving outcome 1 after our first experiment, then we did X, but then when we came back to do Z again - we got only a 50% chance of getting 1 again. You can verify that if we did not do X in

the middle, the probability of getting 1 for Z would have remained at 100% as you would normally expect.

With this we conclude our discussion of this weird experimental result.