Eric Du — Physics C191

Lecture Notes — **Introduction to Quantum Computing** — Whaley, Kolkowitz, Young

Header styling inspired by CS 70: https://www.eecs70.org/

1 Introduction: Axioms of Quantum Mechanics

- Also called postulates.
- Typically four axioms:
  1. Quantum states & superposition
  2. Unitary Evolution: deterministic
  3. Measurements: introduces statistical nature to quantum behavior
  4. Observables: quantities we can measure in the real world

1.1 Quantum States

- A quantum state is denoted by $|\psi\rangle$. It's a vector in a complex-valued vector space, with a particular inner product structure. This combination of the vector space with the inner product structure is called a Hilbert space, denoted by $\mathcal{H}$.
- Vectors in $\mathcal{H}$ are denoted by *kets* $|v\rangle$, and because it's a vector space (hence it's linear), we can make other vectors by adding together two vectors: $|w\rangle = |u\rangle + |v\rangle$.

  We also have a null vector $|0\rangle = |u\rangle - |u\rangle$.
- Linearly independent vectors:
  $$a_1 |u_1\rangle + a_2 |u_2\rangle + \cdots + a_n |u_n\rangle = 0$$
  if the only solution to this is to set $a_1, a_2, \ldots, a_n$ to 0, then the set of vectors $|u_1\rangle, |u_2\rangle, \ldots, |u_n\rangle$ is linearly independent.

  We will only work with finite dimensional vector spaces, for the sake of quantum information
- If the set of vectors $\{|u_i\rangle\}$ spans the space, then they are referred to as a basis. This means that any vector $|w\rangle$ can be written as a linear combination of some $|u_i\rangle$:
  $$|w\rangle = \sum_i a_i |u_i\rangle$$
  It can also be represented as a column vector of n values:
  $$|w\rangle = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$
- An example where $n = 2$, is *spin projection*, which has two possible values: $\pm\hbar/2$. In this case, the general state $|\psi\rangle$ can be written as $|\psi\rangle = a_1 |+\hbar/2\rangle + a_2 |-\hbar/2\rangle$.

  We'll be dealing with mostly two-state systems in this class, and any other two-state system that we choose is sometimes called "pseudo-spin" since the math is nearly identical.
- In all cases, we should have $\sum_i |a_i|^2 = 1$; we call the states that follow this behavior (and they should) to be **normalized to 1**.

1.2 Inner Product

- Given $|w\rangle = \sum_i a_i |u_i\rangle$ and $v = \sum_i b_i |u_i\rangle$, then the complex-valued inner product $\langle v|w\rangle = \sum_i b_i^* a_i$. It can be real-valued, but in general it's considered complex.

  This gives a way for us to talk about how far apart two vectors are from one another, similar to a dot product.

- If the inner product is 0 and our vectors are not the zero vector themselves, then we call these two vectors **orthogonal**.
- An **orthonormal basis** is one where all the vectors are orthogonal, and also normalized to 1. In other words, we have $\langle u_i|u_j\rangle = \delta_{ij}$, where δ_{ij} is the Kronecker delta.
- So what is $\langle u|$? $\langle u|$ lives in the *dual space*, and is defined as follows: if $|w\rangle = \sum_i a_i |u_i\rangle$, then $\langle w| = \sum_i a_i^* \langle u_i|$.

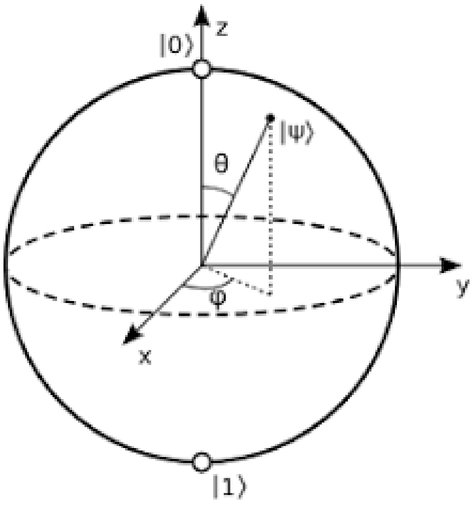
  So if $|w\rangle$ is represented as a column vector (earlier), then $\langle w|$ is represented as a row vector:

$$\langle w| = \begin{bmatrix} a_1^* & \dots & a_n^* \end{bmatrix} = w^{\top *} = w^\dagger$$

- The properties of the inner product:
  - $\langle u|v\rangle = \langle v|u\rangle^*$
  - Antilinearity: $\langle u|av\rangle = a\langle u|v\rangle$, but $\langle au|v\rangle = a^* \langle u|v\rangle$.
  - Norm of $|v\rangle$: $\langle v|v\rangle = \|v\|^2$. Hence, $\|v\| = \sqrt{\langle v|v\rangle}$.
- Conventionally, although we denote $|w\rangle = \sum_i^{n-1} a_i |u_i\rangle$, we generally deal with $n = 2$, so we have $|0\rangle$ and $|1\rangle$ as our states. This is called the **computational basis**.

1.3 Geometric Interpretation

- For $n = 2$, there is a nice geometric interpretation called the **Bloch sphere**:



  The sphere has radius 1, and all points on the sphere represent quantum states. A general state $|\psi\rangle$ is written as

$$|\psi\rangle = e^{i\gamma}\left[\cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle\right] = \alpha|0\rangle + \beta|1\rangle$$

  If we want our state to be normalized, then we want $\|\alpha\|^2 + \|\beta\|^2 = 1$.

- There are also other orthonormal bases we can choose:
  - x-basis: $|+x\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle, |-x\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.
  - y-basis: $|+y\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle$, and $|-y\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)$.

1.4 Unitary Evolution

- All equations we'll deal with are relations in $\mathcal{H}$, and these operations form a group called SU(2). This is called the *Special unitary group.*
- This unitary transformation takes our vectors $|0\rangle$ and $|1\rangle$ and does the following:

$$|0\rangle \xrightarrow{U} a\,|0\rangle + b\,|1\rangle$$
$$|1\rangle \xrightarrow{U} c\,|0\rangle + d\,|1\rangle$$

  In this case, we can write U as a 2x2 matrix:

$$U = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \; U^{\dagger} = \begin{pmatrix} a^* & b^* \\ c^* & d^* \end{pmatrix}$$

  Recall that $U^{\dagger}$ is the conjugate transpose. If U is a unitary operator, then $U\dagger U = I = UU^{\dagger}$. This implies that $U\dagger = U^{-1}$

- On a qubit, we will apply many gates throughout this semester. Some of these are listed below:
  - X-gate: $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
  - Z-gate: $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
  - Hadamard gate: $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

  All of these operations can be interpreted as a series of rotations on the Bloch sphere.

1.5 Observables

- An operator A, and its Hermitian conjugate is denoted by $A^{\dagger} = (A^{\top})^*$.
- In QM, Hermitian operators are related to real observables we can measure in the lab, and because they are measurable, they must have real eigenvalues.
- They will also have mutually orthogonal eigenvectors.
- As an example, the X gate is Hermitian, with eigenvectors of $|+x\rangle$ and $|-x\rangle$. This is also sometimes called the Hadamard basis, because acting the Hadamard gate on $|0\rangle$ gives us $|+x\rangle$, and acting it on $|1\rangle$ gives $|-x\rangle$.

2 Entanglement & Bell Inequalities

2.1 Projection Operators

- The basic form of an operator is that it takes one vector and spits out another: $|c\rangle = |c\rangle\,\langle a|a\rangle$. So, the outer product $|c\rangle\,\langle a|$ is the operator.
- Consider a state $|w\rangle = \sum_{i=1}^{n} a_i\,|u_i\rangle$, where $\{|u_i\rangle\}$ form an orthonormal basis. If we want to find any one of the a_j, then we compute $\langle u_j|w\rangle$:

$$\langle u_j|w\rangle = \sum_{i=1}^{n} a_i \underbrace{\langle u_j|u_i\rangle}_{\delta_{ij}} = a_j$$

  Alternatively, this allows us to write $|w\rangle$ in terms of:

$$|w\rangle = \sum_{i}^{n} \langle u_i|w\rangle\,|u_i\rangle = \sum_{i=1}^{n} |u_i\rangle\,\langle u_i|w\rangle$$

Now, the term $|u_i\rangle \langle u_i|$ is an operator, and is called the **projection operator**. If we act the operator on one of the basis vectors:

$$|u_i\rangle \langle u_i|u_i\rangle = |u_i\rangle$$

whereas if we do it on an arbitrary vector $|w\rangle$:

$$|u_i\rangle \underbrace{\langle u_i|w\rangle}_{a_i} = a_i |u_i\rangle$$

- The projection operator is written as $P_i = |u_i\rangle \langle u_i|$, which has the property that $P_i^2 = |u_i\rangle \langle u_i|u_i\rangle \langle u_i| = P_i$. It also has the property that

$$\sum_i^n P_i = \sum_i^n |u_i\rangle \langle u_i| = I$$

2.2 General Operators

- A general operator is defined as $A = IAI$. Now, we're going to express the identity matrices in terms of the projection operators:

$$\begin{aligned} A &= \sum_i \sum_j |u_i\rangle \overbrace{\langle u_i|A|u_j\rangle}^{A_{ij}} \langle u_j| \\ &= \sum_{i,j} A_{ij} |u_i\rangle \langle u_j| \end{aligned}$$

The term A_{ij} represents a *matrix element*, represented in the $|u_j\rangle$ basis.

What does the $|u_i\rangle \langle u_j|$ operator represent?

- One basis that we'll use very frequently is to express A in terms of the eigenbasis. That is, the set $|a_i\rangle$ of vectors such that

$$A |a_i\rangle = a_i |a_i\rangle$$

In this basis, then A is written as:

$$\begin{aligned} A &= IAI \\ &= \sum_{ij} |a_i\rangle \langle a_i|A|a_j\rangle \langle a_j| \\ &= \sum_{i,j} a_j |a_i\rangle \langle a_i|a_j\rangle \langle a_j| \end{aligned}$$

Here we've used the property that $A|a_j\rangle = a_j |a_j\rangle$. Then, if we choose the eigenvectors to be orthogonal (which is okay for a Hermitian A), then $\langle a_i|a_j\rangle = \delta_{ij}$, so:

$$A = \sum_i a_i |a_i\rangle \langle a_i|$$

Why can we choose the $\{ |a_i\rangle\}$ to be orthogonal?

- We choose A to be Hermitian (which is the only way we were able to make this simplification). Since they have real eigenvalues, they have mutually orthogonal eigenvectors.

2.3 Measurement Postulate

- An observable A can be measured by a set of operators $\{M_m\}$ with outcomes (observable values) m.
- For example, a qubit (so any 2-level system) with states $|0\rangle$ and $|1\rangle$, we can make a general state $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$ with normalization constraint $\|\alpha\|^2 + \|\beta\|^2 = 1$.

By measuring, we "learn" the value of α and β. Our measurement operators consist of

$$M_0 = |0\rangle\langle 0|, \ \ M_1 = |1\rangle\langle 1|$$

You'll notice that these are projections onto a given state – this is intentional.

- Upon measuring $|\psi\rangle$, we will get one outcome (either 0 or 1), with probability $p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$.
- After measurement, the state "collapses" into the state $\frac{M_m|\psi\rangle}{\sqrt{p(m)}}$. This is a fancy way to say that it will only give us $|0\rangle$ if the outcome was 0. This probabilistic determination of the final state is intrinsic to quantum mechanics.

  As an example, if we have one state $|\psi\rangle$, we either get 0 or 1 but have no information about α or β. However, if we have many identical $|\psi\rangle$, then we get 0 with probability $\|\alpha\|^2$, and we get 1 with probability $\|\beta\|^2$. This is because:

$$\begin{aligned} p(m=0) &= \langle\psi|0\rangle\langle 0|0\rangle\langle 0|\psi\rangle = \|\alpha\|^2 \\ p(m=1) &= \langle\psi|1\rangle\langle 1|1\rangle\langle 1|\psi\rangle = \|\beta\|^2 \end{aligned}$$

  Note that $M_m^\dagger = M$, based on the way we've defined them. If we get 0, then the final state is written as:

$$\frac{|0\rangle\langle 0|\psi\rangle}{\sqrt{\|\alpha\|^2}} = |0\rangle = e^{i\theta}|0\rangle$$

  the $e^{i\theta}$ is just some overall phase factor.

- We introduce an average over many measurements to be the quantity $\langle A\rangle$, which is calculated as:

$$\langle A\rangle = \sum_m p(m)a_m$$

  This is also sometimes called the *average value* of an operator. The measurement basis we choose for a Hermitian A is given by the eigenvectors of A, so we have:

$$M_m = |a_m\rangle\langle a_m|$$

  where $|a_i\rangle$ is the i-th eigenvector of A. Then, this means that $\langle A\rangle = \sum_m p(m)a_m$. Remember that A is represented as:

$$A = \sum_m a_m |a_m\rangle\langle a_m|$$

- Some cool expansion:

$$\begin{aligned} \langle A\rangle &= \sum_m p(m)a_m \\ &= \sum_m a_m \langle\psi|M_m^\dagger M_m|\psi\rangle \\ &= \sum_m a_m \langle\psi|a_m\rangle\langle a_m|a_m\rangle\langle a_m|\psi\rangle \\ &= \sum_m a_m \langle\psi|a_m\rangle\langle a_m|\psi\rangle \end{aligned}$$

  But now let's throw a $\langle\psi|$ to the left:

$$\langle\psi|\sum_m a_m |a_m\rangle\langle a_m|\psi\rangle = \langle\psi|A|\psi\rangle$$

  This is the matrix element we've come across earlier.

2.3.1 Specific Examples

- Suppose we want to measure Z for a qubit. Recall that $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. This has eigenvalues ± 1, with eigenvectors $|0\rangle, |1\rangle$.

- Now, we compute $\langle Z \rangle$ for a general state $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$.

$$\begin{aligned} \langle Z \rangle &= p(+1)(+1) + p(-1)(-1) \\ &= \|\alpha\|^2 - \|\beta\|^2 \end{aligned}$$

  Remember that the equation is (probability of obtaining state) × (eigenvalue of that state).

- Now let's measure $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ on the same state $|\psi\rangle$. It's eigenvalues are ± 1, with eigenvectors $|+\rangle, |-\rangle$. Recall that

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \\ |-\rangle &= \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \end{aligned}$$

  This means that we can solve for $|0\rangle$ and $|1\rangle$:

$$\begin{aligned} |0\rangle &= \frac{1}{\sqrt{2}} (|+\rangle + |0\rangle) \\ |1\rangle &= \frac{1}{\sqrt{2}} (|+\rangle - |-\rangle) \end{aligned}$$

  Therefore, the average $\langle X \rangle$:

$$\langle X \rangle = p_m(+1)(+1) + p_m(-1)(-1)$$

  Then, we expand the probabilities:

$$\begin{aligned} p_m(+1) &= \langle \psi | + \rangle \langle + | + \rangle \langle + | \psi \rangle = \langle \psi | + \rangle \langle + | \psi \rangle \\ p_m(-1) &= \langle \psi | - \rangle \langle - | - \rangle \langle - | \psi \rangle = \langle \psi | - \rangle \langle - | \psi \rangle \end{aligned}$$

  To complete the computation, we have to express $|\psi\rangle$ in the $|\pm\rangle$ basis:

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle = \frac{\alpha}{\sqrt{2}} (|+\rangle + |-\rangle)$$

3 Multiple Qubits, Entanglement

3.1 Multiple Qubits

- Suppose we have two qubits $|\psi_1\rangle = \alpha |0\rangle + \beta |1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ and $|\psi_2\rangle = x |0\rangle + y |1\rangle = \begin{bmatrix} x \\ y \end{bmatrix}$
- Then the combined state, if the two qubits live on their own, is given by $|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$. The $\otimes$ symbol denotes a tensor product.

$$|\psi_1\rangle \otimes |\psi_2\rangle = (\alpha |0\rangle + \beta |1\rangle) \otimes (x |0\rangle + y |1\rangle)$$

  In matrix form, this is represented as:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \otimes \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha x \\ \alpha y \\ \beta x \\ \beta y \end{bmatrix}$$

  the resulting vector lives in $\mathbb{C}^4$, with the basis states $|00\rangle = |0\rangle_1 |0\rangle_2$, $|10\rangle = |1\rangle_1 |0\rangle_2$, $|01\rangle = |0\rangle_1 |1\rangle_2$, $|11\rangle = |1\rangle_1 |1\rangle_2$

- In general given n qubits, there are 2^n basis states, and hence we will be working with superpositions over these 2^n basis states. This fact underscores the power of quantum computers, since they scale much more efficiently than classical computers. This is also sometimes referred to as "quantum parallelism".

- If we measure all qubits, then the outcome is just some sort of bitstring, so we have to be clever about how we are measuring to get the information we want.
- With multiple qubits, operators are also tensor products. Given the two operators:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$
$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

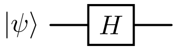
  Then $A \otimes B$, the operator that acts on the multi-qubit state, is given by

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}b \\ a_{21}B & a_{22}B \end{pmatrix}$$

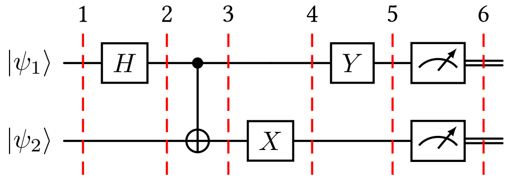
  Note that $A \otimes B$ is not the same as $B \otimes A$.

3.2 Quantum Circuits

- A generic quantum circuit is written as:



  the box with an H denotes a gate (in this case, a Hadamard gate), which corresponds to a rotation on the Bloch sphere.
- Let's analyze the following quantum circuit:



  Let's analyze this in stpes:

  - Initially, we have $|\psi_1\rangle = \alpha_1 |0\rangle + \beta_1 |1\rangle$ and $|\psi_2\rangle = \alpha_2 |0\rangle + \beta_2 |1\rangle$, whose combination can be written as:

$$|\psi_{12}^{(1)}\rangle = \alpha_1\alpha_2 |00\rangle + \alpha_1\beta_2 |01\rangle + \beta_1\alpha_2 |10\rangle + \beta_1\beta_2 |11\rangle$$

  - At step 2, we run the first qubit through a Hadamard gate, and leave the second qubit untouched. This means we act the operator $H \otimes I$ on the state:

$$\begin{aligned} |\psi_{12}^{(2)}\rangle \, H \otimes I \, |\psi_{12}\rangle = \alpha_1\alpha_2 \left(\frac{1}{\sqrt{2}} |00\rangle + \frac{1}{\sqrt{2}} |10\rangle\right) + \alpha_1\beta_2 \left(\frac{1}{\sqrt{2}} |01\rangle + \frac{1}{\sqrt{2}} |11\rangle\right) \\ + \beta_1\alpha_2 \left(\frac{1}{\sqrt{2}} |00\rangle - \frac{1}{\sqrt{2}} |10\rangle\right) + \beta_1\beta_2 \left(\frac{1}{\sqrt{2}} |01\rangle - \frac{1}{\sqrt{2}} |11\rangle\right) \end{aligned}$$

  - At step 3, we apply a CNOT gate, which flips the state of the second bit if the value of the first bit is 1. As a truth table:

Input	Output
00	00
01	01
10	11
11	10

As a matrix, it's written as;

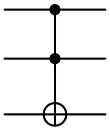
$$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 9 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

We then apply this CNOT gate to each component of $|\psi_{12}^{(2)}\rangle$ to get $|\psi_{12}^{(3)}\rangle$.

- Apply $I \otimes X$ to $|\psi_{12}^{(3)}\rangle \to |\psi_{12}^{(4)}\rangle$
- Apply $Y \otimes I$ tp $|\psi_{12}\rangle^{(4)} \to |\psi_{12}^{(5)}\rangle$
- Measurement in the Z basis, by applying projection operators to the final resulting state.

3.2.1 Other Common Gates

- There are many quantum gates that we'll study, here's a list of them that will be useful:
- CPHASE, or controlled Z gate
- Swap gate: swaps the
- S-phase: rotation by 90 degrees, $\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$
- P-phase: a general phase gate $\begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}$
- Toffoli gate: controlled-controlled NOT gate:



- T-gate: $\begin{pmatrix} 1 & 0 \\ 9 & e^{i\pi/4j} \end{pmatrix}$

3.2.2 Universal Gate Sets

- A set G of quantum gates is considered universal if for $\epsilon > 0$ and for any unitary matrix U on n qubits, there is a sequence of gates from G such that

  $$\|U - U_{g_\ell} \cdots U_{g_2} U_{g_1}\| < \epsilon$$

  In this definition, we define $U_g = V \otimes I$, where V is an operator acting on k qubits, and I acts on the remaining $n - k$ qubits. The double bar represnets an operator norm, defined as:

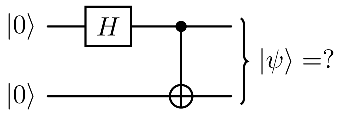
  $$\|U - U'\| = \max_{|v\rangle \text{ unit vectors}} \|(U - U')\,|v\rangle\,\|$$

  where $\|w\| = \sqrt{\langle w|w\rangle}$.

- Examples of universal gate sets:
  - Barenco et al. (1995): CNOT and all single qubit (continuous) gates.
  - CNOT, H, S, T gates
  - Rotation operators $R_x(\theta), R_y(\theta), R_z(\theta)$, the phase operator P_ϕ and CNOT.

3.3 Entanglement

- Consider 2 qubits:



Well, we first start with the state $|00\rangle$, and after passing the first bit through a Hadamard gate, we get the state

$$\frac{|00\rangle + |10\rangle}{\sqrt{2}}$$

Then, running it through the CNOT, then we have:

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} = \Phi^+$$

This is one of four states called the "Bell states", because there is no way to express this state as a product state of two individual qubits.

4 More on Multiple Qubits

- Last time, we looked at multiple-qubit states, and talked about how the combination is the tensor product, written like this:

$$|0\rangle \otimes |1\rangle \otimes |0\rangle \otimes \cdots \otimes |0\rangle$$

- We also talked about how an entangled state is defined as a state where we cannot express as a (tensor) product state. In other words, the state is not separable.
- There are an infinite number of entangled states, called the Bell states:

$$\begin{aligned} |\Phi^+\rangle &= \frac{|00\rangle + |11\rangle}{\sqrt{2}} \\ |\Phi^-\rangle &= \frac{|00\rangle - |11\rangle}{\sqrt{2}} \\ |\Psi^+\rangle &= \frac{|01\rangle + |10\rangle}{\sqrt{2}} \\ |\Psi^-\rangle &= \frac{|01\rangle - |10\rangle}{\sqrt{2}} \end{aligned}$$

- To quantify entanglement, we use a Schmidt decomposition for qubits: ($d = 2$ for qubits)

$$|\psi_{AB}\rangle = \sum_{i=0}^{d-1} c_i \, |i\rangle_A \, |i\rangle_B$$

This state ψ_{AB} is separable if only one $c_i \neq 0$. The number of nonzero c_i is called the schmidt rank, and it's what we use to quantify how entangled a state is. If all c_i are equal, then the state is maximally entangled.

The bell states $\Phi^\pm$ are easily seen to be maximally entangled, since $|00\rangle$ and $|11\rangle$ are the basis states, and they each have a coefficient of $1/\sqrt{2}$.

4.1 Measurement

- Given a state $|00\rangle$ and we measure the first qubit in the Z basis, what happens?
- Recall our measurement operator is a projection operator:

$$\begin{aligned} M_1 &= |1\rangle \langle 1| \\ M_2 &= |0\rangle \langle 0| \end{aligned}$$

- Then, applying the measurement operators, we get an outcome of measuring 0 with probability 1. The state after measurement is given by $|00\rangle$. Note that the second qubit is not affected by this measurement.

  Are these two states identical?

- Now suppose we had a state of the form

$$|\psi\rangle = |0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

  This is the state that results when the second qubit is passed through a Hadamard gate. Now, if we measure the first state, we again certainly get a result of 0, so the measurement is given by: =

$$|\psi\rangle = |0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

  If we measure the second qubit (in the Z basis), then we get the state $|0\rangle \otimes |0\rangle$ with probability $\frac{1}{2}$, and $|0\rangle \otimes |1\rangle$ also with probability $\frac{1}{2}$.

- Another example, given the state:

$$|\psi\rangle = \frac{1}{2}(|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$$

  And now we measure the first qubit, we get 0 and 1 with probability $\frac{1}{2}$, and we get the resulting states:

$$|\psi'\rangle = |0 \text{ or } 1\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

4.2 Measurement with Entangled States

- Suppose we have a qubit in the state $|\Psi^-\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$. Now, we send the first qubit to Alice, and the second one to Bob.

- Alice will measure the first qubit in the Z basis, which will give her 0 or 1 with probability $\frac{1}{2}$.

  The thing is, if alice measures 0, then it means that the state now collapses to the first term in the superposition: $|\psi'\rangle = |01\rangle$, so Bob must get a result of 1 upon measurement. The flip is also true.

- This is an example where the outcomes of the measurements are now correlated!

- Now suppose we change our measurement basis: if we measure in the X basis, where measurements are given by $M_1 = |+\rangle\langle+|$ and $M_2 = |-\rangle\langle-|$.

  The same correlation follows: if Alice measures $|+\rangle$, then Bob will certainly get $|-\rangle$, and if Alice gets $|-\rangle$, Bob will certainly get $|+\rangle$.

  How is the maesurement carried out? Do we express the state $|\Psi^-\rangle$ in terms of the $|\pm\rangle$ basis, and then carry out the probabilities?

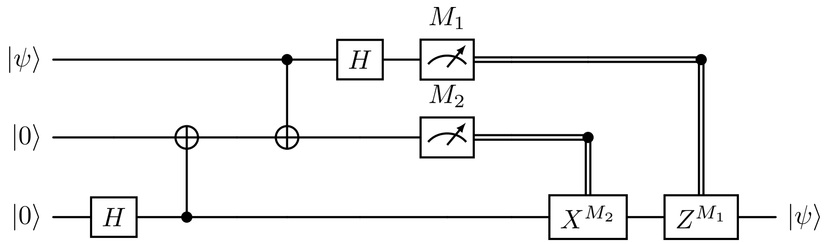
- This idea that you can glean information about a quantum state without making a full measurement was problematic, and led Einstein, Podolsky and Rosen to speculate the presence of "hidden variables".

  John Bell proposed a set of inequalities (now called Bell inequalities) that would tell us for sure whether these hidden variables actually exist. He proposed a set of measurements that can be made called g, and if the systems were truly classical, then we would be able to determine that $\langle g\rangle \leq 2$. Otherwise, $\langle g\rangle > 2$ was possible.

  What we found through experiment was that $\langle g\rangle > 2$ was indeed possible, which leads us to the conclusion that there are no hidden variables are present.

4.3 Quantum Teleportation

- Consider the following circuit:



Initially, the state is in $|\psi\rangle\,|0\rangle\,|0\rangle$. After the third qubit passes through the Hadamard gate, the state is

$$|\psi_2\rangle = |\psi\rangle\,|0\rangle\,\frac{1}{\sqrt{2}}(\,|0\rangle + |1\rangle)$$

After the first CNOT, we have:

$$|\psi_3\rangle = |\psi\rangle\,\frac{1}{\sqrt{2}}(\,|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(\alpha\,|000\rangle + \alpha\,|011\rangle + \beta\,|100\rangle + \beta\,|111\rangle)$$

Note that this is the resulting state because the third qubit is the control bit, which flips the state of the second qubit.

After the second CNOT, then:

$$|\psi_4\rangle = \frac{1}{\sqrt{2}}(\alpha\,|000\rangle + \alpha\,|011\rangle + \beta\,|110\rangle + \beta\,|101\rangle)$$

Finally, we send the first qubit thorugh a Hadamard gate:

$$|\psi_4\rangle = \frac{1}{2}(\alpha\,|000\rangle + \alpha\,|100\rangle + \alpha\,|011\rangle + \alpha\,|111\rangle + \beta\,|010\rangle - \beta\,|110\rangle + \beta\,|001\rangle - \beta\,|101\rangle)$$

The way to compute this is to look at the first qubit, and recall that:

$$H\,|\psi\rangle = H\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\alpha+\beta\\ \alpha-\beta\end{pmatrix}$$

so on the states $|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$, we have:

$$H\begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}(\,|0\rangle + |1\rangle)$$
$$H\begin{pmatrix}0\\-1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}(\,|0\rangle - |1\rangle)$$

so basically, each term gets split up into two terms: the last two qubits aren't changed, and the first qubit is split into a superposition between 0 and 1, with the sign determined by the coefficient.

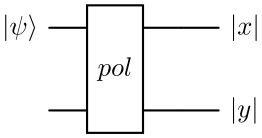
- We then measure this state, and depending on the state of measurement, we either apply (or don't) the X gate or Z gate onto the third qubit.

  What we will find is that upon performing this procedure, the state $|\psi\rangle$ is now encoded in qubit 3 rather than qubit 1.

- Note however, that even though this feels like the state has "teleported", no information is being sent here that violates causality. Because we will only know what gate to apply based upon the measurement results, it means that we need to transfer this information over thorugh some *classical* channel, meaning that the information could not possibly travel faster than light.

5 Quantum Key Distribution

- Qubits used for communication are usually photons, which have a momentum $\vec{k}$, and an electric field E_x and E_y that propagates in the plane perpendicular to $\vec{k}$. The E_x vector is denoted as $|v\rangle$, and E_y as $|H\rangle$, and this means that the general electric field $\overline{E} = \alpha |v\rangle + \beta |H\rangle$.
- We can pass these photons thorugh polarizers, which only transmit light with specific oscillations.
- So as a quantum circuit, it's written as:



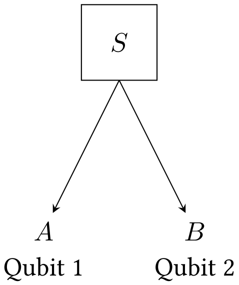
  Only one of these channels can be measured,

5.1 Distributed Entanglement

- One of the ways to do quantum communication and computation
- It includes:
  - teleportation
  - Secure QKD: communication
  - Distributed quantum computation – quantumgatsby teleportation
  - "Blind quantum teleportation"

5.1.1 QKD Secureness

- The way QKD works is a server $|\psi\rangle = |HV\rangle + |VH\rangle$, and the first qubit is sent to Bob, and the second is sent to Alice:



- Classically, if an observer were to say, measure the second qubit, then send an identical copy through that channel, then Alice and Bob won't be able to tell at all that the state has been measured.

  However, if the system was quantum, this measurement is now impossible.

- The proof of this is called the No cloning theorem, whose proof is below:

  Proof. Suppose we have an unknown state $|\phi\rangle = \alpha |0\rangle + \beta |1\rangle$. Now suppose there is a U_{cl} (a "cloning matrix") which can clone $|\phi\rangle$. That is:

  $$|\phi\rangle\, |0\rangle \mapsto |\phi\rangle\, |\phi\rangle = \alpha^2 |00\rangle + \beta\alpha |10\rangle + \alpha\beta |01\rangle + \beta^2 |11\rangle$$

  But if we do this on the initial state $|\phi\rangle$:

  $$(\alpha |0\rangle + \beta |1\rangle) |0\rangle \mapsto \alpha |00\rangle + \beta |11\rangle$$

  We are cloning this exactly based on what we want: we clone the information of the second qubit onto the first qubit, but we see that even if we could "copy", we don't get the desired product state.

But this is not equal to the copied state that we should expect. Therefore, no such U_{cl} can exist. □

5.2 Quantum Algorithms

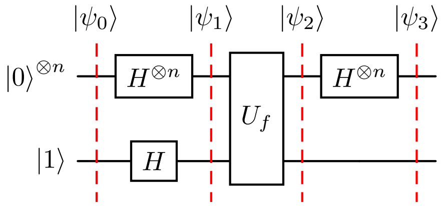
- The Deutsch-Josza is a *promise problem*: we are given a function $f(x)$, and it's one of two types:
  - $f(x)$ is either constant for all x : it is either always 0 or always 1.
  - $f(x)$ is balanced: it is 0 half the time, $f(x)$ is 1 half the time.

  More generally, we can write $f : \{0,1\}^n \mapsto \{0,1\}$, and we ask whether f is constant or balanced.

- For a function on n bits, this implies that the total domain space is of size 2^n. We need to measure a little more than half, or $2^n/2 + 1 = 2^{n-1} + 1$ measurements in order to determine the identity of f.

  Quantumly, we only need a single measurement!

- The quantum circuit is as follows:



  Initially, the state is in $|\psi_0\rangle = |0\rangle^{\otimes n} |1\rangle$. Then, after passing through both Hadamard gates, we have:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} \sum_x |x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

  To explain what's happening here, here's a convenient way to denote H :

$$H = \frac{1}{\sqrt{2}} \sum_{x,y\in\{0,1\}} (-1)^{xy} |y\rangle \langle x|$$

  (check for yourself that this does indeed generate the correct Hadamard matrix). Therefore, the general n-qubit Hadamard gate $H^{\otimes n}$:

$$H^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{x,y\in\{0,1\}^n} (-1)^{x\cdot y} |y\rangle \langle x|$$

  Therefore, we can write:

$$|\psi_1\rangle = H^{\otimes n} |0^{\otimes n}\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = \frac{1}{\sqrt{2^n}} \sum_x |x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

- Now we send $|\psi_1\rangle$ through U_f. What it does is it sends $|y\rangle$ to $|y \oplus f(x)\rangle$. If $y = 0$, then we just output $f(x)$, and if $y = 1$, thne we output the *complement* of $f(x)$, since if $f(x) = 1$ then the addition modulo 2 would return us 0, and vice versa. Therefore, we can write $|\psi_2\rangle$ as:

$$|\psi_2\rangle = \frac{1}{\sqrt{2^n}} \sum_x |x\rangle \otimes \frac{1}{\sqrt{2}}(|f(x)\rangle - |\overline{f(x)}\rangle)$$

  If $f(x) = 0$, then the ancilla (the last qubit) is $|0\rangle - |1\rangle$, and if $f(x) = 1$, then the ancilla is $|1\rangle - |0\rangle = -(|0\rangle - |1\rangle)$. So in general, the ancilla is $(-1)^{f(x)}(|0\rangle - |1\rangle)$. Therefore, we can write:

$$|\psi_2\rangle = \frac{1}{\sqrt{2^n}} \sum_x (-1)^{f(x)} |x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

  We can write $(-1)^{f(x)}$ because when $f(x) = 0$ then $(-1)^{f(x)} = 1$, which doesn't change the product at all, but it changes when $f(x) = 1$, which is what we want.

- Finally, we act $H^{\otimes n}$ on the data register. Note that $H^{\otimes n}\left|x\right\rangle = \sum_y (-1)^{xy}\left|y\right\rangle$, so this gives us:

$$\left|\psi_3\right\rangle = \frac{1}{\sqrt{2^n}}\sum_y \sum_x (-1)^{f(x)+(x\cdot y)}\left|y\right\rangle \otimes \frac{1}{\sqrt{2}}(\left|0\right\rangle - \left|1\right\rangle)$$

- Now we measure all n qubits. If $f(x)$ is constant,

6 Complexity Classes

- Last lecture we saw the Deutsch-Josza algorithm, which required only a single query to determine the identity of $f(X)$. However, in order to show that quantum computation is truly beneficial, then what we really need to show is that no classical algorithm could *possibly* do as well as a quantum one.
- In the case of the Deutscsh-Josza problem, a simple randomized algorithm that samples $f(x)$ at random also achieves a similar accuracy with comparable queries. For instance, 2 queries already give us a probability of error of less than 1/3.
- In general, for a probability of error less than $\frac{1}{2^n}$, we only need $n+1$ queries.
- What we want is a problem in which the quantum algorithm *exponentially* speeds up the computation process.
- Some algorithms which show an exponential gap:
  - Bernstein-Vazirani
  - Simons algorithm

  We won't really go over these too in depth, but they are described in great detail in many textbooks.

6.1 Quantum Complexity

- This is the study of computation using information encoded in quantujm bits, or quantum states $\left|\psi\right\rangle = \alpha\left|0\right\rangle + \beta\left|1\right\rangle$.
- The quantum circuit is the description of the algorithm we apply to get a result.
- We need to determine the number of gates needed, and precisely the number of 2-qubit gates.
- For an arbitrary circuit with n qubits, we need no more than 2^n gates. This is because 2^n denotes a situation where every bit is connected with every other gate. However, for an efficient circuit, we only need a polynomial n^c number of gates, where c is some real, finite constant.
- So really, what we're studying is the class of algorithms which are efficiently implementable on n qubits, which is described by a sequence of $O(n)$ qubit gates drawn from a universl set. Mathematically, we write:

$$U \approx U_{g_k}\cdots U_{g_2}U_{g_1}, \;\; k = O(\text{poly}(n))$$

- There's an approximation symbol because of the **Solovay-Kitaev** theorem, which states that the accuracy of a combination of 1-qubit gates is approximated by $O(\log^c(1/\epsilon)$ with $c \approx 2$.

  By extension, this means that 1 qubit and m 2-qubit gates requires at most $O(m\log^c(1/\epsilon)$.
- This means that we have one universal gate set and we wanted to switch to another, we only have to pay a constant factor of change.

  The Wikipedia Article on this does a better job explaining this.

6.2 Quantum Complexity Clasification

- Generally, we will describe algorithms with respect to the size of the problem (generally the number of qubits n).

  Here, the number of qubits refers to the number of data qubits.

- We also might want to consider the resource scaling, so for instance the number of ancillas, the number of measurements, and also the number of gates.
- **Classical Church-Turing thesis:** Any computable function (classical algorithm) can be modeled by the running of some Turing machine.

  A Turing machine is a machine with a finite set of states, and an infinitely long input tape to which it can read and write from. The Turing machine also defines a transition matrix to transition between states. Note that this thesis does not talk about efficiency at all; the program could run infinitely and still be valid.
- **Strong Church-Turing Thesis (ECT):** A probabilitstic Turing machine, which has the same input and output tape but instead jumps randomly at each step, can efficiently (in polytime) simulate any realistic model of computation.
- Quantum computation challenges this fact – Shor's algorithm factors numbers in $O(\text{poly}(n))$, but there is no known polynomial time for factoring large numbers.
- **Strong Quantum Church-Turing thesis:** A quantum Turing machine can efficiently simulate any realistic (includes quantum) model of computation.

  Anything that's polynomial in L (the input size in bits) is considered easy, and anything that's superpolynomial is considered hard.
- A review on different algorithms and their classical runtimes:
  - Matrix multiplication: $O(n^3)$, with optimizations $O(n^{2.37})$
  - Sorting: $O(n \log n)$, also theorized to be a lower bound.
  - Factoring: number field sieve, has $O(e^{n^{1/3}(\log(n))^{2/3}}$
- Classical complexity classes:
  - P: problems which are solvable in polynomial time.
  - NP: problems which are verifiable in polynomial time.

    Here, the verifier takes in the proposed input from the solver, and we ask whether the verifier can check the solution in polynomial time.
  - It's believed that P $\subset$ NP.
  - NP-Hard: any problem in NP can be converted to an NP-Hard problem within polynomial time.
  - NP-Complete: A problem which is NP-Hard and is also in NP.
  - PSPACE: problems which are solvable in polynomial number of bits, with no constraints on time. It's trivial to show that P $\subset$ PSPACE, but we don't know how close the two sets are to each other.
  - BPP: problems which are efficiently solvable by a randomized algorithm up to some error.

    Does this refer to polynomial in the error?
  - BQP: problems which are efficiently solvable on a quantum computer with an allowed error.
  - Quantum Merlin Arthur: a decision of the following form:

    If the answer is YES, then Merlin (the solver) has a quantum state to which the verifier can verify in polynomial time on a quantum computer.
- What is known:
  - P $\subseteq$ BPP $\subseteq$ BQP $\subseteq$ PSPACE

7 Quantum Fourier Transform

- This is perhaps the most important lecture in all of quantum information.
- Classically, the Fourier Transform is defined as:

$$y_k \equiv \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} x_j e^{2\pi i jk/2^n}$$

- The important thing about this definition is the phase factors. If we let $N = 2^n$, the form looks a bit more familiar: =

$$y_k = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \omega^{-jk} f_k$$

  where $\omega^{-jk} = \exp\left(-2\pi i/N\right)$.
- Quantumly, the Quantum fourier transform is:

$$|j\rangle \to \frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} e^{2\pi i jk/2^n} |k\rangle$$

  The states $|k\rangle$ are analogous to the phases in the classcal Fourier transform, and are represented by an integer basis state.
- So an arbitrary state $|\psi\rangle$ can be written as:

$$|\psi\rangle = \sum_{j=0}^{2^n-1} y_j |k\rangle$$

- Here, we'll represnet j as a binary number: $j = j_{12}^{n-1} + j_2 2^{n-2} + \cdots + j_n 2^0$, and let k be defined similarly.
- Now, we're going to write k out in terms of this binary expansion, and we get:

$$\frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} e^{2\pi i j k_1 2^{-1}} \times e^{2\pi i j k_2 2^{-2}} \times \cdots \times e^{2\pi i j k_n 2^{-n}} |k_1 k_2 \ldots k_n\rangle$$

- After a bit more math (see notes), we ultimately get the following expansion:

$$\frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i j 2^{-1}} |1\rangle)(|0\rangle + e^{2\pi i j 2^{-2}} |1\rangle) \cdots \Big(|0\rangle + e^{2\pi i j 2^{-n}} |1\rangle\Big)$$

  So this is basically just a tensor product of a bunch of states, each with a different phase factor $e^{2\pi i j/2^k}$ where k ranges from 1 to n.
- To put this in a form which we can exploit, we'll rewrite the quantity $\frac{j}{2^k}$ in binary fractional form, so that it's easy to work with. Specifically, write $\frac{j}{2^k}$ as a "decimal":

$$\begin{aligned} \frac{j}{2^k} &= \sum_{\nu}^{n} j_\nu 2^{n-\nu-k} \\ &= j_1 j_2 \ldots j_{n-k} . j_{n-k+1} \ldots j_n \end{aligned}$$

- Now we use this representation on the phase factors we found earlier. For the values of $\frac{j}{2^k}$ that are larger than 1, notice we can just pull it out and instead represent it entirely by a binary fraction:

$$e^{2\pi i \frac{j}{2^k}} = 1 \times e^{2\pi i (0.j_1 j_2 \ldots j_n)}$$

  Doing this for every term, we get:

$$\frac{1}{\sqrt{2^n}} \Big(|0\rangle + e^{2\pi i (0.j_n)} |1\rangle\Big) \Big(|0\rangle + e^{2\pi i (0.j_{n-1} j_n)} |1\rangle\Big) \cdots \Big(|0\rangle + e^{2\pi i (0.j_1 j_2 \ldots j_n)}\Big)$$

  Why does the decimal representation get more complex as we go down the list?

- To see how to implement this, first notice that the $n-l+1$-th qubit is determined by the expression:

$$\frac{1}{\sqrt{2}}\left(|0\rangle + e^{2\pi(0.j_l\dots j_n)}\right)$$

We'll first pull out the first component of the phase:

$$\frac{1}{\sqrt{2}}\left(|0\rangle + e^{2\pi i(0.j_l)} \times e^{2\pi(0.0j_{l+1}\dots j_n)}\right)$$

This first component can be rewritten as:

$$\frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i(0.j_l)} = \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i j_l/2}|1\rangle) = \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{j_l}|1\rangle)$$

where we've used the definition of the binary fraction and also the fact that $e^{i\pi j_l} = (-1)^{j_l}$. This looks exactly like what we would get if we acted the Hadamard gate on $|j_l\rangle$!

This is becuase $(-1)^{j_l} = 0$ if $j_l = 0$, so we get the $+$ in the middle, and the $-$ in the middle if we had $j_l = 1$.

- With this first bit taken care of, we repeat the same process to get the second gate, and so on. Therefore, the rest of the components can be done by implementing a series of rotations:

$$R_k = \begin{bmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{bmatrix}$$

with each value being controlled by the value of the j_k-th qubit. This is becuase we only want the rotation to be applied when $j_k = 1$.

- So in general, we just apply a bunch of controlled H gates to get the final state.

- However, when the state is prepared, notice that the states are in reverse order of what we want them to be. That's okay, however, since we have SWAP gates that allow us to swap the states so that we get them in the right order. Recall that a SWAP gate is basically just three CNOT gates put alongside each other in reversed order.

- In terms of runtime complexity, the total number of R gates needed is $n(n+1)/2$, which means that the Quantum Fourier transform is $O(n^2)$. Compared to the classical algorithm which completes the same task in $O(n\log n)$, this is much faster.

- The QFT is also unitary, just like the Classical FT. This makes sense, since had it not been unitary then we'd have massive issues.

8 Schor's Algorithm

8.1 Modular Exponentials

- A modular exponential is a function of the form $f(x) = a^x \pmod{N}$, hwere a and N are assumed to be coprime (i.e. they only share 1 as a common factor).

- The order of $f(x)$ is given by the smallest integer r such that $f(r) = a^r \pmod{N} = 1$. Since r is defined this way, r can also be thought of as the period of $f(x)$.

Proof. Since $f(r) \equiv 1 \pmod{N}$, we know that $a^r = kN + 1$, so

$$a^{r+1} = kNa + a \equiv a \pmod{N}$$

And as a result, $a^{r+x} \equiv a^x \pmod{N}$, so in general $f(r+x) = f(x)$, so r is indeed the period. □

- Classically, there are three possible solutions for the value of r:

  i) r is odd. This is not useful for finding nontrivial factors of N.

ii) r is even and $a^{r/2} \pmod N \equiv -1$. This implies that $a^{r/1} + 1 = kN$; this is considered a "trivial factor" and isn't useful to us for finding factors.

iii) r is even and $a^{r/2} \not\equiv 1$. This implies (from classical number theory) that at least one of N or $a^{r/2} \pm 1$ (specifically the gcd of the two) is a non-trivial factor of N.

- Therefore, we want to basically find a good value of r, but how do we find one in the first place?

  Our strategy will be to evaluate $f(x)$ at many values in parallel via quantum superposition, and use QFT (quantum fourier transform) to detect the period in the sequence/distribution of values.

8.2 Quantum Period Finding

- We will make use of 2 registers: the first will store K qubits such that $Q = 2^K$, and that $N^2 \le Q \le 2N^2$. The second qubit will store at least $n = \log_2 N$ qubits.
- This is a six step process:

1. **Initialization:** we start with a state $|0\rangle^{\otimes K} \otimes |0\rangle^{\otimes N}$
2. **Transform to equal superposition:** Use the Hadamard gate to make everything into a a superposition:

$$H^{\otimes K}|0\rangle = \frac{1}{\sqrt{2^K}} \sum_y |y\rangle$$

The same superposition can also be constructed by using the quantum fourier transform:

$$|q\rangle \to \sum_{q=0}^{Q-1} \exp\left(\frac{2\pi i q q'}{Q}\right)|q'\rangle$$

then set $q = 0$ to get the same state. This confirms the earlier statement that QFT and the Hadamard gate are very similar.

3. **Apply Quantum U_a that implements modular exponential:** In essence, it takes $q \to f(q) = a^q \pmod N$, and a is a random number between 1 and N.

This function is easy to compute classically; it turns out that we can model every reversible classical computation with a quantum circuit Is this a result of the quantum church-turing thesis?

Note that all the values $f(q)$ are distinct for all $q \in [0, r-1]$; this is a result of a and N being coprime.

We now apply f to register 1, and store the result in register 2. This means our state is of the form:

$$\frac{1}{\sqrt{Q}} = \sum_{q=0}^{Q-1} |q\rangle\, |a^q \pmod N\rangle$$

So the second state is in the state given by $f(q)$.

4. **Measure Register 2:** Note that even though we have Q values, there are only r distinct ones due to the periodicity of $f(q)$.

We'll simplify by assuming that Q is an integer multiple of r. This tells us that there are exactly m differnet q with the same value for $f(q)$.

This is not a necesary condition?.

Therefore, there are $m = \frac{Q}{r}$ different states in register 1 which contribute to the state after measuring register 2 (remember that we've collapsed the state into those that are $f(q) \equiv m \pmod N$), of which there are m states in regsiter 1 that contribute to this state. We can write this as:

$$\frac{1}{\sqrt{Q/r}} \sum_{j=0}^{Q-1} |jr + q_0\rangle\, |f(q_0)\rangle$$

Note now that the periodicity has shifted to register 1, since register 2 now measures a defined value $f(q_0)$. At this point, we no longer need register 2, so we'll stop writing it down.

5. **Periodic Superposition of states in Register 1:** Now we apply the quantum Fourier transform modulo Q, so we get:

$$\frac{1}{\sqrt{Q/r}}\sum_{j=0}^{Q-1}|jr+q_0\rangle \to \frac{1}{\sqrt{r}}\sum_{k=0}^{r}\omega^{kq_i}\left|\frac{kQ}{r}\right\rangle$$

But now we've reduced the number of states from Q/r to r terms:

$$|\Phi_{q_0}\rangle = \frac{1}{\sqrt{Q}}\sum_{a=0}^{Q-1} g(a)\,|a\rangle$$

where $g(a)=\sqrt{\frac{r}{Q}}$ if $a-q_0$ is a multiple of r, and 0 otherwise. Basically, although this looks like a superpositoin of Q states, it is in reality only a superposition of r states, since a lot of them are zero.

We apply QFT again and we get:

$$\sum_c \frac{\sqrt{r}}{Q}\sum_{j=0}^{Q/r-1}\exp\left(2\pi i(jr+q_0)c\right)\,|c\rangle = \sum_c \frac{\sqrt{r}}{Q}\left[\sum_{j=0}^{Q/r-1}\exp\left(2\pi ij\left(\frac{rc}{Q}\right)\right)\exp\left(\frac{2\pi i q_0 c}{Q}\right)\right]|c\rangle$$

Then, since $Q=mr$ is measured, then the state is:

$$\sum_c \sqrt{\frac{r}{Q}}\frac{Q}{r}\exp\left(\frac{2\pi i q_0 c}{Q}\right)|c\rangle$$

If we then write $c=\frac{kQ}{r}$

9 Quanutm Gates

- The discussion here is the preamble to decoherence and why quantum computers "don't work" (yet).
- A classical NOT gate can be implemented using a transistor and some resistors. Such a circuit cannot be built quantumly, and we store information in qubits.
- **Qubit:** There's two different definitions that we use:
  - A two-dimensional hilbert space – this is the theoretically ideal picture of what a qubit is.
  - A physical objects that acts a lot like an ideal object – this is what experimentalists deal with.

  What we will try to investigate is the actual physics of quantum computers.
- Now let's start talking about the former: let $\mathcal{H}$ be a Hilbert space, spanned by the set $\{\,|0\rangle\,,\,|1\rangle\}$, which is also written as the set $\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}$.
- If we were to fully characterize the quantum system, then the quantity we would like to know the most is the Hamiltonian H. In a 2-level system, then Hamiltonian is written as:

$$H = E_a\,|a\rangle\,\langle a| + E_B\,|b\rangle\,\langle b|$$

  Further, the Hamiltonian is an example of an *observable*. Observables have the property that their eigenvalues are real, which implies that they must be Hermitian operators.

9.1 Time Evolution of Quantum Systems

- The time evolution of quantum systems is given by the Schrödinger equation, writtne as:

$$i\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = H(t)\left|\psi(t)\right\rangle$$

  In introductory quantum mechanics, we're usually interested in the stationary states, where we're given the Hamiltonian and asked to solve for the Eigenstates. Here, we'll be interested more in the dynamics of the system. So, this means that we're more interested in the PDE in this equation more than $\left|\psi\right\rangle$ itself.

- Let's look at a simpler example first, the ODE counterpart to this equation (in a sense):

$$\frac{\mathrm{d}y}{\mathrm{d}t} = hy \implies y = Ce^{ht}$$

  Why is this the case? Well, because for every small time step, we're saying that the amount that we increase by is equal to the previous value. This implies that we are growing exponentially over time, which is why we have the exponential term. Specifically, we can then write the increase as

$$\lim_{N\to\infty}\left(1 + h\frac{t}{N}\right)^N$$

  This trick also works with matrices! So, by analogy, the solution to $\left|\psi(t)\right\rangle$ can be written as:

$$\left|\psi(t)\right\rangle = e^{-iHt/\hbar}\left|\psi(0)\right\rangle = \mathcal{U}(t)\left|\psi(0)\right\rangle$$

  Here, the exponentiation can be resolved by taking a Taylor expansion.

- We can also take functions of matrices: consider $f(H) = f(\mathcal{U}\Lambda\mathcal{U}^\dagger) = \mathcal{U}f(\Lambda)U^\dagger$. This can be proven easily by taking a Taylor expansion of f. This is also easy to calculate, since if Λ is a diagonal matrix, then we have the identity:

$$f(\Lambda) = \begin{pmatrix} f(a) & & \\ & f(b) & \\ & & \ddots \end{pmatrix}$$

  basically, we apply f to all the elements on the diagonal.

- Now, let's work with an example: consider a Hamiltonian $H = -\frac{1}{2}g\mu_B\vec{B}\cdot\vec{\sigma}$, where $\vec{B}$ is the magnetic field, and $\vec{\sigma}$ is the vector characterized by the Pauli spin matrices. We can expand this out as:

$$H = -g\mu_B(B_x\sigma_x + B_y\sigma_y + B_z\sigma_z)$$

  If our magnetic field is only in the x direction (for simplicity), then we can write

$$H = -g\mu_B B_x\sigma_x$$

- We want to find time evolution, and as discussed earlier, the unitary matrix that describes the time evolution is given by e^{-iHt}, and given H, therefore we have:

$$\mathcal{U}(t) = \exp\left(i\alpha B_x\sigma_x t\right)$$

  Now, we want to write this in a nicer form, so first we leverage the fact that $\sigma_x = H\sigma_z H$, and since H is Hermitian, then this satisfies the relation $f(\mathcal{U}\sigma_z\mathcal{U}) = \mathcal{U}f(\sigma_z)\mathcal{U}$, so we have:

$$\mathcal{U}(t) = H\exp\left(i\alpha B_x t\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right)H$$

  Applying the exponent to the diagonal matrix, we eventually get:

$$\mathcal{U}(t) = \begin{pmatrix}\cos(\alpha B_x t) & i\sin(\alpha B_x t)\\ i\sin(\alpha B_x t) & \cos(\alpha B_x t)\end{pmatrix}$$

- This result allows us to conclude a much more general fact. Recall De Moivre's formula $e^{i\phi} = \cos(\phi) + i\sin(\phi)$, and more generally with matrices:
$$e^{i\phi\hat{n}\cdot\vec{\sigma}} = \cos(\phi)\mathbb{1} + i\sin(\phi)\hat{n}\cdot\sigma$$
- Now that we've found $\mathcal{U}(t)$, let's see what happens when we act it in $|\psi(0)\rangle$:
$$\begin{aligned}|\psi(t)\rangle &= \mathcal{U}(t)\,|\psi(0)\rangle \\ &= \begin{pmatrix}\cos(\omega t) & i\sin(\omega t) \\ i\sin(\omega t) & \cos(\omega t)\end{pmatrix} \\ &= \begin{pmatrix}\cos(\omega t) \\ i\sin(\omega t)\end{pmatrix}\end{aligned}$$
If we measure our state, the probabiilty that we get the state $|0\rangle$ is given by $\cos^2(\omega t)$, which is also known as *Rabi Oscillation.*
- The Hamiltonian we wrote down earlier is nice, but it's impractical because it's *too* ideal. What we'll look at next is how to deal with small imperfections in our Hamiltonian.

9.2 Energy Splitting

- Now, consider a Hamiltonian $H_0 = \frac{1}{2}\omega_0\sigma_z = \begin{pmatrix}\omega/2 & \\ & -\omega/2\end{pmatrix}$. Under time evolution, the basis states evolve as:
$$\begin{aligned}\begin{pmatrix}1 \\ 0\end{pmatrix} &\to \begin{pmatrix}e^{-i\omega t/2} \\ 0\end{pmatrix} \\ \begin{pmatrix}0 \\ 1\end{pmatrix} &\to \begin{pmatrix}0 \\ e^{i\omega t/2}\end{pmatrix}\end{aligned}$$
So, we can write out the time evolution operator
$$\mathcal{U}(t) = \begin{pmatrix}e^{-i\omega t/2} & \\ & e^{i\omega t/2}\end{pmatrix}$$
What H does to the basis states are easily computable, but they also happen to be the most boring state evolutions. What's more interesting is what happens to, say, the $|+\rangle$ state:
$$|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1 \\ 1\end{pmatrix} \to \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\omega t/2} \\ e^{i\omega t/2}\end{pmatrix} = \begin{pmatrix}\cos(\omega t/2) - i\sin(\omega t/2) \\ \cos(\omega t/2) + i\sin(\omega t/2)\end{pmatrix} = \cos(\omega t/2)\,|+\rangle - i\sin(\omega t/2)\,|-\rangle$$
So if we measure this state in the $|+\rangle$, $|-\rangle$ basis, then we find that there are moments where the state is entirely in the $|+\rangle$ state, and also moments where we're entirely in the $|-\rangle$ state. Here, the frequency ω is also referred to as the **Larmor frequency**.

18 Fault Tolerance

Suppose we have a quantum computer that encodes the following qubits:
$$\begin{aligned}|\bar{0}\rangle &= |000\rangle \\ |\bar{1}\rangle &= |111\rangle\end{aligned}$$
As we've seen before, this corrects against Z-errors (bit flips) very well, but is susceptible to X-errors. Now suppose that this code goes into a channel with a probability p of flipping the bit, and a proability $1 - p$ of it staying the same. There are three states that exist with one error:
$$|001\rangle\,,\; |010\rangle\,,\; |100\rangle$$

since each one has a probability p of occurring, then the probability we end up in either one of these states is $3p$. There are also three states with two errors:

$$|101\rangle\,,\ |011\rangle\,,\ |110\rangle$$

these codes, when checked with our error correction, will send these states to $|111\rangle$, which results in a logical error. The probability of this happening is $3p^2$, since that's the probability we get sent to any one of these qubits. Therefore, we say that this code sends $p \to 3p^2$.

why is this not $3p^2 + p^3$?

18.1 Concatenation

What if we wanted to correct against more errors? We can concatenate the code

19 Phsyical Realization

- While the physical relaization for multiple qubit systems differ, the approach to creating a single-qubit gate is almost identical across all methods.
- The simplest two-level system is an electron subject to a magnetic field $\vec{B} = B_z \hat{z}$. The Hamiltonian of this system is given by $H = -\vec{\mu} \cdot \vec{B} = -g\mu_B B_z \hat{S}_z = -\frac{g\mu_B}{2} B_z \sigma_z$. This g-factor is very approximately equal to 2 (with very small corrections due to quantum field theory).

  μ_B is called the *Bohr Magneton*, which is made up of other fundamental constants:

  $$\mu_B = \frac{e\hbar}{2m} \approx 9.27 \times 10^{-24}\ \mathrm{J/T}$$

- The eigenstates are going to be either aligned or antialigned to the magnetic field, as we would naturally expect. We will define the "spin down" to be the anti-aligned state, and it is separated from the "spin up" (aligned state) state by an energy ΔE.
- If we prepare a superposition of $|\uparrow\rangle$ and $|\downarrow\rangle$, then the state will actually precess around the magnetic field. Specifically, we can calculate the rate or precession:

  $$\frac{\mathrm{d}\vec{L}}{\mathrm{d}t} = \vec{\tau} = \vec{\mu} \times \vec{B}$$

- Starting with a state of the form $|\psi\rangle = \alpha |\downarrow\rangle + \beta |\uparrow\rangle$, then based on the Schrödinger equation, this means that the time evolution of the system is:

  $$i\hbar \frac{\partial}{\partial t} |\psi\rangle = H |\psi\rangle$$

  Recall that:

  $$H = -g\mu_B B_z \hat{S}_z = -\frac{\Delta E}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

  This gives us:

  $$i\hbar \left(\frac{\mathrm{d}\alpha}{\mathrm{d}t} |\downarrow\rangle + \frac{\mathrm{d}\beta}{\mathrm{d}t} |\uparrow\rangle \right) = -\frac{\Delta E}{2} |\downarrow\rangle + \frac{\Delta E}{2} |\uparrow\rangle$$

  Matching the evolution of states, this gives us the following differential equations:

  $$i\hbar\dot{\alpha} = -\frac{\Delta E}{2}\alpha \quad i\hbar\dot{\beta} = \frac{\Delta E}{2}\beta$$

  This gives solutions

  $$\alpha(t) = e^{i\Delta E t/2\hbar}\alpha_0 \quad \beta(t) = e^{-i\Delta E/2\hbar}\beta_0$$

  So, our state is:

  $$|\psi(t)\rangle = e^{i\Delta E t/2\hbar}\alpha_0 |\downarrow\rangle + e^{-i\Delta E t/2\hbar}\beta_0 |\uparrow\rangle = \alpha_0 |\downarrow\rangle + \beta_0 e^{-i\Delta E t/\hbar} |\uparrow\rangle$$

  where α_0, β_0 refer to the initial distribution.

- The frequency of precession is given by the Larmor frequency $\omega_L = \Delta E/\hbar = \gamma_e B_z$.
- So how do we realize someting like an X-gate? We need to be able to flip the state from the $|\uparrow\rangle$ state into the $|\downarrow\rangle$ state. To do this, we apply an additional, much weaker alternating magnetic field along either $\pm\hat{x}$ or $\pm\hat{y}$. This method of flipping is called Rabi flopping.

19.1 Rabi flopping

- The first thing we'll do is abstract away from the spin picture, and instead label the state $|g\rangle$ and $|e\rangle$, separated by an energy $\hbar\omega_0$. We now send in a plane wave with energy $\hbar\omega$.
- We will first break up our Hamiltonian based on perturbation theory: $H = H_0 + H'$. We will also define the ground state energy to be 0, so $|e\rangle$ has energy $\hbar\omega_0$. Hence, we can write $H_0 = \hbar\omega_0 |e\rangle\langle e|$.
- H' consists of a dipole operator, which quantifies how strongly the particle interacts with the surrounding field. For an electric interaction, then we'd have $H' = -e\hat{r}\cdot\vec{E}$, and with a magnetic interaction we could have $H'_x = g\mu_B B_x \sigma_x$. In general, we can write $H' = \hat{d}\cdot\vec{E}$.

  To abstract away from this, we will insetad consider the coherence between $|e\rangle$ and $|g\rangle$ by looking at the off-diagonal elements of the density matrix:
  $$\langle e|\hat{d}|g\rangle = \mu_{eg}^* \quad \langle g|\hat{d}|e\rangle = \mu_{eg}$$
  So, the dipole operator may be written as:
  $$\hat{d} = \mu_{eg} * |g\rangle\langle e| + \mu_{eg}|e\rangle\langle g|$$
  As for $\vec{E}$, we will choose a plane wave, so that the oscillation at the location of the particle would be just a sine or cosine wave:
  $$\vec{E} = \frac{\mathcal{E}}{2}e^{-i\omega t} + \frac{\mathcal{E}^*}{2}e^{i\omega t}$$
- So, our full hamiltonian is written as:
  $$H = \hbar\omega_0 |e\rangle\langle e| - (\mu_0 g|e\rangle\langle g| + \mu_{eg}^*|g\rangle\langle e|)\cdot\left(\frac{\mathcal{E}}{2}e^{-i\omega t} + \frac{\mathcal{E}^*}{2}e^{i\omega t}\right)$$
  with a wavefunction $|\psi\rangle = \alpha|g\rangle + \beta|e\rangle$. Based on the Rabi oscillation example we had earlier, we can write α and β as a function of time. To be consistent with the previous notes, we will use $c_g(t) = \alpha(t)$, and $c_e(t) = \beta(t)e^{i\omega t}$.

  With this redefinition, we can write
  $$|\psi\rangle = c_g(t)|g\rangle + c_e(t)e^{-i\omega t}|e\rangle$$
- Now, we do the Schrödinger equation, and we will split this into two equations by projecting the Schrödinger equation into the two basis states:
  $$\begin{aligned}\langle g|i\hbar\frac{\partial}{\partial t}|\psi\rangle &= \langle g|H|\psi\rangle \\ \langle e|i\hbar\frac{\partial}{\partial t}|\psi\rangle &= \langle e|H|\psi\rangle\end{aligned}$$
  The first equation is relatively easy:
  $$i\hbar\frac{\mathrm{d}c_g}{\mathrm{d}t} = -\mu_{cg}^* e^{-i\omega t}c_e(t)\vec{E} = -\mu_{eg}^* c_e(t)\left(\frac{\mathcal{E}}{2}e^{-2i\omega t} + \frac{\mathcal{E}^*}{2}\right)$$
  The second equation, due to the extra phase factor in $c_e(t)$, we get:
  $$\hbar\omega c_e(t)e^{-i\omega t} + i\hbar\frac{\mathrm{d}c_e(t)}{\mathrm{d}t}e^{-i\omega t} = \mu_{eg}c_g(t)\vec{E} + \hbar\omega_0 e^{-i\omega t}c_e(t)$$
  Combining the terms:
  $$i\hbar\frac{\mathrm{d}c_e(t)}{\mathrm{d}t} = -\mu_{eg}e^{i\omega t}c_g(t)\left(\frac{\mathcal{E}}{2}e^{-i\omega t} + \frac{\mathcal{E}^*}{2}e^{i\omega t}\right) - \hbar\Delta c_e(t) = -\mu_{eg}c_g(t)\left(\frac{\mathcal{E}}{2} + e^{2i\omega t}\frac{\mathcal{E}^*}{2}\right) - \hbar\Delta c_e$$
  Here, we define $\Delta = \omega - \omega_0$, known as the detuning frequency.

- Now, we see that in both cases, the differential equation has a very rapidly oscillating term at a frequency 2ω. Becuse they oscillate so fast, it's relatively safe to ignore them, which leaves us with the following differential equations:

$$\frac{\mathrm{d}c_g}{\mathrm{d}t} = \frac{i\Omega^*}{2}c_e$$
$$\frac{\mathrm{d}c_e}{\mathrm{d}t} = i\Delta c_e + \frac{i\Omega}{2}c_g$$

  we define $\Omega = \frac{\mu_{eg}\mathcal{E}}{\hbar}$ and $\Omega^* = \frac{\mu_{eg}^*\mathcal{E}^*}{hbar}$.

- If we start with the initial conditions $c_g(0) = 1$ and $c_e(0) = 0$, then the probability as a function of time is:

$$|c_g(t)|^2 = \cos^2\left(\frac{\Omega' t}{2}\right) + \frac{\Delta^2}{|\Omega|^2 + \Delta^2}\sin^2\left(\frac{\Omega' t}{2}\right)$$
$$|c_e(t)|^2 = \frac{|\Omega|^2}{|\Omega|^2 + \Delta^2}\sin^2\left(\frac{\Omega' t}{2}\right)$$

  Here, we define $\Omega' = \sqrt{\Omega^2 + \Delta^2}$. When the detuning frequency is zero (i.e. we send in our pulses, or light, at exactly ω_0), then we get even simpler expressions:

$$|c_g(t)|^2 = \cos^2\left(\frac{\Omega t}{2}\right)$$
$$|c_e(t)|^2 = \sin^2\left(\frac{\Omega t}{2}\right)$$

  We can also stop applying this external perturbation at defined times, and specifically if we stop at $t = \frac{\pi}{\Omega}$, then this corresponds to a complete transition from $|g\rangle$ to $|e\rangle$! This is also called a π-pulse.

19.2 Rotation Axis

- The rotation axis that the particle takes along the bloch sphere depends on the phase of the radiation we apply.

$$H_{r,f} = \frac{\hbar\Delta}{2}\hat{\sigma_z} + \frac{\hbar\Omega}{2}\hat{\sigma_x}$$

- Here, depending on Δ alone, we have full control over what our qubit does.

20 Quantum Computing Platforms

- Trapped ions: qubits are single atoms, but we've removed one of the electrons so they're positively charged. We do this so that we have better control over them. This is the platform that's pursued by Honeywell/ Quantinuum, AQT, among other companies.

  Pros:

  - These have long coherence times $T_2 \sim 1$ minute
  - They operate at room temperature – basically it's just a big vaccuum chamber sitting in a room without the need for cryogenics. Nothing about the qubits require low temperature, we just happen to involve cryogenics in order to achieve a better vaccuum.
  - Highest fidelity gates so far, and have been one of the first to discover quantum gates.

  Cons:

  - Gates operate typically at 50 microseconds.
  - Requires lasers, optics
  - The coulomb interaction between ions makes scaling more difficult

- Neutral Ions: basically the same trapping techniques, except the atoms are neutral instead of charged. We can't trap them using fields because they are neutrally charged.

  Pros:

  - Long qubit coherence times
  - Room temperature operation
  - Inherently somewhat scalable, since neutral atoms don't interact
  - Optical interface.

  Cons:

  - Experiments usually looks like a mess (requires a lots of lasers), and much of the effort goes into managing the lasers
  - Requires an ultrahigh vaccuum (so we need cryogenics basically)
  - Trapping is inherently more difficult and requires high laser power

- Superconducting qubits: these are man-made qubits instead of atoms.

  Pros

  - Chip-based architecture. It's something that we can imagine scaling up to a chip, and we interact with it electronically
  - Fast gate times (on the order of 50 nanoseconds)
  - Previously leading the field commercially, but starting to recognize that superconducting qubits might not be the way. (some comapnies are starting to invest in atomic-based approaches)

  Cons:

  - Requires dilution refrigerators, so need cryogenic temperatures
  - Short coherence times, though this is getting much better
  - Control lines needed for every individual qubit – every single qubit you want to add means an extra set of lines you need to connect to your system.
  - Not very anharmonic.

- Quantum Dots: creating a 2D electron gas, and is possibly the most similar to the structure that we studied in the previous lecture.

  Pros:

  - Semiconductor chip based architecture
  - High fidelity, fast-qubit and two qubit gates.
  - Controlled by microwaves and electronics, there are no lasers required in this process.

  Cons:

  - Requires cryogenic temperatures
  - Short coherence times
  - Lots of tuning required for each device – very sensitive architecture
  - Scaling up to many qubits is still an open problem. We can do 2-qubit gates, but it's not clear how we would scale up beyond that.

- Photonics: none of the stuff that we've talked about so far really applies here. States are single photons, where the information is either stored in the polarization or which rail the photon lives on.

  Pros:

  - Silicon chip based architecture
  - Room temperature operation, what this basically means is that in principle we can do this at room temperature, but the best photon detectors still require cryogenics.
  - Fairly easy to scale up, since photons naturally fly around
  - "Measurement" based – some of the gates are very easy to physically implement. For instance, if you wanted to change the polarization then all you'd do is just introduce a wave plate

  Cons:

  - Gate operations are very difficult to achieve, and are inherently probabilistic. Preparing the state itself is a challenge (trying again and figuring out when you succeed), then performing the desired computation afterwards.
  - Requires identical photonic states and photonic elements
  - Low gate fidelities

20.1 Trapped Ions

- Ideally we'd like to use the simplest atom possible, which in our case would be hydrogen-like atoms. We can't use hydrogen itself, because transitions for hydrogen are in the UV spectrum, which poses some challenges (what specifically?)
- Commercially, we normally use alkaline earth metals and then strip one electron off so that we end up with a hydrogen-like atom.
- Hydrogen-like atoms have a principal quantum number n, whose energy scales with:

  $$E_n = -\frac{z^2 \frac{\mu}{m} E_H}{2n^2}$$

  E_H is a physical constant, which is written as $E_H = mc^2\alpha^2$
- We also have angular momentum ℓ, which has possibilties between 0 to $n - 1$.
- Transitions between these obey selection rules, which tells us that $\Delta\ell = \pm1$ and $\Delta m = \pm1$.