2.3: Measurement and Probability

From the second postulate we have seen that the possible outcomes of a measurement are the eigenvalues of the operator corresponding to the measured observable. The question of which one of the eigenvalue we will obtain is still open. This question is resolved by the third postulate, which gives a recipe on how to predict which outcome will be observed. However, this mathematical recipe does not tell us (in general) with absolute certainty which one of the eigenvalue will be the outcome. It only provides us with the probability of obtaining each eigenvalue.

In postulate 3.b we consider an observable $\mathcal{A}$ and a system in the state $|\psi\rangle$. The eigenvalue equation for the operator A (corresponding to $\mathcal{A}$) can be expressed as

$$A|n\rangle = a_n|n\rangle$$

where a_n are the eigenvalues and $|n\rangle$ the eigenvectors . The postulate states that the probability of obtaining a_n as the outcome of the measurement of A is $p\left(a_n\right) = |\langle n \mid \psi\rangle|^2$.

We want to re-express the postulate in terms of the wavefunction $\psi(\vec{x})$. To do so, we need to define the inner product in the Hilbert space of the wavefunctions. Given two wave functions $\psi(\vec{x})$ and $\varphi(\vec{x})$, the inner product $\langle\varphi \mid \psi\rangle$ is given by:

$$\langle\varphi \mid \psi\rangle = \int d^3\vec{x}\varphi(x)^*\psi(x)$$

(where $*$ indicates the complex conjugate).

We first rewrite the eigenvalue problem for the operator A in terms of the eigenfunctions $u_n(\vec{x})$ and the associated eigenvalues a_n:

$$A\left[u_n(\vec{x})\right] = a_n u_n(\vec{x})$$

Then, we have seen that any function can be expressed in terms of the eigenfunctions $u_n(\vec{x})$. We can as well express the wavefunction in terms of these eigenfunctions:

$$\psi(\vec{x}) = \sum_n c_n u_n(\vec{x}), \text{ with } c_n = \int d^3\vec{x} u_n^*(\vec{x})\psi(\vec{x})$$

Finally, according to postulate 3.b the probability of obtaining the outcome a_n if the system is in the state $\psi(\vec{x})$ is given by the inner product :

$$\boxed{p\left(a_n\right) = \left|\int d^3\vec{x} u_n^*(\vec{x})\psi(\vec{x})\right|^2 = |c_n|^2}$$

where the last equality follows from the orthogonality of the eigenfunctions $\int d^3\vec{x} u_n^*(\vec{x})u_m(\vec{x}) = 0$, for $m \neq n$. Since $|c_n|^2 = p\left(a_n\right)$ is a probability, the coefficients c_n of the wavefunction expansion are called *probability amplitude*s. We now confirm that the wavefunction contain all information about the state of the system, since given the wavefunction we can calculate all the probabilities of each outcome for each possible observable with the following procedure:

1. Find the eigenfunctions of the observable's operator. E.g., given an operator O, we will calculate its eigenfunctions $w_n(x)$, such that $O\left[w_n(x)\right] = o_n w_n(x)$.
2. Find the probability amplitude of the wavefunction with respect to the eigenfunction of the desired eigenvalue outcome. E.g., if the outcome is o_m, such that $O\left[w_m(x)\right] = o_m w_m(x)$ we will calculate $c_m = \int d^3\vec{x} w_m^*(\vec{x})\psi(\vec{x})$.
3. The probability of obtaining the given eigenvalue in the measurement is the probability amplitude modulus square. E.g. $p\left(o_m\right) = |c_m|^2$.

Wavefunction collapse

The third postulate states also that after the measurement the system is left in the eigenstate corresponding to the eigenvalue found (more generally, if more than one eigenstate is associated to the same eigenvalue, the state is projected on the subspace of the eigenvalue a_n, that is, the subspace spanned by all the eigenstates associated with a_n).

This is the **wavefunction collapse**, a concept that is usually quite puzzling in quantum mechanics. We can make this statement at least a bit less puzzling by taking into account the following two considerations.

The wavefunction collapse is puzzling because it predicts an instantaneous evolution of the system from its pre-measurement state $\psi(x)$ to its post-measurement state $u_n(x)$ (when we measure a_n). This type of evolution is very different than the usual evolution predicted by the fourth postulate (that we will see in a later lecture). However, this weird behavior arises from considering the measurement apparatus (and hence the measurement) as a classical system, outside the realm of quantum mechanics. Although this view gives most of the time a correct answer – and thus we will use it in this class – it is a quite imprecise description. More rigorous descriptions of the measurement process, invoking for example **decoherence**[4], can give a better picture of what actually happens (e.g. the wave-function collapse can take a finite time and be measured experimentally in some cases).

More pragmatically, the wavefunction collapse is needed in order to make experiment consistent. What the collapse entails is that if I make a measurement and I obtain as an outcome the eigenvalue a_n, I can check that result again, by repeating the measurement just after the first one (with no time for any change in the system between the two measurement). If I could not make this second check, I could never be able to be confident that I got the correct answer the first time (e.g. my detector could be wrong) and so I could never gain any knowledge at all on my system.

Obs.: I want to clarify the meaning of "subspace of the eigenvalue a_n". If there is a set of eigenstates associated with the eigenvalue a_n, $|n_j\rangle$, then the state $|\psi\rangle$ is projected onto a superposition of these eigenstates $|\psi\rangle \rightarrow |n\rangle = \sum_j c_j |n_j\rangle$.

Note

4 Decoherence is the phenomenon by which an *open* quantum system, interacting with the environment, undergoes an irreversible evolution that often leaves it in a state best described by the rules of classical mechanics.

Position measurement

We have already calculated the eigenvalues and eigenfunctions of the position operator. The eigenfunctions were $u_n(x) = \delta(x - x_n)$ with eigenvalues x_n. We also calculated the expansion of a function in terms of the position eigenfunctions. Repeating the calculation for the wavefunction we find:

$$c(x_n) = \int dx \delta(x - x_n)\psi(x) = \psi(x_n),$$

from which we obtain that the probability of finding a particle in the position x_n is given by:

$$p(x_n) = |\psi(x_n)|^2$$

More generally, since x is continuous, we can drop the subscript n, as any value of x is an eigenvalue. Then, generalizing to the 3D case, we can say that the probability of finding a particle described by the wavefunction $\psi(\vec{x})$ at the position $\vec{x}$ is given by the modulus square of the wavefunction itself:

$$\boxed{p(\vec{x}) = |\psi(\vec{x})|^2}$$

We can also say that the wavefunction is the probability amplitude for the position measurement. More precisely, we should say that the probability of finding a particle between x and $x + dx$ is $p(x)dx = |\psi(x)|^2 dx$ while $|\psi(x)|^2$ is a **probability density** (per unit length). In 3D, $|\psi(\vec{x})|^2$ is the probability density per unit volume and the probability is given by $|\psi(\vec{x})|^2 d^3x$.

Given this interpretation of the wavefunction, it becomes natural to require that the wavefunction be normalized. We require that integrating the probability of a particular position over all possible position we obtain 1 (i.e. certainty, the particle **has** to be somewhere!). Then

$$\int d^3\vec{x} p(\vec{x}) = 1 \quad \rightarrow \quad \int d^3\vec{x} |\psi(\vec{x})|^2 = 1$$

From being a very abstract notion, the wavefunction has thus assumed a very physical meaning. We can still say that the wavefunction describes the state of a system and contains all the information about this system. In addition, and more concretely, the absolute value of the wavefunction tells us where it is more probable to find the system.

Momentum measurement

We calculated the eigenvalues and eigenfunctions of the momentum operator to be

$$p = \hbar k \quad \text{and} \quad u_k(x) = \Upsilon e^{ikx}$$

(Notice that we could have labeled the wave-numbers as k_n —and the momentum p_n— to have the eigenvalue equation: $\hat{p}u_n = p_n u_n$, but we omitted the subscript n since momentum is a continuous variable; then we also simply label the eigenfunctions by k instead of n).

As usual, we would like the eigenfunctions to be normalized. However notice that $\int u_k^*(x)u_k(x)dx = \int |\Upsilon|^2 dx = \infty$, so we cannot fix Υ such that the result is normalized as usual. For convention we set $\Upsilon = -\frac{1}{\sqrt{2\pi}} : u_k(x) = -\frac{1}{\sqrt{2\pi}} e^{ikx}$ (we are considering the 1D case). Now we can calculate the probability amplitude for the momentum measurement, by calculating the coefficients of the expansion of the wavefunction in terms of the momentum eigenfunctions basis. Here we rename the coefficients $c(k)$ of the expansion $\varphi(k)$. This is given by:

$$c(k) \equiv \varphi(k) = \int u_k(x)^* \psi(x) \rightarrow \varphi(k) = \frac{1}{\sqrt{2\pi}} \int e^{-ikx} \psi(x)$$

Notice that this last equation is simply stating that the probability amplitude for the momentum measurement is the Fourier transform of the wavefunction, $\varphi(k) = \mathcal{F}[\psi(x)]$. Then

$$p(k \rightarrow k + dk) = |\varphi(k)|^2 dk$$

is the probability of finding that the particle has a momentum between $\hbar k$ and $\hbar(k + dk)$ when measuring the momentum.

Flux of particles

The choice of the coefficient Υ implies that:

$$\int u_k^*(x)u_{k'}(x)dx = \delta(k - k')$$

which is not the usual normalization for the wavefunction.

Why is it not possible to normalize the momentum eigenstates?

We saw that for the wavefunction, the normalization was related to its interpretation as the probability amplitude for the position. If a wavefunction is in a momentum eigenstate $\psi(x) = u_k(x)$, then we cannot really talk about a particle, but rather the system is better described by a wave. In fact, the probability of finding the system at any position x is constant and equal to Υ. Thus the coefficient Υ can be better linked to a flux of particles rather than a particle density.

We can set $v|\psi|^2 = \Gamma$ where $v = \frac{p}{m} = \frac{\hbar k}{m}$ is the velocity and Γ correspond to a flux of particle, as described by the plane wave e^{ikx}. Then $\frac{\hbar k}{m}|\Upsilon|^2 = \Gamma$ fixes the value of Υ to $\Upsilon = \sqrt{\frac{m\Gamma}{\hbar k}}$.

Expectation values

We have just seen that the outcome of a measurement is a random quantity (although it is of course limited to a given set of values –the eigenvalues– from which we can choose from). In order to know more about the state of a system we need then to repeat the measurement several times, in order to build a statistics of the observable.

For example, we could be interested in knowing what is the average of the measurements of a particular observable. This quantity is usually called in QM the **expectation value** of an observable.

How do we usually calculate the average of a given quantity? Consider for example the average number obtained by throwing a dice. In an experiment, I would have to repeatedly throw the dice, record the number that comes out (say n_i) and then calculate the sum: $\langle n \rangle = \frac{1}{N} \sum_{i=1}^{N} n_i$. Equivalently, I could count the number of times t_n that each number n appears and calculate the average number from the frequencies $\nu_n = t_n/N : \langle n \rangle = \sum_{n=1}^{6} \nu_n n$. In the limit of N → ∞, the frequencies ν_n approach the probabilities $\nu_n \rightarrow p_n$. Then for the dice we have $p_n = 1/6 (\forall n)$ and the average is just calculates from the sum $\frac{1}{6}(1 + 2 + \ldots 6) = 3.5$.

The procedure for calculating the average (or expectation value) of a quantity is very general. We have for discrete and continuous probability distribution functions respectively

$$\langle x \rangle = \sum_i p_i x_i \qquad \langle x \rangle = \int dx p(x) x$$

In QM we just need to replace the probability p by its value as given by the third postulate. For example, the expectation value of the position can be expressed in terms of the probability density function given by the modulus square of the wavefunction:

$$\boxed{\langle x \rangle = \int_{-\infty}^{\infty} x |\psi(x,t)|^2 dx}$$

How can we in practice obtain this expectation value? (that is, by performing real experiments). If we make a first measurement on a single particle and then repeat the measurement over and over again this is not what we measure. In fact, in that case we know (from the postulates) that after the first measurement we expect always to get the same result. In fact, these repeated successive measurement are only a way to check that yes, we got the first answer correct. (otherwise we could never be certain of anything, since we would not even know that our experimental apparatus works).

Instead what we can do is adopt one of two strategies. Either we can repeat the same *experiment* (not measurement) many times on the same system. This implies first preparing the system in a given state with a reproducible procedure and then performing a measurement. The second option is to make the experiment on a set (an *ensemble*) of identical systems. In order to obtain the exact expectation value we would need an infinite number (of repetitions or systems), however a large enough sample is usually practically enough.

Obs.: Notice that we can rewrite the expression above for the expectation value of the position as

$$\langle x \rangle = \int_{-\infty}^{\infty} \psi(x,t)^* x \psi(x,t) dx = \int_{-\infty}^{\infty} \psi(x,t)^* \hat{x}[\psi(x,t)] dx$$

This last form is a more general one that is valid for any operator.

Definition: Expectation value

The expectation value of an observable $\hat{\mathcal{O}}$ is

$$\boxed{\langle \hat{\mathcal{O}} \rangle = \langle \psi | \hat{\mathcal{O}} | \psi \rangle = \int_{-\infty}^{\infty} \psi(x,t)^* \hat{\mathcal{O}}[\psi(x,t)] dx}$$

where we first used the Dirac notation to express the expectation value, which is an even more general expression.

Example

The expectation value for the momentum operator is given by:

$$\int_{-\infty}^{\infty} \psi(x,t)^* \hat{p}[\psi(x,t)] dx = \int_{-\infty}^{\infty} \psi(x,t)^* \left(-i\hbar \frac{\partial \psi(x,t)}{\partial x} \right) dx$$