Notes on Scott Aaronson's *Quantum Information Science*

by Paulo Alves

Overview

Lecture 1 (January 17)

An introduction to Quantum Information Science. A few important concepts are introduced (**Probability**, **Locality**, **Local Realism**, the **Church-Turing Thesis** and its extended variation) to contextualize how quantum mechanics affects our understanding of physics.

Lecture 2 (January 19)

Quantum mechanics challenges **Computational Universality**.

The **Double Slit Experiment** introduces **Decoherence** and **Interference**. It motivates us to use **Amplitudes** to measure quantum chance, which are related to probabilities through the **Born Rule**.

Linear Algebra can model classical probabilities using **Stochastic Matrices** and **Tensor Products**.

Lecture 3 (January 24)

Quantum States and the **Qubit** warrant using **Bra-Ket Notation**. Linear Algebra can model quantum states too, but for that we need **Unitary** and **Orthogonal Matrices**, as well as **Unitary Transformations**.

Lecture 4 (January 26)

Several examples of **Quantum Gates** get us working in multiple bases. The compatibility (or lack thereof) between **Unitary Transformations** and **Measurement** is explored. **Quantum Circuit Notation** is introduced, along with phenomena occurring with a single qubit (**Quantum Zeno Effect**, **Watched Pot Effect**, **Elitzur-Vaidman Bomb**).

Lecture 5 (January 30)

Our first quantum protocol distinguishes between a fair and biased coin. The distinguishability of quantum states is explored.

Considering the state of two qubits with linear algebra and quantum circuit notation introduces the **Partial Measurement Rule**, the **Controlled NOT**, and the **Bell Pair**. **Entanglement** comes into play, and we see why it need not require the existence of faster-than-light communication.

Lecture 6 (February 2)

Density Matrices are introduced to represent **Mixed States**. We see the properties of density matrixes including **Trace** and **Rank**, as well as processes we may want to do with them like applying unitary transformations, performing **Eigendecomposition**, and **Tracing Out**.

Lecture 7 (February 7)

The **Bloch Sphere** is introduced as a useful representation of possible states of a qubit.

The **No Communication Theorem** and the **No Cloning Theorem** limit what can be done with quantum information. These limits allow for the creation of **Quantum Money** schemes, such as **Wiesner's Scheme**.

Lecture 8 (February 9)

Attacks on Wiesner's Scheme are explored, including an **Interactive Attack**, and an **Attack Based on the Elitzur Vaidman Bomb.**

BB84 is a **Quantum Key Distribution** scheme allowing two parties to generate a shared secret.

Lecture 9 (February 14)

Using entanglement as a resource allows for **Superdense Coding**, transmitting two classical bits via one qubit, and **Quantum Teleportation**, transmitting a qubit via classical communication.

Lecture 10 (February 16)

Quantum Teleportation is further explored and extended to arbitrary quantum states. Quantifying entanglement leads us to **Entanglement Swapping**, the **GHZ State** and the **Monogamy of Entanglement**, as well as **Schmidt Decomposition**.

Lecture 11 (February 21)

Measuring entropy of a quantum state with **Von Neumann Entropy** and **Entanglement Entropy**. The **Measurement Problem** leads us into interpretation of quantum mechanics, the **Copenhagen Interpretation** and **S.U.A.C.**, as well as useful though experiments, **Schrödinger's Cat** and **Wigner's Friend**.

Lecture 12 (February 23)

Dynamic Collapse theories such as **GRW** and the **Penrose Interpretation** suggest that we're still missing part of the puzzle. **Everett's Many Worlds Interpretation** suggests that the universe branches every time a measurement happens.

Lecture 13 (February 28)

A further discussion of Many Worlds tackles the practicality of an unfalsifiable interpretation and the **Prefered Basis Problem**.

Hidden Variable Theories such as **Bohmian Mechanics** lead to a search for a local hidden variable theory which proves to be impossible given the **Bell Inequality,** leading us to the **CHSH Game**.

Lecture 14 (March 2)

The optimality of our strategy for the CHSH Game is discussed and proven through **Tsirelson's Inequality**. The implications of experimentally **Testing the Bell Inequality** lead us to **Superdeterminism** and modern skepticism of quantum mechanics.

Other non-local games (**The Odd Cycle Game** and **The Magic Square Game)** are covered.

Lecture 15 (March 9)

The CHSH Game can be applied to **Generating Guaranteed Random Numbers**, and many other tasks, which brings us to **Quantum Computing**. We discuss the intellectual origins of the field and a few conceptual points.

Lecture 16 (March 21)

The roles of interference and entanglement in quantum computing lead us to cover the construction of both classical and quantum **Universal Gate Sets**. We discuss **Quantum Complexity** and see our first quantum algorithm, **Deutsch's Algorithm**.

Lecture 17 (March 23)

We finish our discussion of Universal Gate Sets and the usage of black-box functions, which leads us to **Uncomputing**. Revisiting Deutsch's Algorithm, we see it's generalization the **Deutsch-Jozsa Algorithm**, as well as an introduction to the **Bernstein-Vazirani Problem**.

Lecture 18 (March 28)

Lecture 19 (March 30)

Lecture 20 (April 4)

These are coming.

Lecture 21 (April 6)

Lecture 22 (April 11)

Lecture 23 (April 13)

Lecture 24 (April 18)

Lecture 25 (April 20)

Lecture 26 (April 25)

Lecture 27 (April 27)

Lecture 28 (May 2)

Further discussion of the reliability of qubits leads us to **Stabilizer Sets** and their compact representations through **Generator Sets** of **Pauli Matrices**. The **Gottesman-Knill Theorem** explains why stabilizer sets aren't universal, and leads to the use of **Tableau Representation**.

Lecture 29 (May 4)

Quantum error correction codes with **Transversality** are prefered.

Practical implementations of of quantum computing are discussed, including the important speedups it could provide, leading to a discussion of the **HHL Theorem**. **DiVincenzo Criteria** could be satisfied with **Trapped Ions** or **Superconducting Qubits**, as well as **Photonics** (bringing us to the **KLM Theorem** and **Boson Sampling**) or **Non-abelian Anyons**.

Lecture 1: Tues Jan 17

- Quantum Information Science is an inherently interdisciplinary field (Physics, CS, Math, Engineering, Philosophy)
- About clarifying the workings of quantum mechanics.
  - We use it to ask questions about what you can and can't do with quantum mechanics
  - Can help solve problems about the nature of quantum mechanics itself.
- Professor Aaronson is very much on the theoretical end of research.
  - Theorists inform what practicalists make which in turn informs theorists' queries

There are several self-evident truths in the physical world. Quantum mechanics leaves some in place, and slashes others. To start with…

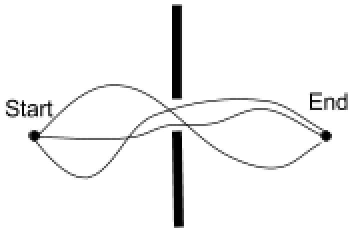
Probability ($p \in [0,1]$) is the standard way of representing uncertainty in the world. Probabilities have to follow certain obvious axioms like:

$P_1 + \ldots + P_n = 1$ mutually exclusive exhaustive possibilities sum to 1

$P_i \geq 0$

As an aside:

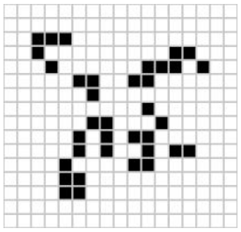
> There's a view that "Probabilities are all in our heads". Which is to say that if we knew everything about the universe (let's say position/velocity of all atoms in the solar system") that we could just crunch the equations and see that things either happen or they don't.



Let's say we have two points separated by a barrier with an open slit, and we want to measure the probability that a particle goes from one point to the other. It seems obviously true that increasing the number of paths (say, by opening another slit) should increase the likelihood that it will reach the other end.

We refer to this property by saying that probabilities are ***monotone***.

Locality is the idea that things can only propagate through the structure of the universe at a certain speed.



When we update the state of a little patch of space, it should only require knowledge of a little neighborhood around it. Conway's Game Of Life (left) is an apt comparison: things you do to the system *can* affect it, but they propagate only at a certain speed.

Einstein's Theory of Relativity explains that a bunch of known physics things are a direct result of light's speed. Anything traveling past the speed of light would be tantamount to travelling back in time.

Local Realism says that any an instantaneous update in knowledge about far away events can be explained by correlation of random variables.

For example, if you read your newspaper in Austin, you can instantly collapse the probability of your friend-in-San-Francisco's newspaper's headline to whatever *your* headline is.

Some Pop Science articles may talk about seeing one particle's spin instantaneously as a result of knowing another particle's spin, but that's basically the same as the newspapers.

The **Church-Turing Thesis** says that every physical process can be simulated by a Turing machine to any desired precision.

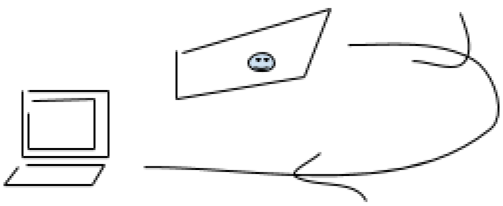
The way that Church and Turing understood this was as a definition of computation, but we think of it instead as a falsifiable claim about the real world. You can think about this as the idea that the entire universe is a video game: You've got all sorts of complicated things like quarks and whatnot, but at the end of the day, you've got to be able to simulate it in a computer.

Theoretical computer science courses can be seen as basically math courses.
So what *does* connect them to reality? The Church-Turing Thesis.

The **Extended Church-Turing Thesis** says that there's at most a polynomial blow-up for simulating reality.

The Church-Turing Thesis *seems* to be True.
The Extended Church-Turing Thesis *seems* to be False.

Lecture 2: Thurs Jan 19



Using time dilation, you could travel billions of years in the future and get results to hard problems. Fun! But you'd need a *LOT* of energy, and if you have that much energy in one place you basically become a black hole. Not so fun!

Computational Universality says that there aren't any computers that could exist which could solve a problem that ours can't already.

The Extended Church-Turing thesis says that if you can't solve a problem in polynomial time on today's computers then no one will ever be able to. Quantum mechanics challenges this. With quantum computers you could solve some problems faster that with a classical computer. With that said, however, there could still be a quantum equivalent to the ECTT.

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

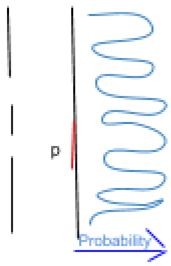
In the double slit experiment, you shoot photons through a wall with two narrow slits. Where the photon lands is probabilistic. If we plot where photons appear on the back wall, some places are very likely, some not.

Note that this itself isn't the weird part, we could totally justify this happening. What's weird is as follows. For some interval:

Let P be the probability that the photon lands on the interval.

Let P_1 be the probability that the photon lands on the interval if only slit 1 is open.

Let P_2 be the probability that the photon lands on the interval if only slit 2 is open.

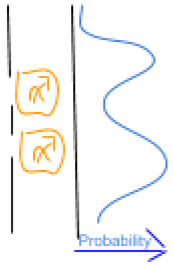


You'd think that $P = P_1 + P_2$, but that's not the case. Dark fringes that exist with two slits end up being hit by photons if only one slit is open.

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 this isn't really a matter of having 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 is called **Decoherence**.

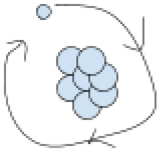
Decoherence is why the usual laws of probability look like they work in everyday life. A cat isn't in superposition because it interacts with normal stuff every day. These interactions essentially leak information about the 'cat system' out.

It's important to note that this relates to particles in isolation. Needing particles to be in isolation is why it's so hard to build a quantum computer.

The story of physics between 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 thing without connecting it to much else. That is, until Schrodinger, etc. came up with quantum mechanics.

A normal quantum physics class would go through this process of experimental proof to arrive at quantum mechanics, but we're just going to accept the rules as given and see what we can do from there.

For example take the usual high school model of the electron, rotating around a nucleus in a fixed orbit. Scientes realized that this model would mean that the electron would need to be constantly losing energy until it hit the nucleus. To explain this (and many other phenomenon) scientists modified the laws of probability.



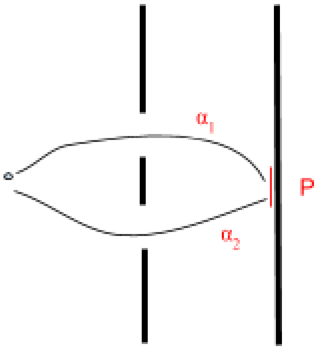
Instead of using probabilities $p \in [0,1]$ they started using **Amplitudes** $\alpha \in \mathbb{C}$. Amplitudes can be positive or negative and can have an imaginary part.

The central claim of quantum mechanics is that to explain a system you'd need to give one amplitude for each particle for each possible configuration of the particles.

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

$P = |\alpha|^2$
$= |\text{The real amplitude}|^2 + |\text{The imaginary amplitude}|^2$

So let's see how amplitudes being complex leads them to act differently from probabilities. Lets 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 going through the first slit α_1
plus the amplitude of it going through the second slit. α_2

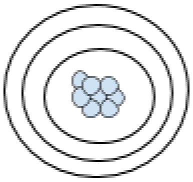
$P = |\alpha|^2 = |\alpha_1 + \alpha_2|^2$
$= |\alpha_1|^2 + |\alpha_2|^2 + 2\ |\alpha_1\alpha_2|$

If $\alpha_1 = ½$ and $\alpha_2 = -½$, then interference means that if both slits are open P = 0, but if only one of them is open, P = ¼.

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 their amplitudes are all positive and negative and they end up canceling each other out.

With some physics we won't cover in this class, you discover that all possibilities where amplitudes don't cancel each out leads to discrete shells where electrons can sit in.



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

$$M\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} = \begin{pmatrix}\alpha_1'\\ \alpha_2'\end{pmatrix}$$

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

$$\begin{pmatrix}p\\ q\end{pmatrix}\begin{matrix}\text{tails}\\ \text{heads}\end{matrix} \qquad \begin{matrix}p, q > 0\\ p + q = 1\end{matrix}$$

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

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

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

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:

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

We could also flip the coin fairly.

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

Which means regardless of previous position, both possibilities are 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}.

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

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

½ 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

½ if (it lands on heads and we flip again).

So we sum those values.

So what matrices CAN be used as transformations?

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

We can also say that all columns must add to 1, since we need the sum of initial probabilities to equal the sum of the transformed probabilities (both should equal 1).

$$A\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} = (i^{th}\text{ column of }A)$$

We can see this clearly by using basis vectors.

A matrix of this form 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(getting 0), P(b) = P(getting 1). For the second coin we'll use P(c) and P(d).

$$\begin{pmatrix}a\\b\end{pmatrix}\begin{array}{c}{}^0\\{}^1\end{array}\quad\begin{pmatrix}c\\d\end{pmatrix}\begin{array}{c}{}^0\\{}^1\end{array}$$

To combine the two probabilities we'll use the **Tensor Product**.

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

It's worth noting that not all combinations are possible.

For example:

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

Would mean that

$(\frac{1}{2})(\frac{1}{2}) = abcd = (0)(0)$

Therefore it can't be a tensor product.

Let's say that if(the first bit is 1){we want to flip the second bit}

$$\begin{array}{c}00\\01\\10\\11\\ \\ \end{array}\begin{array}{c}\left(\begin{array}{cccc} & & & \\ & & & \\ & & & \\ & & & \end{array}\right)\\ \begin{array}{cccc}00 & 01 & 10 & 11\end{array}\end{array}\begin{array}{c}\begin{pmatrix}\frac{1}{2}\\0\\\frac{1}{2}\\0\end{pmatrix}\\ \\ \end{array}\begin{array}{c}00\\01\\10\\11\\ \\ \end{array}$$

We'd do:

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

This is called the **Controlled NOT** it comes up in quantum mechanics.

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

$$\begin{pmatrix} & & \\ & U & \\ & & \end{pmatrix}\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix} = \begin{pmatrix}B_1\\B_2\\B_3\end{pmatrix}$$

Where $\sum_{i=1}^{n}|A_i|^2 = 1 = \sum_{i=1}^{n}|B_i|^2$ and you're measuring with probability $|\alpha_i|^2$

Lecture 3: Tues Jan 24

Tensor Products are a way of building bigger vectors out of smaller ones.

Let's apply a NOT operation to the first bit, and do nothing to the second bit. That's really the same as defining function f as f(00) = 10, f(01) = 11, f(10) = 00, f(11) = 01. So we can fill in the tensor product as follows:

$$\begin{pmatrix}0&1\\1&0\end{pmatrix}\otimes\begin{pmatrix}1&0\\0&1\end{pmatrix}=\begin{array}{c}00\\01\\10\\11\\ \end{array}\begin{array}{cccc}\left(0\right.&0&1&\left.0\right)\\\left(0\right.&0&0&\left.1\right)\\\left(1\right.&0&0&\left.0\right)\\\left(0\right.&1&0&\left.0\right)\\00&01&10&11\end{array}\begin{array}{c}(\quad)\\(\quad)\\(\quad)\\(\quad)\\ \end{array}\begin{array}{c}00\\01\\10\\11\\ \end{array}$$

A **Quantum State** is a unit vector in $\mathbb{C}^N$ referring to the state of a quantum system.

Formally a quantum state could exist in any dimension. Physics courses cover infinitely dimensional vectors, but we'll stick to discrete systems (which is to say that when we make a measurement, there's a discrete number of variables to be read (with continuous outcomes).

- What does quantum mechanics say about the universe being discrete or continuous at the base level? It suggests a strange, hybrid picture. There's an infinite number of possibilities, but a discrete outcome. Formalisms of quantum mechanics technically contain infinite possibilities, like a system with two variables ($\frac{\alpha}{\beta}$) has uncountably infinite possible amplitudes (given the only restriction is that $|\alpha|^2 + |\beta|^2 = 1$), but you could do that in classical mechanics as well by just making a complex formation about the probabilities of flipping coins.

The **Qubit** is the simplest quantum system.

It's a two-state system (we label these '0' and '1') whose amplitudes sum to 1.

A one-state quantum system would just be (1). Not very interesting!

As an alternative to vector notation, we have **Ket Notation**.

$\begin{pmatrix}\alpha\\\beta\end{pmatrix}\begin{array}{c}0\\1\end{array} = \alpha|0\rangle + \beta|1\rangle$

Note that $|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$

and that $|\Psi\rangle$ is the variable we'll usually use for kets.

Why do we use ket notation?

One main advantage is that practically speaking, we'll usually care mostly about really sparse vectors (where most values are 0), so it's easier to represent only the values we are talking about.

It's really just a formalism to make life easier, we can put anything in ket notation. Look, this is Shrodinger's Cat in ket notation: | 〉 + | 〉.

Often you'll need to take the transpose of a vector $\begin{pmatrix}\alpha\\\beta\end{pmatrix} \to \begin{pmatrix}\alpha & \beta\end{pmatrix}$ or for complex values $\begin{pmatrix}\alpha\\\beta\end{pmatrix} \to \begin{pmatrix}\alpha^* & \beta^*\end{pmatrix}$

Using the complex conjugate allows you to define a norm
$||v||^2 = v^Tv$

Then we get

$$\begin{aligned} v^Tv &= \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\ &= \alpha^*\alpha + \beta^*\beta \\ &= |\alpha|^2 + |\beta|^2 \end{aligned}$$

What does this look like in ket notation?

Just like we have the **ket** $|\boldsymbol{\Psi}\rangle = \alpha|0\rangle + \beta|1\rangle$ for $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$

We define the **bra** $\langle\boldsymbol{\Psi}| = \alpha^*\langle 0| + \beta^*\langle 1|$ for $\begin{pmatrix} \alpha^* & \beta^* \end{pmatrix}$

And we define $\langle x|y\rangle$ as the inner product of ket $|x\rangle$ with ket $|y\rangle$

Therefore $\langle\Psi|\Psi\rangle = 1$.
So $\langle v|w\rangle = \langle w|v\rangle^*$.

Remember: the way we change quantum states is by applying linear transformations. $(U) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \alpha' & \beta' \end{pmatrix}$

A linear transformation is **Unitary** if $|\alpha|^2 + |\beta|^2 = |\alpha'|^2 + |\beta'|^2$

Unitary Matrices correspond to unitary transformations.

We've got the identity $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and permutation matrices $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ which are the only stochastic UMs.

Others include

$\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$ <- which maps $|0\rangle \to |1\rangle$, $|1\rangle \to |0\rangle$ $\quad \begin{pmatrix} 0 & -i \\ -1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 1 & e^{i\theta} \end{pmatrix}$ <- Note: Euler's Equation says $e^{i\theta} = \cos\theta + i*\sin\theta$

All real possible states of a qubit define a circle and all complex possibile states define a sphere. That's because these states are all the quantum vectors of length 1.

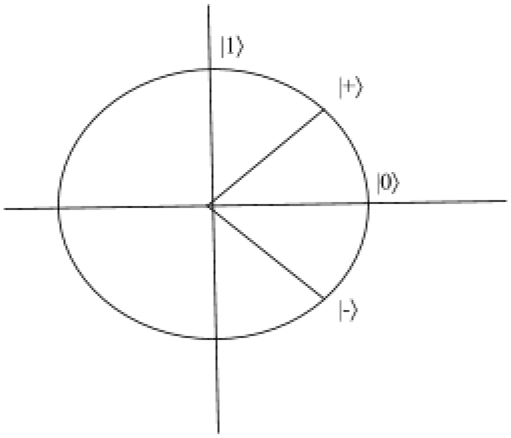
We define:

$|+\rangle = \dfrac{|0\rangle + |1\rangle}{\sqrt{2}}$

$|-\rangle = \dfrac{|0\rangle - |1\rangle}{\sqrt{2}}$

$|i\rangle = \dfrac{|0\rangle + i|1\rangle}{\sqrt{2}}$

$|-i\rangle = \dfrac{|0\rangle - i|1\rangle}{\sqrt{2}}$



Unitary Transformations are norm-preserving linear transformations.

For any angle θ you could have $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ which grabs a vector and rotates it θ degrees.

For example $R_{\pi/4} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$

To get the transform of $(ABv)^\dagger$ $v^\dagger B^\dagger A^\dagger$

<u>What does it mean that a unitary matrix preserves the 2-norm?</u>

It means applying a unitary transformation retains $\langle\Psi|\Psi\rangle$

$$\langle\Psi|\Psi\rangle = (U|\Psi\rangle)^\dagger U|\Psi\rangle = \langle\Psi|U^\dagger U|\Psi\rangle$$

So for this to always hold, $U^\dagger U$ has to be I. Which means $U^{-1} = U^\dagger$

That in turn implies that the rows of U must be an orthogonal unit basis .

So you can check if the rows or columns form an orthogonal unit basis (this isn't part of the definition of unitary matrices or anything, but because unitary matrices will always preserve the inner products).

An **Orthogonal Matrix** is both unitary and real.

They are the product of rotations and reflections.

Some examples:

$R_{\pi/4}|0\rangle = |+\rangle$

$R_{\pi/4}|+\rangle = |1\rangle$ You'll get a full revolution after applying $R_{\pi/4}$ eight times.

$R_{\pi/4}|1\rangle = -|-\rangle$

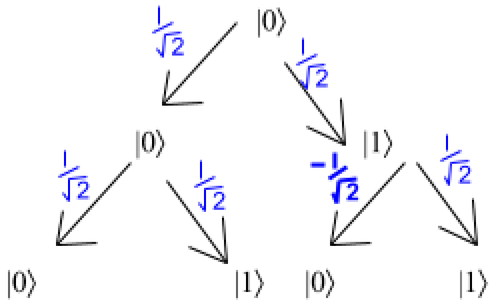
In the classical world

½ probability of a random event + ½ probability of a random event = just random

But in the quantum world

You can apply a transformation to a superpositioned state and get a specific answer

Anything interesting in quantum mechanics can be explained in terms of **interference**.



The |0⟩ amplitude can go to states 0 and 1 equally.

There were two different amplitudes on the 0 state but they cancel each other out.

|0⟩ states interfere destructively

|1⟩ states interfere constructively

No matter what unitary transformation you apply: If |0⟩ goes to U|0⟩, then -|0⟩ goes to -U|0⟩.

The zero state and the minus zero state are indistinguishable mathematically, which is to say:

Global phase is unobservable.

Multiplying your entire quantum state by a scalar is like if last night someone moved the entire universe twenty feet to the left. We can only really measure things relative to other things:

Relative phase *is* observable.

To distinguish between the states |+⟩ and |-⟩ we can rotate and then measure them.

There are no second chances. Once you measure, the outcome is set.
So you can distinguish some states via repeated measurement.

Lecture 4: Thurs Jan 26

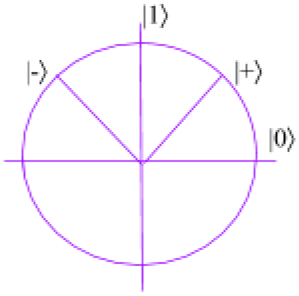
We call the matrix $R_{\pi/4} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ the $\sqrt{NOT}$ gate, as $R_{\pi/4}{}^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ aka the NOT Gate.

The **Hadamard Gate** is $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

It's useful because it represents a mapping between the |0⟩,|1⟩ basis to the |+⟩,|-⟩ basis.

$H|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \end{pmatrix} = |+\rangle$ Similarly $H|+\rangle = |0\rangle$, $H|1\rangle = |-\rangle$, and $H|-\rangle = |1\rangle$

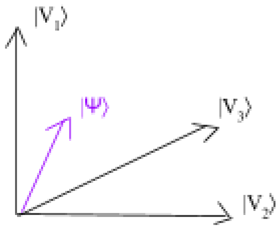
Note that we've got two orthogonal (complementary) basis: being maximally certain in the |+⟩,|-⟩ basis means that you're maximally uncertain in the |0⟩,|1⟩ basis and vice versa.



Why would we want to use 2 different bases?

We like to think of vectors existing abstractly in vector space, but to use it meaningfully, we've got to get it to a basis. We're not really going to get a satisfactory answer until we start talking about quantum protocols.

Side note, when talking about the Born Rule, we've been using a special case for one particular basis for simplicity.



We can think about measurement more generally. Measuring in the orthonormal basis $\{|V_1\rangle,...,|V_N\rangle\}$, you'll get the outcome $|V_i\rangle$ with probability $|\langle V_i|\Psi\rangle|^2$.

So the probability of the outcome $|V_3\rangle$ is the projection onto the basis vector.

$|\langle V_i|\Psi\rangle|^2 = |\alpha_1|^2$

We use bases |0⟩ and |1⟩ arbitrarily as a nice convention.

To do operations in a different basis use unitary transformations to convert.

So for $\{|V_1\rangle,...,|V_N\rangle\}$ use $U|V_1\rangle = |1\rangle,...,U|V_N\rangle = |N\rangle$ to use the basis $\{|1\rangle,...,|N\rangle\}$

There's an extreme point of view in quantum mechanics that unitary transformations are the only thing that really exist, and measurements don't really exist. And the converse also exists: the view that measurements are the only thing that really exist, and that unitary transformations don't.

Unitary Transformations are :

- *Invertible*. This should be clear, since preserving the two norm means that $U^\dagger U = 1$ which means $U^\dagger = U^{-1}$.
  - Reversible. The transformation $|\Psi\rangle \rightarrow U|\Psi\rangle$ can be reversed with $U^{-1}U|\Psi\rangle = |\Psi\rangle$.

    Interestingly this implies that unitary evolution can never destroy information, which should imply that the universe is reversible. We've known that the microscopic laws of physics are reversible since galileo times (i.e. observing a falling object backwards follows gravity backwards). So for example burning a book shouldn't necessarily destroy the information within, as physics says that you can get all the information from the smoke and ash left over.

- *Deterministic*
- *Continuous*
  i.e. you can always apply them in a time-continuous way. That's why it's important that unitary matrices are complex. If the transformation $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ took place in 1 sec. $\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$ took place over the first half of the second.

By the way, there is a 3x3 matrix that squares to $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Which means that you could apply $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ on $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ by using $\begin{pmatrix} \alpha \\ \beta \\ 0 \end{pmatrix}$ on it. without ever needing complex numbers! That's because using complex numbers works in the same way as adding a new dimension to your vector. Just like you could reflect your three-dimensional self by rotating yourself in the fourth dimension.

Important: Never eat anything in the fourth-dimension. It'll mess with the chirality of your molecules.

Measurements break all three rules of unitary transformations. They are:

- *Irreversible*
  - Whatever information about qubit you didn't capture is now lost.
- *Probabilistic*
  - Everything in quantum mechanics is deterministic *until* measurement (or information leaves the system).
- *Discontinuous*

So how can we reconcile these two sets of rules?

That's the **Measurement Problem**. We'll talk about points of view on it later.

Despite the philosophical conflict, unitary transformations and measurement sync up well because:

unitary transformations preserve the 2-norm and
Measurement gives probabilities given by the 2-norm

- We used to think everything was based on the 1-norm, until we found that quantum mechanics was based on the 2-norm. This got researchers looking for things based on the 3-norm, 4-norm, etc. They didn't really find anything though (the extra credit problem on the homework on norm preserving linear transformations sheds light on why).
  - Making quantum mechanics a bit of "an island in theory space". If you try to adjust anything about it in any way you get gunk. You could alternatively say that there's "nothing nice near quantum mechanics".

- There are many reasons why complex numbers work better than the reals or quaternions.

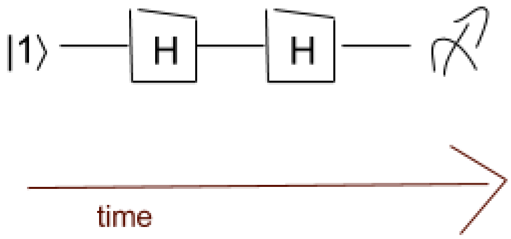
One more example of a linear transformation.

$\sum|\alpha_i|^4$

for $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$ maps $|0\rangle$ -> $|i\rangle$ $\quad \frac{|0\rangle + i|1\rangle}{\sqrt{2}}$

and $|1\rangle$ -> $|-i\rangle$ $\quad \frac{|0\rangle - i|1\rangle}{\sqrt{2}}$

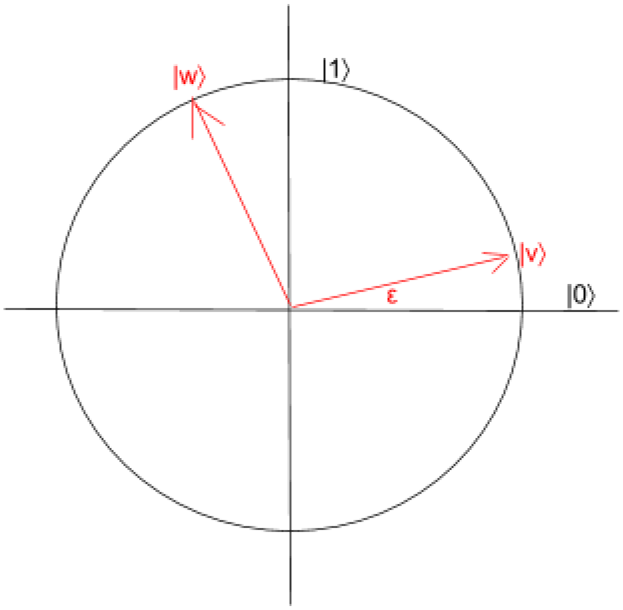
Quantum Circuit Notation keeps track of what qubits we have and the linear transformations we apply.



So to the left we start with $|1\rangle$, apply a Hadamard Gate, apply a Hadamard Gate, then measure (implied to be in the $|0\rangle$,$|1\rangle$ basis)

We'll never brach in a quantum circuits, since that can't correspond to a unitary transformation. To enlarge a system we can use a new $|0\rangle$ qubit, an **ancilla** qubit.

There are several interesting phenomena that already happen in the quantum mechanics of one qubit.



Suppose you have a qubit in the $|0\rangle$ state. We can know this because it's staying 0 over and over in measurements. Let's say we want to put it in the $|1\rangle$ state without using any unitary transformations.

For some small ε, we can measure the qubit on the ε basis. The probability of getting the qubit to move to ε increases as ε decreases.

$\text{Prob}(|v\rangle) = \text{KO}||v\rangle|^2 = \cos^2\varepsilon$

$|v\rangle = \cos\varepsilon|0\rangle + \sin\varepsilon|1\rangle$

Over ε measurements we could inch the qubit from $|0\rangle$ to $|1\rangle$.

$\text{Prob}(|w\rangle) = \sin^2\varepsilon \sim \varepsilon^2$

What's the likelihood that we'd get a measurement that *isn't* the one we want?

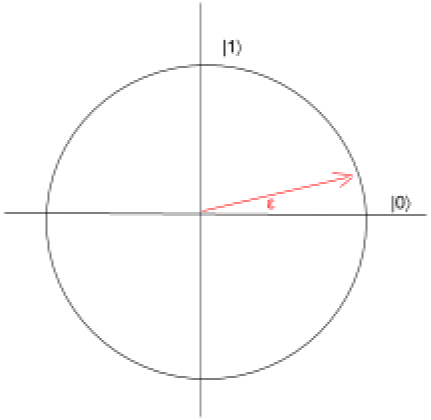
By union bound $1/\varepsilon * \varepsilon^2 = \varepsilon$

This is called **The Quantum Zeno Effect**

It was discovered by Alan Turing.

Perhaps a real world analog would be asking a stranger to marry you, getting coffee and then asking them to marry you, etc. You could refer to this situation as moving between bases for the 'yes'/'no' measurement.

Another interesting variant of the same kind of effect is as follows:

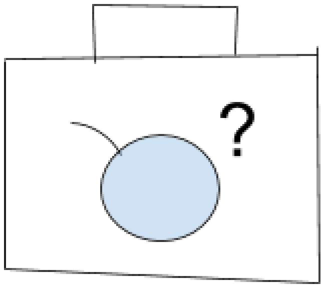


Say we want to keep a qubit at $|0\rangle$ but it keeps rotating towards $|1\rangle$ (it's *drifting*).

If we keep measuring it on the $|0\rangle,|1\rangle$ basis the odds of it rotating to $|1\rangle$ is ε^2.

This is called **The Watched Pot Effect.**

Another interesting phenomenon is the **Elitzur-Vaidman Bomb**.
A quantum effect discovered in the 1990's.



Say we're at a quantum airport and there's a piece of unattended luggage which could be a bomb, but opening the suitcase would trigger it.

How do we disarm the bomb without opening the suitcase?

We could make a query with a classical bit:

$b \in \{0$ (don't make query), 1 (make query)$\}$

But we only either find nothing or blow up the bomb. Not good!

Instead, we can upgrade to a qubit:

$|b\rangle = \alpha|0\rangle + \beta|1\rangle$

Now: If there's no bomb $|b\rangle$ gets returned to you.

If there is a bomb [$|0\rangle$ returns $|0\rangle$] and [$|1\rangle$ explodes it].

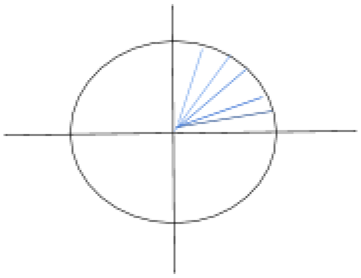
What we can do is apply the rotation $R_\varepsilon = \begin{pmatrix} \cos\varepsilon & -\sin\varepsilon \\ \sin\varepsilon & \cos\varepsilon \end{pmatrix}$. Giving us:

$\cos\varepsilon|0\rangle + \sin\varepsilon|1\rangle$

If there's a bomb, the probability it explodes is $\sin^2\varepsilon \sim \varepsilon^2$, otherwise you get $|0\rangle$

If there's no bomb, $\cos\varepsilon|0\rangle + \sin\varepsilon|1\rangle$

So repeating about $\pi/2$ times makes the probability of setting off the bomb as $1/\varepsilon * \varepsilon^2 = \varepsilon$

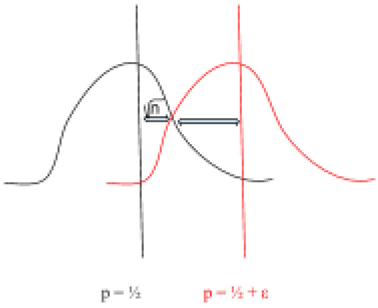


Lecture 5: Tues Jan 30

Say you have a coin, and you want to figure out if it's fair ($p = ½$) or if it's biased ($p = ½ + \varepsilon$). How would you go about doing so?

The classical approach to solving this problem would be to flip the coin a lot (about $1/\varepsilon^2$ times), keeping track of heads and tails until you have a strong degree of certainty that randomness isn't affecting your results. Standard probability stuff.

This requires $\log 1/\varepsilon^2$ of memory to store the running totals. In fact, there's a theorem by Hellman and Cover from the 70's that says that any protocol to solve this problem requires that much storage.



What if instead we used quantum computing?

We can start with a qubit in the |0⟩ state, and consider two rotations R_ε and $R_{-\varepsilon}$. We can repeated flip the coin, and if it lands tails apply R_ε (rotating clockwise) and if it lands heads apply $R_{-\varepsilon}$ (rotating counterclockwise). After many flips (about $1/100\varepsilon^2$) we can measure the qubit and know with a reasonable degree of certainty that if it's in the |0⟩ state, the coin is fair, if it's in the |1⟩ state, the coin is biased.

- Won't counting that high require plenty of storage?
  - No. We can write a protocol with a half-life (some probability that it'll halt at each step) causing it to repeat approximately the number of times we want it to.
- What about if the qubit drifts by a multiple of π, won't that make a biased coin look fair?
  - That's possible, but we can make it so that a bias coin will very likely land on |1⟩.

Quantum information protocols are like baking souffles.
Opening the oven will *literally* collapse the souffle.

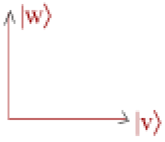
This is our first example of a quantum protocol getting a resource advantage:

the quantum version takes **1 qubit of storage** as opposed to the classical solution's **$\log 1/\varepsilon^2$ bits**.

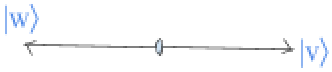
This result was shown by Professor Aaronson and a student of his. It wasn't a particularly hard problem, but no one had asked the question before. There's still "low hanging fruit," even in the mechanics of a single qubit.

Distinguishability of Quantum States

Given two orthogonal quantum states |v⟩ and |w⟩ there's a basis that distinguishes them.



These on the other hand are indistinguishable.

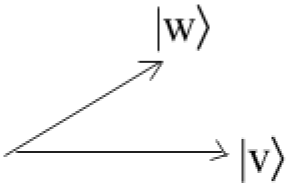


|⟨v|w⟩| gives a good measure of the distinguishability of arbitrary states.

|⟨v|w⟩| = 1

|⟨v|w⟩| = 0

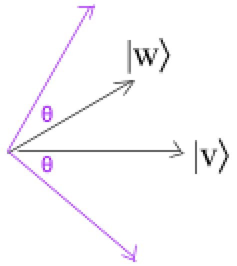
What about these?



More specifically: What measurement would minimize the chance of making a mistake in differentiating |v⟩ and |w⟩?

You may want to measure in the |v⟩,|something else⟩ basis, as it would eliminate one kind of error completely (not getting |v⟩ ensures the state was |w⟩), but there's a better way:

Take the bisector of |v⟩ and |w⟩, and get the angles 45° to either side, ensuring each original vector is the same distance to its closest basis vector.



The general state of **2 Qubits** is:

$\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$

The probability of getting $|00\rangle = |\alpha|^2$ Note that $|00\rangle$ is the same as $|0\rangle|0\rangle$ or $|0,0\rangle$ or $|0\rangle \otimes |0\rangle$

$|01\rangle = |\beta|^2$

$|10\rangle = |\gamma|^2$

$|11\rangle = |\delta|^2$

In principle there's no distance limitation between qubits. You could have one on Earth, and the other could be with your friend on the moon.

You'd only be able to measure the first bit:

The probability of getting $|0\rangle = |\alpha|^2 + |\beta|^2$ because those are the amplitudes compatible with 0 in the 1st bit.

$|1\rangle = |\gamma|^2 + |\delta|^2$

Suppose I measure the first qubit to |0⟩. What can I say about the second qubit?

Well we've narrowed down the possibilities to $\alpha|00\rangle$ and $\beta|01\rangle$. The state of the system is thus now in the superposition $|0\rangle \otimes \frac{(\alpha|0\rangle + \beta|1\rangle)}{\sqrt{|\alpha|^2 + |\beta|^2}}$ ←---- Don't forget to normalize

This is called the **Partial Measurement Rule**

Systems collapse minimally to fit your measurements.

This is actually the last rule of quantum mechanics that we'll cover in the course. Everything else is just a consequence of rules we've already covered.

This $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ is the **Controlled NOT**.

Remember: it flips the 2nd bit if the 1st bit is 1.

What if we wanted to always do NOT on the 2nd bit:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

This is **I ⊗ NOT**

/ |

(nothing on 1st bit) with (NOT on 2st bit)

It can be decomposed as: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ which makes it a tensor product unitary.

What if we want **NOT ⊗ I**?

$$\begin{array}{c|cccc} 00 & 0 & 0 & 1 & 0 \\ 01 & 0 & 0 & 0 & 1 \\ 10 & 1 & 0 & 0 & 0 \\ 11 & 0 & 1 & 0 & 0 \\ \hline & 00 & 01 & 10 & 11 \end{array}$$

Remember that rows and cols represent the transformation

f(row) = col so the prob the input is 00

is the prob that the output is 10

Very often in quantum information we'll want to take a group of qubits and perform an operation to one of them, say Hadamard the 3rd qubit.

What that means in terms of the matrices is applying $I \otimes I \otimes H \otimes \ldots \otimes I$

What's $H \otimes H$?

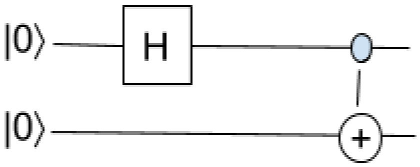
$$\frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

Why should it look like this?

Let's look at the first row: $H|00\rangle = |{+}{+}\rangle$. Which means for each qubit there's an equal prob it's output lands on $|0\rangle$ or $|1\rangle$.

All of these are examples of using tensor products to build bigger unitary matrices, except for the Controlled NOT, where the 1st bit affects the 2nd. We'll need operations like that in order to have one qubit affect another.

2 Qubits In Quantum Circuit Notation



Start with 2 qubits at $\vert 0\rangle$		Apply Hadamard to 1st bit		Apply a Controlled NOT with the 1st bit as the **control** and the 2nd as the **target**.	
(1) (0) (0) (0)	$\vert 00\rangle$	($\frac{1}{\sqrt{2}}$) (0) ($\frac{1}{\sqrt{2}}$) (0)	$\frac{\vert 00\rangle + \vert 10\rangle}{\sqrt{2}}$ $= \vert +\rangle \otimes \vert 0\rangle$	($\frac{1}{\sqrt{2}}$) (0) (0) ($\frac{1}{\sqrt{2}}$)	$\frac{\vert 00\rangle + \vert 11\rangle}{\sqrt{2}}$

The Controlled NOT can also be shown as $|x, y\rangle$ -> $|x, y \otimes x\rangle$

The state that this circuit ends on, $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$ is called the **Singlet** or the **Bell EPR Pair**

This state is particularly interesting because measuring the 1st qubit collapses the 2nd qubit. It can't be factored into a tensor product of the 1st qubit's state and the 2nd's.

An **Entangled** state cannot be decomposed into a tensor product, while an **Unentangled** state can.

The basic rules of quantum mechanics force these properties to exist. They were noticed fairly early in the history of of the field. It turns out that most states are entangled.

As we mentioned earlier, entanglement was what troubled Einstein about quantum mechanics. He thought that it meant that quantum mechanics must ential faster than light communication.

That's because particles need to be close to become entangled, but once they're entangled you can separate them to an arbitrary distance and they'll stay entangled. This has actually been demonstrated experimentally for distances of up to 150 miles.

Let's say that Alice and Bob entangle a pair of particles by setting their state to $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$, then Alice brings her particle to the moon while Bob stays on Earth. If Alice measures her particle, she can *faster-than-the-speed-of-light* know what position Bob's particle is in.

This bothered Einstein, but others thought that it wasn't that big a deal since Alice sees $|0\rangle$ and $|1\rangle$ in equal probability, which means it can be explained as a correlation between two random variables. However, a famous 1935 paper brought up a further problem: there's other things Alice could do instead of measuring in the $|0\rangle,|1\rangle$ basis.

What happens if Alice measures in the $|+\rangle,|-\rangle$ basis?

She'll get $|+\rangle$ as you might expect.

But what if before that, Alice takes this state and Hadamards the 1st bit?

Well it maps $|00\rangle$ to $|00\rangle + |10\rangle$ and $|11\rangle$ to $|01\rangle - |11\rangle$ (ignoring normalization).

That gives us: $\frac{|00\rangle + |10\rangle + |01\rangle - |11\rangle}{2}$ Remember $H|0\rangle = |+\rangle$, etc.

So now, applying the *Partial Measurement Rule* what is Bob's state?

If Alice sees $|0\rangle$, then Bob's qubit collapses to the possibilities where Alice sees $|0\rangle$:

$$\frac{|00\rangle + |01\rangle}{2} = |+\rangle$$

Conversely, if Alice sees $|1\rangle$:

$$\frac{|10\rangle - |11\rangle}{2} = |-\rangle$$

The paper goes on to talk about how this is more troubling than before. Alice's choice to measure in the $|+\rangle,|-\rangle$ basis is affecting Bob's qubit when he measures in the $|+\rangle,|-\rangle$ basis. And *that* looks a lot like faster than light communication.

How can we explain this?

One thing we can do is as “what happens if Bob makes a measurement?”

- In the case where Alice measured in |0⟩,|1⟩, Bob will see |0⟩ or |1⟩ with equal probability.
- In the case where Alice Hadamards her bit, then measures in |+⟩,|-⟩…
  - Bob will still see |0⟩ or |1⟩ with equal probability (measuring in the |0⟩,|1⟩ basis)

So the probability that Bob sees |0⟩ or |1⟩ is the same regardless of what Alice does.

People decided that it looked like there was something more general going on here, though. And so a different description should exist of Bob’s part of the state that’s unaffected by Alice’s measurements. Which brings us to…

Mixed States

We’ve only talked about **Pure States** so far (isolated quantum systems), but you can have quantum uncertainties layered together with regular, old uncertainty. This becomes important when we talk about states where we’re only measuring one part. If we look at the whole Alice-and-Bob-system together, it’ll look like a pure state.

Lecture 6: Thurs Feb 2

Last time we discussed the **Bell Pair**, and how if Alice measures her qubit in any basis, the state of Bob's qubit collapses to whichever state she got for hers. That being said, there's a formalism that tells us that Bob can't do anything to distinguish which basis Alice makes her measurement in, and thus no information travels instantaneously. This brings us to…

Mixed States

Which are probability distributions over quantum superposition.

We define a mixed state as a distribution over quantum states, so $\{p_i, |\Psi_i\rangle\} = p_1, |\Psi_1\rangle, \dots, p_n, |\Psi_n\rangle$

^

Thus, we can think of a pure state as a degenerate state of a mixed state where all probabilities are 1.

Note that these don't have to be orthogonal

The tricky thing about mixed states is that they have to preserve the property we discussed above (that the basis Alice measures in doesn't affect Bob's state), which is to say that if we used the $\{p_i, |\Psi_i\rangle\}$ notation, we'd be allowing multiple instances of the notation to represent the same state. For example $\frac{|00\rangle + |01\rangle}{2}$ could be represented in the $|0\rangle,|1\rangle$ basis or the $|+\rangle,|-\rangle$ basis. To avoid this, we'll use…

Density Matrices

represented as $\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|$

$|\Psi_i\rangle\langle\Psi_i|$ is the **outer product** of Ψ with itself.

It's the matrix you get by multiplying

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} \begin{pmatrix} \alpha_1^* & \dots & \alpha_N^* \end{pmatrix} = \begin{pmatrix} |\alpha_1|^2 & & \alpha_i\alpha_j^* \\ & \dots & \\ \alpha_i^*\alpha_j & & |\alpha_N|^2 \end{pmatrix}$$

Note that $\alpha_i\alpha_j^* = \alpha_i^*\alpha_j$ which means that the matrix is it's own conjugate transpose

$A = A^\dagger$ That makes it a **Hermitian Matrix**.

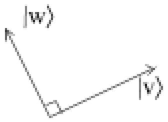
Some examples: $|0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ $|1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

Therefore an even mixture of them would be $\frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} = \begin{pmatrix} ½ & 0 \\ 0 & ½ \end{pmatrix} = \frac{I}{2}$.

Similarly: $|+\rangle\langle +| = \begin{pmatrix} ½ & ½ \\ ½ & ½ \end{pmatrix}$ $|-\rangle\langle -| = \begin{pmatrix} ½ & -½ \\ -½ & ½ \end{pmatrix}$

And $\frac{|+\rangle\langle +| + |-\rangle\langle -|}{2} = \begin{pmatrix} ½ & 0 \\ 0 & ½ \end{pmatrix} = \frac{I}{2}$.

Note that a mixture of $|0\rangle$ and $|1\rangle$ is different from a superposition of $|0\rangle$ and $|1\rangle$ (aka $|+\rangle$), and so they have different density matrices. However, the mixture of $|0\rangle$ and $|1\rangle$ and the mixture of $|+\rangle$ and $|-\rangle$ have the same density matrix: which makes sense because Alice converting between the two bases in our example above should maintain Bob's density matrix representation of the state.



In fact, this is true of whichever basis Alice chooses, and so for orthogonal vectors $|v\rangle$ and $|w\rangle$ we have that $\frac{|v\rangle\langle v| + |w\rangle\langle w|}{2} = \frac{I}{2}$.

Measuring ρ in the basis $|1\rangle, \dots, |N\rangle$ gives us the probability of $|i\rangle$ to be:

$$\Pr[|i\rangle] = \rho_{ii} = \langle i|\, \rho\, |i\rangle$$

Which is represented by the diagonal entries of the density matrix.

You don't need to square the value or anything because the Born Rule is already encoded in the density matrix (i.e. $(\alpha_1)(\alpha_1{}^*) =)|\alpha_1|^2$

That means that a density state which is a diagonal matrix is just a fancy way of writing a classical probability distribution.

$$\begin{pmatrix} p_1 & & \\ & \dots & \\ & & p_N \end{pmatrix}$$

While a pure state would look like $|\Psi\rangle\langle\Psi| = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

What if we want to measure a density matrix in a different basis?

Measuring ρ in $\{|v\rangle, |w\rangle\}$ will give $\Pr[|v\rangle] = \rho_{ii} = \langle v|\, \rho\, |v\rangle$

You can think of density matrices as encoding not one but infinite probability distributions because you can measure it in any basis.

The matrix I/2 we've encountered above, as the even mixture of $|0\rangle$ and $|1\rangle$ (and also that of $|+\rangle$ and $|-\rangle$) is called the **Maximally Mixed State**. This state is basically just the outcome of a classical coin flip, which gives it a special property:

Regardless of the basis we measure it in, both outcomes will be equally likely.

So for some basis $|v\rangle, |w\rangle$ you get the probabilities $\langle v|\, I/2\, |v\rangle = \frac{1}{2} \langle v \,|\, v\rangle = \frac{1}{2}$

$\langle w|\, I/2\, |w\rangle = \frac{1}{2} \langle w \,|\, w\rangle = \frac{1}{2}$

This explains why Alice is unsuccessful in sending a message to Bob: the maximally mixed state in any other basis is *still the maximally mixed state*.

So how do we handle unitary transformations with density matrices?

Since $\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|$, Applying U to ρ means:

$$\sum_i p_i (U|\Psi_i\rangle)(U|\Psi_i\rangle)^\dagger = \sum_i p_i U|\Psi_i\rangle\langle\Psi_i|U^\dagger = U\rho U^\dagger$$

You can pull out the U's since it's the same one applied to each mixture.

It's worth noting that getting n^2 values in the density matrix isn't some abstraction, you really need all those extra parameters. What do the off-diagonal entries represent?

$|+\rangle\langle+| = (\ \tfrac{1}{2}\ \tfrac{1}{2}\ )$

$(\ \tfrac{1}{2}\ \tfrac{1}{2}\ )$ These are where all the 'quantumness' resides.

It's where the interference between qubits is represented.

They can be different depending on relative phase:

$|+\rangle$ has positive off-diagonal entries

$|-\rangle$ has negative off-diagonal entries

$|i\rangle\langle i| = (\ \tfrac{1}{2}\ \ -i/2)$

$(\ i/2\ \ \tfrac{1}{2}\ )$

Later we'll see that as a quantum system interacts with the environment, the off-diagonal states get pushed down.

$(\ \tfrac{1}{2}\ \varepsilon\ )$

The density matrices in experimental quantum papers look like $(\ \varepsilon\ \tfrac{1}{2}\ )$.

The bigger the off-diagonal values, the better the experiment: because it represents them seeing more of the quantum effect.

Which matrices can arise as density matrices?

We're effective asking: What constraints does the form $\sum_i p_i |\Psi_i\rangle\langle\Psi_i|$ put on the matrix?

It must be:

- Square
- Hermitian
- $\sum_i p_{ii} = 1$ (which is to say: the **trace**, $Tr(\rho) = 1$)

Could $M = (\tfrac{1}{2}\ {-10})$ be a density matrix?

$(-10\ \ \tfrac{1}{2})$

No. Measuring this in $|+\rangle,|-\rangle$ would give $\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix} M \begin{pmatrix}\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\end{pmatrix} = 19/2$. Bad!

Remember that you can view each ρ as $U\rho U^\dagger$, whose diagonal has to be a probability distribution for all U. If we want that condition to hold, then in linear algebra terms, we need to add the restriction:

- All eigenvalues are non-negative (aka being **PSD: Positive Semidefinite**)

As a refresher: For the matrix ρ, the eigenvectors $|\Psi\rangle$ hold the equation:

$\rho|\Psi\rangle = \lambda|\Psi\rangle$ for some eigenvalue λ

If we had a negative eigenvector

$\langle\Psi|\rho|\Psi\rangle = \lambda$ would be < 0, which is nonsense.

Could we have missed a condition? Let's check.

We claim: any square, Hermitian, PSD matrix arises as a density matrix of a quantum state.

For such a ρ, find a representation of it in the form $\sum_i p_i |\Psi_i\rangle\langle\Psi_i|$

Then there exist eigenvectors $p\,|\Psi_i\rangle = \lambda_i |\Psi_i\rangle$ for each row $\lambda_i \geq 0$

$\langle\Psi_i|\rho|\Psi_i\rangle = \lambda_i$ so $\Sigma\lambda_i = \Sigma\,\langle i|\rho|i\rangle = \rho_{ii} = 1$

So you can say that for $\Sigma\, \lambda\, |\Psi_i\rangle\langle\Psi_i|$:

λ are the eigenvalues and $|\Psi_i\rangle$ are the eigenvalues.

This process of obtaining eigenvalues and eigenvectors is called **eigendecomposition**.

We know eigenvalues will be real because the matrix is Hermitian,
They're non-negative because the matrix is PSD.

One quantity you can always compute for density matrices is:

Rank

rank(ρ) = the number of non-zero λ_i's
(the number of rows with no eigenvectors)

A density matrix of rank(n) must look like $\begin{pmatrix} p_1 & & 0 \\ & \dots & \\ 0 & & p_n \end{pmatrix}$

And a density matrix of rank(1) represents a pure state.

Rank being at most n means that every mixed state
can be written as a mixture of at most n pure states.

In general, rank tells you the number of pure states that you have to mix to reach this mixed state.

Now, consider the pure 2 qubit state $\frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}}$.

We'll give the first qubit to Alice and the second to Bob.

How does Bob calculate his density matrix?

By picking some orthogonal basis for Alice's side.

You can rewrite Alice's part as $\sqrt{2/3}\,|0\rangle|+\rangle + \sqrt{1/3}\,|1\rangle|0\rangle$, which lets you calculate Bob's d.m.:

$$\tfrac{2}{3}\,|+\rangle\langle+| + \tfrac{1}{3}\,|0\rangle\langle 0|$$
$$= \tfrac{2}{3}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + \tfrac{1}{3}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

In general, if you have a bipartite pure state, it'll look like $\sum_{i,j=1}^{N} \alpha_{ij}\,|i\rangle|j\rangle = |\Psi\rangle$

And you can get Bob's local density matrix

$$(\rho_{Bob})_{j,j'} = \sum_i \alpha_{ij}\, \alpha_{ij'}^*$$

This process of going from part of a mixed state to a whole pure state is called **Tracing Out**.

The Key Points:

1) A density matrix encodes all and only what is physically observable

- 2 quantum states will lead to different probabilities *iff* they have different d.m.'s

2) No-Communication Theorem
    - If Alice and Bob share an entangled state, nothing Alice chooses to do will have any effect on Bob's density matrix.
        - In other words, there's no observable effect on Bob's end. Which is the fundamental reason that quantum mechanics *is* compatible with the physical limitations of reality.

Lecture 7: Tues Feb 7

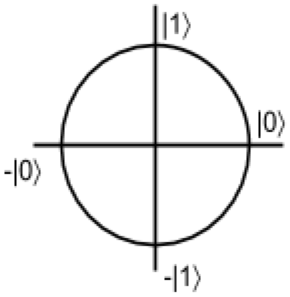
The No Communication Theorem says that if Alice and Bob share an entangled state $|\Psi\rangle = \sum_{i,j=1}^{N} \alpha_{ij} |i\rangle_{Alice}|j\rangle_{Bob}$ there's nothing that Alice can do to her subsystem that can affect Bob's density matrix.

We have the tools to prove this: just apply a tensor product to Alice's side, then see if Bob's density matrix changes. Or have Alice measure her qubit, the see if Bob's density matrix changes, etc.

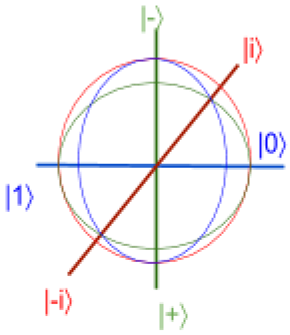
Note that if we condition on the outcome of Alice's measure (i.e. say that if Alice sees i then Bob will see j), we may need to update Bob's density matrix, but that's also true in the classical world.

Bloch Sphere

is a geometric representation of all possible states of a qubit. We've often drawn the state of qubits as a circle, which is already a little awkward: half of the circle is going to waste since $|0\rangle = -|0\rangle$ (both represent the same density matrix).



Instead, what if vectors that pointed in opposite directions were orthogonal? We get the Bloch Sphere...



We can see that $|+\rangle$ and $|-\rangle$ should be between $|0\rangle$ and $|1\rangle$. Then we can add $|i\rangle$ and $|-i\rangle$ as a new dimension.

In this representation, points on the surface of the sphere are pure states, such that

if they're 180° apart, they're orthogonal,

and if they're 90° apart, they're conjugate.

What about mixed states?

Well we know that the maximally mixed state, I/2, can be defined as $\frac{|0\rangle + |1\rangle}{2}$, $\frac{|+\rangle + |-\rangle}{2}$, or $\frac{|i\rangle + |-i\rangle}{2}$. The sum of any two of these vectors on the sphere is the origin.

We can in this way represent any mixed state as any point inside of the sphere.

The mixture of any states $|v\rangle$ and $|w\rangle$ represented as points in or on the sphere can be said to be a point between the two.

We can show geometrically that every mixed state can be written as a mixture of only two pure states because you can always draw a line that connects any pure state you want to some point in the sphere representing a mixed state, and then see which other pure state that the line intersects on the way out. By some vector math, the point can be described as some linear combination of the vectors representing pure states.

Experimentalists love the bloch sphere, because it works almost identically to how spin works with electrons.

With these things called **Spin-½ Particles** you can measure the electron spin relative to any axis of the sphere. You see if the electron is spinning clockwise or counterclockwise relative to the axis. And that behaves just like a qubit, in that the measurement collapses a more complex behavior into a binary result.

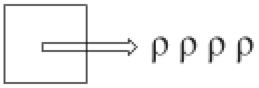
The weird part about Spin-½ Particles is that you *could have* asked the direction of the spin relative to any other axis. So what's really going on: What's the real spin direction? It turns out that it's some point on the Bloch Sphere. So if the state of the electron is that it's spinning in the (1,0,0) direction, we can say that it's in the |0⟩ state, and if it's spinning in the (0,1,0) direction, we can say that it's in the |+⟩ direction, and so forth.

The No Cloning Theorem

We've seen how entanglement seems to lead to non-local effects, like for the state $\frac{|00\rangle + |11\rangle}{2}$ if Alice measures her qubit's state, she can figure out Bob's. The reason that Alice isn't communicating faster than light boils down to Bob not being able to tell if his qubit's state is in the |0⟩,|1⟩ basis or on the |+⟩,|-⟩ basis.

But what if Bob could make unlimited copies of his qubit? He could figure it out through repeated measurements, and so he'd be able to tell what basis Alice measured in. FTL communication!

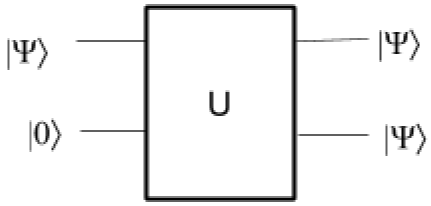
This is called **Quantum State Tomography**,
we'll see it later.



It turns out that we can prove that a procedure to reliably copy an unknown quantum state cannot exist. It's fairly easy to prove, but it's a fundamental fact about quantum mechanics.

Let's try to clone a single quantum bit, $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$

In our quantum circuit we want to apply some unitary transformation that takes $|\Psi\rangle$ and outputs $|\Psi\rangle$, and takes an ancilla from $|0\rangle$ to $|\Psi\rangle$.



Algebraically, a cloner would need to do:

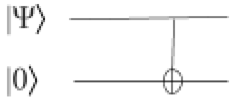
$$(\alpha|0\rangle + \beta|1\rangle) \otimes |0\rangle \dashrightarrow (\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha|0\rangle + \beta|1\rangle)$$
$$= \alpha^2|00\rangle + \alpha\beta|01\rangle + \alpha\beta|10\rangle + \beta^2|11\rangle$$

The cloner would need to look like:

$$\begin{pmatrix}\alpha^2\\ \alpha\beta\\ \alpha\beta\\ \beta^2\end{pmatrix} = \begin{pmatrix} & & \\ & U & \\ & & \\ & & \end{pmatrix}\begin{pmatrix}\alpha\\ \alpha\\ \alpha\\ \alpha\end{pmatrix}$$

The problem: this transformation **isn't linear** so it can't be unitary!

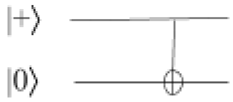
To clarify, a procedure that outputs some $|\Psi\rangle$ can be rerun to get $|\Psi\rangle$ repeatedly. What the No Cloning Theorem says is that if the $|\Psi\rangle$ is unknown, then you can't make a copy.



cNOT seems like a copying gate [as it maps $|00\rangle \to |00\rangle$, $|10\rangle \to |11\rangle$]
why isn't it in violation of the No Cloning Theorem?

Because it only works for $|0\rangle$ and $|1\rangle$. Classical information CAN be copied. Just ask Stallman!

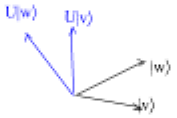
In general, for any orthonormal basis you can clone basis vectors.



Doing cNOT on produces the Bell Pair: $\frac{|00\rangle + |11\rangle}{2}$. Which sort of copies the first qubit in an entangled way, but that's different making a copy of $|+\rangle$.

Having two qubits be I/2, I/2 is not the same as $|+\rangle$,$|+\rangle$.

Since the No Cloning Theorem is so important, we'll present another proof of it:



A unitary transformation can be defined as a linear transformation that preserves inner product. Which is to say that the angle between $U|v\rangle$ and $U|w\rangle$ is the same as the one between $|v\rangle$ and $|w\rangle$.

Thus $\langle w|U^T U|v\rangle = \langle w|v\rangle$.

What would a cloning map do to this inner product?

$$|\langle v|w\rangle|^2 = C$$
$$|\ (\langle v|\otimes\langle v|)\ (|w\rangle\otimes|w\rangle)|^2 = C^2$$

C only ever equals C^2 if the inner product is 0 or 1: so the transformation is only linear if the v and w are in the same orthonormal basis.

There's a problem in classical probability that's a nice analog to the No Cloning Theorem.

If we have a coin of some probability heads, can we produce another coin with the same probability distribution? [Assuming the coin was made to have a certain probability distribution through some process unknown to us]

You'd need

$$\begin{pmatrix} p^2 \\ p(1-p) \\ p(1-p) \\ (1-p)^2 \end{pmatrix} = \begin{pmatrix} & & \\ & & \\ & S & \\ & & \end{pmatrix} \begin{pmatrix} p \\ p \\ p \\ p \end{pmatrix}$$

to be true for some stochastic matrix.

But this transformation isn't stochastic (the result matrix doesn't sum to 1).

The No Cloning Theorem has all sorts of applications to science fiction, because you can't make arbitrary copies of a physical system (say for teleporting yourself) if any of the relative relative information (say, in your brain) is encoded in quantum states.

Quantum Money

is an application of the No Cloning Theorem. In some sense it was the first idea in quantum information, and was involved in the birth of the field. The original quantum money scheme was proposed by Wiesner in 1969, though it was only published in the 80s.

Wiesner had left research by then. He eventually became a sheep herder.

Wiesner realized that uncloneability is useful for money to prevent counterfeiting. In practice, mints use special ink, watermarks, etc., but that's essentially just an arms race with the counterfeiters. So Wiesner proposed using qubits to make physical uncounterfietable money.

The immediate problem is that money systems need *cloneability* and *verifiability*.

Wiesner's Scheme

Have quantum bills (WLOG all are the same denomination). Each has:

- A classical serial number $S = \{0,1\}^n$
- A quantum state $|\Psi_{f(s)}\rangle$ (of n qubits)
  - The qubits in this state are unentangled and will always be in one of four states:
    - $|\Psi_{00}\rangle = |0\rangle$ $|\Psi_{01}\rangle = |1\rangle$ $|\Psi_{10}\rangle = |+\rangle$ $|\Psi_{11}\rangle = |-\rangle$

In order to decide the state of a given bill, the bank maintains a giant database that stores for all bills in circulation:

The classical serial number, and a function that takes the serial number as input and decides which basis to measure each qubit in (and which basis vector it should be).

S_1, $f(S_1)$
S_2, $f(S_2)$ | \
S_3, $f(S_3)$ | /

Wiesner's scheme has an important engineering problem though: you need to ensure that qubits don't lose their state (coherence). With current technology, qubits in a lab decohere in like an hour, tops.

There's two basic things needed for a scheme like this: verifiability and uncloneability.

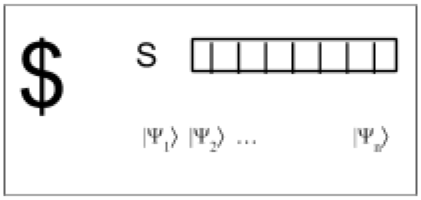
To verify a bill: bring it back to the bank. Bank verifies the bill by looking at the serial number, looking at how each qubit in the bill was supposed to be prepared. If the qubit was supposed to be prepared in $|0\rangle,|1\rangle$ measure in that basis.

Consider a counterfeiter that doesn't know what basis each qubit is supposed to be in, and they encode each qubit in a random allowable state. They only have a $\tfrac{1}{2}^n$ chance of guessing all the right bases.

The security of this scheme wasn't considered when it was proposed. Professor Aaronson asked about it on Stack Exchange a few years ago which prompted someone to write a paper on it.

Lecture 8: Thurs Feb 9

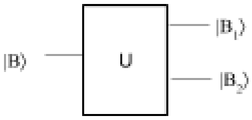
Guest Lecture by Supartha Podder



Continuation of Quantum Money

Last time we covered how classical money is copyable and showed a scheme for making money uncounterfietable through an application of the No Cloning Theorem.

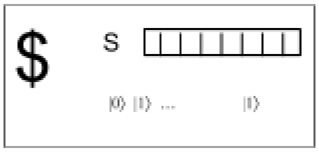
Let's consider a counterfeiter.

Wants to take a legitimate bill B and do 
such that both new bills pass the verification scheme.

Say the counterfeiter decides to measure all qubits in the |0⟩,|1⟩ basis.

Their new bill becomes:

- S gets copied
  - (classical information)
- Puts |0⟩ or |1⟩ as each qubit.



So the bank will check each quantum state, the ones that should be in the |0⟩|1⟩ basis are correct ½ the time. The ones that should be in |+⟩|-⟩ are correct ¼ of the time.

The odds of success of the counterfeiter (bank reading all states correctly is (⅝)^n.

This sort of attack has an upper success bound of (¾)^n.

Interactive Attack

There's an attack on this scheme based around the fact that verification involves giving the bank the bill, then the bank returns the bill and whether or not it's valid.

We can repeatedly go to the bank, ask them to verify the bill.

For some qubit that we set to |0⟩

if the bank measured it correctly, we know it's not |1⟩

if the bank measured it incorrectly, we know it's not |0⟩

We can similarly distinguish out |+⟩ and |-⟩

So running the verification scheme over each possibility for that quantum state allows us to get a strong picture of what state the bank is verifying it against.

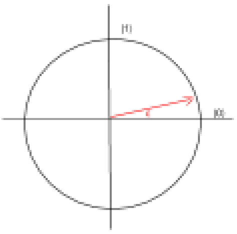
Running this procedure O(log(n)) times and you can copy the note with probability O(1=1/n^2).

Can we come up with another attack?

Recall the **Elitzur Vaidman Bomb**. The general idea is that through repeated applications of a unitary transformation to a state that starts at either |0⟩ or |1⟩, we can with a high probability of success get it to measure as |1⟩. Applying this sort of procedure to quantum money gives us an…

Attack Based on the Elitzur Vaidman Bomb

```
Set $|c\rangle$ to $|0\rangle$
Repeat $\pi/2\varepsilon$ times:
    Apply $R_\varepsilon$ to $|c\rangle$
    Apply cNOT to $|c\rangle|\Psi_1\rangle$
Then send the bill back to the bank.
```



Each time we apply cNOT given $|\Psi_1\rangle = |0\rangle$, we get $(\cos\varepsilon|0\rangle + \sin\varepsilon|1\rangle)|0\rangle = \cos\varepsilon|00\rangle + \sin\varepsilon|11\rangle$

Most of the time $|c\rangle$ will stay at $|0\rangle$.

Which means at each step the probability of getting caught is $\sin^2\varepsilon$.

Thus Prob[getting caught at all] is bounded at $\leq \pi/(2\varepsilon)\ \sin^2\varepsilon = O(\varepsilon)$

The same holds for $|1\rangle$ and $|-\rangle$.

But if $|\Psi_1\rangle = |+\rangle$, cNOT doesn't have the same effect

$(\cos\varepsilon|0\rangle + \sin\varepsilon|1\rangle) \otimes \frac{|0\rangle + |1\rangle}{\sqrt{2}}$ will eventually rotate the qubit to $|1\rangle$.

So when we measure at the end, we can distinguish $|+\rangle$ from the other states, because it's the only one that will be measured at $|1\rangle$.

We can similarly distinguish the other three states by starting $|c\rangle$ to the other three values.

What solution exists for this vulnerability?

The bank can just return a new copy of the money instead of the one that was verified.

This scheme still has a fundamental problem, which is that to make a transaction, you need to go to the bank. If you have to go to the bank, you might as well do an account transfer instead. The point of currency is that anyone should be able to verify it. Which brings us to...

Public Key Quantum Money

The bank produces money that can be verified by anyone.

Proposed by (Aaronson 2009), (Aaronson, Christiano 2012).

With this sort of scheme you'll always need computational assumptions on the counterfeiter, because technically they could always just try *every* possible quantum state with infinite computational power.

Quantum Key Distribution

Proposed by (Bennett, Brassard 1984) and thus called BB84

Key distribution is a fundamental task in cryptography. There's a technique in classical cryptography we can use for this called the **One-Time Pad**.

Given shared $k \in \{0,1\}^n$

Alice has secret message $m \in \{0,1\}^n$

Alice produces, sends c, for which $c_i = m_i \oplus k_i$

Bob decodes the message m as $m_i = c_i \oplus k_i$

As the name implies, this technique can only be used once securely, and it requires Alice and Bob to share some initial knowledge. In fact, it's been proven that Alice and Bob either need initial secret information in common or you must make computational assumptions on an eavesdropper Eve.

So we want a scheme with no assumptions on Eve in which to share a secret (presuming we have a classical authenticated channel: cannot be tampered by Eve, can be read)

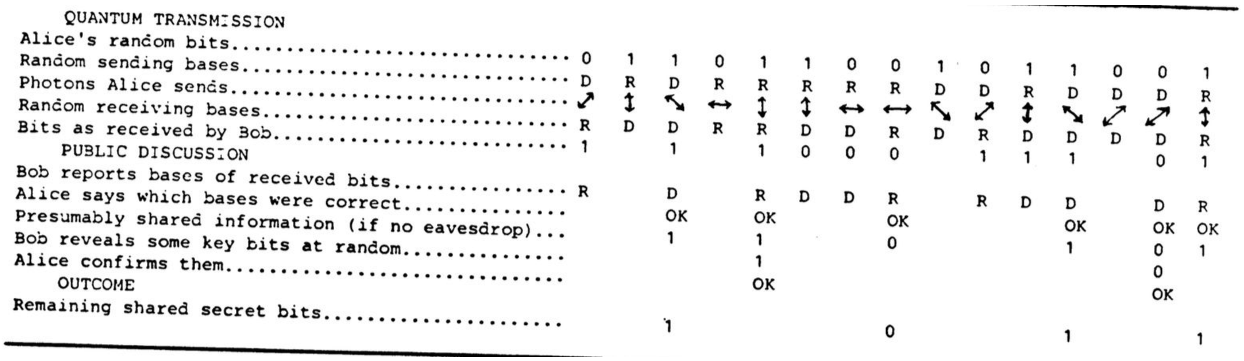
In cryptography we want secrecy and authentication.
This protocol is only going to deal with secrecy.

BB84

This quantum encryption scheme was already there in Wiesner's paper and was later formalized by B&B. It circumvents the issues we've seen in maintaining a qubit, because it only requires coherence for the time it takes for communication between Alice and Bob.

There are companies that are currently already doing quantum key distribution through fiber optic cables over up to 10 miles. There are people trying it working from ground to satellite which would get around the limitations of fiber optics, basically letting you do quantum key distribution over arbitrary distances. China actually has a satellite up for this express purpose.

Here's a diagram from the original paper that shows how BB84 works.



QUANTUM TRANSMISSION															
Alice's random bits..............................	0	1	1	0	1	1	0	0	1	0	1	1	0	0	1
Random sending bases.............................	D	R	D	R	R	R	R	R	D	D	R	D	D	D	R
Photons Alice sends..............................	⤢	↕	⤡	↔	↕	↕	↔	↔	⤡	⤢	↕	⤡	⤢	⤢	↕
Random receiving bases...........................	R	D	D	R	R	D	D	R	D	R	D	D	D	D	R
Bits as received by Bob..........................	1		1		1	0	0	0		1	1	1		0	1
PUBLIC DISCUSSION															
Bob reports bases of received bits...............	R		D		R	D	D	R		R	D	D		D	R
Alice says which bases were correct..............			OK		OK			OK				OK		OK	OK
Presumably shared information (if no eavesdrop)...			1		1			0				1		0	1
Bob reveals some key bits at random..............					1									0	
Alice confirms them..............................					OK									OK	
OUTCOME															
Remaining shared secret bits.....................			1					0				1			1

The basic idea is that you're trying to establish some shared secret knowledge and you want to know for certain that no eavesdroppers on the channel can uncover it. You've got a channel in which to transmit quantum information, and a channel in which to transmit classical information. In both, no one can impersonate Alice or Bob (authenticity) by eavesdroppers may be able to listen in (no secrecy).

- So Alice chooses a string x of random bits $\epsilon \{0,1\}^n$
- And another string y of random bits $y \epsilon \{0,1\}^n$ which she uses to decide which basis to encode each bit from x in.
- She then encodes the qubits in the $|0\rangle|1\rangle$ basis (in the diagram it's R) or the $|+\rangle|-\rangle$ basis (D)
- Then she sends over the qubits to Bob.
    - Bob picks his own random string $y' \epsilon \{0,1\}^n$ and uses y'_i to decide which basis
        - To decode the ith qubit send over (picking again between D and R)

Now Alice and Bob share which bases they picked to encode and measure in (the Y's) and discard any instances where they didn't pick the same one (about half the time).

At this point we consider an eavesdropper Eve who was listening in to the qubits that were sent over. The whole magic of using qubits is that if Eve listened in on the transmission she inherently changed the qubit's that Bob received. Sure, if she measured a $|0\rangle|1\rangle$ qubit in that axis, the qubit didn't change, but what if she measured a $|+\rangle|-\rangle$ qubit in the $|0\rangle|1\rangle$ basis?

If Alice sent $|+\rangle$, then Eve measured $|0\rangle$ and passed that along to Bob. Then Bob has a 50% chance of measuring $|+\rangle$ *or* $|-\rangle$.

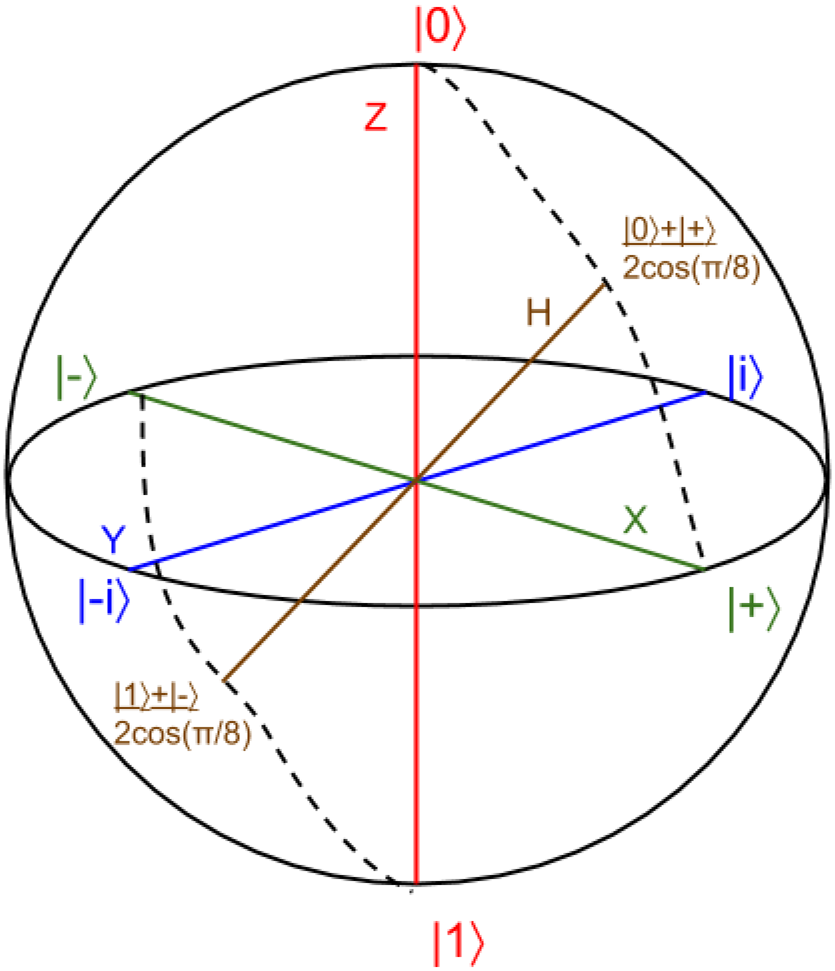
So Alice and Bob can verify that no one listened in to their qubit transmission by making sure that some portion of their qubits that they believe match do match. Of course these qubits aren't going to be secret anymore, but they've still got all the others.

If any of the qubits didn't match, then Eve eavesdropped and they can just try again and again until they can get an instance where no one listened in.

The idea is that now Alice and Bob share some initial information and can thus use some classical encryption scheme, like a 1 Time Pad.

Recitation Session

(Patrick)



Applying gates X,Y,Z or H is the same as doing a half turn on their respective axis.
S corresponds to a quarter turn around Z. [in the |+⟩ to |1⟩ direction]
$T^2 = S$, so T corresponds to an eighth turn around Z.
$R_{\pi/4}$ corresponds to a quarter turn (i.e. $\pi/4$) on Y.

Lecture 9: Tues Feb 14

To review: We've seen 3 different types of states in play:

- Basis States
  - exist in a computational basis $|i\rangle$
- Pure States
  - superpositions of basis states $|\Psi\rangle = \Sigma \alpha_i|i\rangle$
- Mixed States
  - classical probabilities over pure states $\rho = \Sigma \rho_i|\Psi_i\rangle\langle\Psi_i|$

Which represents the actual physical reality: pure or mixed states?
It's complicated. Sometimes we use density matrices to represent our probabilistic ignorance of a state, but other times (i.e. entangled states) they represent the maximal truth that exists of the state. We'll generally just focus on what these representations are useful for.

Wiesner's Scheme, as we've seen it, requires the bank to hold a lot of information. The paper (BBBW 82) circumvents this by basically saying: Let f be a pseudorandom function, so that for any state S_k the bank can compute $f(S_k)$.

Why is this secure?

We use a reduction argument. Suppose that the counterfeiter can copy money by some means. What does that say about f_k? If it is pseudorandom, then f_k is distinguishable from a random function, so it's not very good at being pseudorandom.

Superdense Coding

is the first protocol we'll see that involves entanglement. Basic information theory (Shannon) tells us that "with n bits you can't send more than n bits of information."

Now we'll see how Alice can send Bob *two* classical bits by sending *one* qubit, though there is a catch: Alice and Bob must share entanglement ahead of time.

In the scenario with no prior entanglement, you can't send more than one bit per qubit.

If Alice sends $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$ to Bob, he can only measure it once in some basis and then the rest of the information in $|\Psi\rangle$ is lost.

Instead, let's suppose that Alice and Bob share a Bell Pair: $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$

We claim that Alice can manipulate her half, then send her qubit to Bob, then Bob can measure both qubits and get two bits of information.

The key is to realize that Alice can get a state orthogonal to the Bell Pair by applying the following gates to her bit:

- NOT $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ which gives us $\frac{|01\rangle + |10\rangle}{\sqrt{2}}$

- A phase change $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ which gives us $\frac{|00\rangle - |11\rangle}{\sqrt{2}}$
- And applying both NOT and a phase change $\frac{|01\rangle - |10\rangle}{\sqrt{2}}$

More specifically, any pair of these four states is orthogonal.

Say Alice wants to transmit two bits X, and Y:

If X = 1, she applies the NOT gate.
If Y = 1, she applies a phase change
Then she sends her bit to Bob.

We can derive her encoding matrix as:

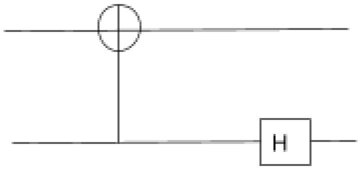
$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 0 & 0 \end{pmatrix}$$

Which makes sense, because each column corresponds to one of the four states we describe above.
(e.g. the second column corresponds to $\frac{|00\rangle - |11\rangle}{\sqrt{2}}$)

For Bob to decode this transformation, he'll want to use its matrix transform:

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

Which corresponds to the gates:
cNOT (2nd controls 1st)
then Hadamard (2nd qubit)



The idea is that Alice transforms the Bell Pair into one of the four entangled states above, then Bob decodes that two-qubit state into one of the four possible combinations of |0⟩ and |1⟩ which correspond to the variables X and Y.

So if Bob receives $\frac{|01\rangle - |10\rangle}{\sqrt{2}}$ applying cNOT gets him $|1\rangle \otimes |-\rangle$, and Hadamard gets him $|1\rangle \otimes |1\rangle$.
if Bob receives $\frac{|00\rangle - |11\rangle}{\sqrt{2}}$ applying cNOT gets him $|0\rangle \otimes |+\rangle$, and Hadamard gets him $|0\rangle \otimes |1\rangle$.

Naturally, we may want to ask: if Alice and Bob had even more preshared entanglement, could Alice send an arbitrarily large amount of information through one qubit?

There's a theorem which answers: No.

It turns out that for every qubit, and any amount of entangled qubits (ebits), you can send two bits of classical information. We show this through the inequality:

1 qubit + ebits ≥ 2 bits

As far as quantum speed-ups go, this isn't particularly impressive, but it is pretty cool that it goes against the most basic rules of information theory established by Shannon himself.

Quantum Teleportation

is a result from 1991 that came as a great surprise. You'll still see it in the news sometimes given its irresistable name. In this lecture we'll over what it can and can't do.

Firstly, what does teleportation mean?

You might think it implies sending qubits instantaneously over distances, but that can't be done (as it violates the causal structure of the universe). "Moving something", or "Putting it on a bus" are less-sexy, but more apt ways of describing it. Fundamentally it means:

It is possible for Alice and Bob to use entanglement plus only classical communication to perfectly transmit a qubit.

The inequality here is almost the converse of the one for superdense coding:

1 ebit + 2 bits $\geq$ 1 qubit

Which is to say, you need one pair of entangled qubits plus two classical bits in order to transmit one qubit.

A more in depth explanation is given in the next lecture, but the gist of it is:
Alice has $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$.
Alice applies some transformation to $|\Psi\rangle$, then measure it.
Alice tells Bob some classical information on the phone.
Bob does some transformations (to his qubit of the entangled pair).
Bob now has $|\Psi\rangle$
At the end, will Alice also have $|\Psi\rangle$?

No. A logical consequence of the No Cloning Theorem is that there can only be one copy of the qubit.

Could we hope for a similar protocol *without* sending classical information?

No. Because of the No Communication Theorem.

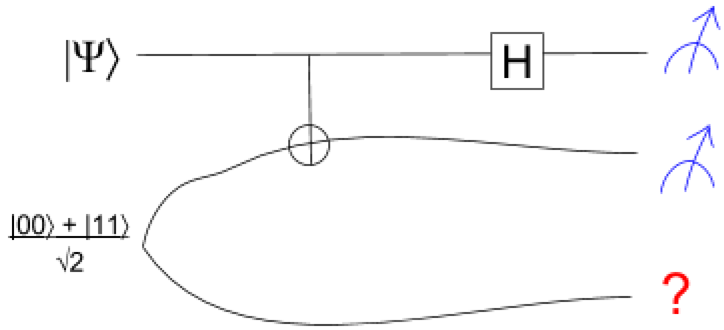
Lecture 10: Thurs Feb 16

Quantum Teleportation (Continued)

So let's say Alice wants to get a qubit over to Bob, but they do not share a quantum communication channel. They do, however, have a classical channel and preshared entanglement.

How should Alice go about this?

You can play around with testing different combinations of operations, and you'd eventually discover that what works is:



The qubit Alice wants to transmit is

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

The entangle qubits form a Bell Pair.

The state starts at:

$$(\alpha|0\rangle + \beta|1\rangle) \otimes \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

Then Alice applies cNOT ($|\Psi\rangle$ controls her entangled qubit):

$$\frac{1}{\sqrt{2}}[\alpha|000\rangle + \alpha|011\rangle + \beta|110\rangle + \beta|111\rangle]$$

Alice Hadamard's her first qubit:

$$\frac{1}{2}[\alpha|000\rangle + \alpha|100\rangle + \alpha|011\rangle + \alpha|111\rangle + \beta|010\rangle - \beta|110\rangle + \beta|001\rangle - \beta|101\rangle]$$

At which point Alice measures both her qubits in the $|0\rangle,|1\rangle$ basis.
This leads to four possible outcomes:

If Alice Sees	00	01	10	11
Then Bob's qubit is	$\alpha|0\rangle + \beta|1\rangle$	$\alpha|1\rangle + \beta|0\rangle$	$\alpha|0\rangle - \beta|1\rangle$	$\alpha|0\rangle - \beta|1\rangle$

We're deducing information about by Bob's state using the partial measurement rule. If Alice sees 00, then we narrow down the state of the entire system to the possibilities that fit, i.e. $|000\rangle$ and $|001\rangle$.

What is Bob's state, if he knows that Alice measured, but not knowing the measurement? It's an even mixture of all four possibilities, which is the Maximally Mixed State. This makes sense given the No Communication Theorem. Until Alice sends information over, Bob's qubit doesn't depend on $|\Psi\rangle$.

Now, Alice sends Bob her measurements via a classical channel.

If the first bit is 1, he applies $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

If the second bit is 1, he applies $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

These transformations will bring Bob's qubit to the state $\alpha|0\rangle + \beta|1\rangle = |\Psi\rangle$.
That means they've successfully sent over a qubit without a quantum channel!

This protocol works *even if* Alice doesn't know what $|\Psi\rangle$ is.

For this protocol to work, Alice had to measure her **syndrome** bits. These measurements were destructive (since we can't ensure that they'll be made in a basis orthonormal to $|\Psi\rangle$, and thus Alice doesn't have $|\Psi\rangle$ at the end.

Something to think about: Where is $|\Psi\rangle$ after Alice's measurement, but before Bob does his operations?

How do people come up with this stuff? I can't picture how anyone trying to solve this problem would even begin their search…

Well it's worth pointing out that quantum mechanics was discovered in 1926 and that quantum teleportation was only discovered in the 90's. These sorts of properties *can* be hard to find. Oftentimes someone tries to prove that something is impossible, and in doing so eventually figures out a way to get it done.

Aren't we fundamentally sending infinitely more information than two classical bits if we've sent over enough information to perfectly describe an arbitrary qubit, since the qubit's amplitudes can be encoded in an arbitrarily complex way?

I suppose, but you only really obtain the information that you can measure, which is significantly less. Amplitudes may exist physically, but they're different from other physical properties like length, in that they seem to act a lot more like probabilities.

For some $\alpha|0\rangle + \beta|1\rangle$ you could say that β is a binary expansion that encodes the complete works of Shakespeare—the rules of quantum mechanics don't put a limit on the amount information that it takes to encode a qubit. With that said, you could also encode the probability of a classical coin to do that.

If we can teleport one qubit, the next question we may want to ask is:

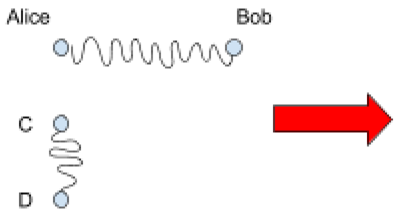
Can we go further? What would it take to teleport an arbitrary quantum state?

There's two things we need to note about the protocol first.

- It destroys Alice's version of $|\Psi\rangle$ (as expected from the No Cloning Theorem)
- It destroys the entanglement (this can be phrased as "Alice and Bob *used* a unit of entanglement")

First, we can notice that the qubit that's transmitted doesn't have to be unentangled.

You could run the protocol and have $|\Psi\rangle$ be half of another Bell Pair. That



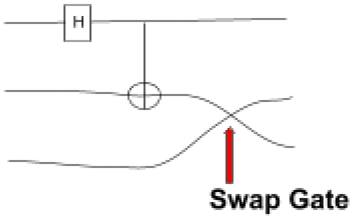
would entangle the fourth qubit to Bob's qubit (You can check this via calculation). That's not a particularly interesting operation, since it lands you where you started, with one qubit of entanglement between Alice and Bob, but it does have an interesting implication.

It suggests that it should be possible to transmit an n-qubit entangled state, by sending each over at a time, thus using n ebits of preshared entanglement.

One further crazy consequence of this is that two qubits don't need to interact directly to become entangled.

A simple example would be this:

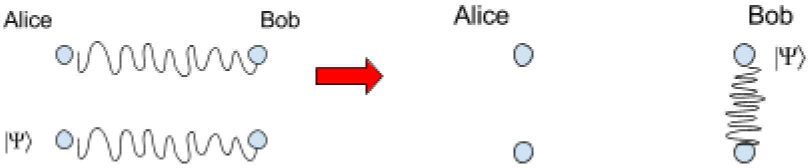
In this circuit the 1st and 3rd qubit become entangled without direct contact.



What does it take for Alice and Bob to get entangled anyways?
The obvious way is for Alice to create a Bell Pair and send one of the qubits to Bob.
In most practical experiments, the entangled qubits are created somewhere between Alice and Bob and are then sent off to them.

An even more surprising consequence is…

Entanglement Swapping

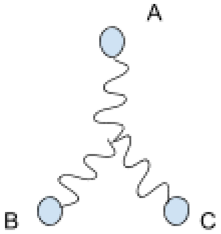


If Alice has a qubit |Ψ⟩ that's entangled with Bob, she can send it over by using an ebit of entanglement.

This only requires measuring Alice's qubits and applying local transformations to Bob's qubits. This process is often used in practical experiments.

By the way, quantum teleportation has been demonstrated experimentally plenty of times.

A few more comments on the nature of entanglement:



We've seen the Bell Pair, and what it's good for. There's an analogue of it to three-party entanglement called **The GHZ State**: $\frac{|000\rangle + |111\rangle}{\sqrt{2}}$. We'll see applications of it later in the course, but for now we'll use it to show an interesting conceptual point.

Let's say that Alice, Bob, and Charlie share 3 classically correlated states. If all three of them get together, they can see that their qubits are classically correlated, and the same can be said if only two of them are together.

But now suppose that Charlie is gone. Can Alice and Bob use the entanglement between them to do quantum teleportation?

No. The trick here is that Charlie can measure without Alice and Bob knowing, which would remove their qubits from superposition, and thus would make the quantum teleportation protocol fail.

A different way to see this is to look at the density matrix of shared by Alice and Bob

$$\rho_{AB} = \begin{pmatrix} ½ & & & \\ & & & \\ & & & \\ & & & ½ \end{pmatrix}$$

And notice that it’s different than the density matrix of a Bell Pair shared by Alice and Bob

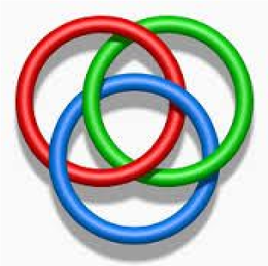
$$\rho_{AB} = \begin{pmatrix} ½ & & & ½ \\ & & & \\ & & & \\ ½ & & & ½ \end{pmatrix}$$

Remember: This gets derived by $|\Psi\rangle\langle\Psi|$

This is actually represents a generalization of…

The Monogamy of Entanglement

Simply put, this means that if Alice has a qubit that is maximally entangled with Bob, then it can’t also be maximally entangled with Charlie.



With GHZ, you can only see the entanglement if you have all three together. This is often analogized to the Borromean Rings (right), a grouping of three rings in a way that all three are linked together, without any two being linked together.

There are other 3-qubit states which aren’t like that…

In the W State, $\frac{1}{\sqrt{3}}$ ($|100\rangle + |010\rangle + |001\rangle$)), there’s *<u>some</u>* entanglement between Alice and Bob, and there’s *<u>some</u>* entanglement between Alice and Charlie, but neither pair is *<u>maximally entangled</u>*.

<u>So how do you quantify how much entanglement exists between two states?</u>

It’s worth noting that we sort of get to decide what we think a measure of entanglement *ought* to mean. We’ve seen how it can be useful to think of quantities of entanglement as a resource, so we can phrase the question as “How many ‘Bell Pairs of entanglement’ is this?”

It’s not immediately obvious whether different kinds of entanglement would be good for different things. That’s actually the case for large multi-party states, but with just Alice and Bob, it turns out that you can just measure in ‘number of Bell Pairs of entanglement’.

Given $\sum_{i,j} \alpha_{ij}|i\rangle_A|j\rangle_B$, how many Bell Pairs is this worth?

Our first observation here should be that given any bipartite state, you can always find a representation of it with Alice and Bob representing their qubits in bases that are orthonormal. So we can write the state as $\Sigma\, \lambda_i|v_i\rangle|w_i\rangle$

such that all $|v_i\rangle$’s are orthonormal,

and all $|w_i\rangle$’s are orthonormal.

We get vectors in this form through…

Schmidt Decomposition

Given a the matrix A = $\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ & \ddots & \\ \alpha_{n1} & \dots & \alpha_{nn} \end{pmatrix}$ representing the entire quantum state.

We can multiply by two unitary matrices to get a diagonal matrix:

$UAV = \Lambda$ U and V can be found efficiently using linear algebra

Essentially this means that we're rotating Alice's and Bob's states into an orthogonal basis.

We then have $\begin{pmatrix} |\lambda_i|^2 \\ : \\ |\lambda_n|^2 \end{pmatrix}$ and we can just ask for the Shannon entropy of this to figure out how many Bell Pairs that's equal to.

Lecture 11: Tues Feb 21

For a classical probability distribution D = (P_1, ..., P_n), we say its **Shannon Entropy** is

$$H(D) = \sum_{i=1}^{n} P_i \, log_2 \frac{1}{P_i}$$

Von Neumann Entropy is generalization of Shannon Entropy from distributions to mixed states.
We say that the Von Neumann Entropy of a mixed state ρ is

$$S(\rho) = \sum_{i=1}^{n} \lambda_i \log_2 1/\lambda_i$$

You could say that Von Neumann Entropy *is* the Shannon Entropy of the vector of eigenvalues of the density matrix of ρ. If you diagonalize the density matrix, it represents a probability distribution over n orthogonal outcomes, and taking the Shannon Entropy of *that* gives you the Von Neumann Entropy of your quantum state.

Another way to think about it:

Say you took all the possible changes in bases of some quantum state. The Von Neumann Entropy of the quantum state would be the minimum of their Shannon Entropies.

$$S(\rho) = \min \begin{cases} H(U\rho U^\dagger) \\ H(U\rho U^\dagger) \\ H(\ \vdots\ ) \end{cases} \qquad \text{each } U\rho U^\dagger \text{ looks like } \begin{pmatrix} x_1 & 0 & \dots \\ 0 & x_2 & 0 \\ 0 & \vdots & \end{pmatrix}$$

Why? Because any measurement basis results in some amount of uncertainty in the result. Most bases will have a degree of probabilistic uncertainty in the measurement, but the basis with the minimum Shannon Entropy can be said to be measurement basis that will provide the maximum amount of information about the quantum state.

So the Von Neumann Entropy of any pure state is 0, because there's always some measurement basis with a certain outcome.

You could choose to measure $|+\rangle$ in the $|0\rangle,|1\rangle$ basis and you'll have complete uncertainty, and an entropy of 1. But if you measure $|+\rangle$ in the $|+\rangle,|-\rangle$ basis, you have an entropy of 0, because you'll always measure it at $|+\rangle$.

So $S(|+\rangle) = 0$.

The Von Neumann Entropy of I/2 is 1.

Similarly, the maximum Von Neumann Entropy of an n-qubit state is N.

We can now talk about how much entropy is in a bipartite pure state.

Entanglement Entropy

Given Alice and Bob share a bipartite, pure state $|\Psi\rangle = \sum_{i,j} \alpha_{ij}|i\rangle_A|j\rangle_B$

To quantify the entanglement entropy, we'll trace out Bob's part, and look at the Von Neumann Entropy of Alice's side, $S(\rho_A)$, by asking: If Alice made an optimal measurement, how much could she learn about Bob's state?

$$S(\rho_A) = S(\rho_B) = H \{ \lambda_i \}$$

^ This is the shannon entropy of these λ's, which you can get by diagonalizing Alice's (or Bob's) matrix, or by putting $|\Psi\rangle$ in schmidt form.

The Entanglement Entropy of $|\Psi\rangle \otimes |\Psi\rangle$ is 0.
The Entanglement Entropy of $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$ is 1.

You can think of entanglement entropy as either:

- The number of Bell Pairs it would take to create the state
- The number of Bell Pairs that you can extract from the state

It's not immediately obvious that these two values would be the same.

A sample calculation...
$|\Psi\rangle = \tfrac{3}{5}\, |0\rangle_A |+\rangle_B + \tfrac{4}{5}\, |1\rangle_A |-\rangle_B$ This is in Schmidt Form: Alice is in the X basis, Bob is in Y.
$E = (\tfrac{3}{5})^2 \log_2 ((5/3)^2) + (\tfrac{4}{5})^2 \log_2 ((5/4)^2)$
$= \sim .942$
That means that if Alice and Bob share 1000 instances of $|\Psi\rangle$, they'd be able to teleport about 942 qubits.

So for any bipartite, pure state we may want to know how many ebits of entanglement it corresponds to. There are two values to consider:

<u>The Entanglement of Formation</u> $E_F(\rho_{AB})$
Which is the number of ebits Alice and Bob need to create one copy of the state
<u>The Distillable Entanglement</u> $E_D(\rho_{AB})$
Which is the number of ebits Alice and Bob could extract from one copy of the state

It turns out that $E_F >> E_D$, which is to say that there exist bipartite, pure states which take a lot of entanglement to make and but that you can only extract a fraction of the entanglement you put in.

We say that a mixed state ρ_{AB} is *separable* if it can be written as a mixture of product states.

i.e. $\rho_{AB} = \sum_i p_i\, |v_i\rangle\langle v_i| + |w_i\rangle\langle w_i|$

The paper (Gurvits, 2003) proves a pretty crazy fact:
If you're given a density matrix, deciding whether ρ_{AB} is separable or entangled is NP-Complete. As a result, there's no nice characterization for telling apart mixed and unmixed states (Since that would prove P = NP).

There are endless paper writing opportunities in trying to classify types of entanglement, since looking at $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$ is much different from looking at $\frac{|000\rangle + |111\rangle}{\sqrt{2}}$

Interpretations of Quantum Mechanics

Now we're in a position to step back and ask,"What is quantum mechanics telling us about the nature of reality?" It should be no surprise that there isn't a consensus, but it's still worth looking at the philosophical debate, as its positions have often corresponded to breakthroughs in quantum mechanics (we'll see an example of this with the Bell Inequality).

Most discussions about the implications of quantum mechanics to our understanding of reality center around The Measurement Problem.

In most physics texts (and in this class), measurement is introduced as an unanalysed primitive that we don't question. There's a fundamental weirdness about it that stems from the fact that quantum mechanics seems to follow both:

1. Unitary Evolution
   when no one is watching $|\Psi\rangle \to U|\Psi\rangle$
2. Measurement
   Which collapses states to a single possibility $|\Psi\rangle \to i$ with probability $|\langle\Psi|i\rangle|^2 = |\alpha|^2$

In other words, quantum mechanics generally seems to work in a way that's gradual, continuous, and reversible most of the time (1), except for during (2), which is the only time we see it work in a way that's probabilistic, irreversible, and sudden. So we can alternatively phrase the question as:

"How does the universe know when to apply unitary evolution and when to apply measurement?"

People have argued about this for about 100 years, and the discussion is perhaps best compared to the discussion surrounding the nature of consciousness (which has gone on for millennia) in that they both devolve into people talking in circles about each other.

It's worth discussing the three main schools of thought, starting with…

The Copenhagen Interpretation

The prefered interpretation of most of the founders of quantum mechanics and was proposed by Bohr (hence the name) and Heisenberg.

It basically says that there are two different worlds: the quantum world and the physical world. We live in the physical world, which only has classical information, but in doing experiments we've discovered that there also exists the quantum world "beneath" it, which has quantum information.

Measurement, in this view, is the operation that bridges the two worlds.

It lets us "peek under the hood" into the quantum world and see what's going on.

Bohr wrote long tracks saying that just to make statements about the quantum world in the classical world is to suppose that there exists a boundary between them, and that we should never make an error in trying to conflate the two. His point of view essentially says "if you don't understand this, then you're just stuck in the old way of thinking, and you need to change".

The next interpretation, which is closely related is…

S.U.A.C. : "Shut Up And Calculate!"

The prefered interpretation of most current current researchers, academics in the field.

It says that at the end of the day, quantum mechanics works (it correctly predicts the results of experiments). If something seems confusing about it, then that's because there's something wrong with our current understanding of it.

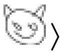
You could say that the Copenhagen interpretation is basically just S.U.A.C. without the S.U. part. After seeing something weird, instead of shutting up, they'll write volumes and volumes about how we can't find a deeper truth.

The popularity of this point of view corresponds to most researchers thinking, "yes, this is how we do things in practice". It seems likely that the popularity of this view isn't going to last forever, because at the end of the day, people will want to understand more about what physical states are truly made of.

Schrödinger's Cat

There were physicists in the 30s and 40s who never accepted the Copenhagen interpretation, namely Einstein and Schrödinger, and they came up with plenty of examples to show just how untenable it is to have a rigid boundary between worlds if you think hard about it.

The most famous of these is Schrödinger's Cat, which first appears with Einstein saying that if you think of a pile of gunpowder as being inherently unstable, you could model it as a quantum state which looks like | ⟩ + | ⟩

Then Schrödinger comes along and adds some flair by asking, "What happens if we create a quantum state that corresponds to a superposition of a state in which a cat is alive and one where the cat is dead?" He allows for the assumption that the cat is isolated by putting it in a box. |⟩ + | ⟩

The point of the thought experiment is that the formal rules of quantum mechanics should apply whenever you have distinguishable states, and thus you should also be able to have linear combinations of such states. It seems patently obvious that at some point we're implicitly crossing the boundary between the worlds, and thus we should have to say something about the nature of what's going on before measurement. Otherwise we'd devolve into extreme solipsism in saying that the cat only exists once we've opened the box to observe it.

Wigner's Friend

Is similar thought experiment. It says that Wigner could be put in a superposition of thinking one thought or another, modeled as $\frac{1}{\sqrt{2}} (|\text{Wigner}_0\rangle + |\text{Wigner}_1\rangle)$.

We can look at the state of him and a friend that's not aware of his state.

$$|\text{Friend}\rangle \otimes \frac{1}{\sqrt{2}} (|\text{Wigner}_0\rangle + |\text{Wigner}_1\rangle)$$

Whichever branch Wigner is in is what he believes (either one thought or the other) after the experiment has been performed. But from his friend's point of view, the experiment hasn't been performed. Then the two can talk, making the state

$$\frac{1}{\sqrt{2}} (|\text{Friend}_0\rangle|\text{Wigner}_0\rangle + |\text{Friend}_1\rangle|\text{Wigner}_1\rangle)$$

But then what happens if another friend comes along, and then another?

The point is to highlight the incompatibility of the perspectives of two observers: one ascribes a pure state other mixed state. We need some way of regarding measurement as fictitious *or* believing in only local truth.

Lecture 12: Thurs Feb 23

Last time we discussed a few interpretations of quantum mechanics and today we cover a few more. The first of these isn't so much an interpretation, but rather a proposal for a new physical theory.

Dynamic Collapse

Says that maybe quantum mechanics isn't a complete theory. It does a good job of describing microscopic systems, but maybe we're not looking at all of the rules that govern reality.

The idea is that there exist some physics rules that we haven't discovered which say that qubits evolve over unitary transformations, but that the bigger the system is, the more likely it will collapse. Thus, we can view this collapse as being a physical process that turns pure states into mixed states.

$\sum_i \alpha_i |i\rangle$ ~> $|i\rangle$ with probability $|\langle\Psi|i\rangle|^2 = |\alpha|^2$

So in the Schrödinger's Cat example, Dynamic Collapse would say that it doesn't matter how isolated the box is. There exists some physical law that says that a system that big would eventually evolve into a mixed state.

$\frac{1}{\sqrt{2}}$ (| ⟩ + | ⟩) ---> $\frac{1}{\sqrt{2}}$ (| ⟩⟨ | + | ⟩⟨ |)

So if you measured in the alive/dead basis, you should be able to distinguish between these two states.

Theoretically you could implement a measurement in any basis of a multi-qubit system. What this means for our cat, is that there should exist a unitary transformation to get the "cat system" into a basis where we can measure any of it's qubits and get 0 if the cat is alive and 1 if the cat is dead.

Professor Aaronson is currently doing research into what other problems you'd have a solution for if you solve this problem (are able to measure in an arbitrary basis). There's already a theorem which says that if you can distinguish between this and that state, then you must have the technological ability to rotate between them.

Which means implementing the Schrödinger's Cat experiment in real life need not involve animal cruelty: if you were able to distinguish between the alive state and dead state, you should be able to rotate the dead cat back into the alive state!

The idea of these Dynamic Collapse theories is that even if you had the technology to distinguish between the two states, a system as big as a cat wouldn't maintain itself in a pure state for a significant amount of time.

The trouble with this is that it's not *really* interpreting quantum mechanics, it's just proposing new laws of physics. Physicists have a high bar for such proposals, and the burden of proof is on you to explain exactly how big a system needs to get to collapse. Fundamentally, there should be implications which we're able to measure the effects of.

The point is that if you propose a Dynamic Collapse theory, the burden is on you to clarify how it works mathematically. Some suggestions include:

- Collapse happens when some number of atoms get involved
    - which is contradictory to our understanding of atoms, which relies on reductionism
- Collapse happen after a certain mass is reached

One famous proposal is the…

Ghirardi-Rimini-Weber Theory (GRW)

which says that each atom has some small probability of collapsing at any point, and that if one atom collapses, the entire system collapses. Thus, the bigger the system, the more likely it collapses.

Just like measuring one qubit of $\frac{1}{\sqrt{2}}(|00...0\rangle + |11...1\rangle)$ will resolve all of the qubits to 0 or 1.

another proposal is the…

Penrose Interpretation

which says that superpositions collapse when enough mass gets involved.

Why mass? mass here ▼ or ▼ mass there

Say we have the superposition of $|*\quad\rangle + |\quad *\rangle$. General relativity tells us that mass curves space-time. Specifically, we know that space-time can be bent like a mattress. That means a mass in one location would make spacetime curve differently than having it somewhere else.

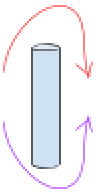
The thing is, no one really knows how to combine general relativity and quantum mechanics, it's one of the biggest open problems in physics. What the Penrose Interpretation is suggesting is that this could be the place where we do so.

The trouble with these theories is that they need to keep adjusting their answers to questions like "How much mass is enough to collapse it?" based on experimental evidence, which keeps producing examples of bigger and bigger states in superposition.

Early on, we discussed the significance of the Double Slit Experiment as performed with photons. People eventually tested it with protons, then molecules, and in 1999 Zeilinger performed it with Buckyballs: molecules large enough to be seen with the naked eye.

To go even further...

Superconducting Qubits



If you take a coil, about 10mm across, and cool it to almost absolute zero, you'll see a current that's in superposition of electrons rotating clockwise or counterclockwise about it.

This constitutes a superposition of billions of particles!

We'll come back to these in time, as they're an important technology for quantum computers.

Penrose has a specific prediction for the scale at which collapse happens, which may be testable in our lifetime, but with GRW, the prediction retreats every time superposition is shown to be possible at a news scale.

A popular position among people who want nature to be simulatable in a classical computer (and thus don't want quantum computers to work) says that:

A frog can be in a superposition of two states. However, a complex quantum computer wouldn't work because systems lose superposition after *sufficient complexity*.

This position is interesting because it could be falsified by building a quantum computer, and reaching falsifiable theories is what moves these discussions from philosophy to science.

<u>What happens if we keep doing experiments and quantum mechanics keeps perfectly describing everything we see?</u>

i.e. we want to not add any new physical laws, but we insist on being realists (saying that there exists a real state of the world without believing that unitary transformations and measurement are separate).

This gets you to…

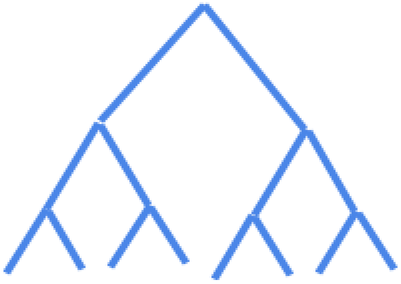
Everett's Many Worlds Interpretation (1957)

Says that the entire universe is a single state, and that the entire history of the universe is the vector $|\Psi\rangle$ that represents reality going through unitary evolution.

You can think of measurement as a special case of entanglement. It's just your brain becoming entangled with the system that you're measuring. A cNOT gate is applied from the system you're observing onto you.

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}}|\text{You}\rangle \quad \text{->} \quad \frac{|0\rangle|You_0\rangle + |1\rangle|You_1\rangle}{\sqrt{2}}$$

Essentially you've now branched into one of the two possibilities.



The universe branches every time that a macroscopically detectable effect occurs. If we were to write down the state of the Earth a month from now, you'd have P(Austin is sunny) + P(Austin is rainy), etc.

We perceive only one branch, but there exist countless other branches where one month later every possible thing that could happen happens.

Some versions of this interpretation chose words carefully to avoid sounding like there exist several physical worlds, but they all imply it. When Everett came up with this as a grad student at Princeton, his advisor told him to remove references about the physical existence of several worlds, because it wouldn't chime with the physics establishment at the time, so he published without it.

Eventually Everett left physics for nuclear work. The only lecture he gave on the topic was at UT decades later when people were finally coming around to the idea. Deutsch, the biggest current advocate of the Many Worlds Interpretation, was there.

One important point to consider is interference between branches.

We don't expect different branches to interfere with one another, because what has happened, happened, and can't be changed. $|0\rangle|You_0\rangle$ shouldn't affect $|1\rangle|You_1\rangle$

This shouldn't need to be a problem. To get the current world, you apply unitary transformations representing every branching between the beginning of time and now. Interference would only happen if two states are reached by applying different unitary transformations. Quantum mechanics says that this is less likely to happen than an egg unscrambling itself (it's thermodynamically disfavored).

But if we take this seriously, keeping in mind:

- Branches never collapse
- The universe is finitely large

Then eventually branches are going to start colliding with one another.
Many Worlds says that this will happen in the timescale of 10^{100} years.

We've said that measurement is the one irreversible part of quantum mechanics, but Many Worlds says it's not. In principle we could apply U^{-1} to get a measurement to unhappen, though like unscrambling an egg, thermodynamics isn't going to make it easy.

The next question we may ask is:
"Where to probabilities come from?"

It's not enough to say that sometimes we see 0 and sometimes we see 1. Quantum mechanics gives very specific probabilities that each will occur, but if the world is just branching once for each observation, then how can we justify these probabilities correlating to anything meaningful?

Everett circumvents this by saying that if the universe split several times, then the probability is connected with the percentage of times it would go to either branch, but many people in the past 50 years don't buy this argument, and have looked for other explanations.

Lecture 13: Tues Feb 28

Everett's Many Worlds Interpretation (Continued)
Everett's Many Worlds Interpretation raises many questions.
Today we'll tackle two of the most important:

1) Where do the (Born) probabilities come from?

In practice we see probabilistic results to experiments. It's the reason that we know that quantum mechanics works in the first place. So people tend to be hesitant about the Everett Interpretation because it's not abundantly clear why these probabilities would arise.

Many Worlders say that there exists a "splitting of the worlds" in such a way that amplitudes of ⅗ and ⅘ would correspond to 9/25th "volume of worldness" going one way, and the other 16/25th going the other.

Some philosophers don't really buy this because if worlds are equal, why wouldn't they just occur with even probabilities? Why bother with amplitudes at all? Many Worlders say that probabilities are just "baked into" how quantum mechanics works. They justify this by arguing that we already agree that density matrixes bake the Born Rule in (since the main diagonal represents Born Rule probabilities).

There's all sorts of other technical arguments that come into play, which boil down to "if nature is going to pick probabilities, they might as well be these," lest we get faster-than light communication, cloning, etc.

There's also been plenty of discussion surrounding the meta-question…

"If there's no experiment that could differentiate the Copenhagen Interpretation from Many Worlds, why bother arguing about it?"

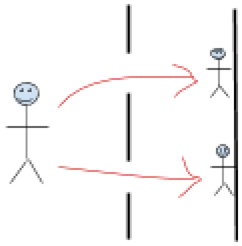
Many Worlders say that the opponents of Galileo and Copernicus could also claim the same about Copernican vs Ptolemaic versions of observations of the planets, since Copernican heliocentrism made no difference to the predictions of celestial movement.

Today we might say that the Copernican view is better because you could fly outside of the solar system and see the planets rotating around the sun; it's only our parochial situation of living on Earth that motivated geocentrism. On that note, it may be harder to think up a physically possible analog for the Many Worlds interpretation, since we can't really get outside of the universe to the see the branching.

There is one neat way you could differentiate the two, though...

Last time we talked about increasing the scope of the Double Slit Experiment. Bringing that thread to its logical conclusion, what if we could run the experiment with a person?

It would then be necessary to say that observers can branch, and that a person is a quantum system. That means it would no longer be enough to use the Copenhagen interpretation.



If you talk to modern Copenhagenists about this they'll take a quasi-solipsistic view, saying that if this experiment were run, "the person being behaving quantumly doesn't count as an observer, only I, the experimenter do."

Another place to consider the differences of interpretations is their relationships with special relativity.

Both the Copenhagen Interpretation and Dynamic Collapse appear to be in some tension with special relativity.

If Alice and Bob share a Bell Pair, and Alice measures her qubit in some basis, Bob's qubit instantaneously collapses to that basis. Sure, Bob won't immediately know the result of Alice's measurement, and thus describes his state as I/2, but that's still a problem.

Simultaneousness for far away things isn't well defined in special relativity, so people argue that Alice's measurement immediately causing a change in Bob's qubit conflicts with it.

You can see this more clearly by taking a frame of reference where Bob's change happens first. How can we say that Alice's measurement caused it?

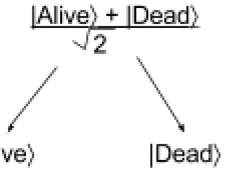
The Many Worlds Interpretation doesn't have to deal with this snag because it doesn't assert that collapse actually happens in the first place. It's ok to view Bob's change as happening first because Alice's measurement didn't cause it, it was just a branching of the universe.

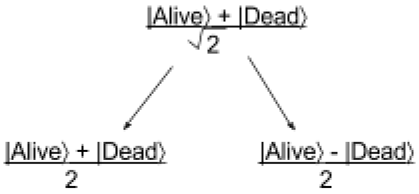
The second question we want to tackle is the **Prefered Basis Problem**. It says:

"Let's say I buy into the argument that the universe keeps branching, well then…"

2) <u>In what basis is this branching occurring?</u>

With a Schrodinger's cat, you can say that the world branches into either the alive state or the dead state.



But it could equally have branched into  instead.

There's a whole field of physics that tries to answer questions like these, called...

Decoherence Theory

which says that there are certain bases that tend to be robust to interactions with the environment, but that most aren't.

So for the example above, decoherence theory would say that an alive cat doesn't easily decohere if you poke it, but that a cat in the ½ (|Alive⟩ + |Dead⟩) state does, because the laws of physics pick out certain bases as being special.

From the point of view of decoherence theory we say that an event has definitely happened only if there exist several records of it scattered all over the place (where it's not possible to collect them all).

This is perhaps best compared to putting an embarrassing picture on Facebook. If only a few friends share it, you can still take it down. On the other hand, if the picture goes viral, then the cat is out of the bag, and deleting all copies becomes an intractable problem.

This is as far as we'll cover Many Worlds/Decoherence.
To pick up the broader conversation about interpretations of quantum mechanics…

You may think that all the options we've seen so far are bizarre and incomprehensible (Einstein certainly did), and wonder if we could come up with a theory that avoids all of the craziness. This leads us to…

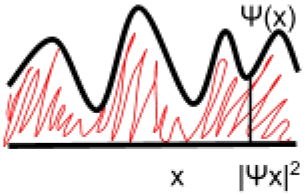
Hidden Variable Theories

which try to supplement quantum state vectors with some sort of hidden ingredients. The idea is to have $\alpha|0\rangle + \beta|1\rangle$ represent a calculation to make a prediction on what the universe *has already* set the qubit to be: either $|0\rangle$ or $|1\rangle$.

The first of these is...

Bohmian Mechanics

which was proposed by Bohm in the 50s.



Normal quantum mechanics says that a particle has some probability of being found at several locations as an amplitude wave. But we now want to also say that there exists a real place where the particle is. Bohm tries to explain how if a particle follows the wave function, that it may continue to do so even if it only truly exists in one place.

There are many rules that could satisfy this property, so there's no experimental way to know which is correct.

Given a quantum state represented as an amplitude vector, when we multiply by a unitary transformation, we want to be able to say "this is the state we are *really in* after the unitary" with probabilities represented as:

$$\begin{pmatrix}\beta_1\\ :\\ \beta_n\end{pmatrix} = \begin{pmatrix}U_{11} & & \\ & \dots & \\ & & U_{nn}\end{pmatrix}\begin{pmatrix}\alpha_1\\ :\\ \alpha_n\end{pmatrix}$$

(β_1)

There are many, many such matrixes. For example you could put (β_n) in every column, which would say that you're always jumping randomly over time in such a way that preserves the Born Rule. You could have been in a different galaxy a planck time ago.

The big selling point of Bohmian Mechanics is that there's only one random decision that has to be made. "God needs to use a RNG to place the hidden variables" at the beginning of time, but afterwards we're just following the Born Rule.

Bohm and others noticed lots of weird consequences of Bohmian Mechanics. It looks nice with just one particle, but problems start to arise when you look at a second. Bohmian Mechanics says that you need to give a definite position for both particles, but people noticed that you can only get that with faster-than-light influence in hidden variables (since Alice's local transformation moves Bob's qubit).

This wouldn't be useful for faster-than-light communication or the like, since hidden variables are explicitly defined as unmeasurable.

When Bohm proposed this, he was super eager for Einstein to accept the interpretation, but Einstein didn't really go for it, because of the sort of things listed above.
What Einstein really wanted (in modern terms), is a…

Local Hidden Variable Theory

where hidden variables can be localized to specific points in space and time.
The idea is that when entanglement is created, the qubits flip a coin and decide, "if anyone asks, let's both be 0," coming up with such answers for all questions that could be asked (infinite bases and whatnot), and that each qubit carries a copy around independently.

This is <u>not</u> Bohmian Mechanics: in 1963 John Bell actually wrote a paper that points out the non-locality of Bohmian Mechanics. Bell says that it would be interesting to show that all hidden variable theories must be non-local, and in fact the paper has a footnote that says that since publication, a proof of this has been found.

This proof is the…

Bell Inequality / Bell Theorem

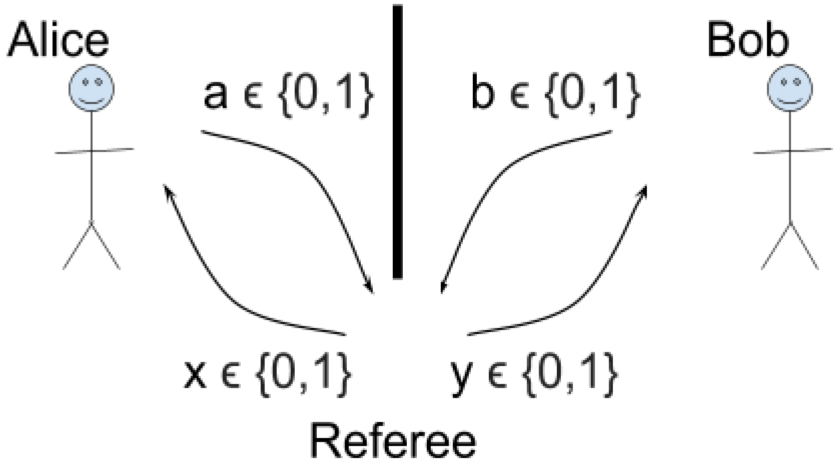
which has changed people's understanding of quantum mechanics perhaps more than anything since the field's inception. It came as a result of Bell philosophizing about the question "Is there an experiment that could follow the rules of quantum mechanics, but would violate the possibility of local hidden variables?"

Bell came up with such an experiment. We'll describe it differently from how Bell did—more computer science-y—as a game between Alice and Bob, where the win probability can be improved through shared entanglement. It's called…

The CHSH Game

named after four people who in 1999 wrote a paper saying "this is what Bell was trying to say." The game doesn't involve quantum mechanics, but quantum mechanics can help us win.

It's a bit of a precursor to quantum computing in that it's one of the first instances of looking to see what basic information processing tasks quantum mechanics can help us solve better.



The idea is that Alice and Bob are placed in separate rooms, and are each given a challenge bit (x and y, respectively) by a referee. Then Alice sends back bit *a*, and Bob bit *b*.

They win the game iff $a + b = xy \ (mod\ 2)$

So if either x or y is 0: *a*, *b* should be the same bit

If $x = y = 1$: *a*, *b* should be different bits

Alice and Bob are allowed to agree on a strategy in advance.

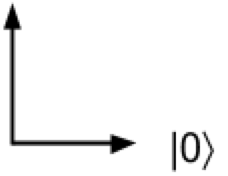
The classical strategy to maximize winning probability is sending the referee $a = b = 0$. They win 75% of the time, losing only if both x and y are 1.

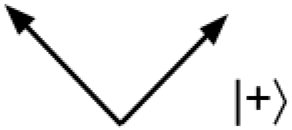
To prove that this is optimal, you'd want to show that introducing randomness isn't going to help. Basically you'd write $a(x) + b(y) = xy \ (mod\ 2)$ such that *a* is a function on x (and *b* on y), and prove that this is going optimal when they're constant functions.

The Bell Inequality is just the statement that the maximum classical win probability for this is 75%.

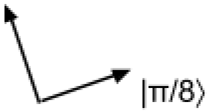
Bell noticed an additional fact though. If Alice and Bob had a pre-shared Bell Pair, there's a better strategy. In fact, the maximum win probability for a quantum strategy is $\cos^2(\pi/8)$ ~ 85%.

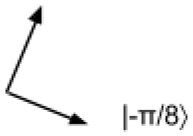
The strategy involves Alice and Bob measuring their entangled qubit based on whether x and y are 0 or 1.

If $x = 0$, Alice measure in 

and if $x = 1$, Alice measures in 

She sets *a* to 0 if she measures $|0\rangle$ or $|+\rangle$
and 1 if she measures $|1\rangle$ or $|-\rangle$

If y = 0, Bob measures in the X basis rotated by π/8 clockwise. 

and if y = 0 rotated by -π/8. 

He sets *b* to 0 if he measures $|\pi/8\rangle$ or $|-\pi/8\rangle$
1 if otherwise

This strategy has the amazing property of making Alice and Bob win with probability $\cos^2(\pi/8)$ for all possible values of x and y.

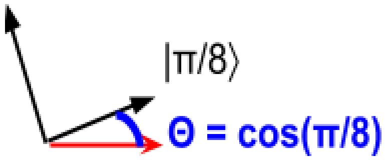
Lecture 14: Thurs March 2

The Bell/CHSH Game (Continued)

Last time we talked about the CHSH Game and how we can use entanglement to create a better strategy than the classical one.

So why does this strategy work 85% of the time?

Lets consider the case where Alice gets $x = 0$ and measures $|0\rangle$. She'll set $a = 0$, and they'll win the game if Bob sets $b = 0$.

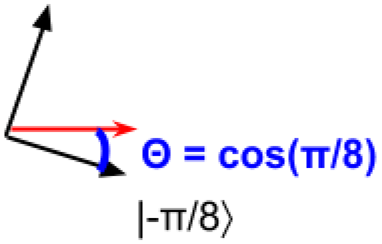


So what are the odds that Bob outputs 0?

Given that Alice measured her qubit already, Bob's qubit collapsed to the $|0\rangle$ state.

So if $y = 0$, Bob measures the $|0\rangle$ state in a basis rotated by $\pi/8$ clockwise. He outputs 0 if he measures $|\pi/8\rangle$. We know the probability of measuring a quantum state in a different basis is the cosine of the angle between the two vectors. Thus, the odds that Bob outputs 0 is $\cos^2(\pi/8) \approx 85\%$.

The same calculation is done for the case where $y = 1$. The angle between vectors is still $\pi/8$. In fact, you can extrapolate this result for all the cases where either x or y is 0.

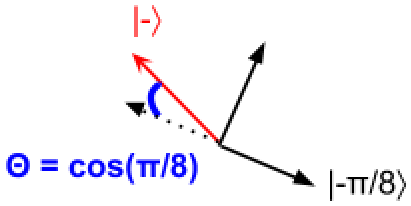


Note that we can assume Alice measured first because of the No Communication Theorem.

The interesting case is where a and b are both set to 1.

This case requires that Alice measured in the $|+\rangle,|-\rangle$ basis and got $|-\rangle$. So what is Bob's probability of getting $|1\rangle$?

Still $\cos^2(\pi/8)$, because the angle between $|-\rangle$ and $-|\pi/8\rangle$ is $\pi/8$, and global phase doesn't matter.



The reason this game relates to hidden variable theories is that if all correlation between particles could be explained as "if anyone asks, we're both 0," you'd predict that Alice and Bob would win only ¾'s of the time (because that's how good they can do by pre-sharing arbitrary amounts of classical information). So you could refute local realism by running this experiment repeatedly—without having to presuppose that quantum mechanics is true.

Does Alice and Bob's ability to succeed more than ¾ of the time mean that they are communicating?

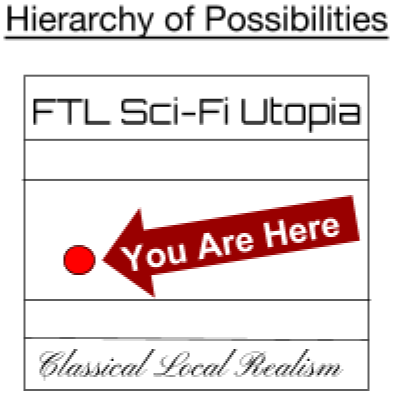
No, we know that's not possible (No Communication Theorem). We can more explicitly work out what Alice and Bob's density matrixes look like over time to check this.

Bob's initial density matrix is $\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ and after Alice measures it's still $\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$.

So in that sense, no signal has been communicated from Alice to Bob. Nevertheless, if you know Alice's measurement and outcome you can predict Bob's measurement to update his density matrix. That shouldn't worry us though, since even classically if you condition on what Alice sees you can change your predictions.

Imagine a hierarchy of possibilities within physics of what the universe allows. You'd have Classical Local Realism at the bottom, where you can determine all outcomes of all measurements you make, and you only need to use probability when you have incomplete information about local objects.

At the top of the hierarchy is a Faster-Than-Light Science-Fiction Utopia where Alice and Bob can communicate instantaneously, you can travel faster than light, and so forth.



A priori people tend to believe that reality must be one or the other, and so reading pop-science articles that negate classical local realism, they think, "Okay, then we must live in a FTL sci-fi utopia."

Instead, the truth is a subtle midterm, which is perhaps so subtle that no science fiction writer would have the imagination to create, where there are no hidden variables, but there's no faster-than-light communication either.

Maybe no science fiction writer ever nailed how our universe works because it's hard to come up with a plot that requires Alice and Bob to win the CHSH game 85% of the time instead of 75%.

If we ran the experiment and Alice and Bob were winning CHSH more than 75% of the time, *and* we kept the assumption that the world is classical, then we would have to suppose that faster-than-light communication is occurring. Instead we suppose the likelier alternative: quantum mechanics is at play.

So where is that cos(π/8) coming from anyways? That seems so arbitrary…

It may seem like that value is simply coming from our particular approach to the problem. Maybe if we came at it another way we could improve on the $\cos^2(\pi/8)$ probability.

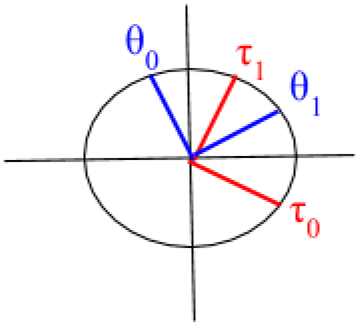
This was answered by...

Tsirelson's Inequality

A cousin of the Bell Inequality from the late 1980s.

Says that even if Alice and Bob share arbitrary amounts of entanglement, quantum mechanics can truly only help you win CHSH up to $\cos^2(\pi/8) \approx 85\%$.

It requires a bit too much machinery to prove here.
What we can do is show that out of strategies similar to the one we used, ours is optimal.



Let’s say that Alice has two angles:

$\theta_{0,}$ the angle she outputs if she receives a 0 and

θ_1, the one she outputs if she receives a 1.

Similarly, Bob has τ_0 and τ_1.

The same rules apply from the solution we constructed earlier for the CHSH game. All we’re doing here is changing the chosen vectors into variables to try and show that there’s no better vectors to chose than the ones we did.

We can then say that the probability of success for Alice and Bob is:

$$P[\text{success}] = \tfrac{1}{4}\,[\,\cos^2(\theta_0 - \tau_0) + \cos^2(\theta_0 - \tau_1) + \cos^2(\theta_1 - \tau_0) + \sin^2(\theta_1 - \tau_1)]$$

^ [1] ^ [2] ^ [3]

Why?

1. We assume each outcome has an equal chance of occurring.
2. Alice and Bob win (in most cases) if they output the same bit, so we measure the cosine between their output angles.
3. Unless, both receive a 1. In this case we measure the chance of their angles being different, which is their sine.

Now we use some high-school trigonometry to get

$$= \tfrac{1}{2} + \tfrac{1}{8}\,[\cos(2(\theta_0 - \tau_0)) + \cos(2(\theta_0 - \tau_1)) + \cos(2(\theta_1 - \tau_0)) - \cos(2(\theta_1 - \tau_1))]$$

And we can abstract out the 2’s on the cosines by understanding that we could adjust our original vectors to account for them.

We can also think of these cosines as the inner product of two vectors.

$$= \tfrac{1}{2} + \tfrac{1}{8}\,[U_0 \cdot V_0 + U_0 \cdot V_1 + U_1 \cdot V_0 - U_1 \cdot V_1]$$

$$= \tfrac{1}{2} + \tfrac{1}{8}\,[U_0\,(V_0 + V_1) + U_1\,(V_0 - V_1)]$$

Since these are all unit vectors, they’re bounded by the norms

$$\leq \tfrac{1}{2} + \tfrac{1}{8}\,[||V_0 + V_1|| + ||V_0 - V_1||]11$$

And from here, we can use the parallelogram inequality to bound it further

$$\leq \tfrac{1}{2} + \tfrac{1}{8}\sqrt{2(||V_0 + V_1||^2 + ||V_0 - V_1||^2}$$

Which equals

$$= 1/2 + (\sqrt{2}/8)\sqrt{4}$$

$$= \tfrac{1}{4}\,(2 + \sqrt{2})$$

Which wouldn’t you know it, brings us to

$$\mathbf{= \cos^2(\pi/8)} \qquad \approx 85\%$$

So $\cos^2(\pi/8)$ really is the maximum winning percentage for the CHSH game.

There's been a trend in the last 10-15 years to study theories that would go past quantum mechanics (past Tsirelson's Inequality), but that would still avoid faster-than-light travel. In such a reality, it's been proven that if Alice and Bob want to schedule something on a calendar, they could agree on a date over only one bit on communication. That's better than can be done under the rules of quantum mechanics!

Testing the Bell Inequality

When Bell proposed his inequality, it was meant only as a conceptual point about quantum mechanics, but by the 1980s it was on it's way to becoming a feasible experiment. Alan Aspect (and others) ran the experiment, and his results were consistent with quantum mechanics.

He didn't quite get to 85% given the usual difficulties that affect quantum experiments, but he was able to reach a high statistical confidence that he was producing wins greater than 80% of the time.

This showed that you can use entanglement to win the CHSH game. Perhaps more impressive is that winning the CHSH game at > ¾ probability provides evidence that entanglement is there.

Most physicists shrugged, already sold on quantum mechanics (and the existence of entanglement), but others looked for holes in the experiment, because it refutes the classical view of the world.

They pointed out two loopholes in the existing experiment, essentially saying "if you squint enough, classical local realism might still be possible":

1. Detector Inefficiency

Sometimes detectors fail to detect a photon or they detect non-existent photons. Enough noise in the experiments could skew the data.

2. The Locality Issue

Taking the measurement and storing it on a computer takes microseconds, which by physics standards isn't negligible. Unless Alice and Bob and the referee are *very* far away from each other, there could be a sort of "local hidden variable conspiracy" going on, where as soon as Alice measures, some particle (unknown to physicists) flies over to Bob and says "hey, Alice's qubit measured to 0. You should measure to 0 too."

Aspect was able to close [2], but only in experiments still subject to [1].
By the 2000s, others were able to close [1], but only in experiments still subject to [2].
In 2016, a bunch of teams were finally able to close both loopholes simultaneously.

There are still people who deny the existence of entanglement, but through increasingly solipsistic arguments. For example…

Superdeterminism

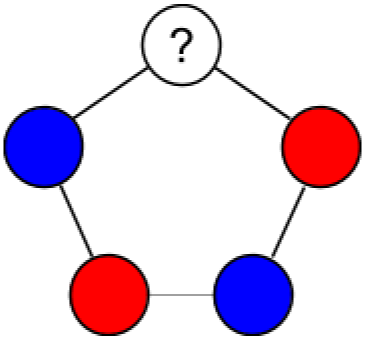
is a theory that says classical local realism is still the law of the land.

Explains the results of CHSH experiments by saying "We only *think* Alice and Bob can choose bases randomly," and that there's a grand cosmic conspiracy involving all of our minds, our computers, and our random number generators with the purpose of ensuring that Alice and Bob win the CHSH game at > ¾ probability by rigging the measurement bases. That's all it does.

Nobel Laureate Gerard 't Hooft advocates superdeterminism, so it's not like the idea lacks serious supporters, but Professor Aaronson is on board with entanglement.

Now we'll look at other non-local games to see what other tasks the Bell Inequality can help with. First, we have…

The Odd Cycle Game



There's a cycle with an odd number of vertices.

Alice and Bob claim that they have a two-coloring of the cycle, but basic graph theory tells us that this isn't possible.

Alice and Bob will agree on a strategy in advance (pre-sharing an arbitrary number of bits/ebits) to try to convince the referee that they've found one anyways.

The referee asks two obvious consistency checks:

- He can ask them both the color of vertex v (in the two-coloring they've found).
    - They pass if $v_A = v_B$
- He can ask Alice the color of vertex v and Bob the color of adjacent vertex w.
    - They pass if $v \neq w$

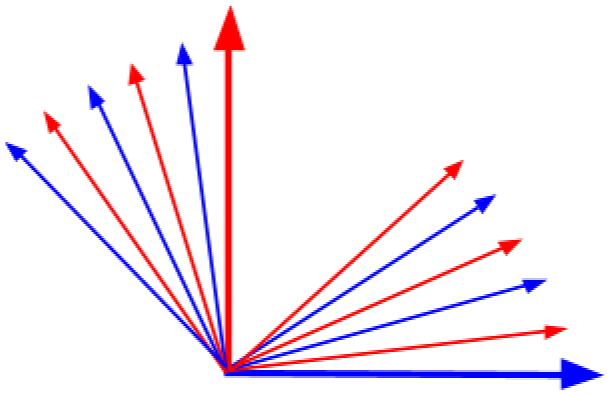
We take one run of the game to mean the referee asking a question once, and getting a response. Without loss of generality, answers are always RED or BLUE, and the cycle has size n.
What strategy provides the best probability that Alice and Bob will pass the test and win the game?

We know that the classical strategy has Pr[win] < 1, because for Alice and Bob to agree on a perfect solution ahead of time, they'd have to find a two-coloring (impossible). The best they can do is agree on a coloring for all but one of the vertices, which gets them $\Pr[\text{win}] \leq 1 - \frac{1}{2n}$.

We claim that with the quantum strategy has $\Pr[\text{win}] \approx 1 - \frac{1}{n^2}$.
First, Alice and Bob share a bell pair, $\frac{|00\rangle + |11\rangle}{2}$.
Alice and Bob each measure their qubit on a basis depending on the vertex they're asked about.

The measurement bases each differ by 2π/n, so they're evenly spaced between $|0\rangle$ and $|1\rangle$.

The first basis has 0 map to answering BLUE and 1 to answering RED. The second has 0 mapped to RED, and 1 to BLUE. They continue alternating.



So when Alice and Bob are asked about the same vertex, they both measure in the same basis, and thus both answer the same color.

When Alice and Bob are asked about adjacent vertices, we get a similar situation to the CHSH game, where the probability of Bob measuring his qubit to the same value as Alice's is the distance between the two vectors. So they answer incorrectly with probability $\sin^2\theta = \sin^2(1/n) \approx \frac{1}{n^2}$.

Another such game is…

The Magic Square Game

Alice and Bob claim that they can fill a 3x3 grid with 0's and 1's such that:

- Every row has an even sum
- Every column has an odd sum

The referee asks Alice to provide a random row of the grid, and Bob to provide a random column.

		0
0	1	1
		0

You can see that this grid can't actually be created by examining the total sum of the grid. The first rule requires it to be even, the second requires it to be odd. That means there's no classical strategy where Alice and Bob always win.

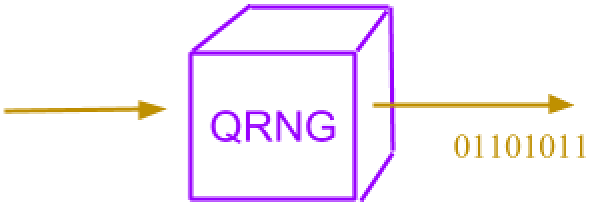
Mermin (the author of our textbook) discovered a quantum strategy where Alice and Bob can always win with only 2 ebits.

We wont write out this strategy.

Lecture 15: Thurs March 9

Until recently, the Bell Inequality was taught exclusively for being historically important, without having any practical applications. Sure, it establishes that you can't get away with a local hidden variable theory, but practically speaking, no one *actually* wants to play the CHSH game. In the last 10 years, however, it's found applications in…

Generating Guaranteed Random Numbers



This is one of the most basic important tasks in computing (or at least in cryptography), and you might think the solution is trivial. After all, you can get a random bit by measuring $|+\rangle$ in the $|0\rangle,|1\rangle$ basis. Easy, right? But this solution isn't good enough for cryptography. Cryptographers are paranoid people, and they want the ability to maintain security, even if the hardware they're on was designed by their worst enemy.

These sorts of assumptions aren't just academic speculation, especially since Snowden. For example, NIST (the National Institute of Standards and Technology) put out a standard for pseudo-random number generation based on elliptic curves to be used for encryption a while back. This standard was later discovered to have a backdoor created by the NSA that would allow them to characterize the output numbers, thus being able to break systems encrypted under this standard.

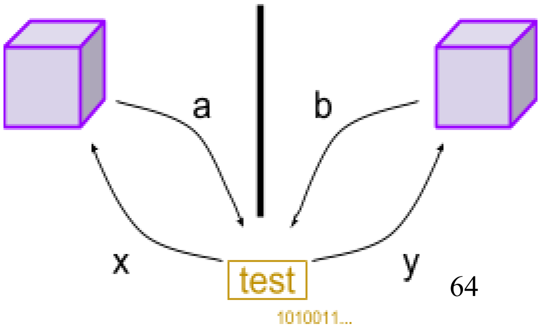
Thus cryptographers want to base their random number generation on the most minimal set of assumptions possible. They want systems that are guaranteed to be truly random, and to be sure that no one had added predictability to the number generation through some sort of backdoor.

You might think that, logically, one can never prove that numbers are truly random, and that the best one can say is that "I can't find any patterns here." After all, you can't prove a negative, and if not the NSA, who's to say that God himself didn't insert a pseudo-random function the workings of quantum mechanics?

Though presumably, if God wanted to read our emails he could do it some other way.

Interestingly, the Bell Inequality lets you certify that numbers are truly random under very weak assumptions, which basically boil down to "No faster-than-light travel is possible." Here's how:

You have two boxes that share quantum entanglement, which presumably were designed by your worst enemy. We'll assume they can't send signals back and forth (say you put them in Faraday Cages).
A referee sends them numbers.
They each return numbers.
If the returned numbers pass a test, we can say that they are truly random.



Vazirani calls this Einstein-Certified Randomness.

The usual way to present the CHSH game is that Alice and Bob prove that they share entanglement, and thus the universe is quantum mechanical. However, winning the game (better than 75% of the time) also establishes that *a* and *b* have some randomness, that there was some amount of entropy generated.

If *a* and *b* were deterministic functions—which is to say that they could be written as $a(x, r)$ and $b(y, r)$, in terms of their input and pure randomness—then you'd have a local hidden variable theory. If *x* and *y* were random, then there must exist some randomness in the outputs.

To put it another way: If Alice has a non-deterministic outcome *and* Alice's state isn't affected by Bob's, then some randomness must be in play.

What is the random result from Alice and Bob? What do you get out?

You can just take the stream of all *b*'s. The measure of entropy is just the Shannon Entropy.

$\{p_x\}_x$ if string x occurs with probability p_x

The total is $\Sigma\ p_x \log_2 1/p_x$

But each output *b* doesn't represent an entire bit of randomness. You'd take these bits and run them through a ***randomness extractor*** which would crunch them down from many sort-of-random bits to fewer very random bits.

David Zuckerman (here at UT) is an expert on this.

There's a problem here:

We need *x* and *y* to be random (CHSH assumes it), which means we're putting in two random bits and getting out less than one. The entropy we put in is greater than the entropy we get out.

In a 2006 paper, Roger Colbeck addresses this by saying that you don't have to give Alice and Bob randomness every time the game is run. You can just input $x = y = 0$ most of the time, and occasionally stick some purely random *x*'s and *y*'s in to prevent Alice and Bob from using hidden variables. If in the test cases Alice and Bob win with classical probability, then discard the results.

So how much entropy needs to be put in?

There was a race to solve this question, first with upper bounds like O($\frac{c*n}{\sqrt{n}}$ for some c< 1) and O($\log^2$n) proposed. Eventually someone asked, "Why not just use a constant amount of randomness to jumpstart the randomness generation, and then feed the randomness outputted by Alice and Bob back in as an input?"

It turns out that this doesn't work because randomness generated by Alice and Bob can be exploited by them if you feed it back to them as input, making further outputs not random.

Remember: We're working under the assumption that
Alice and Bob were designed by our worst enemy!

What you can do instead, if you don't have a limit on the number of devices used, is to feed Alice and Bob's output to two other machines Charlie and David. Using this technique it's possible to generate randomness with a constant amount of seed randomness and only four machines.

You're essentially using the extra devices as "random laundering machines".

Coudron and Yuen did a student project to figure out the number of seed bits necessary, and were able to establish an upper bound of ~200,000 random bits. That's likely still far from the truth: it may be possible with as few as 50.

It's a pretty amazing conceptual fact that playing the CHSH game can create certified randomness, and it's worth mentioning that you can currently go out and buy a quantum RNG from the internet. So you may ask…

What else could you certify about the boxes playing the CHSH game?

It turns out: an *enormous* amount of things.

You can certify that Alice and Bob did a specific sequence of local quantum transformations (up to a change in bases). So just by making them play the CHSH game, you can guarantee they do *any unitary transformation* of your choice. Reichardt and Vazirani describe this as a "classical leash for a quantum system."

One of the main current ideas for how a classical skeptic could verify a quantum solution also appears here. For prime factoring, we can easily verify the solution of a quantum algorithm, but this isn't the case for all problems. Sometimes the only way to verify the solution to a quantum algorithm is by testing the solution on a quantum computer. With this application of a CHSH game, you can guarantee that the quantum computer is behaving as expected.

This brings us nicely to…

Quantum Computation

Having seen all these protocols, we're finally ready to address the holy grail of the field: a programmable quantum computer that can do any short series of operations.

Quantum computation has two distinct intellectual origins:

One comes from Deutsch, who was thinking about experimentally testing Many Worlds (of which he was a firm believer) during his time as a postdoc here at UT. He imagined creating an equal superposition of a brain in configurations where it measured a qubit as |0⟩, and where it measured it as |1⟩. If you measured several times, and always got out the |1⟩ possibility, then we'd have to discard the Copenhagen Interpretation.

But how could we test this? Step 1 would have to be to take a complete description of a human in quantum mechanical terms, and upload it to a computer.

You could presumably instead make an AI that's able to perceive the qubit,
but what would a computer made of quantum mechanics even look like?

The other, less crazy, path to the same association came from Feynman, who gave a famous lecture in 1982 concerned with the question, "how do you simulate quantum mechanics on a classical computer?"

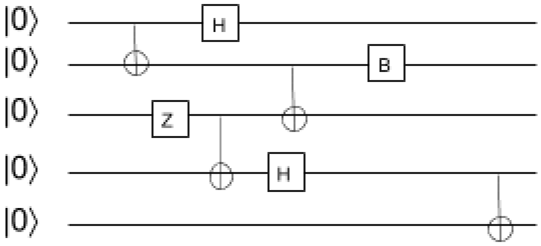
Chemists and Physicists had known for decades that this is hard, because the number of things you need to keep track of increases exponentially with the number of particles. This is the case because, as we know, an n-qubit state can be maximally entangled.

The state $|\Psi\rangle = \sum_{x \varepsilon \{0,1\}^n} \alpha x_n |x\rangle$ must be described by the vector $\begin{matrix}(\alpha_{00.....0})\\ (\alpha_{00.....1})\\ \vdots\end{matrix}$ of length 2^n

Even to solve for the energy of the system, or for the state of some particular qubit, there's no shortcut for reasoning with this enormous vector. So he raised the question, "Why don't we build computers *out of qubits* to simulate qubits?" No one knew if this would be useful for classical tasks as well.

We already have all the tools we need to discuss quantum computers.

The basic picture of a quantum computer is that it's just a quantum circuit, but we're jumping from working with 1, 2, or 3 qubits at a time to *n* (where *n* could equal a million). You apply a sequence of gates on these qubits, each gate acting on a few qubits, then measure some or all of them.



We have a few conceptual points to address:

1. Can we solve anything on a quantum computer that can't be solved on a classical computer?

No. Anything that can be done on a quantum computer can be done on a classical computer too by storing the exponential number of variables that arise when working with qubits.

Quantum computing "only" may violate the Extended Church-Turing Thesis.

2. Why does each gate act only on a few qubits? Where is this assumption coming from?

It's similar to how classical computers don't have gates act on arbitrarily large quantities of bits, and instead use small gates like AND, NOT to build up complex circuitry.

For the quantum case, you could imagine a giant unitary U, which takes qubits, encodes on them the decision version of Travelling Salesman, which is then cNOT'ed to another qubit to get an answer. But given such a definition, how would you go about building U?

Difficulty arises because there exists a staggeringly large amount of possible unitary matrices. You can decompose any U, but it might result in an exponential number of small gates (just like deconstructing an arbitrary Boolean string may require an exponential number of classical gates). We can sort of circumvent this with the…

Accounting Argument

which says that we don't need to consider all unitary matrices, just all of the diagonal ones where the diagonal entries are either 1 or -1. And that's great, because it means you don't need to keep track of 2^(2^n) variables to keep track of U.

Shannon proved that the number of bits it takes to describe a circuit is roughly linear to the number of gates. So almost every unitary matrix would take exponential gates to build.

Interestingly enough, we don't know any examples of such unitary matrices.
But we do know that they're out there!

This tells us something important. In quantum computing, we're not interested in all unitary matrices, only the ones that can be encoded in small circuits requiring a polynomial number of gates.

Lecture 16: Tues March 21

Guest Lecture by Tom Wong

Last time we addressed a few conceptual points about quantum computing. Today we cover two more:

3. What is the role of interference in quantum computing?

   Since quantum amplitudes can cancel out (unlike classical probabilities), we can construct scenarios where the amplitudes for incorrect solutions cancel out with each other, leaving only amplitudes representing the correct solution.

4. What is the role of entanglement in quantum computing?

   We can write a pure state of n qubits as the product,

   $$(\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle) \otimes \ldots \otimes (\alpha_n|0\rangle + \beta_n|1\rangle)$$

   This requires keeping track of only 2n amplitudes, so we can store it efficiently. But an entangled state of n qubits $\sum_{x \,\varepsilon\, \{0,1\}^n} \alpha_x |x\rangle$ requires 2^n amplitudes, which quickly becomes intractable.

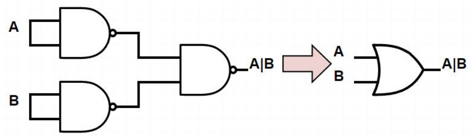
With 300 qubits you'd need more atoms than are available in the universe.

So arbitrarily entangled states can't be simulated well classically. This task requires a quantum computer.

In order to start talking about the construction of quantum computers through quantum gates, we need to cover…

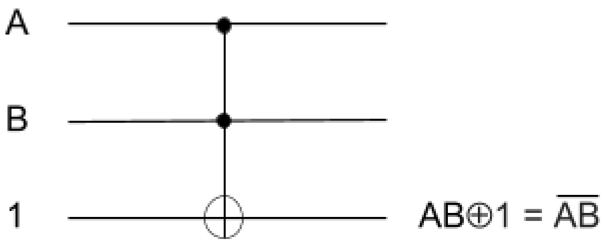
Universal Gate Sets

Classically, you're probably familiar with all of the standard gates (AND, OR, NOT, NAND, etc). A (classical) universal gate set is a grouping of such gates from which you can construct all of the others.



For example, NAND by itself is universal. The diagram on the right shows how you'd construct an OR gate out of NANDs, and the others can all be worked out too.

Similarly, the **Toffoli Gate** universal. The Toffoli Gate, also known as the controlled-controlled-NOT is a three bit gate where if A and B are 1, you flip C. To show that Toffoli is universal, we construct a NAND gate out of one (in the diagram on the right). If a Toffoli can create a universal gate set it must, too, be universal.



By making input C always 1, the output of C is 0 only if both A and B were 1 and the bit was flipped. If A or B or both were 0, the bit is not flipped and the gate returns 1. Thus, we've got a NAND gate.

It's worth noting that since Toffoli is reversible—given the outputs A, B, $AB\oplus C$ we can recover inputs A, B, C—which means we can use it as a quantum gate too. Thus you can see that a quantum computer can do anything a classical computer can do, because one can implement a classical universal gate set.

Now let's talk about *Quantum* Universal Gate Sets:

which we define as a set of gates that allows you to approximate any unitary to any desired precision. An important theorem on the subject is the…

Sololvay-Kitaev Theorem

which says that with any universal gate set, we can approximate a unitary on n qubits to precision using $O(2^n \text{ polylog}(1/\varepsilon))$ gates.

There are plenty of ways that a gate set can fail to be universal.

1. Your gate set doesn't create interference/superposition

   Ex: {cNOT} can only flip between |0⟩ and |1⟩. It can maintain superposition, but it can't create any.
2. Your gate set has superposition, but is missing entanglement

   Ex: {Hadamard} can create superposition, but it should be obvious that the gate can't create entanglement since it only acts on one qubit.
3. Your gate set only has real gates

   Ex: {cNOT, Hadamard} is getting closer, but neither can reach positions with non-real values.
4. Your gate set is "only a stabilizer set"

   We're not going to go in depth with the concept of stabilizer sets. What's important to know is that a set like $\{\text{cNOT, Hadamard, } P = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}\}$ fails because it's efficiently simulated by a classical computer (by the **Gottesman-Knill Theorem**). This property prevents it from getting speedups relative to a classical computer.

Are there any other ways to fail to be universal?

That's an open question!

So what *is* universal?

It turns out that if you replace Hadamard in the above example with almost anything else, the set becomes universal.

So $\{\text{cNOT}, R_{\pi/8} = \begin{pmatrix} \cos(\pi/8) & -\sin(\pi/8) \\ \sin(\pi/8) & \cos(\pi/8) \end{pmatrix}, P\}$ is universal.

Also, {Toffoli, Hadamard, P} is universal.

Any two unitaries picked at random will likely be universal.

Quantum Complexity

There's two major ways we look at the complexity of quantum algorithms

The circuit complexity of a unitary is the size of the smallest circuit that implements it. We like unitaries with polynomial circuit complexity. This can be difficult to find: it's a gate-set-dependent measure. At best we usually only get upper/lower bounds, so instead we tend to use…

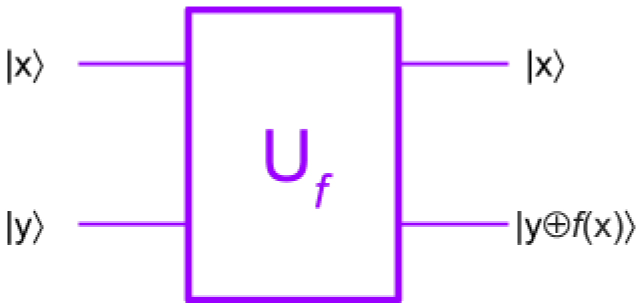
Query complexity, the number of calls the algorithm makes to an oracle (or black box function). The idea is that your oracle takes a bit and outputs a bit $f : \{0, 1\} \to \{0, 1\}$. Classically you'd have a bit go $x \to f(x)$, but we replace this with quantum states $|x\rangle \to |f(x)\rangle$.

Or rather, we *want* to replace it with quantum states, but we run into a bit of trouble because such a transformation is not unitary. What we have to do instead is use an extra answer/target qubit.

So we give the black box two qubits: *x*, which stays the same, and *y*, which receives the answer.

$$|x, y\rangle \;\to |x, y \oplus f(x)\rangle$$

A lot of times we ignore the answer qubit by moving the phases around. So let's say we prepare the answer qubit as |-⟩.



We start with $|x,-\rangle = \frac{1}{\sqrt{2}}(\,|x,0\rangle - \;|x,1\rangle)$

Applying U_f gets us $\frac{1}{\sqrt{2}}(\,|x,0 \oplus f(x)\rangle - \;|x,1 \oplus f(x)\rangle)$

Which equals $\{ \frac{1}{\sqrt{2}}(\,|x,0\rangle - \;|x,1\rangle)$ if $f(x) = 0$

$\{ \frac{1}{\sqrt{2}}(\,|x,1\rangle - \;|x,0\rangle)$ $\quad f(x) = 1$

Which we can rewrite as $(-1)^{f(x)}\,|x,-\rangle$

This lets us avoid dealing with the answer qubit and just use the "phase oracle".

$$|x\rangle \;\to (-1)^{f(x)}|x\rangle$$

Now we're finally ready to tackle our first quantum algorithm.

Deutsch's Algorithm

We're given two unknown bits, b_0 and b_1.

Given an index x ϵ {0,1}, our oracle returns the bit. i.e. $f(x) = b_x$

What we want to know is, "What is the parity of these bits?"

Parity is whether the bits have different values, so $b_0 + b_1$ (mod 2) or $b_0 \oplus b_1$

Classically, this would take two queries since we need to know both bits.

Quantumly, Deutsch's Algorithm can do it in one.

Start with a qubit at |0⟩, Hadamard it, then do a query which applies a phase change to each part depending on the value of the function.

$$|0\rangle \longrightarrow \frac{1}{\sqrt{2}}(\,|0\rangle + \;|1\rangle) \longrightarrow \frac{1}{\sqrt{2}}(\,(-1)^{f(x)}\,|0\rangle + \;(-1)^{f(x)}\,|1\rangle)$$

We can substitute in the bits.

$$= \tfrac{1}{\sqrt{2}}(\,(-1)^{b_0}\;|0\rangle \;+\; (-1)^{b_1}\;|1\rangle)$$

Then drag out b_0.

$$= \tfrac{1}{\sqrt{2}}(-1)^{b_0}\ (\,|0\rangle\ +\ (-1)^{b_1-b_0}\,|1\rangle)$$

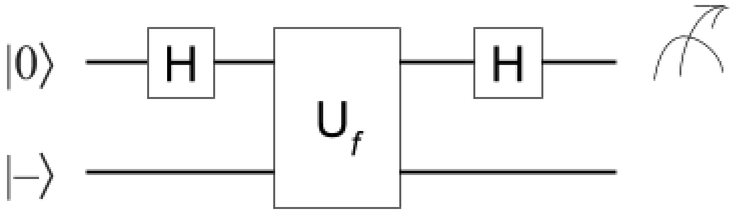
So now if we have $b_0 = b_1$ we get $(-1)^{b_0}\ \tfrac{1}{\sqrt{2}}\,(\,|0\rangle + |1\rangle)$

$b_0 \neq b_1$ $\quad (-1)^{b_0}\ \tfrac{1}{\sqrt{2}}\,(\,|0\rangle - |1\rangle)$

We can ignore the phase out front since global phase doesn't affect measurement, and then Hadamard again to get our quantum states back in the $|0\rangle$, $|1\rangle$ basis.

Now the $b_0 = b_1$ case becomes $|0\rangle$

and the $b_0 \neq b_1$ case becomes $|1\rangle$



The complete quantum circuit is drawn to the right. If the bits had parity we measure 1, if they don't we measure 0.

Lecture 17: Thurs March 23

People often want to know where the true power of quantum computing comes from.

- Is it the ability of amplitudes to interfere with one another?
- Is it that entanglement gives us 2^n amplitudes to work with?

But that's sort of like dropping your keys and asking "what made them fall?"

- Is it their proximity to the Earth?
- Is it the curvature in space-time?

You could come up with all sorts of answers that are perfectly valid.

Last time Tom demonstrated the existence of universal gate sets.

It's worth mentioning that we don't have a criteria to characterize which sets of gates are universal and which aren't. Not many people in the field care about figuring out this particular open problem, since "we have universal gate sets that work, so just roll with it," but it would be nice to know, and you should go figure it out anyways.

It seems like our rules for universal gate sets are just avoiding *certain* bad cases. Do we have formal proof that they work?

Yes. There's a paper from the 90s by Yaoyun Shi on the subject, but it's out of scope for this class.

In designing quantum algorithms, we're ultimately looking to minimize the number of gates required to implement them. That problem turns out to be insanely hard for reasons that have nothing to do with quantum mechanics.

"What's the smallest circuit that solves Boolean satisfiability?"
is a similarly hard problem, for reasons related to P vs NP.

So people design quantum algorithms that center around query complexity. This abstracts away part of the problem by saying:

"There's some Boolean function f : $\{0,1\}^n$ —> $\{0,1\}$ and we're trying to learn something about f."

You might want to learn:

Is there some input x where $f(x) = 0$?
Is there some symmetry in the solution?
Etc.

More importantly, we want to know how many queries it takes to solve such a problem.

In this model we abstract out the cost of gates that don't do queries.

To be precise, we map queries as

$|x, a\rangle \rightarrow |x, a \oplus f(x)\rangle$ since the transformation must be unitary.

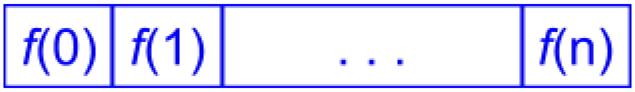
But it can also be thought of as

$|x\rangle \rightarrow (-1)^{f(x)}|x\rangle$

Before we jump into a few quantum algorithms, it's worth asking,
"Why do we care about this model? You're debating how you'd phrase your wishes if you found a genie. Who cares?"

You can think of a black box as basically a huge input. Querying *f*(i) means looking up the ith number in the string.

That allows us to break a problem down to "if you want to do an unordered (or ordered) search, how many queries do you need?"
This is much more reasonable to compute that the alternative.



Another way to think about it:

Imagine I'm writing code, and I have a subroutine that computes *f*(x). How many times do I need to call the subroutine to find some information about *f*?

Under the assumption that we can only learn information about *f* by looking at it's output.

There's a technical question we need to answer…
Suppose we know that there exists a fast algorithm to implement *f*. Could you then implement the black box behavior $|x, a\rangle \rightarrow |x, a \oplus f(x)\rangle$?

Keep in mind that quantum circuits have to be reversible. We're essentially asking what constraints arise from that.

We know *f* must be injective.

What if *f* was a one-way injective function? i.e. there's a unitary C such that $C|x\rangle = |f(x)\rangle$
The problem with this is that a small circuit for C implies that there's a small circuit for C^{-1}, which would imply that there's a small circuit for *f*(*x*), such that $C^{-1}||f(x)\rangle = |x\rangle$.

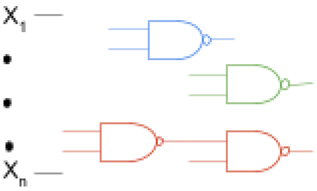
It's worth noting that even though quantum computing can break a few supposedly one-way functions, like finding prime factors, it doesn't provide evidence against the existence of *any* one-way functions.

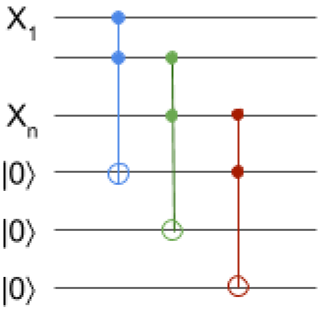
Mapping *f*(*x*) and erasing *x* is much harder for a computer that necessitates reversible circuits. You *could* do $|x, 0\rangle \rightarrow |x,\ f(x)\rangle$ which doesn't let you reverse *f*.

To invert *f* you would need to know that *x* is $|x,0\rangle = C^{-1}|x, f(x)\rangle$.

Tom showed that a reversible circuit can always simulate a non-reversible circuit, since Toffoli can simulate NAND. However, in reversible computing erasing is expensive.

Imagine a classical circuit (without loss of generality, let's say it's a cluster of NAND gates).





You could simulate this as a reversible circuit by having each NAND replaced with a suitable Toffoli. The problem with this is that you'd get all sorts of undesired results in the intermediate bits—the technical name for this is *garbage*. Yes, really.
A truly universal algorithm must produce no garbage, because garbage can prevent the desired interference pattern from showing up.

Garbage is the bane of quantum computing, because the point of quantum algorithms is to create these patterns of interference.

For example, what's the difference between having $\frac{|0\rangle + |1\rangle}{\sqrt{2}}$ and having $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$, but treating the second qubit as garbage?

The garbage creates unwanted entanglement. Looking at a part of the superposition basically turns it into a mixed state.

Suppose you have a circuit to compute *f*. How do we get a circuit that maps $\sum_x \alpha_x |x, 0\rangle \rightarrow \sum_x \alpha_x |x, f(x)\rangle$ without all the garbage? In the 70s, Bennett invented a trick for this called…

Uncomputing

Let's say I have some circuit that maps

$C\,|x, 0, \ldots, 0\rangle = |x, gar(\mathrm{x}), f(x)\rangle.$

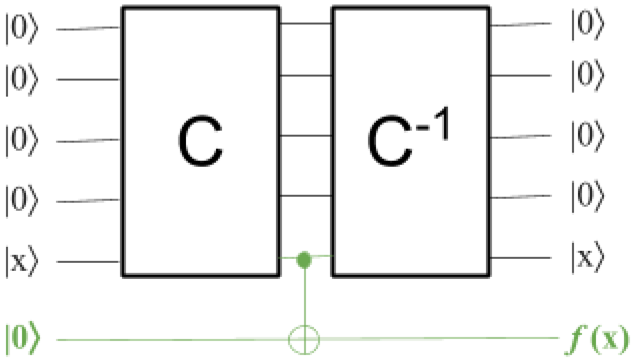
First, run the circuit, C.

Then cNOT x. (make a copy of it in a safe place)

Then run the inverse circuit, C^{-1}.

The reason we can copy *x* in spite of the No Cloning Theorem is that we're assuming that there's a classical answer. This wont work if the output is a general quantum state.

This justifies the quantum query model because if we can compute *f* at all, then we do have the ability to map $|x, a\rangle = |x, f(x)\rangle$.



With that out of the way, we're ready to talk about some quantum algorithms.

Deutsch's Algorithm

computes the parity of two bits with one query. (the parity of n bits would require n/2 queries).

It basically involves making a state like $\frac{1}{\sqrt{2}}\left((-1)^{f(x)}\,|0\rangle + (-1)^{f(x)}\,|1\rangle\right)$ and querying it in the |0⟩,|1⟩ basis.

It uses the *phase kickback trick* to measure phase change.

The basic idea of the phase kickback trick is that we have a quantum oracle that does $\sum \alpha_x |x, y\rangle \rightarrow \sum \alpha_x |x, y \oplus f(x)\rangle$ but we'd rather get a final state in the form $\sum \alpha_x (-1)^{f(x)} |x\rangle$. To accomplish this we put |-⟩ in second register. U_f gives us $\frac{|0\rangle|-\rangle - |1\rangle|-\rangle}{\sqrt{2}}$ and we can interchange |0⟩ - |1⟩ and |1⟩ - |0⟩.

There's a generalization of this, called…

The Deutsch-Jozsa Algorithm

Assume a black box computes f: $\{0,1\}^n \rightarrow \{0,1\}$,and that *f* is either:

- a constant function All outputs are 0 or all outputs are 1
- a balanced function Same number of 0's and 1's in output

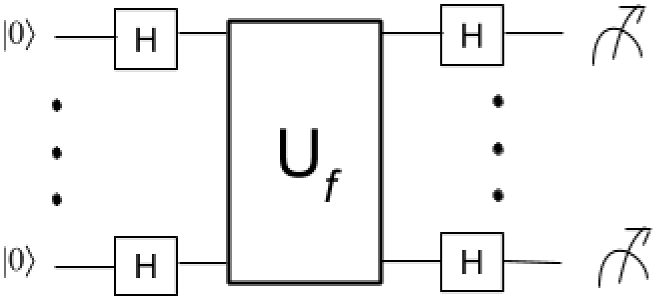
The problem is to decide out which.

Classically, you could look at $2^{n-1}+1$ cases of the function. If all inputs match, then the function is constant. You can improve this through random sampling. On average, you'd need about 5 or 6 queries to get an answer with a sufficiently small probability of error.

We'll see how a quantum algorithm can solve this perfectly with only one query.

Truth is, this isn't a hard classical problem, and so we wont get that big of a speed up. This is why, initially, people didn't care about quantum computing. They figured all advantages would be in the same vein.

Here's the quantum circuit for it:



You'll begin to notice that some patterns appear a lot in quantum algorithms.

- you start by putting everything in superposition
- Then query f, mapping each x to $(-1)^{f(x)} |x\rangle$
- Then measure (from the superposition) the information that we want to know.

If you can't figure out what's next in a quantum algorithm, a round of Hadamards is always a good guess.

So we want to know the probability of getting back the state $|00\ldots0\rangle$.

Let's call the circuit C. We can compute $|\langle 00\ldots0|\, C\, |00\ldots0\rangle|^2$

What's the final amplitude of the C state?

Well H maps $|0\rangle \rightarrow |+\rangle$ and $|1\rangle \rightarrow |-\rangle$.

For an arbitrary $|x\rangle$ it'll map $|x\rangle \rightarrow \frac{|0\rangle + (-1)^x |1\rangle}{\sqrt{2}}$

and H maps a given string $|x_1, \ldots, x_n\rangle \rightarrow \frac{|0\rangle + (-1)^x |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle + (-1)^x |1\rangle}{\sqrt{2}} \otimes \ldots \otimes \frac{|0\rangle + (-1)^x |1\rangle}{\sqrt{2}}$.

After the oracle is applied, you get.

$$= \frac{1}{\sqrt{2^n}} \sum_{y\, \varepsilon\, \{0,1\}^n} (-1)^{x \cdot y} |y\rangle$$

Note: x·y is their inner product. Pick up a phase if $x_i = y_i = 1$.

So keeping track of each basis state individually, you get

$$\frac{1}{\sqrt{2^n}} \sum_{x\, \varepsilon\, \{0,1\}} (-1)^{f(x)} |x\rangle$$

And after another Hadamard

$$\frac{1}{\sqrt{2^n}} \sum_{x\, \varepsilon\, \{0,1\}^n} (-1)^{f(x)} \frac{1}{\sqrt{2^n}} \sum_{y\, \varepsilon\, \{0,1\}^n} (-1)^{x \cdot y} |y\rangle$$

A shortcut to simplify this is to ask, “What is the amplitude when $y = |00\ldots0\rangle$?”

It would be $\frac{1}{2^n} \sum_{x \, \varepsilon \, \{0,1\}} (-1)^{f(x)}$

What does the value of this have to do with f being constant?
If f is constant, then this is either 1 or -1.
If f is balanced, then this is 0.

The first problem we’ll see next time is

The Bernstein-Vazirani Problem

Given a black box function f : $\{0,1\}^n \rightarrow \{0,1\}$
and a promise that $f(x) = s \cdot x$ (mod 2) for some secret string $s \in \{0,1\}^n$

The problem is to find s.

Classically, you could get an answer one bit at a time by querying
$f(1000) = s_1$
$f(0100) = s_2$
$f(0010) = s_3$
$f(0001) = s_4$

But there’s no algorithm that can do better, since each query can only provide one bit of information.

The Bernstein-Vazirani Algorithm, however, can solve this quantumly with only one query.

Lecture 28: Tues May 2

Today we'll see a beautiful formalism for quantum error correction that has many roles in quantum computation.

Last time we discussed the **Quantum Fault Tolerance Theorem**, which says that even if *all* qubits in a system have some rate of noise, by:

- doing a bunch of gates in parallel
- applying measurement
- discarding bad qubits and replacing them
- And doing this all hierarchically (i.e. having layers of error atop one another)

we'll *still* be able to do quantum computation, and the cost will be asymptotically reasonable

$T \rightarrow O(T \log^{C} T)$

This theorem set the research agenda for a lot of experimentalists, who began focusing on attempts to minimize error. Once we can decrease error past a certain threshold, we'll be able to push it arbitrarily small by repeatedly applying our error correction techniques.

The best gauge of how research in quantum computing is going is the *reliability of qubits*. Journalists often ask about things like the number of qubits, or "can you factor 15 into 3 and 5?" but more important is crossing the threshold which would allow us to get arbitrarily small error.

We're not there yet, but lots of progress is being made in two fronts:

1. Making qubits more reliable

Initially, ε (each qubit's probability of failing at each time step) was close to 1, and the quantum state would barely hold at all. The decoherence rates of IBM's Quantum Experience, for example, wouldn't have been possible ten years ago.

John Martinez, with Google, has been able to get ε down to 1/1000 with a small number of qubits. That's already past the threshold, but adding more qubits creates more error, so the trick is to find a way to add qubits while keeping error down.

2. Creating better error correction codes

There are many tradeoffs here. If you used a quantum error correction code that used thousands of physical qubits for each logical qubit, you could get decoherence down to 3-5%.

We're likely to soon see quantum error correction used to keep a logical qubit alive for longer that the physical qubits below it. People are close to figuring this out, but it's not *quite* there yet.

Stabilizer Circuits

are circuits that can be made out of only the gates cNOT, Hadamard, and $P = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$

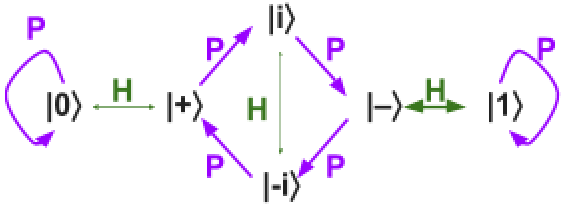
Stabilizer Sets

are sets of states that such a circuit can generate, starting from $|00\ldots0\rangle$.

These came up when we discussed quantum universal gates. It's not obvious that this definition wouldn't cover every quantum state. The Bell Pair is such a state, as are the states arising in Superdense Coding or Quantum Teleportation.

If you play around with these gates, you'll notice that tend to reach a discrete number of states, and never anything between them. You'll also notice that for an arbitrary number of qubits n, when these qubits form superpositions over s pure states, it follows that $|s| = 2^k$ for some k, and s is always a subspace $s \leq F_2^n$.

With only one qubit, you can only reach 6 states, as is shown (right).
We call these the 1-qubit **stabilizer states**.



What about with two qubits?

You'll find that the states you can reach, like $\frac{|00\rangle + i\,|11\rangle}{\sqrt{2}}$ or $\frac{|01\rangle - i\,|10\rangle}{\sqrt{2}}$, follow a specific pattern: For any non-zero α_x, α_y => $|\alpha_x| = |\alpha_y| = \frac{1}{\sqrt{|S|}}$

In other words, all basis states that occur with non-zero amplitudes have the same absolute value. Measuring any of these in the |0⟩,|1⟩ basis will either produce: |0⟩ 100% of the time,
|1⟩ 100% of the time,
or a 50-50 chance of producing |0⟩ or |1⟩.

So what gives?
Before we answer that, we need to define a few things.

A unitary U *stabilizes* a pure state |Ψ⟩ if U|Ψ⟩ = |Ψ⟩.

This only holds for positive eigenstates of U. Global phase matters here!
If U|Ψ⟩ = –|Ψ⟩, it does not stabilize |Ψ⟩.

Notice that if U and V both stabilize |Ψ⟩, then any combination of them *also* stabilizes |Ψ⟩. Also, the identity matrix, I, stabilizes everything.

This means that all the unitaries that stabilize |Ψ⟩ form a group.

We already know that unitaries have inverses and are associative.

The next ingredient we need are the **Pauli Matrices**.

These four matrices come up a lot in quantum physics. For example, you can use them to stabilize the Bloch Sphere. They are:

$$\mathbf{I} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \quad \mathbf{X} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \quad \mathbf{Y} = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} \quad \mathbf{Z} = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

Notice that they match up with the errors that can occur.

No error
$\mathbf{I}|1\rangle = |1\rangle$ $\quad \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix} = \begin{pmatrix}0\\ 1\end{pmatrix}$

Bit flip
$\mathbf{X}|1\rangle = |0\rangle$ $\quad \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix} = \begin{pmatrix}0\\ 1\end{pmatrix}$

Phase flip
$\mathbf{Z}|1\rangle = -|1\rangle$ $\quad \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix} = \begin{pmatrix}0\\ -1\end{pmatrix}$

and Both
$\mathbf{Y}|1\rangle = -i|0\rangle$ $\quad \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix} = \begin{pmatrix}-i\\ 0\end{pmatrix}$

That's not a coincidence!

The Pauli Matrices satisfy several beautiful identities.

$X^2 = Y^2 = Z^2 = I$ $\quad XY = iZ \quad YX = -iZ$

$YZ = iX \quad ZY = -iX$

$ZX = iY \quad XZ = -iY$

If you've seen the quaternions, you may notice that they satisfy the same kinds of relations.
This is also not a coincidence! Nothing is a coincidence in math!

Also, all of them are unitary and Hermitian.

So what does each of them stabilize?

I stabilizes everything
–I stabilizes nothing Remember: global phase matters, so $-I|\Psi\rangle \neq |\Psi\rangle$.
X stabilizes $|+\rangle$
–X stabilizes $|-\rangle$
Z stabilizes $|0\rangle$
–Z stabilizes $|1\rangle$
Y stabilizes $|i\rangle$
–Y stabilizes $|-i\rangle$

So the six 1-qubit stabilizer states each correspond to a Pauli Matrix.

For a given *n*-qubit pure state $|\Psi\rangle$, we define $|\Psi\rangle$'s **stabilizer group** as:

The group of all tensor products of Pauli Matrices that stabilize $|\Psi\rangle$.

We know this is a group since being Pauli (and being a stabilizer) is closed under multiplication.
Additionally, this group is abelian.

For example, the stabilizer group of $|0\rangle$ is { I, Z } closed because $Z^2 = I$
and that of $|+\rangle$ is { I, X }

The stabilizer group of $|0\rangle\otimes|+\rangle$ will be the product of those groups

{ $I\otimes I$, $I\otimes X$, $Z\otimes I$, $Z\otimes X$ } as a convention we omit the ⊗'s { II, IX, ZI, ZX}

For a slightly more interesting example, what's the stabilizer group of a Bell Pair?

We know XX is in it because $\frac{X|0\rangle\otimes X|0\rangle + X|1\rangle\otimes X|1\rangle}{\sqrt{2}} = \frac{|11\rangle + |00\rangle}{\sqrt{2}} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$.

The same argument can be made for –YY.

We can get the last element by doing component-wise multiplication: XX * –YY = –(iZ)(iZ) = ZZ

So the stabilizer group of $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$ is { II, XX, –YY, ZZ }

You can likewise find the stabilizer group of $\frac{|00\rangle - |11\rangle}{\sqrt{2}}$ to be { II, –XX, YY, ZZ }

So now, here's an amazing fact:

The stabilizer states on *n* qubits are exactly the states with a stabilizer group of size 2^n.

We wont see a proof of this, only an intuition for why it’s true.

So the 1-qubit stabilizer states are those with 2 elements in their stabilizer group.
The 2-qubit stabilizer states are those with 4 elements in their stabilizer group.
And so forth.

This is a completely different characterization of stabilizer states, a structural one. It tells us what invariant is being preserved without any mention of quantum mechanics.

Furthermore, you make want to know
What is the size of the generation sets?

i.e. the minimum number of elements it would take to produce all others (using multiplication)

It’s *n* ± the tensor product of the Pauli Matrices
So to specify a stabilizer group, you only need to specify a generator state of size *n*, and this group uniquely determines the state.

(X X)
So for the Bell Pair, you could give (Z Z), which is enough to generate the group { II, XX, –YY, ZZ }

Now we get to a crucial point:
How many qubits are necessary to represent such a stabilizer set?

(Z I X Y)
So given (Y X Z Z) , how many bits of information is this?
(I I X X)

It’s O(n), or more specifically: $2n^2 + n$
bits to represent each Pauli —^ ^ ^ number of ± signs to keep track of
|
| number of Pauli Matrices

Writing out the entire group would have otherwise taken 2^n bits. This is (one part of) why the stabilizer formalism is important.

There’s an important result from 1999,
The Gottesman-Knill Theorem
which says that stabilizer circuits acting on the all-zero state, |00…0⟩, can be simulated classically in polynomial time.

A more cynical way of interpreting it is to say:
stabilizer sets can’t get better-than-polynomial speedups.

This is done by only keeping track of generator sets, and it covers anything you might call “simulating”: predicting a measurement between |0⟩, |1⟩ or 50-50 between them, doing a sample over the distribution of possible measurement outcomes, etc.

The one time that Professor Aaronson (being a theorist) ever wrote code that people actually went out and used, was a project in grad school for a Computer Architecture course. He made a fast simulator for stabilizer sets called CHP, letting a normal computer handle thousands of qubits (limited only by their RAM). He was only trying to pass the class, but incidentally published a paper with Gottesman for a better algorithm to implement this.

Truth be told, it had nothing to do with Computer Architecture.
He's not sure with the professor accepted it.

So for a series of qubits starting at $|00\ldots0\rangle$, how do we find all of its stabilizer states?

We know it contains II…I but we wont put that in the generator. It's implied.

We'll also need { ZIII…I
{ IZII…I
{ IIZI…I
:
{ IIII…Z

But this is starting to get messy.

For Gottesman-Knill, it's useful to have another representation of qubits.

Tableau Representation

which keeps track of two matrices of 1's an 0's.

The X Matrix and The Z Matrix

+ (0 0 0 0 | 1 0 0 0)
+ (0 0 0 0 | 0 1 0 0) <= Each row represents a generator state, as a sequence of Paulis
+ (0 0 0 0 | 0 0 1 0)
+ (0 0 0 0 | 0 0 0 1)
^ ^
1 if X or Y 1 if Z or Y
0 otherwise 0 otherwise

Instead of representing each Pauli in a single matrix, it is specified over two bit in separate matrices.
The above matrix represents { ZIII, IZII, IIZI, IIIZ}.

We're going to provide the rules for Tableau Representation without any formal proof that they work, but you can go through each rule and reason through why it makes sense.

We're also going to cheat a little. Keeping track of the +'s and –'s is tricky and not particularly illuminating, so we'll just ignore them. If we only want to know if measuring a qubit will give a definite answer or not (without figuring out if it's a $|0\rangle$ or $|1\rangle$), we can ignore the signs.

So what are the rules?

The gates available to us are cNOT, H, and P, so we need to figure out how to update the tableau for each.

- To apply H on the i^{th} qubit:
  - Swap the i^{th} column of X for the i^{th} column of Z.

This should be intuitive: Hadamard swaps the X and Z bases.

- To apply P on the i^{th} qubit:
  - Take the bitwise XOR of the i^{th} column of X *into* the i^{th} column of Z

Note that P has no effect on the tableau representation of $|00\ldots0\rangle$.
Coincidence? I think not.

- To apply cNOT from the i^{th} qubit to the j^{th}:
  - Take the bitwise XOR of the i^{th} column of X *into* the j^{th} column of X

    That seems reasonable enough, *but...* remember from the homework how a cNOT from i–>j in the Hadamard basis is equivalent to a cNOT from j–>i? That means we also have to…

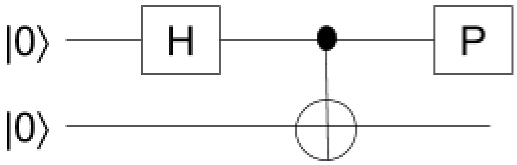
  - Take the bitwise XOR of the j^{th} column of Z *into* the i^{th} column of Z

These rules are enough to establish that measuring the i^{th} qubit in the $|0\rangle,|1\rangle$ basis has a determinate outcome *iff* the i^{th} outcome of the X matrix is all 0's.

Another cool fact:
The number of basis states that our state is a superposition over is 2^k, where k is the rank of the X matrix.

For the above tableau, rank(X) = 0, so it's a superposition of a single state.

Let's test this out, keeping track of the tableau for the following circuit.



We start with

(0 0 | 1 0)
(0 0 | 0 1)

After the Hadamard (swap 1^{st} column of X and Z)

(1 0 | 0 0)
(0 0 | 0 1)

You could convert this back into Paulis by saying the current state is the one generated by (X I) [e.g. top left qubit is 1 in X, 0 in Y => top left is X]
(I Z)

That makes sense since, as we say before, these two are a generator state for $|0\rangle\otimes|+\rangle$

After the cNOT (in X: XOR 1^{st} column into 2^{nd}, in Z: XOR 2^{nd} column into 1^{st})

(1 1 | 0 0)
(0 0 | 1 1)

This is (X X) the stabilizer generator for the Bell Pair.
(Y Y)

After the phase gate (XOR 1^{st} column of X into 1^{st} column of Z)

(1 1 | 1 0)
(0 0 | 1 1)

A phase gate signifies the introduction of *i*'s. This corresponds to $\frac{|00\rangle + i|11\rangle}{\sqrt{2}}$

Most quantum error correction codes are done with stabilizer circuits, making them easy to compute. As a result, the real importance of the stabilizer formalism is letting us keep track of them in a more elegant way.

For example, with Shor's 9-qubit code, we were dealing with qubits in the form $\left(\frac{|000\rangle + |111\rangle}{\sqrt{2}}\right)^{\otimes 3}$. Since you can flip any two qubit in a grouping and retain the form, we can write this state's generator as:

{ **Z Z** I I I I I I I,

I **Z** **Z** I I I I I I,

I I I **Z** **Z** I I I I,

I I I I **Z** **Z** I I I,

I I I I I I **Z** **Z** I,

I I I I I I I **Z** **Z,**

X **X** **X** **X** **X** **X** I I I,

I I I **X** **X** **X** **X** **X** **X,**

± **X** **X** **X** **X** **X** **X** **X** **X** **X** }

The last line can have either a + or –, encoding $|\bar{0}\rangle$ or $|\bar{1}\rangle$ respectively

Now we can finally see the 5-qubit error correction code.
The state is impractical to write out explicitly, so it's usually only represented through the stabilizer formalism

{ XZZXI,
IXZZX,
XIXZZ,
ZXIXZ,
± XXXXX }

Lecture 29: Thurs May 4

For a given quantum error correction code, applying a gate usually entails:

decoding the qubit => applying the gate => re-encoding the qubit

That's why in practice people prefer quantum error correction codes with **transversality**.

We say that the Hadamard gate is **transversal** for a qubit if you can Hadamard the logical qubit by applying the Hadamard gate to each physical qubit separately.

You can work out that Hadamard is transversal for Shor's 9-qubit code.

There are quantum error correction codes where cNOT, H, and P are all transversal.

Unfortunately, there's a theorem that says that arbitrary non-stabilizer gates *can't* be transversal. That means gates like Toffoli or $R_{\pi/8}$ must be implemented through sequences of gates that are much more expensive.

So in practical quantum computing, stabilizer applications cost almost nothing.

A 2004 paper by Aaronson and Gottesman says:

To simulate a circuit with mostly stabilizer gates (*n* gates total, T non-stabilizer gates) requires a runtime that's polynomial in *n*, and exponential in T.

This means that an exponential speedup in quantum computing requires an exponential number of stabilizer gates.

There are various tricks to produce this, like Magic State Distillation. The basic idea is that applying stabilizer gates to some nonuniformly-distributed states called 'magic states', such as $\cos(\pi/8)|0\rangle + \sin(\pi/8)|1\rangle$ lets you escape Gottesman-Knill and reach a universal quantum computer.

We've seen a high-level overview of quantum computing, so it behooves us to take a lecture to discuss practical implementations.

Professor Aaronson has visited quantum computing labs all over the world. They all have one strict rule for theorists: "don't touch anything!"

So to recap: Why would someone want to build a scalable quantum computer?

Proving it's possible is reason enough for many—it would show whether nature defies the Extended Church-Turing Thesis, not to mention disproving the people who say it's not possible.

But there are also important computational speedups to keep in mind, the top five (in order) are:

1. Quantum Simulation

Take a Hamiltonian of a real system, trotterize it, then run.

This would let you compute the effect of any chemical reaction without physically running it, which would be amazing for chemists. We tend not to talk much about this, since it's pretty straightforward, but it's easily the best application of quantum computing.

Even imperfect implementations of quantum computing are enough to see advantages for this.

2. Code Breaking

The sexiest application of quantum computing.

This would be very important for intelligence agencies, nefarious actors, intelligence agencies *who are themselves* nefarious actors.

It would completely change how e-commerce is run, requiring everybody to move to private-key crypto, lattice cryptography, etc. However, advantages here would require a fully fault-tolerant quantum computer.

3. Grover

As we've seen, Grover's Algorithm can only provide a polynomial speedup. However, it would be over a broad range of applications.

It would essentially just "give a little more juice to Moore's Law."

4. Adiabatic Optimization

Might produce speedups better than Grover, but we'll only know once we try.

5. Machine Learning

Very hot in recent years.

It's a good match for quantum computing because many problems in the field (classifying data, creating recommendation systems, etc) boil down to performing linear algebra on large sets of data. Even better, you typically only need an approximate answer.

Over the past ten years, many papers have been published claiming that quantum computing can give up to exponential speedups on such problems. This started in 2007 with…

The HHL Algorithm (Harrow, Hassidim, Lloyd)

which is billed as a quantum algorithm to solve linear systems exponentially faster than a classical computer can.

There is a catch (which Professor Aaronson wrote an article in *Nature* about).

In the fine print of the algorithm, it's assumed that the input and the output are in a quantum format. Usually, we only assume that we have $A\overline{x} = \overline{b}$ with A and b stored in memory.

But this algorithm says:

"Suppose I have qubits encoding $|b\rangle$ and I'm able to apply a Hamiltonian $e^{-iH}|b\rangle = |x\rangle$ which uses matrix A to encode the solution vector. Then (assuming e^{-iH} follows a few conditions), you can do this with an n by n matrix"

Converting into and out of a quantum format may be hard enough that the entire process would have no speedup relative to a classical computer.

It's like the algorithm gets you halfway across the world, but leaves you stranded at the airport.

For example, if you had $|x\rangle = \sum_{i=1}^{n} \alpha_i |i\rangle$ and you want to get all the $|i\rangle$'s out, it may be necessary to run and measure n times, at which point you're not getting an exponential speedup.

Journalists often ask Professor Aaronson, "When will we all have personal quantum computers and qPhones in our pocket?" It's hard to imagine that'll ever happen though, because most things we do on our PCs can be done quickly on classical computers. At most we'll likely see cloud quantum computing, like the IBM Quantum Experience, where a central location deals with the issues of maintaining quantum states while we reap the benefits.

Maybe this'll seem myopic in a hundred years, like a guy from the 70s saying, "I only see a market for five computers in the world, tops." But you could argue that such people were simply ahead of their time. We *are* moving to a world where most computation is done on the cloud in a few centralized locations. Though this might also be shortsighted because our current list of applications of quantum computing may be woefully incomplete.

So what *do* you have to do to build a quantum computer?

There's a famous list of requirements it takes for a system to be able to to quantum operations called the **DiVincenzo Criteria**. There are several, but the four most important are…

- *Long-Lived Qubits*

It's self-evident that you need some system that can maintain quantum states over long periods of time. As we've said before, "the first requirement of quantum computing is the ability to perform I."

- *Universal Gates*

You must be able to apply *some* universal set of gates.

Implicit here is the requirement that qubits can interact with one another.

- *Initialization*

You must be able to get qubits to $|00\ldots0\rangle$.

- *Measurement*

You're familiar with measurement.

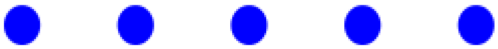
Different architectures have achieved different combinations of these. There are architectures where initialization is hard or measurement is hard.

So what are the major architectures that have been built?

To a theorist, a qubit is a qubit is a qubit.
But experimentalists talk about four main architectures that may work.

The oldest approach, dating back to 1985 is **Trapped Ions**.

The basic idea is that you have a bunch of ions (let's say they're atomic nuclei) with electric charges so that they respond to a magnetic field. You can then manipulate the magnetic field to get them trapped in a line.



Such a lab will have ions, a magnet, and a classical computer that lets them see images of the ions. This method isn't totally reliable, and it takes work to keep the ions penned in.

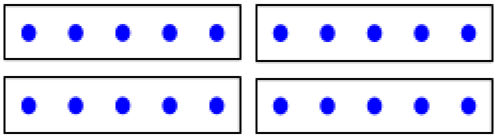
It's a bit like herding cattle.

Atomic nuclei have a spin state that can be clockwise, counterclockwise, or a superposition of both. So we treat the nuclei's spin as their quantum state. If you bring two ions close to each other, the *Coolum Effect* can be used to create something resembling a CNOT gate.

You can manipulate qubits by using a laser to pick them up and move them around. This may sound like it would be a tough balancing act, moving nuclei via laser *while* keeping them all floating magnetically…

Yes. Yes it is.

But it has been demonstrated with up to 10-20 qubits. After that it becomes hard to interact with a single qubit at a time.



Proposals to scale up this method often involve many small ion traps, with some 2-qubit gates that interact between traps. You could use quantum teleportation to communicate between traps.

Many groups are pursuing this at NIST, UMaryland, and Innsbruck (Austria).

In the last few years, several such ventures that were originally academic became start-ups—which tend to give out much less information about how they're doing.

Another approach is **Superconducting Qubits**.

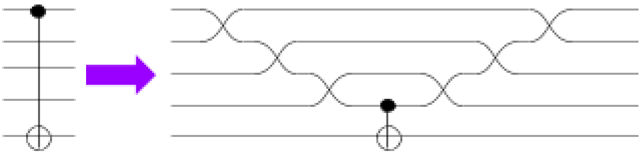
We talked about this earlier with DWave, but there are many others trying this.

In this approach, everything happens on a chip in a refrigerator cooled down enough for the coils to superconduct.

Electrons can flow around the qubit clockwise, counterclockwise, or in a superposition of both. If two of these qubits come close, *some* 2-qubit Hamiltonian happens between them.

Advantages of this setup: – You can make lots of these coils
– Gate operations are very fast
Disadvantages: – Coherence times are much shorter
The big one is that – These qubits can't move around and can only talk to their neighbors.

In designing quantum circuits, we've been implicitly assuming that any two qubits can interact. Of course you could always simulate it with a whole cascade of swaps, but you pay a price for that.

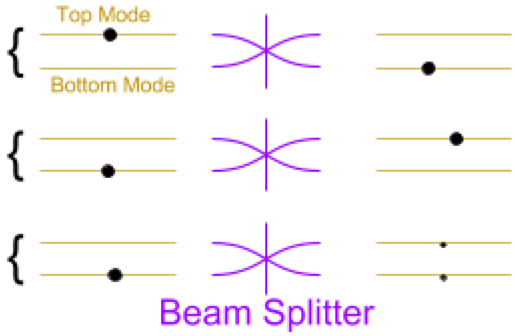


That being said, this is the currently the most popular approach.

Google bought out almost the whole Santa Barbara lab (Martinez's group), and have publicly announced that they expect to have a 50-qubit system in a year. IBM also has a superconducting group with similar claims. In addition there are several startups working with superconducting qubits, like Rigetti—made up of several people who left IBM.

A third approach is **Photonics**

which treats photons as qubits. For example, you could say that the photon is either horizontally polarized, vertically polarized, or in a superposition of being polarized both ways.



There's many ways to generate photons in such a system. One way is to use **Dual Rail**. Each photon has two modes—two places it can be in. A photon can be in the top mode, the bottom mode, or in a superposition of both.

The basic idea is that you generate photons and send them through fiber optic cables. When the two come together, you use a **Beam Splitter**, which corresponds as a 2x2 unitary acting on the qubit, take taking the state to or from superposition.

2-qubit gates are harder to do. There's a set of operations that are easy to implement in such a system. What's not obvious is whether those operations are sufficient to produce a universal quantum computer.

There was a breakthrough in this area in 2001 called

The KLM Theorem (Knill, Laflamme, Milburn)

which says that if I can generate photons and send them through beam splitters, phase shifters, *plus* at any time, for all channels, I can tell if there's a qubit, or not, or a piece of qubit in the channel and feed the answer forward to further operations:

Then LO + FFM = BQP (linear optics + feed-forward measurement = BQP)

Furthermore, the KLM Theorem opens the possibility of a new way to build a quantum computer, where qubits are photons travel at the speed of light.

Photons can maintain superposition indefinitely by flying in a vacuum. The trouble is that they're flying at the speed of light, which makes it hard for them to interact with one another. This may require stalling photons, but that may introduce decoherence.

In 2011 Professor Aaronson and Alex Arkhipov proposed **Boson Sampling**,

The idea was to investigate what can be done with linear optics if you <u>don't</u> have feed-forward measurement. They concluded that this probably can't produce a universal quantum computer, *but* it could do some problems faster than a classical computer.

Not any problems that people actually care about, mind you.
Nevertheless, it would be a good candidate to prove that *any* quantum speedups truly exist.

<u>Could you represent a beam splitter as a 2x2 matrix?</u>

We're sweeping several lectures about photonics under the rug here. The jist of it is that instead of building tensor products, composite systems are created out of smaller systems in a different way.

The last approach we'll cover is fairly esoteric, it's called **Non-abelian Anyons**.

There are two types of fundamental particles: Bosons and Fermions. But in a two-dimensional field, you can have particles that behave as neither Bosons nor Fermions.

Lots of physicists won Nobels in the 80s for this stuff.

If you can make such "quasiparticles" in a two-dimensional surface, just moving them around would be sufficient to create a universal quantum computer. This setup may be naturally resistant to decoherence. The caveat is that we're only now starting to understand how to create the simplest quasiparticles.

Microsoft is the current leader in this approach, and has hired several experts in this field recently.

As a path forward…

Professor Aaronson's thinks about what to expect from **Quantum Supremacy** in terms of three steps.

Lots of people dislike this term ^ for obvious reasons, but it has stuck for now.

Step 1: Doing *something* faster than a classical computer can.

50 qubits may be enough, and Boson Sampling may be used to achieve this.

Step 2: Doing useful quantum simulations

A Microsoft paper claims that 100 qubits would be enough to simulate one quantum system using another for several useful applications.

Step 3: Creating a full universal quantum computer

Which would let us finally run scalable implementations of Shor's Algorithm.

Step 1 may likely be coming soon, but no promises on steps 2 and 3!