A short introduction to quantum mechanics I: observables and eigenvalues

This article is aimed at students with a solid knowledge of linear algebra and high school physics. It should be suitable for college sophomores studying science or engineering.

In quantum mechanics there is the concept of the 'observable.' An observable is something that can be observed. There are many observables in quantum mechanics, including energy, momentum, angular momentum, position, etc. Each observable corresponds to a linear operator, in other words, a matrix. Thus there exist matrices for energy, momentum, angular momentum, position, etc. These are more commonly referred to as operators: energy operators, momentum operators, angular momentum operators, position operators, etc.

There should be eigenvalues and eigenvectors corresponding to each matrix. It is known that observed values are always eigenvalues.

Let's say you want to measure the energy of an object. The total energy is the sum of potential energy and kinetic energy, and there should be an energy operator (energy matrix) corresponding to this total energy. If you calculate the eigenvalues of this energy matrix you will get certain values. Only these values can be the energy of the object.

For example, let's say that the eigenvalues of the energy matrix are 2J, 3J, and 5J. This implies that the energy of the object can be 2J or 3J or 5J, but never 4J.

For a real example, let's consider a hydrogen atom. If you know some chemistry, you may know that the energy level of a hydrogen atom can be represented by $E = -R/n^2$ where n is a natural number and R is the Rydberg constant. It is a well-known fact that the energy of a hydrogen atom admits only certain values. The reason is that the eigenvalues of a hydrogen atom's energy matrix are given by $-R/n^2$. As there are an infinite number of choices for the value of n, there are an infinite number of eigenvalues. Since the number of eigenvalues coincides with the order of a matrix, we see that the energy matrix of a hydrogen atom is an infinity by infinity square matrix (in fact, most linear operators, or matrices, considered in quantum mechanics are infinity by infinity matrices).

In quantum mechanics there is also a vector corresponding to each object. These are called "state vectors" or "wave functions." I will now explain how they play a role in quantum mechanics.

Let's come back to the example of an energy matrix with eigenvalues 2J, 3J, and 5J and corresponding eigenvectors $|2\mathrm{J}\rangle$, $|3\mathrm{J}\rangle$, and $|5\mathrm{J}\rangle$ (in quantum mechanics, we usually denote vectors using this notation, called Dirac's bra-ket notation; please see my article entitled "Dirac's bra-ket notation: an exposition for science and engineering students" it for more information).

Let's say that an object's state vector (or wave function) is $|2\mathrm{J}\rangle + |3\mathrm{J}\rangle$. This object is in the state of half $|2\mathrm{J}\rangle$ and half $|3\mathrm{J}\rangle$. Upon observation, there is a 50% probability that the object's energy will be 2J and a 50 % probability that it will be 3J (notice that the probability that the energy will be 2.5J is zero). Also upon observation, the state vector changes to the eigenvector of the eigenvalue corresponding to the observed value. For example, if an observer observes the energy of this object to be 2J, then the state vector (or wave function) changes to $|2\text{J}\rangle$, the eigenvector of the eigenvalue 2J.

For a more concrete understanding, let me give you another example. Let's say that an object's state vector is $(0.6)|2\text{J}\rangle + (0.8)|5\text{J}\rangle$. You can see that the square of 0.6 is 0.36 and the square of 0.8 is 0.64 and that the sum of 0.36 and 0.64 is 1. Therefore, the probability of an observer observing the energy to be 2J is 36% while the probability of observing 5J is 64%. Also, upon observation, the state vector changes to $|2\text{J}\rangle$ with 36% probability and $|5\text{J}\rangle$ with 64% probability. If the state vector changes to $|2\text{J}\rangle$ and the energy is measured again, the result will be 2J with 100% certainty. (This probabilistic interpretation of quantum mechanics is called "Copenhagen interpretation.")

The same is true for all observables, including position, angular momentum, and so on. The state vector of an object is expanded out on the basis of eigenvectors, with the probability of observing a given eigenvalue being the ratio of the square of the coefficient of the corresponding eigenvector to the sum of the squares of the coefficients of all the eigenvectors. (We will explain why the probability is proportional to the square of the coefficient rather than the coefficient itself in "Why is the probability proportional to the wave function squared?". At this point, let me just mention that the coefficients can take negative values, while their squares can only take non-negative values. As probability must be always non-negative, it would not make sense for the probability to be proportional to potentially negative coefficient. Furthermore, those of you who studied freshman optics may remember that the intensity is not proportional to the amplitude, but rather the square of the amplitude. Think along this line.)

To show you another aspect of quantum mechanics, let me mention that it is known that position operators and momentum operators don't commute, that is, $XP - PX$ is non-zero, where X is a position operator (or matrix) and P is a momentum operator (or matrix). Therefore, as we will see later in another article, this imply that we cannot determine an object's position and momentum at the same time. In fact, $XP - PX$ is equal to $ih/2\pi$ where h is the Planck constant, and from this one can derive Heisenberg's famous uncertainty principle.

All calculations in Heisenberg's quantum mechanics are done with the formula $XP - PX = ih/(2\pi) = i\hbar$, while all calculations in Schrödinger's quantum mechanics are done with the assumption that the energy matrix is $-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$ where $V(x)$ denotes potential energy. For example if the state vector (or wave function) is $\psi(x)$, the state vector multiplied by the energy matrix is $-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x)}{\partial x^2} + V(x)\psi(x)$ (although seems to be a function rather than a vector, it is actually both; I will explain why in my next article).

Surprisingly, it was proven that Heisenberg's quantum mechanics and Schrödinger's quantum mechanics are equivalent, and always give the same result. However, in practice, it is much easier to use Schrödinger's quantum mechanics for most cases. In my third article on quantum mechanics, we will learn about the simple proof of this equivalency, which initially took months for Schrödinger to figure out.

Problem 1. Let's say that an object's wave function is $3|2\text{J}\rangle - |4\text{J}\rangle$. What will be the probability of an observer observing its energy to be 4J?

Summary

- An observable is something that can be observed. For every observable, there is a corresponding linear operator (i.e. matrix).
- Observed values are always the eigenvalues of such a linear operator.
- For every object, there is a corresponding vector called “state vector” or ”wave function.”
- The probability of observing a given eigenvalue is proportional to the square of the coefficient of the corresponding eigenvector.
- Upon observation, the state vector changes to the eigenvector of the eigenvalue corresponding to the observed value.
- Heisenberg’s quantum mechanics and Schrödinger’s quantum mechanics are equivalent.

A Historical notes

Heisenberg wrote a breakthrough paper in quantum mechanics by discovering the matrix formulation of quantum mechanics in 1925. In that paper, he argued expressions like

$$S(n, n-\gamma) = \sum_{\alpha=-\infty}^{\infty} \sum_{\beta=-\infty}^{\infty} U(n, n-\alpha) U(n-\alpha, n-\alpha-\beta) U(n-\alpha-\beta, n-\gamma). \tag{1}$$

If you know matrix, you will immediately recognize that this is the component expression of $S = UUU$ for matrices S and U. However, Heisenberg didn’t know about matrix, so Born, who knew about matrix, first noticed this. Using Heisenberg’s approach, Born obtained

$$\sum_b X_{ab} P_{bc} - P_{ab} X_{bc} = i\hbar, \qquad \text{if } a = c. \tag{2}$$

where X_{ab} and P_{bc} are components of position matrix and momentum matrix. However, what he wanted was $XP - PX = i\hbar I$ where I is the identity matrix. In other words, he wanted to show the following as well:

$$\sum_b X_{ab} P_{bc} - P_{ab} X_{bc} = 0, \qquad \text{if } a \neq c. \tag{3}$$

As he had difficulties, he asked for help from Jordan, who solved the problem. Born and Jordan immediately wrote a paper that included this result. In other words, $XP - PX = i\hbar I$. Then, Born, Heisenberg, and Jordan, all three wrote a paper together that also includes this result among other ones. The next year, 1926, Born wrote a paper alone that shows that the probability is proportional to the coefficient squared. In 1933, Heisenberg received alone the Nobel Prize in Physics 1932. Recognizing Born and Jordan’s contribution, he wrote Born that receiving the Nobel Prize alone without them depressed him. In 1954, Max Born received the Nobel Prize for the probabilistic interpretation, but Jordan never received the Nobel Prize.

Today, physicists do not interpret $XP-PX = i\hbar I$ as an expression to be derived, but taken for granted. Immediately after Dirac got a copy of Heisenberg’s first paper on the matrix formulation of quantum mechanics, he independently obtained this expression and found a very similar expression in classical mechanics. Physicists today take it for granted that the classical expression is just replaced by its quantum version through the connection Dirac found. We will talk about Dirac's discovery in our later article "Transition from classical mechanics to quantum Mechanics."