Contents

1 Lecture 1

1.1 Motivating Quantum Computing

The classical unit of computation is a **bit**. How small can we shrink bits? Let's conduct a thought experiment. Let's suppose we could shrink them down to the size of a Hydrogen atom. The "state" of $|0\rangle$ being the ground state and $|1\rangle$ first excited state. However, electrons in general exist in superposition states! These states look like:

$$\{\alpha |0\rangle + \beta |1\rangle : \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1\}$$

But it gets weirder. According to quantum theory, when conducting a measurement on such a state, we end up getting:

$$M = \begin{cases} 0 & \text{wp } |\alpha|^2 \\ 1 & \text{wp } |\beta|^2 \end{cases}$$

Furthermore, the act of measurement "collapses" the wavefunction to a state $|0\rangle$ or $|1\rangle$. Subsequent measurements will give that pure state deterministically.

Now, suppose we have a system of two such Hydrogen. There are now 4 basis states:

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

Effective computation now comes from extrapolating to n such **qubits**. Now such a state would look like $\sum_{x \in \{0,1\}} \alpha_x |x\rangle$.

This is pretty profound. Classical computers were designed to use nature (through silicon) in order to work for humans. But with all this effective work that nature is doing behind the scenes, it seems that quantum computing is really the more powerful framework we should've asked for.

1.2 Measurement

Now suppose we do a "partial" measurement, e.g. only measuring the first bit. What will we get? It seems reasonable that the probability should be the sum of the probabilities of getting a 0 in the first qubit, e.g. we get a 0 w.p. $|\alpha_{00}|^2 + |\alpha_{01}|^2$. The state collapses, but it must be renormalized so the coefficients can still be probabilities! So the new state is actually

$$\frac{\alpha_{00} |00\rangle + \alpha_{01} |01\rangle}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}}$$

Now suppose we are given a qubit in the state

$$|\psi\rangle = \frac{1}{\sqrt{2}} |0\rangle + \frac{e^{i\theta}}{\sqrt{2}} |1\rangle$$

How can we figure out θ (phase estimation)? Well if we measure this, we will get either 0 or 1 with probability $1/2$ each. This will tell us nothing about θ. It turns out this is only a special case of measurement.

To understand what general measurement is, we first go back to our state representation. What we really mean by a superposition is a linear combination of two vectors. We fix some basis $|0\rangle$ and $|1\rangle$, and a normalized state $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$ is a unit vector in a 2-dimensional complex vector space. Now we can think about a measurement in the following way:

Definition 1.1
A **measurement** of some state $|\phi\rangle$ in some basis $\mathcal{U}$ is a projection onto one of the basis vectors $|u\rangle$. The value of the measurement is: u with probability of the scalar projection squared, $\left|\frac{\langle u|\psi\rangle}{\langle\psi|\psi\rangle}\right|^2$.

So for example, let's stick to our 2-space and pick a new orthonormal basis $\{|u\rangle, |u^\perp\rangle\}$ and our state $|\psi\rangle$. Suppose $|\psi\rangle$ makes an angle of θ with the $|0\rangle$ axis and makes an angle of μ with the $|u\rangle$ axis. By a simple diagram, it's clear from ψ's projections that measurement in the standard basis yields 0 with probability $\cos^2\theta$ and in our new basis it yields u with probability $\cos^2\mu$.

Note 1.1
There is a bit of a subtlety here. We assumed that the amplitudes we are working with are real, but in general they can be complex. It turns out, all of quantum computing can be formalized with only real amplitudes, but it gets more messy when interfacing with physics. For now, we will assume real amplitudes only, but most results generalize to complex amplitudes.

Another common example of a basis is:

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

Measuring our original phase estimation in this new basis is exactly what we need! We just need to write it in the new basis to figure out the amplitudes:

$$\frac{1}{\sqrt{2}}|0\rangle + e^{i\theta}\frac{1}{\sqrt{2}}|1\rangle = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle\right) + e^{i\theta}\frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|+\rangle - \frac{1}{\sqrt{2}}|-\rangle\right) \tag{1}$$

$$= \frac{1}{2}(1 + e^{i\theta})|+\rangle + \frac{1}{2}(1 - e^{i\theta})|-\rangle \tag{2}$$

$$= \frac{1}{2}(1 + \cos\theta + i\sin\theta)|+\rangle + \dots \tag{3}$$

$$\tag{4}$$

so we get $|+\rangle$ from the measurement with probability

$$\frac{1}{2}|1 + \cos\theta + i\sin\theta|^2 = \cos^2(\theta/2)$$

Now we can repeat the measurement (with other processed inputs) to get statistics and thus a good estimate on θ.

2 Lecture 2

2.1 Axioms of Quantum Mechanics

We list some axioms of Quantum Mechanics. Consider an electron with k energy levels, $|0\rangle, |1\rangle, \ldots, |k-1\rangle$.

Note 2.1 (Superposition Principle)
If there are k distinguishable (eigenstates) of a system, then the state of a system can be written as:

$$|\psi\rangle = \sum_{j=0}^{k-1} \alpha_j |j\rangle$$

where $\alpha_j \in \mathbb{C}$ and $\sum_j |\alpha_j|^2 = 1$.

This forms a Hilbert space, i.e. a Complex inner product space (but we will often think of all amplitudes as real). The $\{|j\rangle\}_{j=0}^{k-1}$ forms a basis for this state space. We can think of

$$|\psi\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{k-1} \end{pmatrix}, |0\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots$$

For inner products, we use Dirac's Bra-Ket notation. As we have already seen, the "kets" are regular vectors and the "bras" $\langle\psi| = |\psi\rangle^\dagger$ are elements of the dual vector space (which can be thought of as conjugate transposes). This means:

$$\langle\psi| = |\psi\rangle^\dagger = \sum_j \left(\alpha_j |j\rangle\right)^\dagger = \sum_j \alpha_j^* \langle j|$$

where $(\cdot)^*$ is the complex conjugate.

Now define $|\phi\rangle = \sum_j \beta_j |j\rangle$. We can take inner products by using the following notation:

$$\langle\psi, \phi\rangle = \langle\psi|\phi\rangle = \left(\sum_i \alpha_i^* \langle i|\right)\left(\sum_j \beta_j |j\rangle\right) = \sum_{i,j} \alpha_i^* \beta_j \langle i|j\rangle = \sum_j \alpha_j^* \beta_j$$

Because $\langle i|j\rangle = 1$ if and only if $i = j$ (they form an orthonormal basis).

We generally use $k = 2$, call the Hilbert space generated $\mathcal{H}$. We typically think about chaining together (tensor-producting) this Hilbert space with itself n times. This is called a n-**qubit** state. A general state can then be written as:

$$|\psi\rangle = \sum_{x \in \{0,1\}^n} \alpha_x |x\rangle$$

with $\alpha_x \in \mathbb{C}$ and $\sum_x |\alpha_x|^2 = 1$.

Note 2.2 (Measurement Principle)
Pick an orthonormal basis $\mathcal{U} = |u_0\rangle, |u_1\rangle, \ldots, |u_{k-1}\rangle$. The outcome of a measurement is j with probability $\left|\langle u_j|\psi\rangle\right|^2$. In this process, the state is also perturbed and turned into the state $\left|u_j\right\rangle$

Look at last lecture for examples of measuring in different bases, with real amplitudes one can think about qubit states geometrically. The basis $\{|+\rangle, |-\rangle\}$ serves us well.

2.2 Bell Inequalities

Let us look more closely at combining two qubits, each with states $\alpha_0 |0\rangle + \alpha_1 |1\rangle$, $\beta_0 |0\rangle + \beta_1 |1\rangle$. We (tensor) product them together, producing a state:

$$|\psi\rangle = \alpha_0\beta_0 |00\rangle + \alpha_0\beta_1 |01\rangle + \alpha_1\beta_0 |10\rangle + \alpha_1\beta_1 |11\rangle$$

but most states are not a product of two states.

The Bell basis states are a common example of states which are **entangled**, e.g. cannot be written as "product states."

$$\left|\Phi^{\pm}\right\rangle = \frac{1}{\sqrt{2}} |00\rangle \pm \frac{1}{\sqrt{2}} |11\rangle, \left|\Psi^{\pm}\right\rangle = \frac{1}{\sqrt{2}} |01\rangle \pm \frac{1}{\sqrt{2}} |10\rangle$$

These four states form an orthonormal basis for two qubits.

Suppose your system was in the state Φ^+ and we did a partial measurement on the first qubit. Then with probability $1/2$ we collapse to $|00\rangle$ and with probability $1/2$ we collapse to $|11\rangle$. Note that we could achieve this in a classical sense too, with correlated ("glued") coin flips.

Furthermore, the Bell states are rotationally invariant.

Theorem 2.1
In any basis, we can write the Bell States as:

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

Let's prove this. Suppose $v = \alpha |0\rangle + \beta |1\rangle$. Then without loss of generality, we can write $v^{\perp} = -\beta^* |0\rangle + \alpha^* |1\rangle$. This means that:

$$\begin{aligned} |vv\rangle + \left|v^{\perp}v^{\perp}\right\rangle &= (\alpha |0\rangle + \beta |1\rangle)(\alpha |0\rangle + \beta |1\rangle) + (-\beta^* |0\rangle + \alpha^* |1\rangle)(-\beta^* |0\rangle + \alpha^* |1\rangle) \\ &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \end{aligned}$$

where some algebra is elided. Note that we could achieve this in a classical sense too, with correlated coin flips that are rotated.

To go beyond classical computation, we consider two qubit measurements. The first player measures in the standard basis and the second player measures in a new basis, $\{|v\rangle, |v^{\perp}\rangle\}$, rotated at an angle θ from the standard basis. The probability that these two measurements are unequal is $\sin^2\theta$ (for example, if the first measurement is 0, then the state is $|00\rangle$, so the component of $|v\rangle$ in the $|0\rangle$ direction is $\cos\theta$).

However, classically, the probability that one observes a different outcome is proportional to θ.

So John Bell's experiment is as follows. Alice is given a uniformly random bit x and Bob is given a uniformly random bit y. They must each report back a bit a and b respectively. Alice and Bob "win" the game if $xy = a + b \pmod 2$.

They can play the game in two ways: either classically or quantumly. Classically, they cannot communicate (apart from maybe the "glued" coin). In the quantum setup, Alice and Bob share a Bell state. If Alice chooses a bit 0, they measure their qubit in the standard basis, otherwise they measure it in a basis rotated by $\pi/4$. If Bob chooses a bit 1, they measure their qubit in a basis rotated by $\pi/8$, otherwise they measure in a basis rotated by $-\pi/8$. Call their measured bits a and b respectively.

We then mention the following two facts:

1. No classical strategy can win with probability $> 75\%$. A randomized strategy can do no better than a deterministic strategy since the opponent's strategy is known. The best deterministic strategy is to report $a = 0$ and $b = 0$ (or $a = 1$ and $b = 1$), because $xy = 0$ with probability 75% (if at least one of the bits is 0); trying to force the answer to be 1 will give you a lower probability of success. You can do no better. The glued coin doesn't help you either; the best it could do is give you a shared source of randomness.

2. In each of the 4 cases, the probability winning in a quantum setup is $\cos^2 \pi/8 \approx 85\%$. For example, take the case when x and y are both 0. Then they need to both measure a 1 or both measure a 0. The probability Alice measures a 0 is $1/2$ and then collapses the state to a $|00\rangle$. The probability that Bob then sees a 0 is $\cos^2 \frac{\pi}{8}$ because of the rotation, giving us $\frac{1}{2} \cos^2 \frac{\pi}{8}$. Likewise, the probability Alice measures a 1 is $1/2$ and then collapses the state to a $|11\rangle$. The probability that Bob then sees a 1 is $\cos^2 \frac{\pi}{8}$, so overall the probability is $2 \cdot \frac{1}{2} \cdot \cos^2 \frac{\pi}{8} = \cos^2 \frac{\pi}{8}$. The other cases are similar.

which clearly shows the quantum setup gives us something not present in the classical one.

3 Lecture 3

Recall the superposition and measurement principles from last lecture. They tell us that quantum states inhabit a Hilbert space $\text{span}\{|0\rangle, |1\rangle, \ldots, |k-1\rangle\}$ and we can "measure" in orthonormal basis in this Hilbert space, randomly projecting it onto a basis vector. These were two axioms of quantum mechanics.

3.1 Unitary Evolution

A third axiom of quantum information is the ability to apply a unitary transform. These are ubiquitous in linear algebra, but nonetheless we give a quantum-tuned introduction here.

For 1 qubit, we can think of a unitary transform as a "rigid-body rotation" (rotation/reflection), which preserves the orthogonality of vectors.

$$U = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

With our canonical representation of $|0\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}^T, |1\rangle = \begin{pmatrix} 1 & 1 \end{pmatrix}^T$, this transform can equivalently be stated as:

$$|0\rangle \mapsto a|0\rangle + b|1\rangle, |1\rangle \mapsto c|0\rangle + d|1\rangle$$

Define the adjoint of a matrix as its conjugate transpose, e.g.:

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

Now, we can interpret this in the 2-by-2 case as the following:

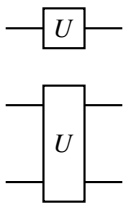
$$UU^\dagger = \begin{pmatrix} a^*a + b^*b & a^*c + b^*d \\ c^*a + d^*b & c^*c + d^*d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The last equality is only true if the columns of U are orthonormal, they are normalized (the top left and bottom right entries are just norms) and have inner product 0.

In general, we have the following definition:

Definition 3.1 (Unitary)
A transform $U \in \mathbb{C}^{n \times n}$ is unitary if and only if $UU^\dagger = U^\dagger U = I$, where I is the n-by-n identity.

Another name for these unitary transform are "quantum gates." We can draw such gates on one or two inputs as the following:



3.2 The Fundamental Quantum Gates

Some simple 1-qubit gates are the following:

Definition 3.2

1. The identity gate, which takes a state and does nothing:

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

2. The rotation gate, which rotates a state by θ:

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

3. The inversion (NOT) gate, which flips a bit:

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

   i.e., $a\,|0\rangle + b\,|1\rangle \mapsto b\,|0\rangle + a\,|1\rangle$.

4. The phase flip gate, which flips the phase of the second bit:

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

   i.e., $a\,|0\rangle + b\,|1\rangle \mapsto a\,|0\rangle - b\,|1\rangle$

5. The Hadamard gate, which converts to the $\{|+\rangle, |-\rangle\}$ basis.

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

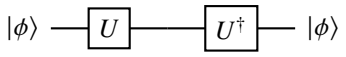
   i.e. $a\,|0\rangle + b\,|1\rangle \mapsto a\,|+\rangle + b\,|-\rangle$. One can view the Hadamard gate as a reflection over the line $\theta = \pi/8$. But hang on, we only allowed rotations, so what gives? It turns out the Hadamard gate is a rotation in $\mathbb{C}^2$, but not in $\mathbb{R}^2$!

Note that $X^2 = Z^2 = H^2 = I$, so they are involutions (and thus their own inverses). Furthermore X and Z are the same under a change of basis (note that $Z\,|+\rangle = |-\rangle$ and $Z\,|-\rangle = |+\rangle$).

$$X = HZH, Z = HXH$$

If you recall the Pauli spin matrices, you may remember X, Z as two of them; however, they are not expressive enough to correspond to all unitaries! Using H is computationally more interesting and is helpful for our analysis.

Let us make a bit of a silly circuit:



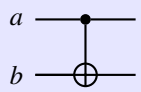
We applied a gate and then applied its adjoint, which is its inverse since it is unitary. Thus, we can always "undo/uncompute" quantum circuits (before measurement).

Now let us define a two-qubit gate,

Definition 3.3 (Controlled NOT)
The CNOT gate is:

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

To draw a CNOT gate, we draw it as the following:

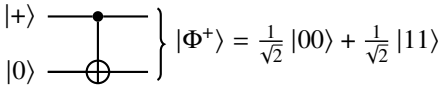


a is called the "control bit" and b is called the "target bit."

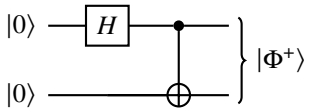
If a and b are pure bits, then we have the following truth table (the first bit controls whether a NOT gate is active, i.e. you XOR the two bits):

a	b	a_o	b_o
0	0	0	0
0	1	0	1
1	0	1	1
1	1	1	0

What if I did the following? I will input $|+\rangle|0\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|10\rangle$.



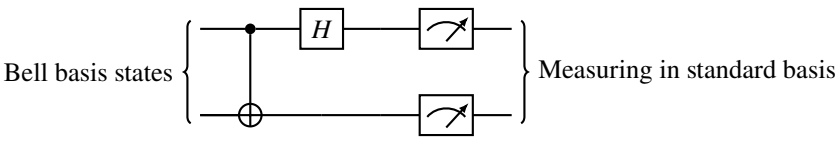
We created a Bell-state. Now to make it from two $|0\rangle$'s, we can add a Hadamard:



Now consider applying this circuit to any two-qubit state. We can do something very similar (the reader can verify the details):

Input	Output
$\|00\rangle$	$\|\Phi^+\rangle$
$\|01\rangle$	$\|\Psi^+\rangle$
$\|10\rangle$	$\|\Phi^-\rangle$
$\|11\rangle$	$\|\Psi^-\rangle$

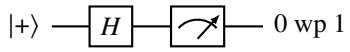
Now, consider turning the circuit backwards. Since both gates are their own inverses (CNOT is made up of a block diagonal of involutions so it is also an involution), this just inverts the circuit. Let us add a measurement apparatus:



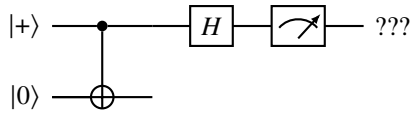
Now, we can measure in the Bell-state basis with *just* using our module for measuring the standard basis. This means without loss of generality we can measure in any basis.

3.3 Intuition for Entanglement

We know that:



But, now what if we entangle the state with a CNOT?



Intuitively, we expected the measurement controlled bit to not get changed, so we should expect the same result as above.

But let's analyze it formally. After the CNOT, the state is $|\Phi^+\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$. Then, the Hadamard doesn't act on the second bit, but the first bit is split, e.g. the final state is:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|10\rangle\right) + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|01\rangle - \frac{1}{\sqrt{2}}|11\rangle\right) = \frac{1}{2}|00\rangle + \frac{1}{2}|01\rangle + \frac{1}{2}|10\rangle - \frac{1}{2}|11\rangle$$

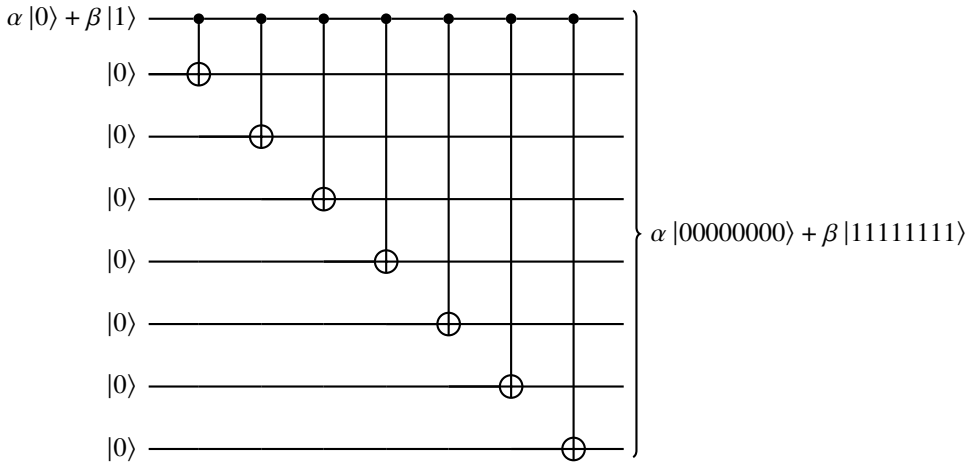
Measuring only the first qubit actually makes it so that we actually have a $\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}$ chance of measuring a 0.

What happened here? In the first case, there is a cancellation of the probability amplitudes:

$$\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \mapsto_H \frac{1}{2}|0\rangle + \frac{1}{2}|1\rangle + \frac{1}{2}|0\rangle - \frac{1}{2}|1\rangle = |0\rangle$$

But in the second case, we have an entanglement which stops this cancellation (the product states cannot just be cancelled).

This gives us a view into what measurement really is. What if we entangle a bunch of bits:



At some macro point, nature cannot support such a large entangled state and collapses it probabilistically into one of the two basis states. This is how measurement is done in practice.

4 Lecture 4

4.1 The Tensor Product

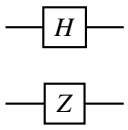
Remember that we "multiplied" two states $|0\rangle |0\rangle = |00\rangle$. This combined two 2-state systems into a 4-state system. But what is this mystical multiplication? The answer is the **tensor product**, denoted by the symbol $\otimes$. Really, it would be proper to say $|0\rangle \otimes |0\rangle = |00\rangle$, in which we "call" the tensor product this nickname. We also saw that we could identify $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ with vectors in $\mathbb{C}^4$, in this sense, we say $\mathbb{C}^2 \otimes \mathbb{C}^2 \cong \mathbb{C}^4$, i.e. the spaces are isomorphic.

In a general setting, letting $\mathcal{H}_1$ be a system with k levels and $\mathcal{H}_2$ be a system with ℓ levels, we say

$$\mathcal{H}_1 \otimes \mathcal{H}_2 \text{ is the vector space spanned by } \{|i\rangle |j\rangle\}_{0 \le i \le k, 0 \le j \le \ell}$$

which is a vector space with dimension $k \cdot \ell$.

We can also talk about the tensor product of operators. When you tensor product two operators, they act on their spaces separately. This is a circuit representation of $H \otimes Z$:



In this circuit, $|00\rangle \mapsto H|0\rangle \otimes Z|0\rangle = |+\rangle |0\rangle$, keeping the tensor products separate. On the other hand, note that the operator can be equivalently written as an element of $\mathbb{C}^{4\times 4}$.

$$H \otimes Z = \begin{pmatrix} |0\cdot\rangle \to |0\cdot\rangle & |1\cdot\rangle \to |0\cdot\rangle \\ |0\cdot\rangle \to |1\cdot\rangle & |1\cdot\rangle \to |1\cdot\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & -\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

The first matrix represents the parts of the subspace that the matrix acts on. The notation $|b\cdot\rangle$ corresponds to the subspace first bit being b and the second bit being anything.

In addition, we will state one more fact about tensor products, which we will not prove:

Theorem 4.1
The inner product multiplies under the tensor product.

$$(\langle u| \otimes \langle v|)(|x\rangle \otimes |y\rangle) = \langle u|x\rangle \langle v|y\rangle$$

4.2 No Cloning Theorem

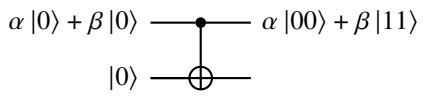
The no cloning theorem states that you cannot create a copy of a qubit. Formally,

Theorem 4.2 (No Cloning)
There is no unitary transform (quantum circuit) U such that

$$U|0\rangle |\psi\rangle = |\psi\rangle |\psi\rangle$$

for all possible states that $|\psi\rangle$ could take on.

But hang on, didn't we see a copying circuit before? Recall the following circuit:



It sent $|0\rangle \mapsto |00\rangle$ and $|1\rangle \mapsto |11\rangle$, which uniquely defines the transformation. But it doesn't work on a general state! By linearity note that:

$$(\alpha|0\rangle + \beta|1\rangle)|0\rangle = \alpha|00\rangle + \beta|10\rangle \mapsto \alpha|00\rangle + \beta|11\rangle \neq (\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha|0\rangle + \beta|1\rangle)$$

This doesn't work.

To prove the no-cloning theorem another way, we then suppose for the sake of contradction that there exists a unitary U that can clone. Then for two vectors:

$$U|\phi\rangle|0\rangle = |\phi\rangle|\phi\rangle, U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle$$

But a unitary cannot change the angle between two vectors. To begin with and end with, the inner product was:

$$\langle\psi|\phi\rangle\langle 0|0\rangle = \langle\psi|\phi\rangle\langle\psi|\phi\rangle \implies \langle\psi|\phi\rangle = 0, 1$$

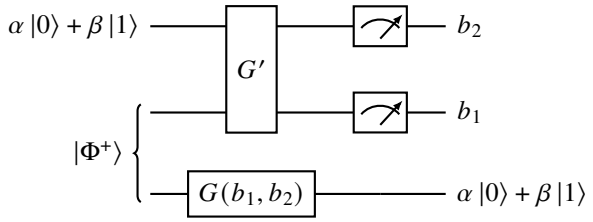
Any unitary can only clone only orthogonal states or the same state. In practice, you can only clone some orthonormal basis of your choosing, like $|0\rangle, |1\rangle$.

4.3 Superdense Coding

Let's say Alice and Bob share a Bell state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$. Alice has two bits of information $b_1, b_2 \in \{0, 1\}$ that she wants to send to Bob. Classically she can send this information in exactly two bits, no more, but with a quantum state, Alice can do it with just one qubit. Based on the bits, she turns the state into one of the Bell states based on the bits. Alice then sends Bob her qubit. Now Bob can just measure the qubit in the Bell basis. But can we do better than rate 2? It turns out we can't.

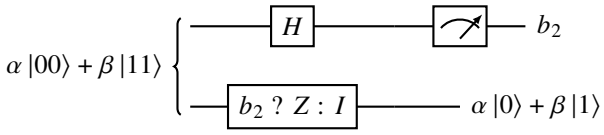
4.4 Quantum Teleportation

Professor Vazirani assures us that Alice has been working hard in her lab and made a state $\alpha|0\rangle + \beta|1\rangle$ that takes a LONG time to make. Unfortunately, she does not have access to an apparatus needed to process this qubit into cool things. Fortunately, she shares a Bell state with Bob and can exchange (classical) information with him, who does have such an apparatus. As we will see, the following circuit allows Alice to share her state with Bob:



where b_1 tells you whether to do a bit flip and b_2 tells you whether to do a phase flip in G.

Let's consider a simpler problem, Alice and Bob have two entangled bits $\alpha|00\rangle + \beta|11\rangle$ and they want Bob to create $\alpha|0\rangle + \beta|1\rangle$. They can use the following circuit to solve this problem:

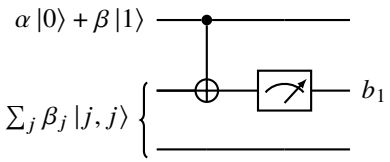


The state becomes:

$$\frac{1}{\sqrt{2}}(\alpha |00\rangle + \beta |01\rangle) + \frac{1}{\sqrt{2}}(\alpha |10\rangle - \beta |11\rangle)$$

Upon measurement, we get 0 and the state is $\alpha |0\rangle + \beta |1\rangle$, or we get 1 and the state is $\alpha |0\rangle - \beta |1\rangle$. She tells Bob this one bit b_2. If it were 1, Alice would tell Bob to apply a phase flip Z to fix the quantum state.

Now, to address the original problem, we can reduce to this case by just getting $\alpha |00\rangle + \beta |11\rangle$. The following subcircuit takes the two states and gives a 0 if the state is $\alpha |00\rangle + \beta |11\rangle$ and a 1 if the state is $\alpha |01\rangle + \beta |10\rangle$. In the second case, a bit flip on the second qubit gets us our target state.



Let's prove this claim, using an index for brevity, where the initial state of the (technically three) qubits is: $\sum_{i,j} \alpha_i |i\rangle \otimes |j, j\rangle$. Under CNOT:

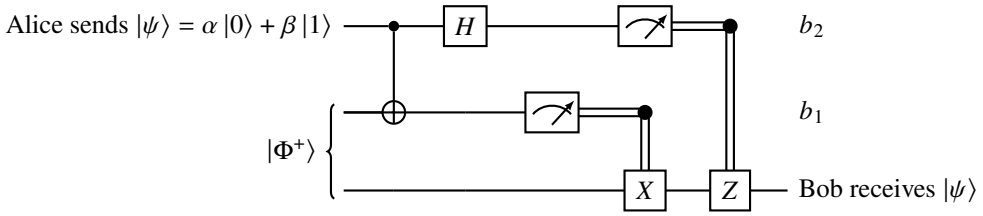
$$\text{CNOT} \sum_{i,j} \alpha_i |i\rangle \otimes |j, j\rangle = \sum_{i,j} \alpha_i |i, i \oplus j, j\rangle$$

But then after measurement of ℓ in qubit 1, we have $i \oplus j = \ell \implies j = i \oplus \ell$ is the only term that survives:

$$\sum_i \alpha_i |i, i \oplus \ell\rangle$$

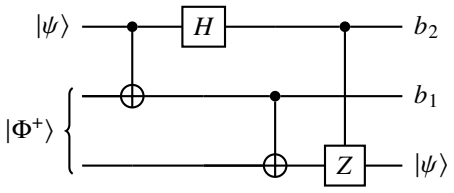
But we want to end up with $\sum_i \alpha_i |i, i\rangle$. So we only have to do a bit flip on the third bit if $\ell = 1$.

Combining all these subcircuits together gives us the following circuit



which allows Bob to successfully receive Alice's qubit. Using this circuit, we can teleport Alice's qubit anywhere!

Note that due to the principle of deferred measurement (covered in the next lecture), this circuit is equivalent to



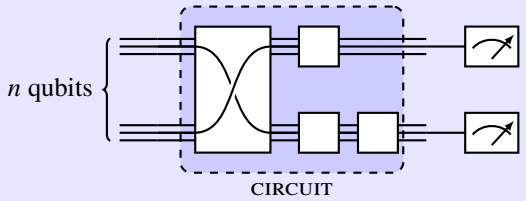
5 Lecture 5

The midterm is coming next Tuesday. It should serve as a checkpoint for what we've learned so far. The material in scope is the first 4 lectures plus the Hadamard transform we will define today.

5.1 Quantum Circuits and Quantum Algorithms

We will now begin quantum computer science in an algorithmic sense.

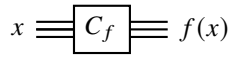
Definition 5.1 (Quantum Circuit)
A quantum circuit on n qubits is a collection of m gates (unitary transforms), with measurement at the end.



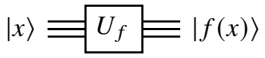
The depth d of a quantum circuit is the maximum amount of gates any one qubit is passed through. At the end of it, we measure in the standard basis in some qubits.

One way to think about a quantum circuit is "rotating" a state a lot of times and then measuring in the standard basis, or we could think about a quantum circuit as measuring in some "rotated" basis. In the quantum world, "programming" is really just designing such a quantum circuit. In a typical case, we want $m = O(\text{poly}(n))$, i.e. on the order of some polynomial of the number of qubits.

Suppose we had a (classical) circuit that could compute a Boolean function $f : \{0,1\}^n \to \{0,1\}^k$:



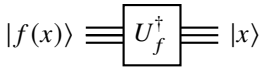
How can we create such a circuit in the quantum world? Well first, we'd like:



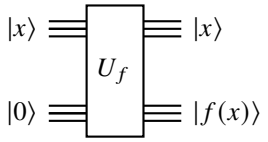
So by linearity,

$$U_f\left(\sum_x \alpha_x |x\rangle\right) \to \sum_x \alpha_x |f(x)\rangle$$

Remember that in our analysis here, since f is a Boolean function, we assume that x is a classical state, i.e. $|x_1, x_2, \ldots, x_n\rangle$ with $x_i \in \{0,1\}$. Furthermore, such a circuit is always invertible (since it's always unitary!):

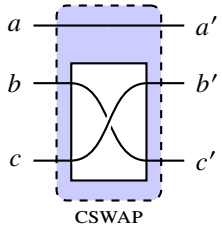


But this is clearly a problem! Suppose $f(x)$ was not injective, i.e. there existed two inputs $x \neq x'$ such that $f(x) = f(x')$. Clearly you cannot "go backwards." This makes us think that maybe all Boolean functions are not representable as quantum circuits. Thus we can only do it for f that are bijections! If we put the input included with the output, then we can go ahead and do this. Note that this does not break no-cloning because we are only cloning a pure binary state $x \in \{0,1\}^n$.



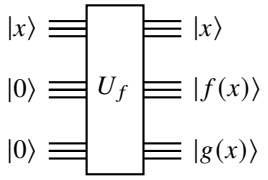
But note that all gates we use must also be reversible in the same. Recall that any classical circuit can be produced from AND and NOT gates. The NOT gate is invertible, and has quantum analogue X. But the AND gate is obviously not reversible.

However, it's very easy to make a reversible AND gate, we just use the same trick: just keep the inputs at the output. However, in practice, the more elegant way to do this is the CSWAP gate.



If $a = 0$, then $b' = b$ and $c' = c$, but if $a = 1$, $b' = c$ and $c' = b$. We claim this implements and AND gate if we set $c = 0$. Then, $c' = 0$ if $a = 0$ and $c' = b$ if $a = 1$, i.e. $c' = a \wedge b$. Furthermore, we can even implement a fanout, to "clone" an input. One can see through a calculation that if $b = 1, c = 0$, then $a' = a, b' = \bar{a}, c = a$.

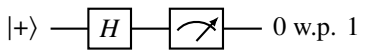
Note that we do not want a lot of these bits to actual AND them, they are useless to us. So we can add some "garbage" bits to the output:



But this is a disaster! Why? We shall explain soon. First, let's discuss how to fix it. To fix this, we can just copy $|f(x)\rangle$, and then run $U_f^\dagger$ on everything else (including the original $|f(x)\rangle$). This eliminates the garbage $g(x)$ (the output of the $U_f^\dagger$ is the just input padded with 0's), while preserving $|f(x)\rangle$.

5.2 Principle of Deferred Measurement

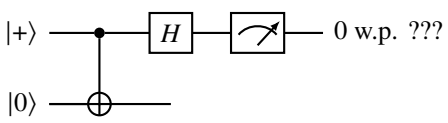
Let's go back to a basic quantum circuit.



This is because:

$$\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \overset{H}{\mapsto} \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right) + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle\right) = |0\rangle$$

Now, let's go to a more complicated circuit:



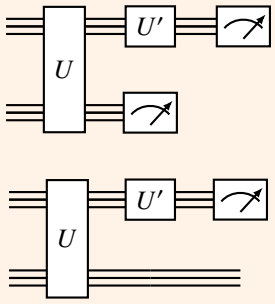
The output of the CNOT is fed into the H:

$$\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle \overset{H}{\mapsto} \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|10\rangle\right) + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|01\rangle - \frac{1}{\sqrt{2}}|11\rangle\right)$$
$$= \frac{1}{2}|00\rangle + \frac{1}{2}|10\rangle + \frac{1}{2}|01\rangle - \frac{1}{2}|11\rangle$$

Now we don't have the interference that happened before where the $|1\rangle$ cancelled out. So now, upon measurement we see 0 and 1 with equal probability.

Let's generalize this.

Theorem 5.1 (Principle of Deferred Measurement)
After qubits stop interacting, the state is the same as if you measured them instead. For instance, the following two circuits will conduct equivalent measurements on the beginning bits:



So now, going back to the garbage bits, we now see why we cannot just "throw them away" or measure them. This will mean you can no longer use those bits for computation! As quantum information scientists, we have to keep our workspace "clean," so to speak.

5.3 Hadamard Transform

Let's send a n-bit classical bit state $|u\rangle$ into a bunch of Hadamards tensored together. Call $H^{\otimes n} = \underbrace{H \otimes H \otimes \cdots \otimes H}_{n \text{ times}}$.

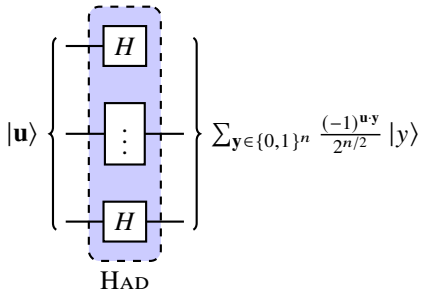
At the level of a singular qubit, we have:

$$H|u_1\rangle = \sum_{y \in \{0,1\}} \frac{(-1)^{u_1 \cdot y}}{\sqrt{2}}|y\rangle$$

which can clearly be seen by the definition (we only have a negative sign if both u and y are 1). Similarly, by nearly "squaring" the above expression, we have:

$$(H \otimes H)(|u_1, u_2\rangle) = \sum_{\mathbf{y} \in \{0,1\}^2} \frac{(-1)^{u_1 y_1}(-1)^{u_2 y_2}}{\left(\sqrt{2}\right)^2}|y\rangle$$

Thus, we can write:



HAD

where $\mathbf{u} \cdot \mathbf{y} = \sum_{i=1}^{n} u_i y_i \pmod 2$ is the dot product of the two bitstrings.

Lastly, one can compute the tensor product using the formula from last lecture:

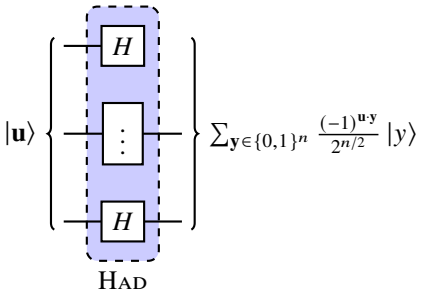
$$H \otimes H = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

One can continue this with higher order products in a similar fashion.

6 Lecture 6

6.1 Hadamard Transform (Continued)

Recall the Hadamard Transform from last lecture, made up of a bunch of Hadamard gates tensored together.



Recall that such a $|u\rangle \in \{0,1\}^n$ is a computational basis vector. So in the standard basis, it has a 1 in some position, and a 0 everywhere else. We will call this position the uth position. Thus, this picks out the uth column of $H^{\otimes n}$. This means the uth column is just $\frac{(-1)^{\mathbf{u}\cdot\mathbf{y}}}{2^{n/2}}$ for all possible y. Therefore, we have

$$[H^{\otimes n}]_{y,u} = \frac{(-1)^{\mathbf{u}\cdot\mathbf{y}}}{2^{n/2}}.$$

Note that these are just bitstrings, but when we use them as indices, we convert them to their decimal counterparts (and use 0-indexing).

One can view this as a Discrete Fourier Transform over $\mathbb{Z}_2^n$, which means it will be very nice for use later. Let's revisit the special case of 2x2.

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

Let's think about how to act on one of our favorite states, $\Phi^+ = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$. The $|00\rangle$ maps to the first column, $\frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$ and the $|11\rangle$ maps to the last column, $\frac{1}{2}(|00\rangle - |01\rangle - |10\rangle + |11\rangle)$. By linearity, this means:

$$H^{\otimes 2}\Phi^+ = \frac{1}{\sqrt{2}}\frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) + \frac{1}{\sqrt{2}}\frac{1}{2}(|00\rangle - |01\rangle - |10\rangle + |11\rangle) = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$$

That means this Bell state through the Hadamard transform just returns the same thing.

With some similar algebra, it turns out another Bell state can be mapped as:

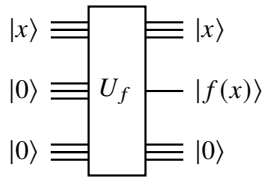
$$H^{\otimes 2}\left(\frac{1}{\sqrt{2}}|01\rangle + \frac{1}{\sqrt{2}}|10\rangle\right) = \frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle$$

We can use the Hadamard to do something interesting: $\sum_x \alpha_x |x\rangle$ — *Had* —

Suppose the state before measurement is $\sum_y \beta_y |y\rangle$, then measuring in the standard basis gives y with probability $|\beta_y|^2$. The question is, what are some interesting states we can put into the Hadamard?

6.2 Building Blocks for Quantum Algorithms

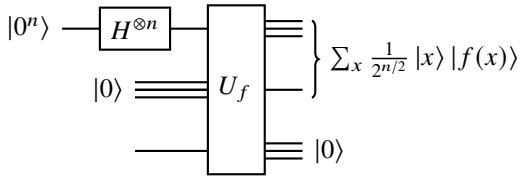
Now how do we program a quantum computer? Let's again try to emulate a classical computer. We saw last lecture, we can emulate a classical circuit C_f using U_f:



Suppose given a function $f : \{0,1\}^n \to \{0,1\}$, we want to make a circuit that produces:

$$\sum_x \frac{(-1)^{f(x)}}{2^{n/2}} |x\rangle$$

Our first attempt is:



But applying a phase flip Z to the middle $|f(x)\rangle$ bit, well for every term, it picks up a minus sign if $f(x) = 1$, e.g., it becomes:

$$\sum_x \frac{(-1)^{f(x)}}{2^{n/2}} |x\rangle |f(x)\rangle$$

Now, to erase $|f(x)\rangle$, we can do the usual trick of running $U_f^\dagger$ on everything to undo the transform. This will yield our intended answer (since $|x\rangle |f(x)\rangle \mapsto |x\rangle$).

Before building up quantum algorithms, we thought about the Extended Church-Turing Thesis, which roughly stated that any "reasonable" model of computation is polynomial-time simulable on a (probabilistic) Turing Machine. Clearly, simulating n qubits with a classical computer requires tracking 2^n amplitudes, which is an exponential-time process. However, quantum computers violate this thesis and early quantum algorithms showcased this fact. Recently, Google performed a "Quantum supremacy" experiment, where they argued there was a **experimental** speedup from a classical computer running the same task.

6.3 Bernstein-Vazirani Algorithm

Suppose there is a secret function $f(x) = u \cdot x$ where we don't know u, but we have oracle access to it. The algorithm to solve this can be stated simply.

1. Create a phase state $\sum_x \frac{(-1)^{f(x)}}{2^{n/2}} |x\rangle = \sum_x \frac{(-1)^{u \cdot x}}{2^{n/2}} |x\rangle$
2. Feed it through the Hadamard Transform (it is its own inverse). This will yield u.

But how would you figure this out classically? To uniquely find f, you'll need to put in a full basis of x elements. So you would need n queries. But in the quantum setting, you only needed two query accesses (assuming you can invert f).

You can generalize this algorithm with one called Recursive Fourier Sampling. This is a similar algorithm that works on multi-dimensional vectors, allowing you to compute gradients (see that here we computed the gradient of f).

7 Lecture 7

7.1 Simon's Algorithm

We discuss Simon's algorithm: another quantum protocol that solves a toy problem. The setup is as follows. Suppose there is a black-box $f : \{0,1\}^n \to S \subseteq \{0,1\}^n$ that is two to one in a specific way: for a fixed string s, $f(x) = f(x \oplus s)$, where the $\oplus$ is vector addition mod 2. Our challenge is to find s.

Classically, there are 2^n possibilities for s, so if we query $x \neq x'$ and get $f(x) \neq x'$ we can only cross out one possibility for s ($x \oplus x'$). So in the worst case, it takes 2^n tries. On average, due to the Birthday paradox, the runtime is actually the square root of this, $2^{n/2}$.

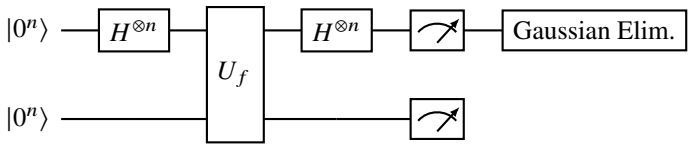
Here is a quantum algorithm that solves the same problem (with some randomness tossed in as well). Let $a \in_R A$ denote that a is uniformly chosen from A. Here is a brief roadmap of the steps we will take.

1. Set up superposition:
$$r \in_R \{0,1\}^n, |\psi\rangle = \frac{1}{\sqrt{2}} |r\rangle + \frac{1}{\sqrt{2}} |r \oplus s\rangle$$
2. Perform Fourier sampling on $|\psi\rangle$, i.e. apply $H^{\otimes n}$, and measure. We claim that this yields uniformly random $a \in \{0,1\}^n$ such that $a \cdot s = 0$.
3. Wait until we get $n-1$ equations. Solve the linear equations for s. This takes time polynomial in n, classically.
4. Check if the solution you got to the system is correct by checking if $f(0) = f(s)$. Repeat the algorithm if not.

In the last step, the reason we use $n-1$ equations is that we will always have at least two solutions, 0 and s, so no n equations can all be independent.

Let's figure out how to set up a suitable superposition for step 1. Let's first apply a Hadamard to $|0\rangle$ to make $\frac{1}{\sqrt{2^n}} \sum_x |x\rangle$. Then, if we apply the function, we recall this just tensor products with the function output. Upon measurement on those function qubits as $f(r)$:
$$\frac{1}{\sqrt{2^n}} \sum_x |x\rangle |f(x)\rangle \mapsto \frac{1}{\sqrt{2}} |r\rangle + \frac{1}{\sqrt{2}} |r \oplus s\rangle$$

This makes the circuit altogether:



Now, let's make sure the claim in step 2 is correct. The Hadamard transforms:
$$\begin{aligned} H^{\otimes n}\left(\frac{1}{\sqrt{2}} |r\rangle + \frac{1}{\sqrt{2}} |r \oplus s\rangle\right) &= \frac{1}{\sqrt{2^{n+1}}} \sum_a \left((-1)^{a \cdot r} + (-1)^{a \cdot (r \oplus s)}\right) |a\rangle \\ &= \frac{1}{\sqrt{2^{n+1}}} \sum_a (-1)^{a \cdot r} (1 + (-1)^{a \cdot s}) |a\rangle \end{aligned}$$

However, note that:
$$(1 + (-1)^{a \cdot s}) = \begin{cases} 0 & \text{if } a \cdot s = 1 \\ 2 & \text{if } a \cdot s = 0 \end{cases}$$

So we finally get the state as:
$$\frac{1}{\sqrt{2^{n-1}}} \sum_{a : a \cdot s = 0} (-1)^{a \cdot r} |a\rangle$$

Measuring ignores the phase, so we thus have a uniform distribution over all the 2^{n-1} vectors orthogonal to s.

Finally, note that the measurement made the math a lot easier, but by the principle of deferred measurement the measurement on the second n qubits is not strictly necessary.

7.2 Quantum Fourier Transform

Recall the roots of unity over $\mathbb{C}$.

Note 7.1 (Roots of unity)
An Mth root of unity is a complex number z such that $z^M = 1$. The primitive Mth root of unity ω is

$$\omega = e^{2\pi i/M} = \cos\frac{2\pi}{M} + i\sin\frac{2\pi}{M}$$

Furthermore, any Mth root of unity can be written as ω^k for $1 \le k \le M$.

The Fourier transform is just the act of applying a polynomial on roots of unity. Suppose you have a polynomial $\alpha(x) = \sum_{j=0}^{M-1} \alpha_j x^j$. Then we can write $\beta_k = \alpha(\omega_k)$, which can be expanded and written as multiplication by a special Vandemonde matrix, called the Discrete Fourier Transform (DFT).

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{M-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega & \dots & \omega^{M-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \omega^{M-1} & \dots & \omega^{(M-1)(M-1)} \end{pmatrix}$$

We added a normalizing factor because we want to use quantum bases, so we'll work with this slightly different definition:

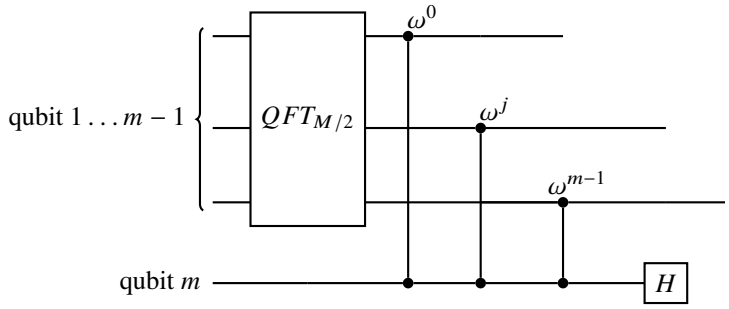
$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{M-1} \end{pmatrix} = \frac{1}{\sqrt{M}} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega & \dots & \omega^{M-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \omega^{M-1} & \dots & \omega^{(M-1)(M-1)} \end{pmatrix}$$

The naive algorithm to do this matrix-vector multiplication is $O(M^2)$, but there is a divide-and-conquer algorithm (the FFT) that can do this much faster, in $O(M \log M)$ time. The Quantum Fourier Transform can do this in $\tilde{O}(\log M)$, where $\tilde{O}$ hides poly-log factors. We will show a simple way to $O(\log^2 M)$. However, there is a big caveat, we know the answer takes M time to even write down. The reason QFT can go faster is that since the QFT gives you a quantum state, you only get a single index j upon any actual measurement; the FFT gives you the entire answer!

Now we discuss the implementation of the QFT, which will be surprisingly similar to the FFT. Without loss of generality, assume M is a power of two and $m = \log_2 M$. By the matrix multiplication above, it's clear that:

$$|k\rangle \mapsto_{QFT} \frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} \omega^{jk} |j\rangle =: |\chi_k\rangle$$

Suppose inductively that we knew how to apply the $QFT_{M/2}$. Then to apply QFT_M we can implement it as:



Note that Unrolling the recursion, and calling QFT_0 the identity, it's clear that the runtime is:

$$1 + 2 + \cdots + (m-1) + m = O(m^2)$$

Now to prove correctness, consider the FT matrix. We reorder the columns so the first $M/2$ columns are the even-indexed columns (0-indexing) and the rest are odd-indexed columns. Then

$$FT_M = \begin{pmatrix} FT_{M/2} & \omega^j FT_{M/2} \\ FT_{M/2} & -\omega^j FT_{M/2} \end{pmatrix}$$

where $\alpha^j A$ means to multiply the jth row of A by α^j. To show this, let's consider the four quadrants. Call j the row, ℓ the column, and k some integer satisfying $0 \le k < M/2$. We will only do two cases for simplicity, the rest are similar. If you're in the top left, e.g. if $0 \le j < M/2$ and $\ell = 2k$, then the entry is $\omega_M^{2jk} = \omega_{M/2}^{jk}$, i.e. the correct entry in the top left. If you're in the bottom right, e.g. if $M/2 \le j < M$ and $\ell = 2k+1$. Let $j' = j - M/2$, then the entry is

$$\omega^{(2k+1)(j'+M/2)} = \omega^{2kj'+kM+j'+M/2} = \omega^{kM}\omega_{M/2}^{kj'}\omega^{M/2+j'} = -\omega^{j'}\omega_{M/2}^{kj'}$$

This proves the claim.

8 Lecture 8

8.1 Factoring

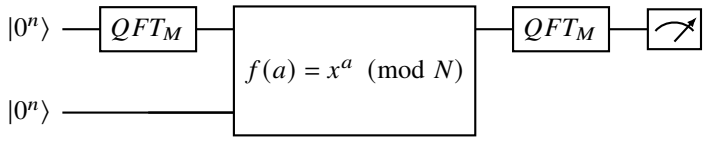
We discuss the classic factoring problem. Given a number N, we wish to find its prime factorization

$$N = p_1^{e_1} p_2^{e_2} \dots p_k^{e_k}$$

The simplest algorithm is to try dividing N by every number up to its square root (its biggest factor). This will take $O(\sqrt{N})$ time. But if you call n the number of bits to write N, e.g. about $\log N$, then $\sqrt{N} = 2^{n/2}$. So this algorithm is still exponential in n. If we are in a cryptographic setting, where n is a 1024-bit or 2048-bit number, $\sqrt{N}$ is gigantic. The best known classial algorithm is the Quadratic Number Sieve, which runs in $\exp(O(\sqrt[3]{n}))$. We would prefer a $\text{poly}(n)$ algorithm (or prefer to disprove its existence to protect cryptography). It turns out we can do better with a quantum computer.

If we could just factor a composite N into two non-unit factors, $N = N_1 \cdot N_2$, then we could easily factor it by repeated calling this algorithm (the divide and conquer perhaps gaining an extra polylogarithmic factor). Thus, the hardest case is really when $N = P \cdot Q$ for two primes of roughly equal size.

Here is a circuit we claim splits N into such factors. Fix $M > N$ and pick x at random satisfying $0 < x < N$.



Then from the measurement, we do a little bit of classical postprocessing.

Let's work with an example to motivate how this circuit works. We have $N = 15, M = 16$, then for chosen $x = 2$:

$$x^0 \equiv x^4 \equiv x^8 \equiv x^{12} \equiv 1 \pmod{15}$$

Furthermore

$$x^1 \equiv x^5 \equiv x^9 \equiv x^{13} \equiv 2 \pmod{15}$$

If $\gcd(x, N) = 1$, then there exists a smallest $r > 0$ such that $x^r \equiv 1 \pmod N$. We call r the order of x. If we picked an x that isn't relatively prime, we could use the Euclidean algorithm to find a common factor of x and N and thus a factor of N.

Suppose we knew the order of x mod 15 was 4 beforehand. Since r is even, let's trying halving it (if it wasn't, we'd need a different x). So, define $y \equiv x^{r/2} \pmod{15}$. This means that it's a second root of unity, e.g. $y^2 \equiv 1 \pmod{15}$. In the special case when N is a product of two primes, there are four roots of unity. In this case we have $y = 2^2 = 4$, which is not trivially plus or minus 1 (if it was, we'd need a different x). Now, this means

$$y^2 - 1 \equiv 0 \pmod{15} \implies 15 \mid (y+1)(y-1)$$

But we know that 15 cannot divide either thing individually, since $y \not\equiv \pm 1 \pmod{15}$. Thus, now we can just compute $\gcd(15, y+1)$ and $\gcd(15, y-1)$; this gives us our two factors. In our example this is $\gcd(15, 5) = 5$ and $\gcd(15, 3) = 3$. We claim that our circuit above will find the order of x.

8.2 Period Finding

How do we extract the order of x (e.g. the period)? Now let's think about what we could get out of the circuit right before the second QFT based on different measurements of $f(a)$ (which we can consider by the principle of deferred measurement).

$f(a)$	State of first register
0	Not possible
1	$\|0\rangle + \|4\rangle + \|8\rangle + \|12\rangle$
2	$\|1\rangle + \|5\rangle + \|9\rangle + \|13\rangle$
3	Not possible
4	$\|2\rangle + \|6\rangle + \|10\rangle + \|14\rangle$

Suppose r divides M for now. Then, the superposition we get (up to normalization) is

$$|\psi\rangle = \sum_{j=0}^{\frac{M}{r}-1} \sqrt{\frac{r}{M}} |jr\rangle \mapsto_{QFT_M} \sum_{\ell=0}^{M-1} \sum_{j=0}^{\frac{M}{r}-1} \frac{\sqrt{r}}{M} \omega^{jr\ell} |\ell\rangle$$

For $\ell = \frac{kM}{r}$, we get a constructive interference of M/r terms, with each amplitude $\frac{\sqrt{r}}{M}$, which gives amplitude $\frac{1}{\sqrt{r}}$. But now notice that there are r such terms, which means the superposition is normalized by these; the other amplitudes must be zero. This means that

$$|\psi\rangle \mapsto_{QFT_M} \frac{1}{\sqrt{r}} \sum_{\ell=0}^{r-1} \left|\ell \frac{M}{r}\right\rangle$$

Now, if we continually measure, we can get the period with high probability, just by finding $\frac{M}{r}$ with the gcd of the measurements (which will be $\frac{k_1 M}{r}$, $\frac{k_2 M}{r}$, etc). One can view this as time-frequency uncertainty principle–increasing the period r in the original domain decreases the period M/r in the new domain.

In the general case, r does not divide M. But with constant probability, we see ℓ such that $|\ell r \mod M| \leq r/2$. In other words,

$$\left|\frac{\ell}{M} - \frac{k}{r}\right| \leq \frac{1}{2M}$$

If you choose $M \approx N^2$, then one can figure out r by tightening this bound. The continued fractions algorithm does the job.