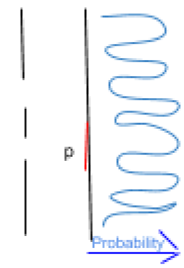
Lecture 2, Thurs Jan 19: Probability Theory and QM

Feynman said that everything about quantum mechanics could be encapsulated in the **Double Slit Experiment.**

In the double-slit experiment, you shoot photons one at a time toward a wall with two narrow slits. Where each photon lands on a second wall is probabilistic. If we plot where photons appear on the back wall, some places are very likely, some not.

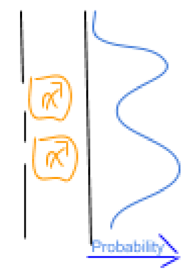


Note that this itself isn't the weird part: we could totally justify this happening, by some theory where each photon just had some extra degree of freedom (an "RFID tag") that we didn't know about, and that determined which way it went. What's weird is as follows. For some interval on the second wall:

Let P be the probability that the photon lands in the interval with both slits open.
Let P_1 be the probability that the photon lands in the interval if only slit 1 is open.
Let P_2 be the probability that the photon lands in the interval if only slit 2 is open.

You'd think that $P = P_1 + P_2$. But experiment finds that that's not the case! Even places that are *never* hit when both slits are open, can sometimes be hit if only one slit is open.

The weirdness isn't that "God plays dice," but rather that "these aren't normal dice"!



You may think to measure which slit the photon went through, but doing so *changes* the measurements into something that makes more sense. Note that it isn't important whether there's a conscious observer: if the information about which slit the photon went through leaks out in any way, the results go back to looking like they obey classical probability.

As if Nature says "What? Me? I didn't do anything!"

This reversion to classical probabilities, when systems are coupled to their environments, is called **decoherence**.

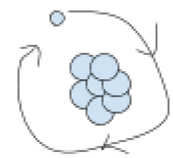
Decoherence is why the usual laws of probability look like they work in everyday life. A cat isn't found in a superposition of alive and dead states, because it interacts constantly with its environment. These interactions essentially leak information about the 'cat system' out.

Quantum superposition is something that happens to particles, or groups of particles, when they're isolated from their environments. Needing the particles to be isolated is why it's so hard to build a quantum computer. (And what if the particles aren't *perfectly* isolated, but merely *mostly* isolated? Great question! We'll come back to it later in the course.)

The story of atomic physics between roughly 1900 and 1926 is that scientists kept finding things that didn't fit with the usual laws of mechanics or probability. They usually came up with hacky solutions that explained a phenomenon without connecting it to much else. That is, until Heisenberg, Schrödinger, etc. came up with the general rules of quantum mechanics.

A traditional quantum physics class would go through the whole series of experiments by which physicists arrived at quantum mechanics, but we're just going to accept the rules as given and see what we can do from there.

Briefly, though, take the usual high school model of the electron, rotating around a nucleus in a fixed orbit. Scientists realized that this model would mean that the electron, as an accelerating electric charge, would be constantly losing energy and spiraling inwards until it hit the nucleus. To explain this, along with the double-slit experiment and countless other phenomena, physicists eventually had to change the way probabilities are calculated.



Instead of using probabilities $p \in [0, 1]$ they started using **amplitudes** $\alpha \in \mathbb{C}$. Amplitudes can be positive or negative, or more generally complex numbers (with real and imaginary part).

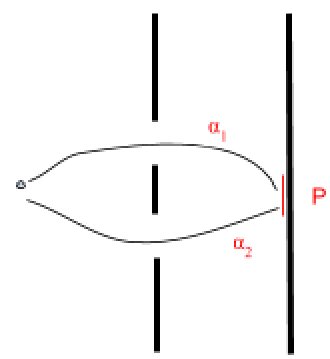
The central claim of quantum mechanics is that to fully describe the state of an isolated system, you need to give one amplitude for each possible configuration that you could find the system in on measuring it.

The **Born Rule** says that the probability you see a particular outcome is the squared absolute value of the amplitude:

$$\begin{aligned} P &= |\alpha|^2 \\ &= (\text{the real part of } \alpha)^2 + (\text{The imaginary part of } \alpha)^2 \end{aligned}$$

So let's see how amplitudes being complex leads them to act differently from probabilities. Let's revisit the Double Slit Experiment considering **interference**. We'll say that:

the total amplitude of a photon landing in a spot, α
is the amplitude of it getting there through the first slit, α_1
plus the amplitude of it getting there through the second slit, α_2



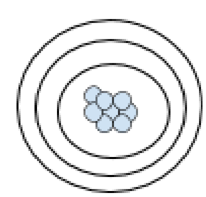
$$\begin{aligned} P &= |\alpha|^2 = |\alpha_1 + \alpha_2|^2 \\ &= |\alpha_1|^2 + |\alpha_2|^2 + \alpha_1^*\alpha_2 + \alpha_1\alpha_2^* \end{aligned}$$

If $\alpha_1 = \frac{1}{2}$ and $\alpha_2 = -\frac{1}{2}$, then interference means that if both slits are open $P = 0$, but if only one of them is open, $P = \frac{1}{4}$.

So then to justify the electron not spiraling into the nucleus:

We say that, yes, there are many paths where the electron does do that, but some have positive amplitudes and others have negative amplitudes and they end up canceling each other out.

With some physics we won't cover in this class, you'd discover that the possibilities where amplitudes don't cancel each other out lead to discrete energy levels, where are the places where the electrons can sit—this phenomenon being what leads to chemistry.



We use **Linear Algebra** to model states of systems as vectors and the evolution of systems in isolation as transformations of vectors.

$$M \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1' \\ a_2' \end{bmatrix}$$

For now, we'll consider classical probability. Let's look at flipping a coin:

$$\begin{array}{cl} \begin{bmatrix} p \\ q \end{bmatrix} & \begin{matrix} \text{tails} \\ \text{heads} \end{matrix} \end{array} \qquad \begin{matrix} p, q \geq 0 \\ p + q = 1 \end{matrix}$$

We model this with a vector listing both possibilities and assigning a probability to each.

We can apply a transformation, like turning the coin over.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} q \\ p \end{bmatrix}$$

Turning the coin over means the prob that the coin *was* heads is now the probability that the coin *is* tails. If it helps, you can think of the transformation matrix as:

$$\begin{bmatrix} P(tails|p) & P(tails|q) \\ P(heads|p) & P(heads|q) \end{bmatrix}$$

We could also flip the coin fairly.

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

Which means regardless of previous position, both possibilities are now equally likely.

Let's say we flip the coin, and if we get heads we flip again, but if we get tails we turn it to heads.

$$\begin{bmatrix} 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} \frac{q}{2} \\ p + \frac{q}{2} \end{bmatrix}$$

Does that make sense?

If we say that p, q are $P(tails)$ and $P(heads)$ after the first flip:

Then the probability the coin will land on tails in the end is:

0 if (it lands on tails on the first flip) and

$\frac{1}{2}$ if (it lands on heads and we flip again).

So we sum those values.

The probability that the coin will land on heads in the end is:

1 if (it lands on tails on the first flip) and

$\frac{1}{2}$ if (it lands on heads and we flip again).

So we sum those values.

So which matrices *can* be used as transformations?

Firstly, we know that <u>all entries have to be non-negative</u> (because classical probabilities can't be negative).

We also know that <u>all columns must sum to 1</u>, since we need the sum of initial probabilities to equal the sum of the transformed probabilities (namely, both should equal 1).

$$A \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \left[\text{i}^{\text{th}} \text{ column of A}\right]$$

We can see this clearly by using basis vectors.

A matrix that satisfies these conditions is called a **Stochastic Matrix**.

Now let's say we want to flip two coins, or rather, two bits. For the first coin $P(a) = P(\text{getting } 0)$, $P(b) = P(\text{getting } 1)$. For the second coin we'll use $P(c)$ and $P(d)$.

$$\begin{array}{cc} \begin{bmatrix} a \\ b \end{bmatrix} & \begin{array}{c} 0 \\ 1 \end{array} \end{array} \quad \begin{array}{cc} \begin{bmatrix} c \\ d \end{bmatrix} & \begin{array}{c} 0 \\ 1 \end{array} \end{array}$$

To combine the two vectors, we need a new operation, called **Tensor Product**.

$$\begin{bmatrix} a \\ b \end{bmatrix} \otimes \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} P_{00} \\ P_{01} \\ P_{10} \\ P_{11} \end{bmatrix} = \begin{bmatrix} ac \\ ad \\ bc \\ bd \end{bmatrix}$$

It's worth noting that not all possible 4-element vectors can arise by tensoring two 2-element vectors. For example:

$$\begin{bmatrix} ac \\ ad \\ bc \\ bd \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \\ \frac{1}{2} \end{bmatrix}$$

would mean that

$$\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) = (ac)(bd) = (ad)(bc) = (0)(0)$$

Therefore the right-hand side can't be a tensor product.

Let's say that if the first bit is 1, we want to flip the second bit:

$$\begin{array}{c|c} & \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array} & \begin{bmatrix} & & & \\ & & & \\ & & & \\ & & & \end{bmatrix} \end{array} \begin{array}{cc} \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{bmatrix} & \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array} \end{array}$$

We'd do:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \\ \frac{1}{2} \end{bmatrix}$$

This 4×4 matrix is called the **Controlled NOT** or **CNOT** matrix; it will also come up often in quantum computing.

Note that we've reached an output distribution that we previously proved can't arise as a tensor product! Such a distribution is called **correlated**: learning one bit tells you something about the other bit. (In this case, the two bits are always equal; with 50% probability they're both 0 and with 50% probability they're both 1.) So, we've learned that the CNOT matrix can *create correlations*: it can transform an uncorrelated distribution into a correlated one.

Quantum mechanics basically follows the same process to model states in quantum systems except that it uses amplitudes instead of probabilities.

$$\begin{bmatrix} & & \\ & U & \\ & & \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_1 \\ b_3 \end{bmatrix}$$

Where we preserve $\sum_{i=1}^{n} |a_i|^2 = 1 = \sum_{i=1}^{n} |b_i|^2$

and $|a_i|^2$ or $|b_i|^2$ represent the probability of measuring outcome *i*.