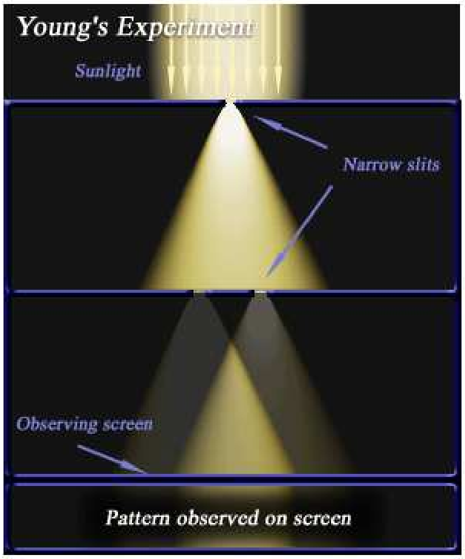
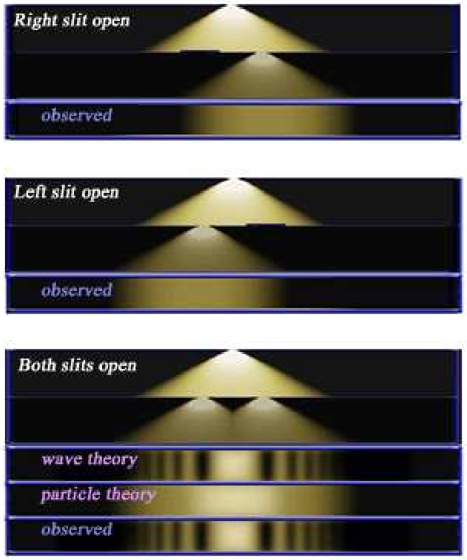
Lecture 1: Axioms of QM + Bell Inequalities

0.1 Young's double-slit experiment

Is light transmitted by particles or waves? The basic dilemna here (which dates as far back as Newton) is to reconcile the evidence that light is transmitted by particles (called photons), with experiments demonstrating the wave nature of light. To be concrete, let us recall Young's double-slit experiment from high school physics, which was used to demonstrate the wave nature of light. The apparatus consists of a source of light, an intermediate screen with two very thin identical slits, and a viewing screen (see picture on next page). If only one slit is open then intensity of light on the viewing screen is maximum on the straight line path and falls off in either direction. However, if both slits are open, then the intensity oscillates according to the familiar interference pattern predicted by wave theory. These facts can be qualitatively and quantitatively explained by positing that light travels in waves (as you did in high school physics).





Let us now introduce the particle nature of light into this experiment. To do so, we turn down the intensity of the light source, until a photodetector clicks only occassionally to record the emission of a photon. As we turn down the intensity of the source, the magnitude of each click remains constant, but the time between successive clicks increases. This is consistent with light being emitted as discrete particles (photons) — the intensity of light is proportional to the rate at which photons are emitted by the source. So now with the light source emitting a single photon every so often, we can ask where this single emitted photon hits the viewing screen. The answer is no longer deterministic, but probabilistic. We can only speak about the probability that a photodetector placed at point x detects the photon. If only a single slit is open, then plotting this probability of detection as a function of x gives the same curve as the intensity as a function of x in the classical Young experiment. What happens when both slits are open? Our intuition would strongly suggest that the probability we detect the photon at x should simply be the sum of the probability of detecting it at x if only slit 1 were open and the probability if only slit 2 were open. In other words the outcome should no longer be consistent with the interference pattern. In the actual experiment, the probability of detection does still follow the interference pattern. Reconciling this outcome with the particle nature of light appears impossible, and that is the dilemna we face.

Let us spell out in more detail why this contradicts our intuition: for the photon to be detected at x, either it went through slit 1 and ended up at x or it went through slit 2 and ended up at x. Now the probability of seeing the photon at x should be the sum of the probabilities in the two cases. To make the contradiction seem even more stark, notice that there are points x where the detection probability is zero (or small) if both slits are open, even though it is non-zero (large) if either slit is open. How can the existence of more ways for an event to happen actually decrease its probability?

Let us now turn to quantum mechanics to see how it explains this phenomenon.

Quantum mechanics introduces the notion of the complex amplitude $\psi_1(x) \in \mathscr{C}$ with which the photon goes through slit 1 and hits point x on the viewing screen. The probability that the photon is actually detected at x is the square of the magnitude of this complex number: $P_1(x) = |\psi_1(x)|^2$. Similarly, let $\psi_2(x)$ be the amplitude if only slit 2 is open. $P_2(x) = |\psi_2(x)|^2$.

Now when both slits are open, the amplitude with which the photon hits point x on the screen is just the sum of the amplitudes over the two ways of getting there: $\psi_{12}(x) = \psi_1(x) + \psi_2(x)$. As before the probability that the photon is detected at x is the squared magnitude of this amplitude: $P_{12}(x) = |\psi_1(x) + \psi_2(x)|^2$. The two complex numbers $\psi_1(x)$ and $\psi_2(x)$ can cancel each other out to produce destructive interference, or reinforce each other to produce constructive interference or anything in between.

Some of you might find this ”explanation” quite dissatisfying. You might say it is not an explanation at all. Well, if you wish to understand how Nature behaves you have to reconcile yourselves to this type of explanation — this wierd way of thinking has been successful at describing (and understanding) a vast range of physical phenomena. But you might persist and (quite reasonably) ask “but how does a particle that went through the first slit know that the other slit is open”? In quantum mechanics, this question is not well-posed. Particles do not have trajectories, but rather take all paths simultaneously (in superposition). As we shall see, this is one of the key features of quantum mechanics that gives rise to its paradoxical properties as well as provides the basis for the power of quantum computation. To quote Feynman, 1985, ”The more you see how strangely Nature behaves, the harder it is to make a model that explains how even the simplest phenomena actually work. So theoretical physics has given up on that.”

0.2 Basic Quantum Mechanics

The basic formalism of quantum mechanics is very simple, though understanding and interpreting (and accepting) the results is much more challenging. There are three basic principles, enshrined in the four basic postulates of quantum mechanics:

- The superpostion principle: this axiom tells us what are the allowable (possible) states of a given quantum system.
- The measurement principle: this axiom governs how much information about the state we can access.
- Unitary evolution: this axoim governs how the state of the quantum system evolves in time.

0.3 The superposition principle

Consider a system with k distinguishable (classical) states. For example, the electron in a hydrogen atom is only allowed to be in one of a discrete set of energy levels, starting with the ground state, the first excited state, the second excited state, and so on. If we assume a suitable upper bound on the total energy, then the electron is restricted to being in one of k different energy levels — the ground state or one of $k-1$ excited states. As a classical system, we might use the state of this system to store a number between 0 and $k-1$. The superposition principle says that if a quantum system can be in one of two states then it can also be placed in a linear superposition of these states with complex coefficients.

Let us introduce some notation. We denote the ground state of our qubit by $|0\rangle$, and the succesive excited states by $|1\rangle, \dots, |k-1\rangle$. These are the k possible classical states of the electron. The superposition principle tells us that, in general, the (quantum) state of the electron is $\alpha_0|0\rangle + \alpha_1|1\rangle + \cdots + \alpha_{k-1}|k-1\rangle$, where $\alpha_0, \alpha_1, \dots, \alpha_{k-1}$ are complex numbers normalized so that $\sum_j |\alpha_j|^2 = 1$. α_j is called the *amplitude of the state* $|j\rangle$.

For instance, if $k = 3$, the state of the electron could be
$1/\sqrt{2}|0\rangle + 1/2|1\rangle + 1/2|2\rangle$ or
$1/\sqrt{2}|0\rangle - 1/2|1\rangle + i/2|2\rangle$ or
$(1+i)/3|0\rangle - (1-i)/3|1\rangle + (1+2i)/3|2\rangle$, where $i = \sqrt{-1}$.

The superposition principle is one of the most mysterious aspects about quantum physics — it flies in the face of our intuitions about the physical world. One way to think about a superposition is that the electron does not make up its mind about whether it is in the ground state or each of the $k-1$ excited states, and the amplitude α_0 is a measure of its inclination towards the ground state. Of course we cannot think of α_0 as the probability that an electron is in the ground state — remember that α_0 can be negative or imaginary. The measurement priniciple, which we will see shortly, will make this interpretation of α_0 more precise.

0.4 The Geometry of Hilbert Space

We saw above that the quantum state of the k-state system is described by a sequence of k complex numbers $\alpha_0, \dots, \alpha_{k-1} \in \mathscr{C}$, normalized so that $\sum_j |\alpha_j|^2 = 1$. So it is natural to write the state of the system as a k dimensional vector:

$$\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ . \\ . \\ \alpha_{k-1} \end{pmatrix}$$

The normalization on the complex amplitudes means that the state of the system is a unit vector in a k dimensional complex vector space — called a Hilbert space.

But hold on! Earlier we wrote the quantum state in a very different (and simpler) way as: $\alpha_0|0\rangle + \alpha_1|1\rangle + \cdots + \alpha_{k-1}|k-1\rangle$. Actually this notation, called Dirac's ket notation, is just another way of writing a vector. Thus $|0\rangle = \begin{pmatrix} 1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix}$ and $|k-1\rangle = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ 1 \end{pmatrix}$.

So we have an underlying geometry to the possible states of a quantum system: the k distinguishable (classical) states $|0\rangle, \dots, |k-1\rangle$ are represented by mutually orthogonal unit vectors in a k-dimensional complex vector space. i.e. they form an orthonormal basis for that space (called the standard basis). Moreover, given any two states, $\alpha_0|0\rangle + \alpha_1|1\rangle + \cdots + \alpha_{k-1}|k-1\rangle$, and $\beta|0\rangle + \beta|1\rangle + \cdots + \beta k-1|k-1\rangle$, we can compute the inner product of these two vectors, which is $\sum_{j=0}^{k-1} \bar{\alpha}_j \beta_j$. The absolute value of the inner product is the cosine of the angle between these two vectors. You should verify that the inner product of any two vectors in the standard basis is 0, showing that they are orthogonal.

The advantage of the ket notation is that the it labels the basis vectors explicitly. This is very convenient because the notation expresses both that the state of the quantum system is a vector, while at the same time explicitly writing out the physical quantity of interest (energy level, position, spin, polarization, etc).

0.5 Measurement Principle

This linear superposition $\left|\psi\right\rangle = \sum_{j=0}^{k-1} \alpha_j \left|j\right\rangle$ is part of the private world of the electron. Access to the information describing this state is severely limited — in particular, we cannot actually measure the complex amplitudes α_j. This is not just a practical limitation; it is enshrined in the measurement postulate of quantum physics.

A measurement on this k state system yields one of at most k possible outcomes: i.e. an integer between 0 and $k-1$. Measuring $\left|\psi\right\rangle$ in the standard basis yields j with probability $|\alpha_j|^2$.

One important aspect of the measurement process is that it alters the state of the quantum system: the effect of the measurement is that the new state is exactly the outcome of the measurement. I.e., if the outcome of the measurement is j, then following the measurement, the qubit is in state $\left|j\right\rangle$. This implies that you cannot collect any additional information about the amplitudes α_j by repeating the measurement.

Intuitively, a measurement provides the only way of reaching into the Hilbert space to probe the quantum state vector. In general this is done by selecting an orthonormal basis $\left|e_0\right\rangle, \ldots, \left|e_{k-1}\right\rangle$. The outcome of the measurement is $\left|e_j\right\rangle$ with probability equal to the square of the length of the projection of the state vector ψ on $\left|e_j\right\rangle$. A consequence of performing the measurement is that the new state vector is $\left|e_j\right\rangle$. Thus measurement may be regarded as a probabilistic rule for projecting the state vector onto one of the vectors of the orthonormal measurement basis.

Some of you might be puzzled about how a measurement is carried out physically? We will get to that soon when we give more explicit examples of quantum systems.

0.6 Qubits

Qubits (pronounced “cue-bit”) or quantum bits are basic building blocks that encompass all fundamental quantum phenomena. They provide a mathematically simple framework in which to introduce the basic concepts of quantum physics. Qubits are 2-state quantum systems. For example, if we set $k = 2$, the electron in the Hydrogen atom can be in the ground state or the first excited state, or any superposition of the two. We shall see more examples of qubits soon.

The state of a qubit can be written as a unit (column) vector $\binom{\alpha}{\beta} \in \mathscr{C}^2$. In Dirac notation, this may be written as:

$$\left|\psi\right\rangle = \alpha\left|0\right\rangle + \beta\left|1\right\rangle \quad \alpha, \beta \in \mathscr{C} \quad \text{and} \quad |\alpha|^2 + |\beta|^2 = 1$$

This linear superposition $\left|\psi\right\rangle = \alpha\left|0\right\rangle + \beta\left|1\right\rangle$ is part of the private world of the electron. For us to know the electron’s state, we must make a measurement. Making a measurement gives us a single classical bit of information — 0 or 1. The simplest measurement is in the standard basis, and measuring $\left|\psi\right\rangle$ in this $\{\left|0\right\rangle, \left|1\right\rangle\}$ basis yields 0 with probability $|\alpha|^2$, and 1 with probability $|\beta|^2$.

One important aspect of the measurement process is that it alters the state of the qubit: the effect of the measurement is that the new state is exactly the outcome of the measurement. I.e., if the outcome of the measurement of$\left|\psi\right\rangle = \alpha\left|0\right\rangle + \beta\left|1\right\rangle$ yields 0, then following the measurement, the qubit is in state $\left|0\right\rangle$. This implies that you cannot collect any additional information about α, β by repeating the measurement.

More generally, we may choose any orthogonal basis $\{\left|v\right\rangle, \left|w\right\rangle\}$ and measure the qubit in that basis. To do this, we rewrite our state in that basis: $\left|\psi\right\rangle = \alpha'\left|v\right\rangle + \beta'\left|w\right\rangle$. The outcome is v with probability $|\alpha'|^2$, and $\left|w\right\rangle$ with probability $|\beta'|^2$. If the outcome of the measurement on $\left|\psi\right\rangle$ yields $\left|v\right\rangle$, then as before, the the qubit is then in state $\left|v\right\rangle$.

0.7 Examples of Qubits

Photon Polarization:

A photon can be described as a traveling electromagnetic wave where the electric field oscillates along an axis that is oriented perpendicular to the photon's forward motion. The orientation of this axis is the ”polarization” of the photon. So, for a given direction of photon motion, the photon's polarization axis might lie anywhere in a 2-d plane perpendicular to that motion. It is thus natural to pick an orthonormal 2-d basis (such as $\vec{x}$ and $\vec{y}$, or ”vertical” and ”horizontal”) to describe the polarization state (i.e. polarization direction) of a photon. In a quantum mechanical description, this 2-d nature of the photon polarization is represented by a qubit, where the amplitude of the overall polarization state in each basis vector is just the projection of the polarization in that direction.

The polarization of a photon can be measured by using a polaroid or a calcite crystal. A suitably oriented polaroid sheet transmits x-polarized photons and absorbs y-polarized photons. Thus a photon that is in a superposition $|\phi\rangle = \alpha|x\rangle + \beta|y\rangle$ is transmitted with probability $|\alpha|^2$. If the photon now encounters another polariod sheet with the same orientation, then it is transmitted with probability 1. On the other hand, if the second polaroid sheet has its axes crossed at right angles to the first one, then if the photon is transmitted by the first polaroid, then it is definitely absorbed by the second sheet. This pair of polarized sheets at right angles thus blocks all the light. An somewhat counter-intuitive result is now obtained by interposing a third polariod sheet at a 45 degree angle between the first two. Now a photon that is transmitted by the first sheet makes it through the next two with probability $1/4$.

To see this first observe that any photon transmitted through the first filter is $|0\rangle$. The probability this photon is transmitted through the second filter is $1/2$ since it is exactly the probability that a qubit in the state $|0\rangle$ ends up in the state $|+\rangle$ when measured in the $|+\rangle, |-\rangle$ basis. We can repeat this reasoning for the third filter, except now we have a qubit in state $|+\rangle$ being measured in the $|0\rangle, |1\rangle$-basis — the chance that the outcome is $|0\rangle$ is once again $1/2$.

Spins:

Like photon polarization, the spin of a (spin-1/2) particle is a two-state system, and can be described by a qubit. Very roughly speaking, the spin is a quantum description of the magnetic moment of an electron which behaves like a spinning charge. We will say much more about the spin of an elementary particle later in the course.

0.7.1 Measurement example I: phase estimation

Consider the quantum state $|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{e^{i\theta}}{\sqrt{2}}|1\rangle)$. If we were to measure this qubit in the standard basis, the outcome would be 0 with probability $1/2$ and 1 with probability $1/2$. Is there any measurement that yields information about the phase θ?

Let us consider a measurement in a different basis: the $\{|+\rangle, |-\rangle\}$-basis. Here $|+\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and $|-\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. What does $|\phi\rangle$ look like in this new basis? This can be expressed by first writing $|0\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$ and $|1\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)$.

Now,

$$\begin{aligned}
\big|\psi\big\rangle &= \tfrac{1}{\sqrt{2}}\big|0\big\rangle + \tfrac{e^{i\theta}}{\sqrt{2}}\big|1\big\rangle) \\
&= \tfrac{1}{2}(\big|+\big\rangle + \big|-\big\rangle) + \tfrac{e^{i\theta}}{2}(\big|+\big\rangle - \big|-\big\rangle) \\
&= \tfrac{1+e^{i\theta}}{2}\big|+\big\rangle + \tfrac{(1-e^{i\theta})}{2}\big|-\big\rangle \quad .
\end{aligned}$$

Writing $e^{i\theta} = cos\theta + isin\theta$, we see that the probability of measuring $\big|+\big\rangle$ is $\frac{1}{4}((1+cos\theta)^2 + sin^2\theta)$. This can be further simplified to $\frac{1}{2}(1+cos\theta) = cos^2\frac{\theta}{2}$. A similar calculation reveals that the probability of measuring $\big|-\big\rangle$ is $sin^2\frac{\theta}{2}$. Measuring in the $\{\big|+\big\rangle, \big|-\big\rangle\}$-basis therefore reveals some information about the phase θ.

In §0.9 we shall show how to analyze the measurement of a qubit in a general basis.

0.8 Bra-ket notation.

The notation $\langle v|$ (pronounced "bra v") denotes the row vector $|v\rangle^\dagger$, the conjugate-transpose of $|v\rangle$. For example, $\langle 0| = (\,1\;0\,)$ and $\langle 1| = (\,0\;1\,)$. More generally, if $\big|\psi\big\rangle = \alpha\big|0\big\rangle + \beta\big|1\big\rangle$, then

$$\langle\psi| = \binom{\alpha}{\beta}^\dagger = (\,\bar{\alpha}\;\bar{\beta}\,) = \bar{\alpha}\langle 0| + \bar{\beta}\langle 1| \quad . \tag{1}$$

Thus, for instance, if $|\psi\rangle = i/\sqrt{2}\big|0\big\rangle + (1+i)/2\big|1\big\rangle$, then $\langle\psi| = -i/\sqrt{2}\langle 0| + (1-i)/2\langle 1|$.

Let

$$\big|v\big\rangle = a_0\big|0\big\rangle + a_1\big|1\big\rangle, \quad \big|w\big\rangle = b_0\big|0\big\rangle + b_1\big|1\big\rangle \quad . \tag{2}$$

Then $\big\langle v\big|w\big\rangle$ (shorthand for $\big\langle v\big|\,\big|w\big\rangle$) is the inner product between $\big|v\big\rangle$ and $\big|w\big\rangle$. It is a matrix product of the 1×2 matrix $\big\langle v\big|$ and the 2×1 matrix $\big|w\big\rangle$:

$$\big\langle v\big|w\big\rangle = (\,\bar{a}_0\;\bar{a}_1\,)\begin{pmatrix} b_0 \\ b_1 \end{pmatrix} = \bar{a}_0 b_0 + \bar{a}_1 b_1 \quad . \tag{3}$$

In the next lecture, we will introduce tensor product spaces, where the advantages of this notation increase.

0.9 Measurement example II.

What is the result of measuring a general qubit state $\big|\psi\big\rangle = \alpha\big|0\big\rangle + \beta\big|1\big\rangle$, in a general orthonormal basis $\big|v\big\rangle, \big|v^\perp\big\rangle$, where $|v\rangle = a|0\rangle + b|1\rangle$ and $|v^\perp\rangle = \overline{b}|0\rangle - \overline{a}|1\rangle$. (check that $\big|v\big\rangle$ and $\big|v^\perp\big\rangle$ are orthogonal by computing $\langle v|v^\perp\rangle 0$).

To answer this question, let us make use of our recently acquired bra-ket notation. Let us start by rewriting $\big|\psi\big\rangle$ in the $\big|v\big\rangle, \big|v^\perp\big\rangle$-basis.

$$\begin{aligned}
|\psi\rangle &= I\big|\psi\big\rangle \quad = \\
\Big(|v\rangle\langle v| + |v^\perp\rangle\langle v^\perp|\Big)(\alpha|0\rangle + \beta|1\rangle) & \\
&= \alpha(|v\rangle\langle v|0\rangle + |v^\perp\rangle\langle v^\perp|0\rangle) + \beta(|v\rangle\langle v|1\rangle + |v^\perp\rangle\langle v^\perp|1\rangle) \\
&= (\alpha\langle v|0\rangle + \beta\langle v|1\rangle)|v\rangle + (\alpha\langle v^\perp|0\rangle + \beta\langle v^\perp|1\rangle)|v^\perp\rangle \\
&= (\alpha\overline{a} + \beta\overline{b})|v\rangle + (\alpha b + \beta a)|v^\perp\rangle \quad .
\end{aligned}$$

The probability of measuring $|v\rangle$ in a measurement in the $v, v^{\perp}$ basis is therefore

$$|\langle v|\psi\rangle|^2 = |\alpha\bar{a} + \beta\bar{b}|^2 \ \ .$$

0.10 Unitary Operators

The third postulate of quantum physics states that the evolution of a quantum system is necessarily unitary. Intuitively, a unitary transformation is a rigid body rotation (or reflection) of the Hilbert space, thus resulting in a transformation of the state vector that doesn't change its length.

Let us consider what this means for the evolution of a qubit. A unitary transformation on the Hilbert space $\mathscr{C}^2$ is specified by mapping the basis states $|0\rangle$ and $|1\rangle$ to orthonormal states $|v_0\rangle = a|0\rangle + b|1\rangle$ and $|v_1\rangle = c|0\rangle + d|1\rangle$. It is specified by the linear transformation on $\mathscr{C}^2$:

$$U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$$

If we denote by $U^{\dagger}$ the conjugate transpose of this matrix:

$$U^{\dagger} = \left(\begin{smallmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{smallmatrix}\right) \ \ .$$

then it is easily verified that $UU^{\dagger} = U^{\dagger}U = I$. Indeed, we can turn this around and say that a linear transformation U is unitary iff it satisfies this condition, that $UU^{\dagger} = U^{\dagger}U = I$.

Let us now consider some examples of unitary transformations on single qubits or equivalently single qubit quantum gates:

- Hadamard Gate. Can be viewed as a reflection around $\pi/8$, or a rotation around $\pi/4$ followed by a reflection.

  $$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

  The Hadamard Gate is one of the most important gates. Note that $H^{\dagger} = H$ – since H is real and symmetric – and $H^2 = I$.

- Rotation Gate. This rotates the plane by θ.

  $$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

- NOT Gate. This flips a bit from 0 to 1 and vice versa.

  $$NOT = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

- Phase Flip.

  $$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

  The phase flip is a NOT gate acting in the $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ basis. Indeed, $Z|+\rangle = |-\rangle$ and $Z|-\rangle = |+\rangle$.

How do we physically effect such a (unitary) transformation on a quantum system? To explain this we must first introduce the notion of the Hamiltonian acting on a system; you will have to wait for three to four lectures before we get to those concepts.

1 Two qubits:

Now let us examine a system of two qubits. Consider the two electrons in two hydrogen atoms, each regarded as a 2-state quantum system:

Since each electron can be in either of the ground or excited state, classically the two electrons are in one of four states – 00, 01, 10, or 11 – and represent 2 bits of classical information. By the superposition principle, the quantum state of the two electrons can be any linear combination of these four classical states:

$$\left|\psi\right\rangle = \alpha_{00}\left|00\right\rangle + \alpha_{01}\left|01\right\rangle + \alpha_{10}\left|10\right\rangle + \alpha_{11}\left|11\right\rangle \ ,$$

where $\alpha_{ij} \in \mathscr{C}$, $\sum_{ij} |\alpha_{ij}|^2 = 1$. Again, this is just Dirac notation for the unit vector in $\mathscr{C}^4$:

$$\begin{pmatrix} \alpha_{00} \\ \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{pmatrix}$$

Measurement:

Measuring $\left|\psi\right\rangle$ now reveals two bits of information. The probability that the outcome of the measurement is the two bit string $x \in \{0,1\}^2$ is $|\alpha_x|^2$. Moreover, following the measurement the state of the two qubits is $\left|x\right\rangle$. i.e. if the first bit of x is j and the second bit k, then following the measurement, the state of the first qubit is $\left|j\right\rangle$ and the state of the second is $\left|k\right\rangle$.

An interesting question comes up here: what if we measure just the first qubit? What is the probability that the outcome is 0? This is simple. It is exactly the same as it would have been if we had measured both qubits: $\Pr\{\text{1st bit } = 0\} = \Pr\{00\} + \Pr\{01\} = \left|\alpha_{00}\right|^2 + \left|\alpha_{01}\right|^2$. Ok, but how does this partial measurement disturb the state of the system?

The answer is obtained by an elegant generalization of our previous rule for obtaining the new state after a measurement. The new superposition is obtained by crossing out all those terms of $\left|\psi\right\rangle$ that are inconstent with the outcome of the measurement (i.e. those whose first bit is 1). Of course, the sum of the squared amplitudes is no longer 1, so we must renormalize to obtain a unit vector:

$$\left|\phi\right\rangle_{new} = \frac{\alpha_{00}\left|00\right\rangle + \alpha_{01}\left|01\right\rangle}{\sqrt{\left|\alpha_{00}\right|^2 + \left|\alpha_{01}\right|^2}}$$

.

Entanglement

Suppose the first qubit is in the state $3/5\left|0\right\rangle + 4/5\left|1\right\rangle$ and the second qubit is in the state $1/\sqrt{2}\left|0\right\rangle - 1/\sqrt{2}\left|1\right\rangle$, then the joint state of the two qubits is $(3/5\left|0\right\rangle + 4/5\left|1\right\rangle)(1/\sqrt{2}\left|0\right\rangle - 1/\sqrt{2}\left|1\right\rangle) = 3/5\sqrt{2}\left|00\right\rangle - 3/5\sqrt{2}\left|01\right\rangle + 4/5\sqrt{2}\left|10\right\rangle - 4/5\sqrt{2}\left|11\right\rangle$

But there are states such as $\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}\left(\left|00\right\rangle + \left|11\right\rangle\right)$ which cannot be decomposed in this way as a state of the first qubit and that of the second qubit. Can you see why? Such a state is called an entangled state.

If the first (resp. second) qubit of $\left|\Phi^+\right\rangle$ is measured then the outcome is 0 with probability $1/2$ and 1 with probability $1/2$. However if the outcome is 0, then a measurement of the second qubit results in 0 with certainty. Furthermore this is true even if both qubits are measured in a rotated basis $\left|v\right\rangle, \left|v^\perp\right\rangle$, where $\left|v\right\rangle = \alpha\left|0\right\rangle + \beta\left|1\right\rangle$ and $\left|v^\perp\right\rangle = -\beta\left|0\right\rangle + \alpha\left|1\right\rangle$.

Claim: $\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}\left(\left|00\right\rangle + \left|11\right\rangle\right)$
$= \frac{1}{\sqrt{2}}\left(\left|vv\right\rangle + \left|v^\perp v^\perp\right\rangle\right).$

Proof: Then $\frac{1}{\sqrt{2}}(\left|vv\right\rangle + \left|v^\perp v^\perp\right\rangle)$
$= \frac{1}{\sqrt{2}}(\alpha^2\left|00\right\rangle + \alpha\beta\left|01\right\rangle + \alpha\beta\left|10\right\rangle + \beta^2\left|11\right\rangle) + \frac{1}{\sqrt{2}}(\beta^2\left|00\right\rangle - \alpha\beta\left|01\right\rangle - \alpha\beta\left|10\right\rangle + \alpha^2\left|11\right\rangle)$
$= \frac{1}{\sqrt{2}}(\alpha^2+\beta^2)(\left|00\right\rangle + \left|11\right\rangle)$
$= \frac{1}{\sqrt{2}}(\left|00\right\rangle + \left|11\right\rangle)$

1.1 Two Qubit Gates

Let us now consider how a system of two qubits evolves in time. Recall that the third axiom of quantum physics states that the evolution of a quantum system is necessarily unitary. Intuitively, a unitary transformation is a rigid body rotation (or reflection) of the Hilbert space, thus resulting in a transformation of the state vector that doesn't change its length.

Let us consider what this means for the evolution of a two qubit system. A unitary transformation on the Hilbert space $\mathscr{C}^4$ is specified by a $4x4$ matrix U that satisfies the condition $UU^\dagger = U^\dagger U = I$. The four columns of U specify the four orthonormal vectors $\left|v_{00}\right\rangle$, $\left|v_{01}\right\rangle$, $\left|v_{10}\right\rangle$ and $\left|v_{11}\right\rangle$ that the basis states $\left|00\right\rangle$, $\left|01\right\rangle$, $\left|10\right\rangle$ and $\left|11\right\rangle$ are mapped to by U.

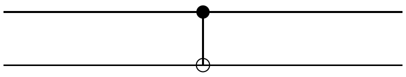
A very basic two qubit gate is the controlled-not gate or the CNOT:

- Controlled Not (CNOT).

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The first bit of a CNOT gate is the "control bit;" the second is the "target bit." The control bit never changes, while the target bit flips if and only if the control bit is 1.

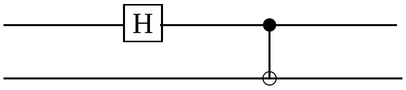
The CNOT gate is usually drawn as follows, with the control bit on top and the target bit on the bottom:



Though the CNOT gate looks very simple, any unitary transformation on two qubits can be closely approximated by a sequence of CNOT gates and single qubit gates. This brings us to an important point. What happens to the quantum state of two qubits when we apply a single qubit gate to one of them, say the first? Let's do an example. Suppose we apply a Hadamard gate to the superposition: $\left|\psi\right\rangle = 1/2\left|00\right\rangle - i/\sqrt{2}\left|01\right\rangle + 1/\sqrt{2}\left|11\right\rangle$. Then this maps the first qubit as follows: $\left|0\right\rangle \rightarrow 1/\sqrt{2}\left|0\right\rangle + 1/\sqrt{2}\left|1\right\rangle$, and
$\left|1\right\rangle \rightarrow 1/\sqrt{2}\left|0\right\rangle - 1/\sqrt{2}\left|1\right\rangle$.
So $\left|\psi\right\rangle \rightarrow 1/2\sqrt{2}\left|00\right\rangle + 1/2\sqrt{2}\left|01\right\rangle - i/2\left|00\right\rangle + i/2\left|01\right\rangle + 1/2\left|10\right\rangle - 1/2\left|11\right\rangle$
$= (1/2\sqrt{2} - i/2)\left|00\right\rangle + (1/2\sqrt{2} + i/2)\left|01\right\rangle + 1/2\left|10\right\rangle - 1/2\left|11\right\rangle$.

Bell states:

We can generate the Bell states $\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}\left(\left|00\right\rangle + \left|11\right\rangle\right)$ with the following simple qauntum circuit consisting of a Hadamard and CNOT gate:



The first qubit is passed through a Hadamard gate and then both qubits are entangled by a CNOT gate.

If the input to the system is $|0\rangle \otimes |0\rangle$, then the Hadamard gate changes the state to

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle = \frac{1}{sqrt2}|00\rangle + \frac{1}{\sqrt{2}}|10\rangle \ \ ,$$

and after the CNOT gate the state becomes $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, the Bell state $|\Phi^+\rangle$.

The state $\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}\left(\left|00\right\rangle + \left|11\right\rangle\right)$ is one of four Bell basis states:

$$\begin{aligned}\left|\Phi^\pm\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|00\right\rangle \pm \left|11\right\rangle\right) \\ \left|\Psi^\pm\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|01\right\rangle \pm \left|10\right\rangle\right) \ \ .\end{aligned}$$

These are maximally entangled states on two qubits. Show how to generate all these states by a simple quantum circuit, and verify that the four Bell states form an orthonormal basis.

1.2 EPR Paradox:

Everyone has heard Einstein's famous quote “God does not play dice”. It is lifted from Einstein's 1926 letter to Max Born where he expressed his dissatisfaction with quantum physics by writing: ”Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not really bring us any closer to the secret of the Old One. I, at any rate, am convinced that He does not throw dice.” Even to the end of his life he held on to the view that quantum physics is just an incomplete theory and that some day we would learn a more complete and satisfactory theory that describes nature. For example, consider coin-flipping. We can model coin-flipping as a random process giving heads 50% of the time, and tails 50% of the time. This model is perfectly predictive, but incomplete. If we knew the initial conditions of the coin with perfect accuracy (position, momentum), then we could solve Newton's equations to determine the eventual outcome of the coin flip with certainty.

Einstein sharpened this line of reasoning in a paper he wrote with Podolsky and Rosen in 1935, where they introduced the famous Bell states. Recall that for Bell state $\frac{1}{\sqrt{2}}(\left|00\right\rangle + \left|11\right\rangle)$, when you measure first qubit, the second qubit is determined. However, if two qubits are far apart, then the second qubit must have had a determined state in some time interval before measurement, since the speed of light is finite. Moreover this holds in any basis. This appears analogous to the coin flipping example. EPR therefore suggested that there is a more complete theory where “God does not throw dice.” Until his death in 1955, Einstein tried to formulate a more complete ”local hidden variable theory” that would describe the predictions of quantum mechanics, but without resorting to probabilistic outcomes. But in 1964, almost three decades after the EPR paper, John Bell showed that properties of Bell (EPR) states were not merely fodder for a philosophical discussion, but had verifiable consequences: local hidden variables are not the answer.

How does one rule out every possible hidden variable theory? Here's how: we will consider an extravagant framework within which every possible hidden variable theory. And then we will show that there is a particular quantum mechanical experiment using Bell states, whose results cannot be duplicated by any theory in this framework. The framework is this: when the Bell state is created, the two particles each make up a (infinitely long!) list of all possible experiments that they might be subjected to, and decide how they will behave under each such experiment. When the two particles separate and can no longer communicate, they consult their respective lists to coordinate their actions.

2 Bell's Thought Experiment

Bell considered the following experiment: the two particles in a Bell pair move in opposite directions to two distant apparatus. A decision about which of two experiments is to be performed at each apparatus is made randomly at the last moment, so that speed of light considerations rule out information about the choice at one apparatus being transmitted to the other. How correlated can the outcomes on the two experiments be? It can be shown that any theory in the classical hidden variable framework above gives a correlation of at most 0.75 whereas the quantum experiments described below give a correlation of about 0.8. Therefore the predictions of quantum mechanics are not consistent with any local hidden variable theory. We now describe the experiment in more detail.

The two experimenters A and B (for Alice and Bob) each receive a random bit r_A and r_B respectively. Each also receives one half of a Bell state, and makes a suitable measurement described below based on the received random bit. Call the outcomes of the measurements a and b respectively. We are interested in the achievable correlation between the two quantities $r_A \times r_B$ and $a+b(mod2)$. We will show that for the particular quantum measurements described below $P[r_A \times r_B = a+b(mod2)] \approx .8$.

What would a classical hidden variable theory predict for this setting? Now, when the Bell state was created, the two particles could share an arbitrary amount of information. But by the time the random bits r_A and r_B are generated, the two particles are too far apart to exchange information. Thus in any experiment, the outcome can only be a function of the previously shared information and one of the random bits. It can be shown that in this setting the best correlation is achieved by always letting the outcomes of the two experiments be $a = 0$ and $b = 0$ (see homework exercise). This gives $P[r_A \times r_B = a+b(mod2)] \leq .75$. This experiment therefore distinguishes between the predictions of quantum physics and those of any arbitrary local hidden variable theory. It has now been performed in several different ways, and the results are consistent with quantum physics and inconsistent with any classical hidden variable theory.

Here is the protocol:

- if $X_A = 0$, then Alice measures in the $-\pi/16$ basis.
- if $X_A = 1$, then Alice measures in the $3\pi/16$ basis.
- if $X_B = 0$, then Bob measures in the $\pi/16$ basis.
- if $X_B = 1$, then Bob measures in the $-3\pi/16$ basis.

Now an easy calculation shows that in each of the four cases $X_A = X_B = 0$, etc, the success probability is $cos^2\pi/8$. This is because in the three cases where $x_A \cdot x_B = 0$, Alice and Bob measure in bases that differ by $/pi/8$. In the last case they measure in bases that differ by $3\pi/8$, but in this case they must output different bits.

We still have to prove that Bell state $\frac{1}{\sqrt{2}}(\left|00\right\rangle + \left|11\right\rangle) = \frac{1}{\sqrt{2}}(\left|v_A v_A\right\rangle + \left|v_A^\perp v_A^\perp\right\rangle)$ Let $\left|v_A\right\rangle = \alpha\left|0\right\rangle + \beta\left|1\right\rangle$ and $\left|v_A^\perp\right\rangle = -\beta\left|0\right\rangle + \alpha\left|1\right\rangle$. Then $\frac{1}{\sqrt{2}}(\left|v_A v_A\right\rangle + \left|v_A^\perp v_A^\perp\right\rangle) = \frac{1}{\sqrt{2}}(\alpha^2\left|00\right\rangle + \alpha\beta\left|01\right\rangle + \alpha\beta\left|10\right\rangle + \beta^2\left|11\right\rangle) + \frac{1}{\sqrt{2}}(\beta^2\left|00\right\rangle - \alpha\beta\left|01\right\rangle - \alpha\beta\left|10\right\rangle + \alpha^2\left|11\right\rangle) = \frac{1}{\sqrt{2}}(\alpha^2+\beta^2)(\left|00\right\rangle + \left|11\right\rangle) = \frac{1}{\sqrt{2}}(\left|00\right\rangle + \left|11\right\rangle)$