Notes 8.370/18.435 Fall 2022

Lecture 2 Prof. Peter Shor

Today, I'll start explaining the quantum mechanics you will need for this course. I'll start with a statement of what physicists call the *Superposition Principle* or the *Linearity Principle* of quantum mechanics. The following statement is expressed close to the way you'll see it physics literature.

The Superposition Principle
Suppose $|A\rangle$ *and* $|B\rangle$ *are two perfectly distinguishable quantum states of a quantum system, and* $\alpha, \beta \in \mathbb{C}$ *with* $|\alpha^2 + \beta^2| = 1$. *Then*

$$\alpha\,|A\rangle + \beta\,|B\rangle$$

is a valid state of the quantum system.

Here $|A\rangle$ and $|B\rangle$ are notation which physicists use for quantum states. It simply means that A and B can be described by complex column vectors.

What *perfectly distinguishable* means here is that there is some experiment that can distinguish between these two states, at least theoretically (all experiments will have some amount of noise in practice).

There is a second part of this principle—if the experiment distinguishing $|A\rangle$ from $|B\rangle$ is applied to the state $\alpha\,|A\rangle + \beta\,|B\rangle$, then the probability of getting the answer $|A\rangle$ is $|\alpha|^2$ and the probability of $|B\rangle$ is $|\beta^2|$.

How can we rewrite this principle using mathematics?

First, the fact that you can add states $\alpha\,|A\rangle + \beta\,|B\rangle$ means that the quantum state space is a complex vector space. What also follows, but is harder to see, is that a quantum state is a unit vector in the vector space. Two states are orthogonal if and only if they are perfectly distiguishable.

In this course, we will be talking about quantum computation. Let us recall that classical computers work on bits, which are systems that are in one of two states. Traditionally, we call the values of these bits 0 and 1, independent of the details of their physical representation. For example, on an integrated circuit, a higher voltage might represent a 1 and a low voltage might represent a 0. Or in a magnetic memory, a region where the magnetic north is oriented upward might be a 0 and where it's oriented downward might be a 1.

The same thing is true of qubits. A qubit is a quantum system that can only take on two distinguishable values. But these values might be the polarization of a photon, or the spin of an electron, or the ground state and an excited state of an ion, or a clockwise and a counterclockwise rotation of a current in a superconductor. We will give two examples in this lecture; the first is the polarization of a photon, and the second is the spin of a spin-$\frac{1}{2}$ particle (such as an electron, a proton, or a neutron). Polarization is probably the example most people are familiar with, so we start with that.

You may be familiar with polarized sunglasses. A polarizing filter only lets vertically polarized (or horizontally polarized) photons pass through. Because sun glare tends to be horizontally polarized, if you use vertical polarizing filters as lenses, they screen out most the glare, but let through most of the rest of the light, and so reduce glare. Of course, you can turn these sunglasses 45°, so they only let right diagonally polarized light through. How do we explain this in terms of the quantum state space?

It is the case that horizontal and vertical polarization are two distinguishable vectors. This means that other polarizations are linear combinations (called superpositions) of these.

Now, we need to introduce a little physics notation. Physicists like to represent iquantum states using a funny notation they call "kets", so a quantum state v would be represented as $|\,v\rangle$. In mathematics notations, kets are column vectors. So $|\leftrightarrow\rangle$ and $|\updownarrow\rangle$ are the horizontal and vertical polarizations of light.

Because we're doing quantum computation and working with qubits, we will represent these abstractly as $|\,0\rangle$ and $|\,1\rangle$, using the convention that $|\,0\rangle = |\leftrightarrow\rangle$ and $|\,1\rangle = |\updownarrow\rangle$. For example, if we have the vector

$$v = \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix},$$

then

$$|\,v\rangle = \frac{1}{\sqrt{2}}|\,0\rangle + \frac{i}{\sqrt{2}}|\,1\rangle = \frac{1}{\sqrt{2}}\left(|\leftrightarrow\rangle + i\,|\updownarrow\rangle\right)$$

Other polarizations are linear combinations (called superpositions) of $|\leftrightarrow\rangle$ and $|\updownarrow\rangle$. For example,

$$\begin{aligned} |\,\nearrow\!\!\!\!\swarrow\,\rangle &= \tfrac{1}{\sqrt{2}}\left(|\leftrightarrow\rangle + |\updownarrow\rangle\right) = \tfrac{1}{\sqrt{2}}\left(|\,0\rangle + |\,1\rangle\right), \\ |\,\nwarrow\!\!\!\!\searrow\,\rangle &= \tfrac{1}{\sqrt{2}}\left(|\leftrightarrow\rangle - |\updownarrow\rangle\right) = \tfrac{1}{\sqrt{2}}\left(|\,0\rangle - |\,1\rangle\right). \end{aligned}$$

We will be using the states $\frac{1}{\sqrt{2}}(|\,0\rangle + |\,1\rangle)$ and $\frac{1}{\sqrt{2}}(|\,0\rangle - |\,1\rangle)$ often enough that it helps to have special names for these; we call these $|\,+\rangle$ and $|\,-\rangle$ respectively.

Since the quantum state space is a complex vector space, we can also use complex coefficients in our superposition. What does this give us? It turns out that these give right and left circularly polarized light, $|\circlearrowright\rangle$ and $|\circlearrowleft\rangle$:

$$\begin{aligned} |\circlearrowright\rangle &= \tfrac{1}{\sqrt{2}}\left(|\leftrightarrow\rangle + i\,|\updownarrow\rangle\right), \\ |\circlearrowleft\rangle &= \tfrac{1}{\sqrt{2}}\left(|\leftrightarrow\rangle - i\,|\updownarrow\rangle\right). \end{aligned}$$

And if you use unequal coefficients in the above equations, you get elliptically polarized light.

Also note that the choice of whether $\frac{1}{\sqrt{2}}(|\leftrightarrow\rangle + |\updownarrow\rangle)$ is $|\,\nearrow\!\!\!\!\swarrow\,\rangle$ or $|\,\nwarrow\!\!\!\!\searrow\,\rangle$ is purely a convention; I'll try to stick with the conventions that the textbook uses in these notes.

There is another half of the ket notation: a *row* vector is represented by a *bra:* $\langle\cdot\,|$. By convention, $\langle v\,|$ is the conjugate transpose of $|\,v\rangle$, so if

$$|\,v\rangle \ \text{ is the column vector} \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix},$$

then

$$\langle v\,| \ \text{ is the row vector} \left(1/\sqrt{2}, -i/\sqrt{2}\right),$$

Together, the "bracket" $\langle v|w\rangle$ is the inner product of $\langle v\,|$ and $|\,w\rangle$. Since quantum states are unit vectors, and going from the ket to the bra involves a complex conjugation, we get

$$\langle\psi|\psi\rangle = 1$$

for any quantum state $|\,\psi\rangle$. The notation $\langle\phi|\psi\rangle$ represents the inner product of two states, and is called a *bracket* (and here you see the wordplay that Dirac used to come up with "bra" and "ket".

How can we change a quantum state? We will be doing quantum computation, so we need to know how to change the state of our quantum computer. Physics says that the evolution of an isolated quantum system is linear, and if we apply a linear transformation to a vector, we have a matrix.

Suppose we have a matrix M that operates on a quantum state $|\,\Psi\rangle$. If this transformation is something that can be implemented physically, the result of applying M to a quantum state $|\,\psi\rangle$, which is $M\,|\,\psi\rangle$, has to be a unit vector, so we need

$$\langle\psi\,|\,M^\dagger M\,|\,\psi\rangle = 1 \quad \text{for all quantum states } |\,\psi\rangle.$$

This implies that $M^\dagger M = I$, the identity matrix. This is the condition for M being a *unitary* matrix. Unitary matrices are essentially rotation matrices in complex space.

There are other conditions for a matrix U being unitary besides $U^\dagger U = I$. A matrix is unitary if and only if any of the following conditions hold:

- $U^\dagger U = I$,
- $UU^\dagger = I$,
- the columns of U are orthonormal vectors,
- the rows of U are orthonormal vectors.

There are three unitary matrices that operate on qubits (and thus are 2×2 which are very useful for quantum computing (and also for quantum physics in general). These are

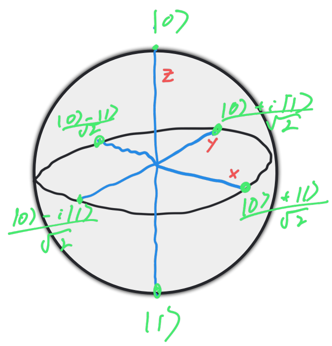
$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

These are called the *Pauli matrices*. A silly mnemonic for remembering which entry has $-i$ in σ_y is that $-i$ is lighter than i, and thus floats to the top.

How do these matrices act on a qubit. Why are they labeled x, y, and z? To answer this, we will use a different representation of a qubit, namely, the spin of a spin $\frac{1}{2}$ particle. A spin $\frac{1}{2}$ particle, such as an electron, only has two completely distinguishable states of its spin, which we will take to be up and down. Of course, because of three-dimensional symmetry, you can prepare a spin $\frac{1}{2}$ particle with its spin axis pointing in any direction. However, these other directions are linear combinations of spin up and spin down. Namely, using the most common convention, Let us take up and down to be the z-axis, and left and right to be the x-axis, so the y-axis goes into and out of the board. Then we will let $|\,0\rangle = |\uparrow\rangle$ and $|\,1\rangle = |\downarrow\rangle$. The standard convention is:

$$
\begin{aligned}
|\rightarrow\rangle &= \tfrac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right), \\
|\leftarrow\rangle &= \tfrac{1}{\sqrt{2}}\left(|\uparrow\rangle - |\downarrow\rangle\right), \\
|\,\mathrm{in}\rangle &= \tfrac{1}{\sqrt{2}}\left(|\uparrow\rangle + i\,|\downarrow\rangle\right), \\
|\,\mathrm{out}\rangle &= \tfrac{1}{\sqrt{2}}\left(|\uparrow\rangle - i\,|\downarrow\rangle\right),
\end{aligned}
$$

We can represent these states on the Bloch sphere (See Figure.) Here, each point on the Bloch sphere represents a quantum state of a spin $\frac{1}{2}$ particle that that has spin in the corresponding direction.



Bloch Sphere

What happens when we apply $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$? One can see that it takes the vector $|\,0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to $|\,1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and vice versa. It also takes the vector $|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ to itself, and $|-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ to $\frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = -|-\rangle$. Except for the fact that $|-\rangle$ is taken to $-|-\rangle$, this is a 180° rotation around the x-axis.

In fact, in some sense this really does represent a 180° rotation around the x-axis. This is because multiplying like a unit complex vector like -1, i, or $e^{i\theta}$ does not change the essential physical state — two quantum states $|\,\psi\rangle$ and $e^{i\theta}\,|\,\psi\rangle$ are indistinguishable by any quantum measurement. We will see this when we get to measurements next week. So the matrix σ_x represents a 180° rotation of the Bloch sphere around the x axis. Similarly, the matrices σ_y and σ_z represent rotations of the Bloch sphere around the y- and z-axes, respectively.

We will come back to the Bloch sphere later and tell you how to find the quantum state pointing in an arbitrary direction along the Bloch sphere, as well as the unitary matrix corresponding to an arbitrary rotation of the Bloch sphere.