Lecture memo for the class of 《Quantum Computing》 by Prof. Man-Hong Yung



# 1. Lecture 01

**Classical bit:**

- logically, '0' or '1'
- physically, voltage (0, 5V) or current
- why use discrete bits instead of continuous variables?
  - error correction
- Boolean logic (good, beautiful mathematical tool)
  - boolean function is "molecular/atomic", it represent any function
- 1 "physical state/configuration" = 1 value of bit
- E.g. N transistors $\Longrightarrow$ N logical states at a time, e.g. 010101010...

**Quantum bit (qubit):**

- We will use dirac notation: $|0\rangle$ and $|1\rangle$ (qubit states)

- No one really undersands quantum mechanics
- A kid can learn how to ride a bicycle, but he/she may have no idea how it really works
- Mysterious concept called "superposition" (entanglement)

If you want to enjoy this course, you have to be in an excited state, not ground state

## 2. Lecture 02

- What does the braket means? In short, it is just a symbol, $|label\rangle$, e.g. $|cat\ alive\rangle$, $|cat\ dead\rangle$, a certain configuration of all of the degrees of freedom.
- A qubit ≡ two-level system (TLS), does not necessarily an eigenstate
- A qubit can be in a superposition state, in general,

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{1}$$

where $|\alpha|^2 + |\beta|^2 = 1 = \Pr(0) + \Pr(1)$ (normalization condition).

---

My recent thinking about quantum superposition:

- Analogy: classical mxiture versus quantum superposition.
- [Classical picture] Suppose we have a teapot, A would put either tea or coffee into the teapot, then A gives the teapot to B. B will then pour the liquid out and decide if it is really tea of coffee.
- [Quantum picture] Suppose now A put both tea and coffee into the teapot. Obviously, inside the teapot, the liquid cannot be considered as tea nor coffee. The point is that when the teapot is given to B, who would pour the liquid out. B will find either tea or coffee probabiltiically.
- The point is that in quantum theory, a superposition describes a physical state when certain property cannot be determined before measurement.

Another story:

- There is something called Durian test: you can either like to eat durian very much or you hate it very much, nothing in between.
- Therefore, if someone have never eaten Durian before, then well you may suppose this property does not exist until you try it (like measurement).
- Suppose we have a twins, who never tried Durian before. Quantum entanglement is like when the twins are separated far apart, if one of them tried Durian, then immediately the other twin brother/sister will be determined in terms of liking or hating Durian.
- This is exactly what EPR paper is talking about (at least you read the abstract).

---

- Probability of measuring "0" is given by $\Pr(0) = |\alpha|^2$, and $\Pr(1) = |\beta|^2$.

- Examples of qubit
  - energy states of a microscopic particle
  - path degrees of freedom of a photon
  - polarization degrees of freedom of a photon
  - electron spin states
- Quantum mechanics is all about linear algebra (by magic), which means matrices and vectors.
- Mathemtically, one just take (basis vectors)

$$|0\rangle \equiv (1,0)^T, \quad |1\rangle \equiv (0,1)^T \tag{2}$$

- In other words, general qubit state

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \equiv (\alpha,\beta)^T \tag{3}$$

- Example, mathematics of superposition

$$|\psi_1\rangle + |\psi_2\rangle = (\alpha_1|0\rangle + \beta_1|1\rangle) + (\alpha_2|0\rangle + \beta_2|1\rangle) = (\alpha_1+\alpha_2)|0\rangle + (\beta_1+\beta_2)|1\rangle \tag{4}$$

- Any state is physically indistinguishable with itself, if we add a global to it, i.e.,

$$|\psi\rangle = e^{i\phi}|\psi\rangle \tag{5}$$

**More about quantum measurement**

- The example about single photon and beamsplitter is a good example of quantum measurement.
- Sometimes, we talk about measurement along different bases.
- For example, if we have only $|0\rangle$, instead of superposition, and if we measure in the computational basis, then we always get 0,0,0,0,0,0,0,0....
- Even if we are in a superposition state,

$$|\pm\rangle \equiv (|0\rangle \pm |1\rangle)/\sqrt{2} \tag{6}$$

- If we measure along the x-basis, $\{|\pm\rangle\}$, then $|+\rangle$ state always gives the same results, e.g. 0,0,0,0,0,....
- If we measure $|+\rangle$ state along the **computational basis**, then we get random bits, e.g. 010100110101010...
- There is a connection between the two bases,

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

which is still a normalized state, please check during the break.

Solution (assume real numbers):

$$\frac{1}{2}(\alpha+\beta)^2+\frac{1}{2}(\alpha-\beta)^2=\frac{1}{2}\left(\alpha^2+\beta^2+2\alpha\beta\right)+\frac{1}{2}\left(\alpha^2+\beta^2-2\alpha\beta\right)=\alpha^2+\beta^2=1 \quad (7)$$

- In general, there are many different ways to characterize a single qubit state, as long as $|\alpha|^2+|\beta|^2=1$ is satisfied.
- Recall,

$$|\alpha|^2+|\beta|^2=\cos^2\theta+\sin^2\theta=1 \quad (8)$$

- We can also write, (easy to miss a factor of 1/2)

$$|\psi\rangle=\alpha|0\rangle+\beta|1\rangle \quad =\cos(\theta/2)|0\rangle+e^{i\phi}\sin(\theta/2)|1\rangle \quad (9)$$

- This one has a geometrical picture: bloch sphere

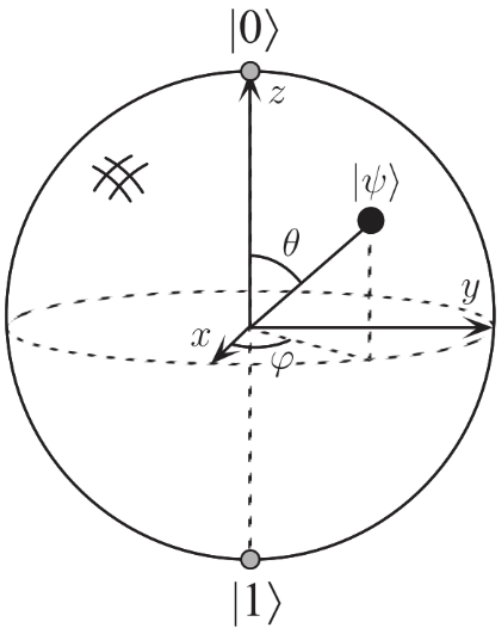



**Figure 1:** Bloch Sphere

Multiple qubits: more is different

- Bad news: no simple bloch sphere (geometrical) picture
- Two qubit computational basis:

$$|00\rangle,|01\rangle,|10\rangle,|11\rangle \quad (10)$$

where $|00\rangle\equiv|0\rangle|0\rangle\equiv|0\rangle\otimes|0\rangle$.

- Tensor product means the following: (no proof)

$$|0\rangle \otimes |0\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv \begin{pmatrix} 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{11}$$

$$|0\rangle \otimes |1\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \begin{pmatrix} 1\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{12}$$

$$|1\rangle \otimes |0\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv \begin{pmatrix} 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{13}$$

$$|1\rangle \otimes |1\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \begin{pmatrix} 0\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 1\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{14}$$

- The most general two-qubit (pure) state,

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle \tag{15}$$

where $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$.

- It means that the probability of e.g. getting 00 is given by $\Pr(00) = |\alpha|^2$.

**Partial quantum measurement**

- If we measure qubit 1 (left), we can get either 0 or 1. If we get 1, then the second qubit becomes,

$$|\psi_2\rangle = \frac{1}{\sqrt{|\gamma|^2 + |\delta|^2}}(\gamma|0\rangle + \delta|1\rangle) \tag{16}$$

## 3. Lecture 03

- Bell states (totally four Bell states) : maximally entangled

$$|\psi_{Bell}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{17}$$

- The other Bell's states are

$$\frac{1}{\sqrt{2}}(|00\rangle - |11\rangle),\ \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle),\ \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \tag{18}$$

- Entangled (pure) states are those states that cannot be decomposed into a product form.
- Multiple-qubit state (n-qubit states, $2^n$ basis vectors)

$$|\psi_n\rangle = \sum_{x\in\{0,1\}^n} c_{x_1x_2\cdots x_n}|x_1x_2\ \cdots\ x_n\rangle = \sum_{x\in\{0,1\}^n} c_x|x\rangle\ , \tag{19}$$

where $\sum_x |c_x|^2 = 1$ .

- If $n = 50$, it is almost same as the maximal memory capacity of supercomputer nowadays.
- Recall that standard quantum computation: 1. state preparation (all qubits to be $|0\rangle$), 2. apply quantum gates, 3. measurement (partial)
- The actual computation part is point 2.
- What is quantum gate? Classical gate, we know like AND, OR, NAND, etc.... logical operations. (elementary Boolean function)
- In quantum mechanics, we know that a quantum state can only be changed by **unitary matrix (transformation)**. (Why?)

**Unitary transformation**

- Recall that quantum state is a (column) vector $|\psi\rangle$ (ket). Also, the amplitude sqaures are probabilities.
- It is natural to introduce a corresponding (row) vector $\langle\psi|$ (bra),

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad \rightarrow \quad \langle\psi| \equiv \alpha^*\langle 0| + \beta^*\langle 1|$$

- Conservation of probabilities means; inner product is called "bracket"

$$\langle\psi|\psi\rangle = |\alpha|^2 + |\beta|^2 = 1 \tag{20}$$

- This is also called normalization condition.
- Because, if a state is a vector, than the change of state is necessarily a matrix (based on linear algebra).

$$|\psi\rangle \rightarrow U|\psi\rangle \tag{21}$$

- In the dual space, we write it as (**the proof is a good exercise**)

$$\langle\psi| \ \rightarrow \ \langle\psi|U^\dagger \tag{22}$$

where $A^\dagger \equiv \left(A^T\right)^*$.

- To preserve probability, we need

$$\langle\psi|\psi\rangle = \langle\psi|U^\dagger U|\psi\rangle = 1 \tag{23}$$

for all possbile $|\psi\rangle$.

- It means that

$$U^\dagger U = I \implies U^\dagger = U^{-1} \tag{24}$$

which is the defining (iff) condition for a unitary matrix.

**Concluding remarks**

- Any change of a quantum state, has to be achieved by a matrix
- Because of the conservation of probabilities, the matrix has to be unitary
- In other words, any quantum dynamcs, or quantum computation, has to be a kind (subset) of unitary transformation.

**Quantum computation as unitary transformation?**

- Then, we should imagine quantum computation is represented by a class of unitary transformation, e.g.,

$$U|\psi_{in}\rangle = |\psi_{out}\rangle \tag{25}$$

- Traditionally, physicists solve problems in this way: given initial state, and Hamiltonian, find the final state.
- For quantum computation, usually, we would be given the relationship between the input and output state, and the goal would be to find the quantum circuit $U$ (quantum algorithm).
- Compare classical computation:

$$f(x_{in}) = y_{out} \tag{26}$$

- We will later show that quantum computation is more general, which means that we can also do the same classical computation with quantum computer.
- The point is that once the logical relation between the input and output is defined. We will need to know how to physically realize the corresponding unitary transformation (called quantum circuit).
- Recall that classical circuits consist of elementary logical operations. (AND, OR, NOT, NAND, etc.)
- Quantum circuits should also consist of elementary logical operations. What are they? (we will find out later)

**Terminology**

- Quantum circuit: one way to realize a quantum algorithms (larger unitary matrix)
- Quantum gates: elements of a quantum circuit (smaller unitary matrix)
- Operator: ~matrix (may or may not be unitary)

**Single-qubit gates ≡ single-qubit unitary matrices (transformation)**

- Let us start with single-bit gate in the classical case.

$$Logical\ NOT\ gate:\ 1 \to 0\ \ and\ \ 0 \to 1$$

- We want to do the following

$$Logical\ NOT\ gate:\ |1\rangle \to |0\rangle\ \ and\ \ |0\rangle \to |1\rangle \tag{27}$$

- This is sometimes called **bit flip**.
- Yes, we know one of the Pauli matrices can do it. It is called Pauli-X,

$$\sigma_x \equiv U_X \equiv X = |1\rangle\langle 0| + |0\rangle\langle 1| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{28}$$

- First, make sure it is a unitary matrix:

$$X^\dagger = X \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$X^\dagger X = I \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{29}$$

- Sometimes, it is easier to use the Dirac form:

$$X|1\rangle = (|1\rangle\langle 0| + |0\rangle\langle 1|)|1\rangle = |1\rangle\langle 0|1\rangle + |0\rangle\langle 1|1\rangle = |0\rangle \tag{30}$$

$$X|0\rangle = (|1\rangle\langle 0| + |0\rangle\langle 1|)|0\rangle = |1\rangle\langle 0|0\rangle + |0\rangle\langle 1|0\rangle = |1\rangle \tag{31}$$

- What happens when we apply Pauli-X to a general qubit state?

$$X|\psi\rangle = X(\alpha|0\rangle + \beta|1\rangle) \ = \alpha X|0\rangle + \beta X|1\rangle = \alpha|1\rangle + \beta|0\rangle \tag{32}$$

$$X\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix} \tag{33}$$

## 4. Lecture 04

- In general, a quantum gate $U$, for a single qubit, would do the following:

$$U|0\rangle = \alpha|0\rangle + \beta|1\rangle \equiv |\phi\rangle \tag{34}$$

and

$$U|1\rangle = \alpha'|0\rangle + \beta'|1\rangle \equiv |\phi_\perp\rangle \tag{35}$$

- Note that inner product,

$$0 = \langle 0|U^\dagger U|1\rangle = \alpha^*\alpha' + \beta^*\beta' \tag{36}$$

- The point is that the coefficients are related, but at the moment, I don't want to completely figure it out (not that difficult, but it takes a bit time).
- We can **always** write

$$U = (\alpha|0\rangle + \beta|1\rangle)\langle 0| + (\alpha'|0\rangle + \beta'|1\rangle)\langle 1| \tag{37}$$

$$= U|0\rangle\langle 0| + U|1\rangle\langle 1|$$
$$= U$$

because $0\rangle\langle 0| + |1\rangle\langle 1| = I$ (completeness relation)

- If you know the action of the gate to all elements of a basis vectors, then you can always construct the matrix form of the gate.
- I learned quantum mechanics for 10+ years. I only knew how to mechanically (error free?) convert between operators in Dirac notation and matrix form couple years ago.
- In the following way (make sure you understand below, may be exam question, who knows?),

$$U = (\alpha|0\rangle + \beta|1\rangle)\langle 0| + (\alpha'|0\rangle + \beta'|1\rangle)\langle 1| \quad (38)$$
$$= \begin{pmatrix} \alpha \\ \beta \end{pmatrix}\langle 0| + \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix}\langle 1| \quad (39)$$
$$= \begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix} \quad (40)$$

- Just in case, you may lost, we are going to see different animals. For each animal, I will describe some features associated with it, without much about the use of them.
- Think about it like introducing to you new friends.
- Let us first consider a rather common single-qubit gate, the Puali-$Z$ gate:

$$Z: \quad |0\rangle \rightarrow |0\rangle \;\; but \;\; |1\rangle \rightarrow -|1\rangle \quad (41)$$

- Generally,

$$Z(\alpha|0\rangle + \beta|1\rangle) = \alpha Z|0\rangle + \beta Z|1\rangle = \alpha|0\rangle - \beta|1\rangle \quad (42)$$

- In the context of error correction, it may be called **phase flip**.
- The explicit form of the $Z$ gate is as follows:

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = |0\rangle\langle 0| - |1\rangle\langle 1| \quad (43)$$

- Of course, we also need to consider Pauli-$Y$ gate (diffiult to remember the minus sign):

$$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ i \end{pmatrix}\langle 0| + \begin{pmatrix} -i \\ 0 \end{pmatrix}\langle 1| = i|1\rangle\langle 0| - i|0\rangle\langle 1| \quad (44)$$

- What I usually remember is the following relation:

$$X\,Y\,Z = iI \quad (45)$$

- With it, then I can define $Y$ by using the identity: $X^2 = Y^2 = Z^2 = I$ (self inverse),

$$Y\,Z = iX$$
$$Y = iXZ \quad (46)$$

- Why are they interesting? Because they are related to basic operation on physical qubits.
- In the laboratory, we can apply external driving field on the qubits, and the field has a "direction". For example, magnetic field along the z-direction (related to things like **spin precession**).
- In the sense that, the elementary operation in the lab, are the following unitary transformations:

$$R_x(\theta) = e^{-i\theta X/2}, \quad R_y(\theta) = e^{-i\theta Y/2}, \quad R_z(\theta) = e^{-i\theta Z/2} \tag{47}$$

- First, we have to explain why a matrix can be an exponent? Because, they can be defined by an infinite series, i.e.,

$$e^A \equiv I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + ... \tag{48}$$

where $A$ is a matrix.

- If $A$ is diagonal, e.g. $A = \begin{pmatrix} a_0 & 0 \\ 0 & a_1 \end{pmatrix}$, then $A^k = \begin{pmatrix} a_0^k & 0 \\ 0 & a_1^k \end{pmatrix}$. Recall that $a_0$ and $a_1$ are called the eigenvalues of $A$.
- Then, we know that

$$e^A \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} a_0 & 0 \\ 0 & a_1 \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} a_0^2 & 0 \\ 0 & a_1^2 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} a_0^3 & 0 \\ 0 & a_1^3 \end{pmatrix} + ... = \begin{pmatrix} e^{a_0} & 0 \\ 0 & e^{a_1} \end{pmatrix} \tag{49}$$

- Furthermore, you can write

$$e^A = e^{a_0}|0\rangle\langle 0| + e^{a_1}|1\rangle\langle 1| \tag{50}$$

- In other words, for diagonal operator $A$, it is quite easy to find the matrix form; you just need to find the exponential form for the eigenvalues.

- In the laboratory, the choice of Pauli matrix, X, Y, Z, depends on the experimental setup.
- The value of $\theta$ depends on the length of time one applies the field, because it is like time evolution.
- To completely understand what I am talking here, you need to study textbooks of quantum mechanics (dynamics of spin-1/2 particle; ask me if you could not figure out).
- Sometimes, we need to know the matrix form for each roation operator.

$$R_z(\theta) = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + ... = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix} \tag{51}$$

where $A = -i\theta Z/2$.

- Recall that the eigenvalues of $Z$, are $\pm 1$.
- For the other matrices, let us leave it as **exercise** (or **homework**),

$$R_x(\theta) = \begin{pmatrix} c_\theta & -is_\theta \\ -is_\theta & c_\theta \end{pmatrix}, \qquad R_y(\theta) = \begin{pmatrix} c_\theta & -s_\theta \\ s_\theta & c_\theta \end{pmatrix} \tag{52}$$

where $c_\theta \equiv \cos\frac{\theta}{2}$, and $s_\theta \equiv \sin\frac{\theta}{2}$ (double check, may be wrong).

## 5. Lecture 05

- Let us consider $R_x(\theta) = e^{-i\theta X/2}$. Let us we recall the definition of the exponential of a matrix, which is a series expansion. Suppose we consider sometime easier,

$$Q_\theta = e^{\theta X} = I + \theta X + \frac{\theta^2}{2!}X^2 + \frac{\theta^3}{3!}X^3 + ... \tag{53}$$

- Now, recall that $X^2 = I$, so we have

$$Q_\theta = e^{\theta X} = I + \theta X + \frac{\theta^2}{2!} + \frac{\theta^3}{3!}X + ... = \left(1 + \frac{\theta^2}{2!} + ...\right)I + \left(1 + \frac{\theta^4}{3!} + ...\right)X \tag{54}$$

- Note that $e^x = \cosh x + \sinh x$, therefore we conclude that $Q_\theta = I\cosh\theta + X\sinh\theta$ .
- If we now replace $\theta$ by $-i\theta/2$, then we get

$$R_x(\theta) = I\cos\frac{\theta}{2} - iX\sin\frac{\theta}{2} = \begin{pmatrix} c_\theta & 0 \\ 0 & c_\theta \end{pmatrix} - i\begin{pmatrix} 0 & s_\theta \\ s_\theta & 0 \end{pmatrix} = \begin{pmatrix} c_\theta & -is_\theta \\ -is_\theta & c_\theta \end{pmatrix}. \tag{55}$$

- Also, as $Y|0\rangle = i|1\rangle$, and

$$R_y(\theta) = I\cos\frac{\theta}{2} - iY\sin\frac{\theta}{2} \tag{56}$$

which means that $R_y(\theta)|0\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle$.

- One should know about the follow identities:

$$XR_z(\theta)X = R_z(-\theta) = R_z(\theta)^{-1} = R_z(\theta)^\dagger \tag{57}$$

- The last relation is reasonable, because $R_z(\theta)$ is unitary. To be sure, ($Z^\dagger = Z$)

$$R_z(\theta)^\dagger = I\cos\frac{\theta}{2} + iZ\sin\frac{\theta}{2} = e^{i\theta Z/2} \tag{58}$$

- The first relation is true, let us check it in two ways:
- (1st way) The simple way, $R_z(\theta)(\alpha|0\rangle + \beta|1\rangle) = \alpha e^{-i\theta/2}|0\rangle + \beta e^{i\theta/2}|1\rangle$

$$\begin{aligned} XR_z(\theta)X(\alpha|0\rangle + \beta|1\rangle) &= XR_z(\theta)(\alpha|1\rangle + \beta|0\rangle) \\ &= X\left(\alpha e^{i\theta/2}|1\rangle + \beta e^{-i\theta/2}|0\rangle\right) \\ &= \alpha e^{i\theta/2}|0\rangle + \beta e^{-i\theta/2}|1\rangle \end{aligned} \tag{59}$$

compare $R_z(-\theta)(\alpha|0\rangle + \beta|1\rangle) = \alpha e^{i\theta/2}|0\rangle + \beta e^{-i\theta/2}|1\rangle$

- (2nd way) The "better" way would be,

$$XR_z(\theta)X = XIX\cos\frac{\theta}{2} - iXZX\sin\frac{\theta}{2} = I\cos\frac{\theta}{2} + iZ\sin\frac{\theta}{2} \tag{60}$$

- Because we know that $XZX = -Z$, or $ZX + XZ = 0$ or $XZ = -XZ$ (True for any combination of Pauli matrices)

| commutators | anticommutators |
|---|---|
| $[\sigma_1, \sigma_2] = 2i\sigma_3$ | $\{\sigma_1, \sigma_1\} = 2I$ |
| $[\sigma_2, \sigma_3] = 2i\sigma_1$ | $\{\sigma_1, \sigma_2\} = 0$ |
| $[\sigma_3, \sigma_1] = 2i\sigma_2$ | $\{\sigma_1, \sigma_3\} = 0$ |
| $[\sigma_1, \sigma_1] = 0$ | $\{\sigma_2, \sigma_2\} = 2I\,.$ |

**Figure 2:** from wikipedia

- Note that $XY - YX = 2iZ$, which means that $XY = iZ$. Two different Pauli matrices multipied together, is proportional to the other Pauli matrix.
- Finally, why X, Y, Z? we can also consider any arbitrary roation. So we need to define,

$$R_n(\theta) \equiv e^{-i(\theta/2)n\cdot\sigma} = c_\theta I - is_\theta(n\cdot\sigma)\,. \tag{61}$$

where $n\cdot\sigma \equiv n_x X + n_y Y + n_z Z$, for $n_x^2 + n_y^2 + n_z^2 = 1$. (the second equality is a good execrise/homework).

- This rotation operator can generate an arbitrary single-qubit rotation $U$ (up to a global phase). In the sense that, given any $U$, one can always choose the angle $\theta$ and the rotation axis $\widehat{n}$ , such that (**think about it**)

$$U = e^{i\alpha}R_n(\theta) \tag{62}$$

---

- Partial proof: first, any 2x2 matrix can be expanded by Paulis matrices (complete basis). Recall that any state can be decomposed by 0,1 basis, i.e., $|\psi\rangle = a|0\rangle + b|1\rangle$. In the beginning, you do not need to know a and b. To determine them, we can apply the inner product, $\langle 0|\psi\rangle = a\langle 0|0\rangle + b\langle 0|1\rangle = a$.
- Similarly, we can also do something like that using Pauli matrices. It means that for any 2x2 matrix $U$, we can always write

$$U = aI + bX + cY + dZ \tag{63}$$

- First of all, X, Y, Z Pauli matrices are traceless. (Trace Tr means taking the sum of the diagonal values; Tr is a linear operation)

$$Tr(U) = aTr(I) + bTr(X) + cTr(Y) + dTr(Z) = 2a \tag{64}$$

which means that $a = Tr(U)/2$.

- Then, for example, let us apply $X$ to $U$, then we have

$$XU = aX + bI + cXY + dXZ \tag{65}$$

- After taking the trace,

$$Tr(XU) = bTr(I) + cTr(XY) + dTr(XZ) = 2b \tag{66}$$

which means that $b = Tr(XU)/2$. Similarly, $c = Tr(YU)/2$, and $d = Tr(ZU)/2$.

- Well if we make $a = |a|e^{i\alpha}$, we are getting close.

---

- Hadamard gate $H$: (I think almost all quantum cirucits has Hadamard gate).

$$H \equiv \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\langle 0| + \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\langle 1| \tag{67}$$

where

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad , H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \tag{68}$$

- Note that $H$ is also self inverse, i.e., $H^2 = I$, and that $H = (X + Z)/\sqrt{2}$.
- There are some identities (sandwiching with Hadamard) which would be relavant later (may be mid-term or final exam questions).

$$HXH = Z \quad HZH = X \quad HYH = -Y \tag{69}$$

- As an example, $HR_x(\theta)H = I\cos\frac{\theta}{2} - iHXH\sin\frac{\theta}{2} = R_z(\theta)$.
- Finally, we also need to know two more single qubit gates:

$$S \equiv \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \quad , \quad T \equiv \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}. \tag{70}$$

where $S = T^2$.

- In reality, in the lab, one may not have the capability of controlling the qubit rotation from an arbitrary direction.
- For example, what if we can only have $R_z$ and $R_y$?

## 6. Lecture 06

**Z-Y-Z decomposition of single-qubit unitary gate**

- Any single-qubit gate $U$ (arbitrary 2x2 matrix), can be decomposed by the following (up to a phase factor):

$$U = e^{i\alpha}R_z(\beta)R_y(\gamma)R_z(\delta) \tag{71}$$

which means that each $U$ can be characterized by 4 numbers $\alpha, \beta, \gamma, \delta$.

---

**Proof:**

- recall that any unitary gate can be decomposed in the following way:

$$U = |\phi\rangle\langle 0| + |\phi_\perp\rangle\langle 1| \tag{72}$$

where we may want to associate $|\phi\rangle$ with some rotation operators, i.e.,

$$|\phi\rangle \equiv U|0\rangle = a|0\rangle + b|1\rangle = e^{i\phi_0}\cos(\gamma/2)|0\rangle + e^{i\phi_1}\sin(\gamma/2)|1\rangle \tag{73}$$

- Recall that $R_y(\gamma)|0\rangle = \cos\frac{\gamma}{2}|0\rangle + \sin\frac{\gamma}{2}|1\rangle$, I can always write, for any quantum state,

$$|\phi\rangle = e^{i\alpha}R_z(\beta)R_y(\gamma)|0\rangle \neq ?\, U|0\rangle \tag{74}$$

- But, we can apply a final $R_z(\delta)$ for any value of $\delta$. So, we can also write

$$|\phi\rangle = e^{i\alpha}R_z(\beta)R_y(\gamma)R_z(\delta)|0\rangle = U|0\rangle \tag{75}$$

- Because $R_z(\beta)R_y(\gamma)|0\rangle = e^{-i\beta/2}\cos(\gamma/2)|0\rangle + e^{i\beta/2}\sin(\gamma/2)|1\rangle$ . In this way, we can just set $\phi_0 = \alpha - \beta/2$, and $\phi_1 = \alpha + \beta/2$, which determines the values of the angles.
- On the other hand, since $0 = e^{i\theta}\langle\phi_\perp|\phi\rangle = \langle\phi_\perp|U|0\rangle$, so it means that (up to a phase factor)

$$U^\dagger|\phi_\perp\rangle = |1\rangle,\ or\ |\phi_\perp\rangle = U|1\rangle \tag{76}$$

$$|\phi_\perp\rangle = e^{i\alpha}R_z(\beta)R_y(\gamma)R_z(\delta)|1\rangle \tag{77}$$

- Putting everything together, $|0\rangle\langle 0| + |1\rangle\langle 1| = I$,

$$U = e^{i\alpha}R_z(\beta)R_y(\gamma)R_z(\delta)(|0\rangle\langle 0| + |1\rangle\langle 1|) = e^{i\alpha}R_z(\beta)R_y(\gamma)R_z(\delta) \tag{78}$$

---

- Note that the Z-Y-Z decomposition can be extended to any pairs of rotatioinal gates, e.g. Z-X-Z or X-Y-X, as along as they are perpendicular to each other.
- However, if the rotation angles are not perpendicular, then it would require more than 3 rotations. The textbook is unfortunately, wrong. (Imagine if m and n are very close to each other, then it is effectively a single rotation operator along n)

*Theorem 4.1*: (**$Z$-$Y$ decomposition for a single qubit**) Suppose $U$ is a unitary operation on a single qubit. Then there exist real numbers $\alpha, \beta, \gamma$ and $\delta$ such that

$$U = e^{i\alpha} R_z(\beta) R_y(\gamma) R_z(\delta). \tag{4.11}$$



*Proof*

Since $U$ is unitary, the rows and columns of $U$ are orthonormal, from which it follows that there exist real numbers $\alpha, \beta, \gamma$,and $\delta$ such that

$$U = \begin{bmatrix} e^{i(\alpha-\beta/2-\delta/2)} \cos\frac{\gamma}{2} & -e^{i(\alpha-\beta/2+\delta/2)} \sin\frac{\gamma}{2} \\ e^{i(\alpha+\beta/2-\delta/2)} \sin\frac{\gamma}{2} & e^{i(\alpha+\beta/2+\delta/2)} \cos\frac{\gamma}{2} \end{bmatrix}. \tag{4.12}$$

Equation (4.11) now follows immediately from the definition of the rotation matrices and matrix multiplication. □

**Figure 3:** Proof given by the textbook, too brief

**Exercise 4.11:** Suppose $\hat{m}$ and $\hat{n}$ are non-parallel real unit vectors in three dimensions. Use Theorem 4.1 to show that an arbitrary single qubit unitary $U$ may be written

$$U = e^{i\alpha} R_{\hat{n}}(\beta) R_{\hat{m}}(\gamma) R_{\hat{n}}(\delta), \tag{4.13}$$

for appropriate choices of $\alpha, \beta, \gamma$ and $\delta$.

**Figure 4:** from N&C textbook (p.176), it is wrong!

- Notes the authors did provide an errata list (but the book is not updated)

pp 176 Exercise 4.11 should be replaced by

> Suppose $\hat{m}$ and $\hat{n}$ are non-parallel real unit vectors in three dimensions. Show that an arbitrary single qubit unitary $U$ may be written as
>
> $$U = e^{i\alpha} R_{\hat{n}}(\beta_1) R_{\hat{m}}(\gamma_1) R_{\hat{n}}(\beta_2) R_{\hat{m}}(\gamma_2) \ldots, \tag{4.13}$$
>
> for appropriate choices of $\alpha$ and $\beta_k$, $\gamma_k$.

- Next, we will also need to know another decomposition of single unitary gate. This is

needed for some practical reason.

## 7. Lecture 07

**ABC decomposition of single-qubit unitary gate**

$$U = e^{i\alpha} AXBXC \tag{79}$$

where $A$, $B$, and $C$ are products of the rotation operators, with $ABC = I$, $X$ is Pauli $X$.

- This decomposition is important for constructing the two-qubit control-$U$ gate.
- The point is to find the explicit form of $A, B, C$, subject to the constriant $ABC = I$.

---

**Proof (I am not 100% satisfied with it):**

- Let us just put $B = A^{-1}C^{-1}$. Then, we have

$$AXA^{-1}C^{-1}XC = R_z(\beta)R_y(\gamma)R_z(\delta) \tag{80}$$

- The solution may not be unqiue (?)
- [Guess] Because $C$ is the last one in the expression, may be we can try to make $C = R_z(c)$?
- It means that we can get rid of the inverse, i.e., $XC^{-1}X = C$. (We showed this trick before)

$$LHS = AXA^{-1}XXC^{-1}XC = AXA^{-1}XC^2 \tag{81}$$

- It may be tempting to put $C^2 = R_z(\delta)$, which means that $c = \delta$, but it turns out not to be a good idea.
- Let us, instead, consider how we can make $AXA^{-1}X = A^2$ ?
- What can $A$ be? $A$ can be $R_z$ or $R_y$ or the product of them?.
- Therefore, let us try to put $A = R_z(a)R_y(b)$, then

$$AXA^{-1}X = R_zR_yXR_y^{-1}XXR_z^{-1}X = R_zR_y^2R_z \tag{82}$$

- Combing with the $C^2$, we have something like

$$LHS = R_z(a)R_y^2(b)R_z(a)R_z(2c) \tag{83}$$

- Obviously, we would put $R_z(a) = R_z(\beta)$, which means $a = \beta$.
- Then, we put $R_z(a)R_z(2c) = R_z(a+2c) = R_z(\delta)$, which means that $c = (\delta - \beta)/2$.
- Also, we have $b = \gamma/2$.
- To summarize, we have $A = R_z(\beta)R_y(\gamma/2)$, $C = R_z\left(\frac{\delta-\beta}{2}\right)$, $B = A^{-1}C^{-1}$.

---

**Alternative proof (a better one):**

- Alternatively, let us consider the following trick:

$$R_y\left(\frac{\gamma}{2}\right)X^c R_y\left(\frac{-\gamma}{2}\right)X^c \tag{84}$$

(i) $c = 1$, we have $R_y\left(\frac{\gamma}{2}\right)XR_y\left(\frac{-\gamma}{2}\right)X = R_y(\gamma)$, (ii) $c = 0$, we have $R_y\left(\frac{\gamma}{2}\right)R_y\left(\frac{-\gamma}{2}\right) = I$.

- This is basically, a control-$R_y$ gate.
- Now, if we look at the expression:

$$R_z(\beta)R_y\left(\frac{\gamma}{2}\right)\cdot X^c R_y\left(\frac{-\gamma}{2}\right)X^c \tag{85}$$

(i) when $c = 1$, we recover $R_z(\beta)R_y(\gamma)$, which is missing a factor of $R_z(\delta)$. (ii) but, when $c = 0$, we have $R_z(\beta)$.

- Therefore we need to construct the $R_z$ gates for theses cases.
- Let us consider the extra part,

$$X^c R_z(a)X^c R_z(b) \tag{86}$$

(i) $c = 1$, it becomes $R_z(-a+b)$, (ii) $c = 0$, it becomes $R_z(a+b)$

- Therefore, we just need to set $\delta = -a+b$ and $-\beta = a+b$. In other words, we have $a = -(\beta+\delta)/2$ and $b = (\delta-\beta)/2$.
- Putting everything together, the follwing expression:

$$R_z(\beta)R_y\left(\frac{\gamma}{2}\right)\cdot X^c R_y\left(\frac{-\gamma}{2}\right)X^c \cdot X^c R_z(a)X^c R_z(b) = R_z(\beta)R_y\left(\frac{\gamma}{2}\right)X^c R_y\left(\frac{-\gamma}{2}\right)R_z(a)X^c R_z(b) \tag{87}$$

should give the correct answer.

---

- I encourage you to think if there is a more elegant proof.

*Corollary 4.2*: Suppose $U$ is a unitary gate on a single qubit. Then there exist unitary operators $A, B, C$ on a single qubit such that $ABC = I$ and $U = e^{i\alpha} AXBXC$, where $\alpha$ is some overall phase factor.

*Proof*
In the notation of Theorem 4.1, set $A \equiv R_z(\beta)R_y(\gamma/2)$, $B \equiv R_y(-\gamma/2)R_z(-(\delta+\beta)/2)$ and $C \equiv R_z((\delta-\beta)/2)$. Note that

$$ABC = R_z(\beta)R_y\left(\frac{\gamma}{2}\right)R_y\left(-\frac{\gamma}{2}\right)R_z\left(-\frac{\delta+\beta}{2}\right)R_z\left(\frac{\delta-\beta}{2}\right) = I\,. \tag{4.14}$$

Since $X^2 = I$, and using Exercise 4.7, we see that

$$XBX = XR_y\left(-\frac{\gamma}{2}\right)XXR_z\left(-\frac{\delta+\beta}{2}\right)X = R_y\left(\frac{\gamma}{2}\right)R_z\left(\frac{\delta+\beta}{2}\right). \tag{4.15}$$

Thus

$$AXBXC = R_z(\beta)R_y\left(\frac{\gamma}{2}\right)R_y\left(\frac{\gamma}{2}\right)R_z\left(\frac{\delta+\beta}{2}\right)R_z\left(\frac{\delta-\beta}{2}\right) \tag{4.16}$$

$$= R_z(\beta)R_y(\gamma)R_z(\delta)\,. \tag{4.17}$$

Thus $U = e^{i\alpha}AXBXC$ and $ABC = I$, as required. □

**Figure 5:** Proof given by the textbook; you can't learn anything

- This finishes all we want to discuss for single-qubit gates. Now, let us move on to look at the two-qubit gates, i.e., 4x4 unitary matrices.
- There are too many degrees of freedom; we will just learn the ones that would be useful later.

**CNOT gates**

- The most famous two-qubit gate would be CNOT, which is applied in the computational basis as

$$|00\rangle \rightarrow |00\rangle$$
$$|01\rangle \rightarrow |01\rangle$$
$$|10\rangle \rightarrow |11\rangle$$
$$|11\rangle \rightarrow |10\rangle$$

- Sometimes, the first qubit is called the control qubit, the second qubit is called the target qubit.
- These can be expressed in a more compact form:

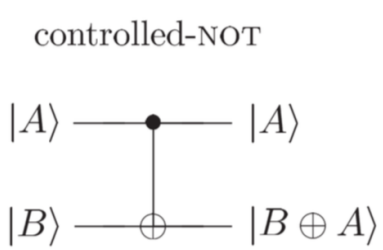
$$|x, y\rangle \rightarrow |x, y \oplus x\rangle \tag{88}$$

where $y \oplus x$ is the same as bitwise product (addition modulo two), same as XOR gate.

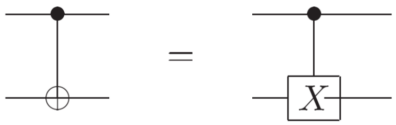
- Let us create the matrix form for CNOT

$$U_{CNOT} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \langle 00| + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \langle 01| + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \langle 10| + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \langle 11| = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{89}$$

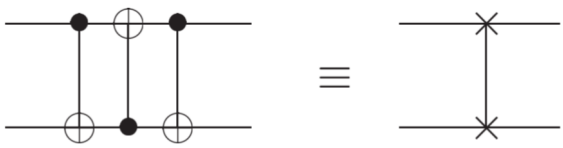
- The quantum circuit symbol of CNOT is like this:




or equivalently,




- CNOT gate is reversible; in fact we have $CNOT\ CNOT\ = I$.
- Later, we will prove that CNOT together with the set of single-qubits gates are universal for quantum computation, which means that with them, we can consstruct all unitary transformation of any number of qubits.
- The question is, how do we realize a CNOT gate in the lab? In general, interaction between two physical qubits is needed to realize two-qubit gates.
- One more thing, it is a common knowledge in quantum computing that three CNOTS forms a SWAP gate:




$$|00\rangle \rightarrow |00\rangle \rightarrow |00\rangle \rightarrow |00\rangle$$
$$|01\rangle \rightarrow |01\rangle \rightarrow |11\rangle \rightarrow |10\rangle$$
$$|10\rangle \rightarrow |11\rangle \rightarrow |01\rangle \rightarrow |01\rangle$$
$$|11\rangle \rightarrow |10\rangle \rightarrow |10\rangle \rightarrow |11\rangle$$

- Look at it. The first and the last rows are not changed, but the second and the thrid rows are flipped. What does it mean? Consider the following:

$$U_{SWAP}(\alpha|0\rangle + \beta|1\rangle)(\gamma|0\rangle + \delta|1\rangle) = (\gamma|0\rangle + \delta|1\rangle)(\alpha|0\rangle + \beta|1\rangle) \tag{90}$$

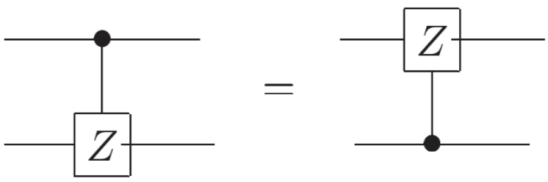
which is clear if we perform an expansion: $\alpha\gamma|00\rangle + \beta\gamma|10\rangle + \alpha\delta|01\rangle + \beta\delta|11\rangle$

$$\begin{aligned} U_{SWAP}(\alpha\gamma|00\rangle + \beta\gamma|10\rangle + \alpha\delta|01\rangle + \beta\delta|11\rangle) &= \alpha\gamma|00\rangle + \beta\gamma|01\rangle + \alpha\delta|10\rangle + \beta\delta|11\rangle \\ &= \gamma|0\rangle(\alpha|0\rangle + \beta|1\rangle) + \delta|1\rangle(\alpha|0\rangle + \beta|1\rangle) \\ &= (\gamma|0\rangle + \delta|1\rangle)(\alpha|0\rangle + \beta|1\rangle) \end{aligned} \tag{91}$$

## 8. Lecture 08

**Quantum circuits**

- A quantum cirucuits (~big unitary transformation) consists of a set of gates (smaller unitary transformation).
- Each line represents a qubit.
- Unfortunately, quantum circuit starts from left to right, but our matrix multiplication starts from right to left.
- In other words, for $AB|\psi\rangle$, we have to draw B first followed by A.
- Quantum circuit is acyclic, meaning that there is no loop
- Quantum circuit is reversible. (Classical circuits are normally irreversible)
- As a further example, we also have controlled-$Z$ ($CZ$) gate:




- This gate is also very popular.
- What is does is the following: (recall that $Z|0\rangle = |0\rangle$, $Z|1\rangle = -|1\rangle$)

$$\begin{aligned} |00\rangle &\rightarrow |00\rangle \\ |01\rangle &\rightarrow |01\rangle \\ |10\rangle &\rightarrow |10\rangle \\ |11\rangle &\rightarrow -|11\rangle \end{aligned}$$

- For completeness, let us also work out the matrix form:

$$U_{CZ} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\langle 00| + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}\langle 01| + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}\langle 10| + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}\langle 11| = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{92}$$

- [Common knowledge] There is a relationship betweem CZ and CNOT gates:

Sandwiching a CZ gate with two Hadamard gates is equivalent to a CNOT gate.

**Controlled-U gate**

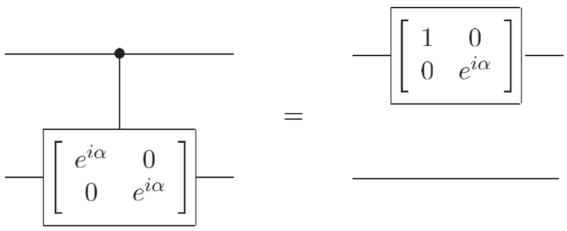
- Let us first consider how to deal with the special case, where $U$ is just a global phase $U = e^{i\alpha} I$.
- Note that if this is just for a single qubit gate, we can just ignore it, for the purpose of quantum computing; this just gives a global phase to a quantum state, which is unphysical.
- However, for controlled operations it is somewhat non-trival. Because,

$$|00\rangle \to |00\rangle,\ |01\rangle \to |01\rangle,\ |10\rangle \to e^{i\alpha}|10\rangle,\ |11\rangle \to e^{i\alpha}|11\rangle \tag{93}$$

- In other words, for a general two qubit state, $|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$, it becomes

$$|\psi'\rangle = a|00\rangle + b|01\rangle + ce^{i\alpha}|10\rangle + de^{i\alpha}|11\rangle = \left(|0\rangle\langle 0| + e^{i\alpha}|1\rangle\langle 1|\right) \otimes I\ |\psi\rangle. \tag{94}$$

- Recall that $(A \otimes B)|\psi\rangle \otimes |\phi\rangle = A|\psi\rangle \otimes B|\phi\rangle$.
- If you look at it closely, you would find out that it is equivalent to a single-qubit gate acting on the first qubit.




- The second part is the $AXBXC$, where $ABC = I$.
- Previously, we already learned how to do a controlled-$R_y$ gate with two additional CNOT gates, labelled as either $U_{CNOT}$ or $C_X$.
- Of course, we can also do the same for the other $R_z$ gates, but it would require like 4 CNOT gates in total.
- Now, let us consider the following circuit identitiy:

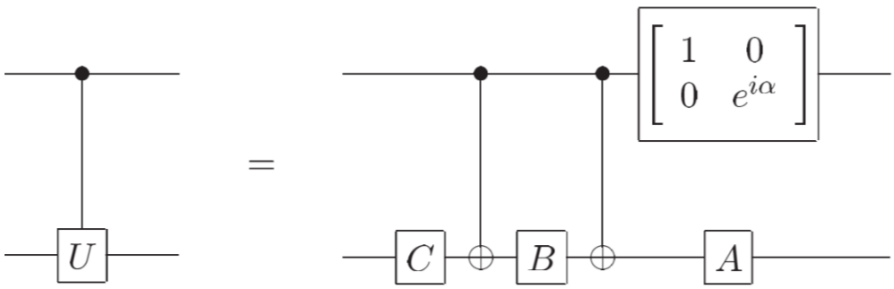



**Figure 6:** prove by pointing fingers

- LHS: Let us consider a general quantum state: $|\psi\rangle = a|0\rangle|\phi_0\rangle + b|1\rangle|\phi_1\rangle$. A controlled-U, $C_U$,operation does the following job:

$$C_U|\psi\rangle = a|0\rangle|\phi_0\rangle + b|1\rangle U|\phi_1\rangle \tag{95}$$

- RHS: let us work on it step-by-step:

(i) $(I \otimes C)|\psi\rangle = a|0\rangle C|\phi_0\rangle + b|1\rangle C|\phi_1\rangle$
(ii) $C_X C|\psi\rangle = a|0\rangle C|\phi_0\rangle + b|1\rangle XC|\phi_1\rangle$
(iii) $BC_X C|\psi\rangle = a|0\rangle BC|\phi_0\rangle + b|1\rangle BXC|\phi_1\rangle$
(iv) $C_X BC_X C|\psi\rangle = a|0\rangle BC|\phi_0\rangle + b|1\rangle XBXC|\phi_1\rangle$
(v) $AC_X BC_X C|\psi\rangle = a|0\rangle ABC|\phi_0\rangle + b|1\rangle AXBXC|\phi_1\rangle$
(vi) $R_\alpha AC_X BC_X C|\psi\rangle = a|0\rangle ABC|\phi_0\rangle + b|1\rangle e^{i\alpha} AXBXC|\phi_1\rangle$

- Recall that $ABC = I$ and $U = e^{i\alpha} AXBXC$, so the final is the same as that from $C_U$.
- To summarize, if we have two $CNOT$ gates, together with single-qubit gates, then we can simulate any controlled-U gates.

**Multiple controlled operations**

- In general, we may want to have multiple controlling qubits, as well as multiple target qubits. Why not?
- Operationally, we want to have something like

$$C_n(U)|x_1 x_2 \ldots x_n\rangle|\psi\rangle = |x_1 x_2 \ldots x_n\rangle U^{x_1 x_2 \cdots x_n}|\psi\rangle \tag{96}$$

which means that if any bit value is zero, then $U^{x_1 x_2 \cdots x_n} = I$

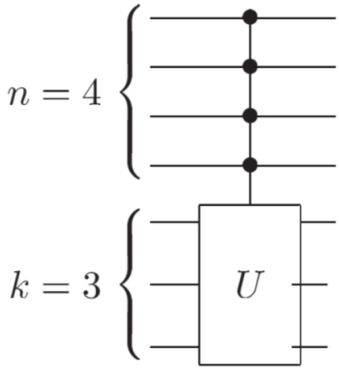



**Figure 7:** the first four qubits need to be all 1, in order to apply the unitary gate

## 9. Lecture 09

- The idea of the following is discussion is to first create some building blocks, which will be needed later. You may find it kind of random. The later goal is to apply these building blocks to prove that universality of CNOT+single-qubit gates.
- As an example of multiple control gate, we consider the following case:

$$U = V^2. \tag{97}$$

- Later, we will just need to consider the case of $U = X$ for the so-called Toffoli gate.
- Let us consider the following:

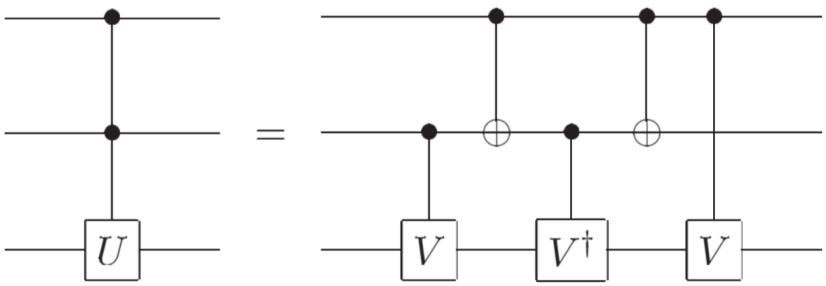



**Figure 8:** CC-U gate can be simulated by CNOT+single qubit gates

- This is a kind of standard decomposition method for simplifying controlled-controlled-U gate with controlled V gates.
- To understand it (not prove it), may be let us forget about the first and the last control operations. In this case, the **middle part** gives the following transformation:

$$\begin{aligned} |00\rangle &\to I & |00\rangle\langle 00| \otimes I &\to |00\rangle\langle 00| \otimes I \\ |11\rangle &\to I & |11\rangle\langle 11| \otimes I &\to |11\rangle\langle 11| \otimes V^2 \\ |01\rangle &\to V^\dagger & |01\rangle\langle 01| \otimes V^\dagger &\to |01\rangle\langle 01| \otimes V^\dagger V \end{aligned}$$

$$|10\rangle \to V^\dagger \qquad\qquad |10\rangle\langle 10| \otimes V^\dagger \to |10\rangle\langle 10| \otimes VV^\dagger$$

- Of course, the actual thing we want for CCU is like

$$CC_U = |00\rangle\langle 00| \otimes I + |01\rangle\langle 01| \otimes I + |10\rangle\langle 10| \otimes I + |11\rangle\langle 11| \otimes U \tag{98}$$

- If we now apply back the first and the last controlled operations, then we get the correct answer.
- Next, may be we should **try to pretend that we did not see the decomposition** and try to figure out ourselves, using only controlled-V, controlled-V^dagger, CNOT. (Why not using controlled-U?) This would be a good homework.
- Note that the textbook did not provide an explanation for the construction:

> Suppose $U$ is a single qubit unitary operator, and $V$ is a unitary operator chosen so that $V^2 = U$. Then the operation $C^2(U)$ may be implemented using the circuit shown in Figure 4.8.
>
> **Exercise 4.21:** Verify that Figure 4.8 implements the $C^2(U)$ operation.

**Toffoli gate**

- When $U = X$, then $CCX$ is called the Toffoli gate.
- In this case, $V = (1-i)(I+iX)/2$ (do you want prove it? homework? hint: any 2x2 matrix can be expanded by the Puali groups, $\{I, X, Y, Z\}$, so you may take $V = aI + bX + cY + dZ$; there may be a smarter way of doing it.)
- Let us quickly check to make sure $V$ is unitary ,as

$$V^\dagger = (1+i)(I-iX)/2 \tag{99}$$

- Then,

$$V^\dagger V = \frac{1}{4}(1-i)(1+i)(I-iX)(I+iX) = \frac{1}{2}(I+I) = I \tag{100}$$

- At least, we now how to simulate Toffoli gate with the CNOT+singel-qubit gates set.
- Let us look at the first application of the Toffoli gate: multiple-control U gate:

$$C_n(U)|x_1x_2x_3x_4x_5\rangle|\psi\rangle = |x_1x_2x_3x_4x_5\rangle U^{x_1x_2x_3x_4x_5}|\psi\rangle \tag{101}$$

- c is the same as x in the figure below:

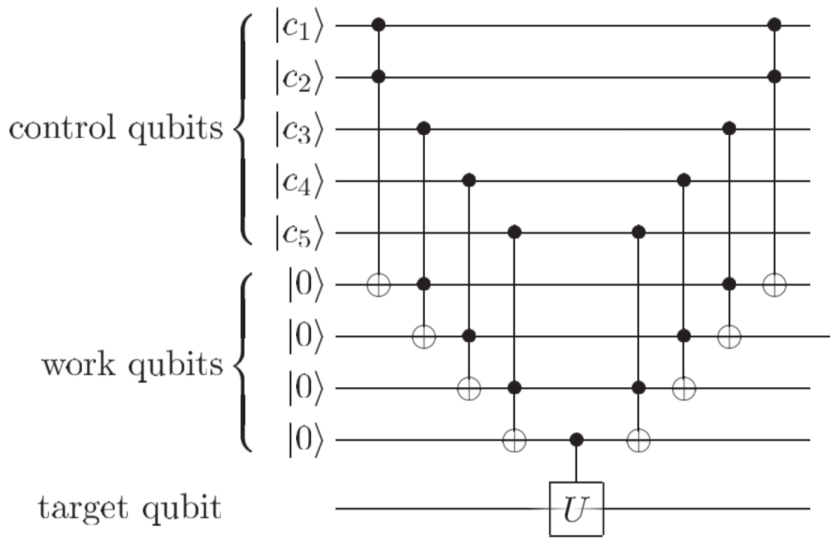



**Figure 9:** can you prove by pointing fingers?

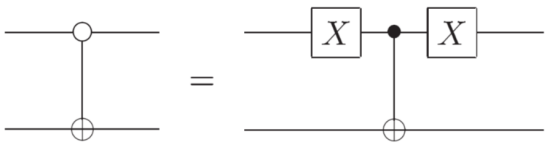
- The first work (ancilla) qubit gets flipped to $|1\rangle$, iff both c1 and c2 are $|1\rangle$.
- The second work qubit gets flipped to $|1\rangle$, iff both c3 and work qubit 1 are $|1\rangle$, which means that all control qubits 1, 2, 3 are $|1\rangle$.
- **Example**, consider the initial state like: $\alpha|00000\rangle|\psi\rangle + \beta|11111\rangle|\psi\rangle$, after inserting the ancilla qubits in the middle, we have $\alpha|00000\rangle|0000\rangle|\psi\rangle + \beta|11111\rangle|0000\rangle|\psi\rangle$. Let us then apply the first half of the gates (until the CU), we will have the following:

$$\alpha|00000\rangle|0000\rangle|\psi\rangle + \beta|11111\rangle|1111\rangle U|\psi\rangle \tag{102}$$

- Here the ancilla qubits are entangled with the system qubits, which is not something we want. Therefore, we need to reset them to all 0, at the end. This is why we need to apply the Toffoli gates in the reverse order.

## 10. Lecture 10

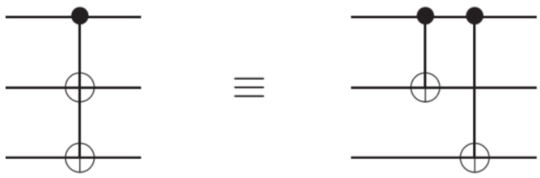
- A couple of remarks: there is also a convention of controlled gates based on the "0" state instead of the "1".




- To see this, note that we can insert $XX = I$ to both sides of the following,

$$|0\rangle\langle 0| \otimes X + |1\rangle\langle 1| \otimes I = X|1\rangle\langle 1|X \otimes X + X|0\rangle\langle 0|X \otimes I = (X \otimes I)U_X(X \otimes I) \tag{103}$$

- Also, don't get confused with the following:



**Creation of Bell states (entangled states)**

- There are four Bell states:

$$|B_{00}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$
$$|B_{01}\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}$$
$$|B_{10}\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$
$$|B_{11}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

which forms a complete basis for two qubits, i.e., any two qubit state can be expanded by them,

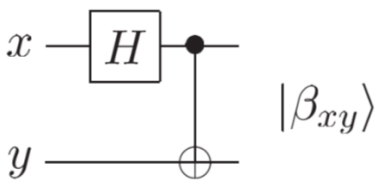
$|\psi\rangle = a|B_{00}\rangle + b|B_{01}\rangle + c|B_{10}\rangle + d|B_{11}\rangle$ .

- They are all maximally entangled, i.e., they are the most entangled.
- To generate the $|B_{00}\rangle$ Bell state, we just need to have two operations:

$$U_{CNOT}(H \otimes I)|00\rangle = \frac{1}{\sqrt{2}} U_{CNOT}(|0\rangle + |1\rangle)|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) \quad (104)$$

- Two ingredients: **superposition** + **entangling gate** (CNOT)
- The corresponding quantum circuit diagram is as follows ($\beta = B$ below):

| In | Out |
|---|---|
| $\|00\rangle$ | $(\|00\rangle + \|11\rangle)/\sqrt{2} \equiv \|\beta_{00}\rangle$ |
| $\|01\rangle$ | $(\|01\rangle + \|10\rangle)/\sqrt{2} \equiv \|\beta_{01}\rangle$ |
| $\|10\rangle$ | $(\|00\rangle - \|11\rangle)/\sqrt{2} \equiv \|\beta_{10}\rangle$ |
| $\|11\rangle$ | $(\|01\rangle - \|10\rangle)/\sqrt{2} \equiv \|\beta_{11}\rangle$ |



- To get the other cases, you may also consider **working on the diagram directly**, using circuit identities.

**Quantum Teleportation**

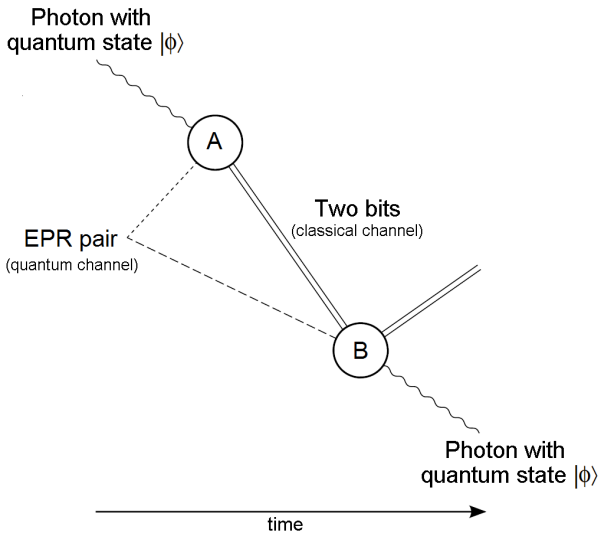



**Figure 10:** from wikipedia

- It does one thing: destroy (unknown) quantum state at point A, and re-create it at point B.
- There is no transfer of matter between A and B.
- We can only achieve it by consuming one pair of entangled state. (Entanglement may be regarded as resource for the purpose of teleportation).
- I think the quantum satelite in China has already demonstrated the quantum teleportation over a long distance.
- To be more precise, we assume (i) Alice is given an **unknown** quantum state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. (ii) Alice and Bob are shared with a Bell state $|B_{00}\rangle$.

**Step 0**: the initial state is as follows:

$$|\psi_0\rangle = |\psi\rangle|B_{00}\rangle = (\alpha|0_A\rangle + \beta|1_A\rangle)(|0_A 0_B\rangle + |1_A 1_B\rangle)/\sqrt{2} \tag{105}$$

**Step 1**: Alice applies a CNOT to her qubits: $|\psi_1\rangle = U_{CNOT}|\psi_0\rangle$

$$|\psi_1\rangle = \alpha|0\rangle|B_{00}\rangle + \beta|1\rangle(X \otimes I)|B_{00}\rangle = \alpha|0\rangle|B_{00}\rangle + \beta|1\rangle|B_{01}\rangle \tag{106}$$

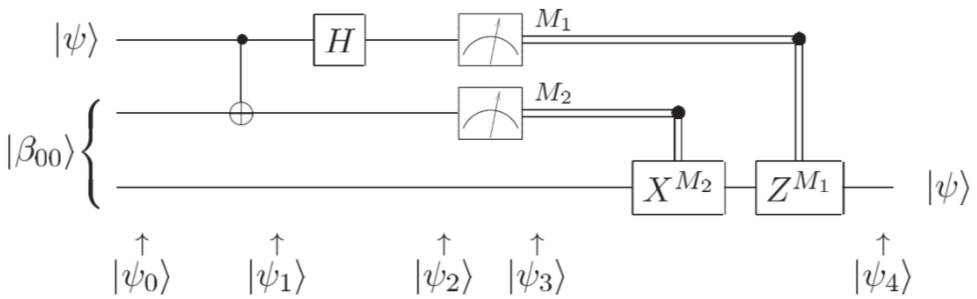
**Step 2**: Alice further applies a Hadamard gate to the first qubit: then we have something like: $\alpha'(|0\rangle + |1\rangle)|B_{00}\rangle + \beta'(|0\rangle - |1\rangle)|B_{01}\rangle$, where $\alpha' \equiv \alpha/\sqrt{2}$ and $\beta' \equiv \beta/\sqrt{2}$.

$$|\psi_2\rangle = |0\rangle(\alpha'|B_{00}\rangle + \beta'|B_{01}\rangle) + |1\rangle(\alpha'|B_{00}\rangle - \beta'|B_{01}\rangle) \tag{107}$$

which is equal to the following (good homework exercise):

$$|\psi_2\rangle = \frac{1}{2}(|00\rangle|\psi\rangle + |01\rangle X|\psi\rangle + |10\rangle Z|\psi\rangle + |11\rangle ZX|\psi\rangle)\ (?) \tag{108}$$

- To check this, let us take a look at:
  $\alpha'|B_{00}\rangle + \beta'|B_{01}\rangle = \alpha''(|00\rangle + |11\rangle) + \beta''(|01\rangle + |10\rangle)$ , where $\alpha'' \equiv \alpha'/\sqrt{2}$ and $\beta'' \equiv \beta'/\sqrt{2}$. Therefore, we have
  $|0\rangle(\alpha''|0\rangle + \beta''|1\rangle) + |1\rangle(\alpha''|1\rangle + \beta''|0\rangle) = \frac{1}{2}(|0\rangle|\psi\rangle + |1\rangle X|\psi\rangle)$ . You may work out the second part.
- If Alice makes a measurement (on the computational basis) on her two qubits, she will get one of the four possibilities. For example, if she gets $|01\rangle$, then Bob has the $X|\psi\rangle$ state.
- However, Bob does not know the answer to the measurement, unless Alice tells Bob about the measurement outcome.
- Teleporation **cannot** achieve superluminal communication.
- Finally, Bob applies the corresponding correction to his qubit, e.g. $X$ when Alice got $|01\rangle$.
- The corresponding quantum circuit diagram is given by:




- Alternatively, we may have a deeper understanding on quantum teleportation based on the following:
- Let us consider only two qubits, initialized as $|\psi\rangle|+\rangle$ (where $|\psi\rangle \equiv \alpha|0\rangle + \beta|1\rangle$ ), now what happens if we apply CZ gate?

$$|\psi\rangle|+\rangle \rightarrow \alpha|0\rangle|+\rangle + \beta|1\rangle|-\rangle \tag{109}$$

- Then, we apply a hadamard gate to the first qubit,

$$\alpha|0\rangle|+\rangle + \beta|1\rangle|-\rangle \propto |0\rangle\big(\alpha|+\rangle + \beta|-\rangle\big) + |1\rangle\big(\alpha|+\rangle - \beta|-\rangle\big) \tag{110}$$

- So, if we apply another Hadamard, we get something like

$$|0\rangle\ |\psi\rangle + |1\rangle\ Z|\psi\rangle \tag{111}$$

## 11. Lecture 11

- What if Alice does not tell Bob the measurement outcome? Is there a change in Bob's state? Let us look at Bob's qubit. Before receiving Alice's classical message, Bob's qubit is a mixture of all four possible states, namely $\{|\psi\rangle, X|\psi\rangle, Z|\psi\rangle, ZX|\psi\rangle\}$.
- In fact, this set of states is eqivalently to $\{|0\rangle, |1\rangle\}$, which comes from the Bell state.
- To see this, for example, Bob performs some measurement on $Z$. With $\{|0\rangle, |1\rangle\}$, we have

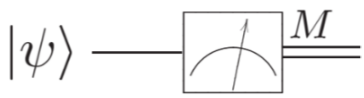
$$\langle Z\rangle = \frac{1}{2}\langle 0|Z|0\rangle + \frac{1}{2}\langle 1|Z|1\rangle = 0 \tag{112}$$

- After Alice measurement, Bob would imagine he has instead of the following set of states $\{|\psi\rangle, X|\psi\rangle, Z|\psi\rangle, ZX|\psi\rangle\}$. (Remember $XZ = -ZX$, and $X^2 = Z^2 = I$)

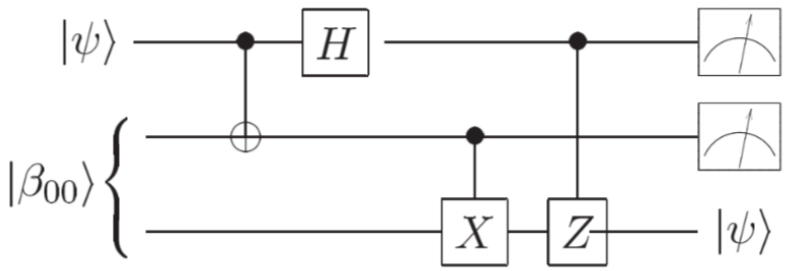
$$\langle Z\rangle = \frac{1}{4}\langle\psi|Z|\psi\rangle - \frac{1}{4}\langle\psi|Z|\psi\rangle + \frac{1}{4}\langle\psi|Z|\psi\rangle - \frac{1}{4}\langle\psi|Z|\psi\rangle = 0 \tag{113}$$

**Measurement**

- A quantum measurement, from quantum computing point of view, is often assumed to be a projective measurement along the computational basis.
- One may express the measurement action in the circuit diagram as follows:




- The single line before the measurement means a qubit; the double line after the measurement means a classical bit.
- There is something quite trivial, and the textbook gave it a name "principle of deferred measurement".
- Let us look at the quantum teleportation circuit as an example:



- It is because, before the measurement, we have something like:

$$|\psi_2\rangle = \frac{1}{2}(|00\rangle|\psi\rangle + |01\rangle X|\psi\rangle + |10\rangle Z|\psi\rangle + |11\rangle ZX|\psi\rangle) \tag{114}$$

- If we perform instead the controlled operations, then we get the following:

$$|\psi_2^{new}\rangle = \frac{1}{2}(|00\rangle|\psi\rangle + |01\rangle|\psi\rangle + |10\rangle|\psi\rangle - |11\rangle|\psi\rangle) \tag{115}$$

- Also, there is something called principle of implicit measurement, which means the last qubits need not be measured.
- If you do not understand these things, it is totally normal; because we did not cover the background materials.

**More on quantum teleportation circuit**

- Recall that the main part of the quantum circuit associated with quantum teleportation is really the reverse of the circuit of Bell state creation.
- We may derive the same results by projecting a Bell state to the first two qubits.
- Explicitly, let us consider the case of $00$, we can first write

$$|\psi\rangle|B_{00}\rangle = \frac{1}{\sqrt{2}}(|\psi, 0\rangle|0\rangle + |\psi, 1\rangle|1\rangle) \tag{116}$$

  where $|\psi\rangle \equiv \alpha|0\rangle + \beta|1\rangle$ and $|\psi, 0\rangle \equiv |\psi\rangle|0\rangle$.
- Therefore, after projecting with $|B_{00}\rangle$, we have

$$\frac{1}{\sqrt{2}}(\langle B_{00}|\psi, 0\rangle|0\rangle + \langle B_{00}|\psi, 1\rangle|1\rangle) = \frac{1}{2}(\alpha|0\rangle + \beta|1\rangle) = \frac{1}{2}|\psi\rangle \tag{117}$$

- You may want to try other possibilities.

**No quantum-copy circuit**

- It is also called no-cloning theorem, which inspired many "no-go" theorem.
- Recall that in classical world, classical information can be copied any times we want. $0 \to 000000000....$ , $1 \to 111111....$
- For example, we can achieve classical cloning by CNOT, for $x \in \{0, 1\}$,

$$U_{CNOT}|x\rangle|0\rangle = |x\rangle|x\rangle \tag{118}$$

- The point is that it is impossible to make such an (**state-independent**) operation for (**unknown**) quantum states, i.e., **perfect** cloning $|\psi\rangle \to |\psi\rangle|\psi\rangle|\psi\rangle|\psi\rangle|\psi\rangle....$ is not allowed in quantum mechanics.
- For example, with $|\psi\rangle \equiv \alpha|0\rangle + \beta|1\rangle$,

$$U_{CNOT}|\psi\rangle|0\rangle = \alpha|00\rangle + \beta|11\rangle \neq |\psi\rangle|\psi\rangle \tag{119}$$

Proof of general cases (by contradiction):

- The following works for quantum systems of any size.
- Assume such a cloning machine exists, i.e.,

$$U_{clone}|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle \tag{120}$$

- It should work for any state, so

$$U_{clone}|\phi\rangle|0\rangle = |\phi\rangle|\phi\rangle \tag{121}$$

- Of course, we will need to assume such an operation is unitary, which means that

$$\langle\psi|\phi\rangle = \langle\psi|\langle 0|U_{clone}^{\dagger}U_{clone}|\phi\rangle|0\rangle = \langle\psi|\phi\rangle^2 \tag{122}$$

Alternative proof (based on linearity):

- Here let us consider qubits, $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$
- Assume such a cloning machine exists, i.e.,

$$U_{clone}|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle \tag{123}$$

- Also assume that it is a linear operation, i.e.,

$$U_{clone}|\psi\rangle|0\rangle = \alpha U_{clone}|0\rangle|0\rangle + \beta U_{clone}|1\rangle|0\rangle = \alpha|0\rangle|0\rangle + \beta|1\rangle|1\rangle \tag{124}$$

- This again leads to a contradiction.

**Universal quantum gates (exact)**

- Classical computation, e.g. NAND gates is unviersal, and also Toffoli gate, together with ancilla (extra) bits.
- The idea is to find out what are the universal gate set for quantum computation; there may be many different choices.
- More mathematically, for $U_j$ acting on $k$ or less qubits; later we will find out that $k = 2$,

$$G = \{U_1, U_2, U_3, .., U_m\} \tag{125}$$

- As an example, we would be interested in the set like

$$G = \{U_{CNOT}, R_z(\theta), R_y(\phi)\} \tag{126}$$

## 12. Lecture 12 (Review+Discussion)

- Is quantum computer (model) more powerful than classical computer (model)? In short, we don't know for sure, which means that we don't have rigourous proof.
- What is means by more powerful? or faster?
- We want to separate the effects from the hardware; so we would be interested in counting the number of steps, instead of the actual running time.
- In computer science, roughly speaking, if a computing machine can solve a problem with a running time scaling as a **polynomial function** poly(n) of the the input size n, then we say the problem can be efficiently solved by the machine.

- In constrast, if the machine (algorithm) can only solve the problem in **exponential** times, then we say the algorithm is not efficient.
- With this understanding, may be we should ask is there a problem, where quantum computing model can solve efficiently, but not classically? Shor's factoring algorithm? The problem is that no one can prove rigorouly that there does not exist a classical algorithm beating the Shor's quantum algorithm.
- Reference on quantum supremacy:

PERSPECTIVES

*National Science Review*
00: 1–2, 2018
doi: 10.1093/nsr/nwy072
Advance access publication 6 July 2018

PHYSICS

Special Topic: Quantum Computing

## Quantum supremacy: some fundamental concepts

Man-Hong Yung[1,2]

## 13. Lecture 13

- So far, we learnt that quantum computation can regarded as a transform of a unitary matrix to some quantum state (vector).
- We also learnt many single qubit gates and some two-qubit gates, and some multi-qubit gates.
- We also learnt some really basical quantum algorithms, using single and two qubit gates.
- Moreover, we have seen that the multi-qubit gates can be decomposed into single qubit and two-qubit gates.
- Are they sufficient for arbitrary unitary transformation for any number of qubits? If so, we can call them (the set of all **single-qubit gates+CNOT**) universal gate set.

**[WARNING]: all maths below (forget about qubits)**

- First of all, we need to define the unitary gate. To do that, let us forget about the physics of quantum mechanics.
- For simplicity, we just label a $N$-dimensional vector by some basis vectors: $\{|0\rangle, |1\rangle, |2\rangle, ..., |N-1\rangle\}$.
- In this way, any arbitrary (unitary) matrix $U$ can be defined by $N$ vectors as follows:

$$U = |\psi_0\rangle\langle 0| + |\psi_1\rangle\langle 1| + ... + |\psi_{N-1}\rangle\langle N-1| \tag{127}$$

where $|\psi_k\rangle$'s form a **orthonormal set** of basis vectors (**important**).

- As an example, recall that $\langle\psi_1|\psi_0\rangle = \langle 1|U^\dagger U|0\rangle = \langle 1|0\rangle = 0$ .
- In other words, the column vectors $|\psi_k\rangle$'s are the input of the problem.

**Goal of part (1/2): decompose $U$ into something called 2x2 unitaries, each of which acts on a small subspace**

- The strategy we will follow, is to step-by-step transform the unitary matrix into a diagonal matrix.
- For example, we will find another unitary matrix (hopefully with a simpler structure) $W_0$, to make the following happen:

$$W_0^{-1}U = |0\rangle\langle 0| + |\widetilde{\psi}_1\rangle\langle 1| + ... + |\widetilde{\psi}_{N-1}\rangle\langle N-1| \tag{128}$$

where $|\widetilde{\psi}_k\rangle \equiv W_0|\psi_k\rangle$.

- Note that the product $W_0^{-1}U$ remains a unitary matrix. Due to the requirement of orthnomal condition, all $\boxed{\langle 0|\widetilde{\psi}_k\rangle = 0}$ for all $k > 0$ (this condition is important and like a magic).
- Then, we keep doing it, so that

$$W_{N-2}^{-1}... W_1^{-1}W_0^{-1}U = |0\rangle\langle 0| + |1\rangle\langle 1| + ... + |N-1\rangle\langle N-1| = I \tag{129}$$

- Finally, we can express,

$$U = W_0 W_1 ... W_{N-2} \tag{130}$$

- Next, we will focus on the construction of these $W$'s.
- Later, we will show how they can be constructed with single and two qubit gates (part 2/2).
- Remark: this is not necessarily the most clever way; there may be much more efficient way for some $U$; we are just trying to argue the completeness instead of efficiency.

**Step 1:**

- Let us consider following

$$U|0\rangle = |\psi_0\rangle \equiv a_0|0\rangle + a_1|1\rangle + ... + a_{N-1}|N-1\rangle \tag{131}$$

where all of the $a$'s are given (assumed to be real numbers for simplicity).

- In the matrix form, we can write ( the $*$ means some value we don't care )

$$U = \begin{pmatrix} a_0 & * & * & \cdots & * \\ a_1 & * & * & \cdots & * \\ a_2 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N-1} & * & * & * & * \end{pmatrix} \tag{132}$$

- We will construct a unitary matrix $W_0$ to have the same result as $U$ does on $|0\rangle$, i.e.,

$$W_0|0\rangle = U|0\rangle = |\psi_0\rangle \tag{133}$$

which means that the combination $W_0^{-1}U|0\rangle = |0\rangle$ does not change $|0\rangle$.

- The matrix form of $W_0$ has one common column vector as $U$:

$$W_0 = \begin{pmatrix} a_0 & * & * & \cdots & * \\ a_1 & * & * & \cdots & * \\ a_2 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N-1} & * & * & * & * \end{pmatrix} \tag{134}$$

- In this way, we can write

$$U = \boxed{W_0|0\rangle\langle 0|} + |\psi_1\rangle\langle 1| + ... + |\psi_{N-1}\rangle\langle N-1| \tag{135}$$

  which implies the previous expression:

  $W_0^{-1}U = |0\rangle\langle 0| + |\widetilde{\psi}_1\rangle\langle 1| + ... + |\widetilde{\psi}_{N-1}\rangle\langle N-1|$ .

- This is, in fact, quite similar to the way we prove the Z-Y-Z decomposition.
- Consequently,

$$U_1 \equiv W_0^{-1}U = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & * & * \end{pmatrix} \tag{136}$$

- But we have to put some restriction to $W_0$: it can only be constructed by dealing with pairs like $\{|k\rangle, |k+1\rangle\}$, i.e.,,

$$W_0 = \omega_{N-2} \, ... \, \omega_1\omega_0 \tag{137}$$

- For example, we can start with $|0\rangle$, then apply $\omega_0$ such that

$$\omega_0|0\rangle = a_0|0\rangle + b_0|1\rangle \tag{138}$$

where $|b_0|^2 = |a_1|^2 + ... + |a_{N-1}|^2$ .

- An explicit form of $\omega_0$ is like (not unique, up to some phase factors)

$$\omega_0 = (a_0|0\rangle + b_0|1\rangle)\langle 0| + (b_0|0\rangle - a_0|1\rangle)\langle 1| + |2\rangle\langle 2| + |3\rangle\langle 3| + ... \tag{139}$$

- In matrix form,

$$\omega_0 = \begin{pmatrix} a_0 & b_0 & 0 & \cdots & 0 \\ b_0 & -a_0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{140}$$

- Note that if we consider switching back to the computational basis, e.g. of 2 qubits, $\{|0\rangle \equiv |00\rangle, |1\rangle \equiv |01\rangle, |2\rangle \equiv |10\rangle, |3\rangle \equiv |11\rangle\}$.
- Then,

$$\omega_0 = (a_0|00\rangle + b_0|01\rangle)\langle 00| + (b_0|00\rangle - a_0|01\rangle)\langle 01| + |10\rangle\langle 10| + |11\rangle\langle 11| \tag{141}$$

$$= |0\rangle\langle 0| \otimes ((a_0|0\rangle + b_0|1\rangle)\langle 0| + (b_0|0\rangle - a_0|1\rangle)\langle 1|) + |10\rangle\langle 10| + |11\rangle\langle 11| \tag{142}$$

- It means that we apply a single-qubit gate to the second qubit, iff the first qubit is in the 0 state, which is basically a controlled operation.

---

- Next, we consider $\omega_1$, acting on $\{|1\rangle, |2\rangle\}$, such that

$$\omega_1\omega_0|0\rangle = a_0|0\rangle + b_0\omega_1|1\rangle = a_0|0\rangle + a_1|1\rangle + b_1|2\rangle \tag{143}$$

where $|b_1|^2 = |a_2|^2 + ... + |a_{N-1}|^2$.

- The matrix form of $\omega_1$ is as follows:

$$\omega_1 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & a_1 & b_1 & \cdots & 0 \\ 0 & b_1 & -a_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{144}$$

- We will eventually reach the condition: $\omega_{N-2}$ ... $\omega_1\omega_0|0\rangle = |\psi_0\rangle$
- These $\omega$ gates are called 2x2 unitary gates.
- Therefore, $W_0 = \omega_{N-2}$ ... $\omega_1\omega_0$ is a product of these 2x2 unitary gates.
- If we can show that any 2x2 unitary gate can be simulated by single-qubit+CNOT, then $W_0$ can also be simulated by this set of gates.

**Step 2:**

- We defined a new unitary matrix by $U_1 \equiv W_0^{-1}U$, where $U_1|0\rangle = |0\rangle$.
- Recall that $U_1 = |0\rangle\langle 0| + |\widetilde{\psi}_1\rangle\langle 1| + ... + |\widetilde{\psi}_{N-1}\rangle\langle N-1|$,
- Then, we will construct something similar, i.e., $W_1$ such that (to make things simple,

let us **recycle** the symbols $a$, $b$, $\omega$), $W_1|0\rangle = |0\rangle$, and that

$$W_1|1\rangle = U_1|1\rangle = |\widetilde{\psi}_1\rangle \equiv a_1|1\rangle + ... + a_{N-1}|N-1\rangle \tag{145}$$

- Again, note that $a_0 = 0$, because $\langle 0|\widetilde{\psi}_1\rangle = 0$.
- So, we can write

$$U_1 = |0\rangle\langle 0| + \boxed{W_1|1\rangle\langle 1|} + ... + |\widetilde{\psi}_{N-1}\rangle\langle N-1| \tag{146}$$

- In this way, we can construct a product of 2x2 unitaries,

$$W_1 = \omega_{N-2} \; ... \; \omega_1 \tag{147}$$

satisying the above condition.

- Consequently, we have

$$W_1^{-1}W_0^{-1}U = |0\rangle\langle 0| + |1\rangle\langle 1| + |\widetilde{\psi}_2\rangle\langle 2| + ... \tag{148}$$

where the skipped part is not diagonal and are orthogonal to $|0\rangle$ and $|1\rangle$, the same time, i.e.,

$$|\widetilde{\psi}_2\rangle = a_2|2\rangle + a_3|3\rangle + ... a_{N-2}|N-2\rangle \tag{149}$$

- In the matrix form, we further reduce the non-trivial part:

$$U_2 \equiv W_1^{-1}W_0^{-1}U = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & * & * \end{pmatrix} \tag{150}$$

## 14. Lecture 14

**Goal of part (2/2): decompose the 2x2 unitaries into single-qubit gates and CNOTs**

- Ok, it means that we now need to worry about these 2x2 unitaries.
- Let us consider an example, assuming the 2x2 matrix is between these two basis vectors (for the case of $W_0$, it is related to the last one $\omega_{N-2}$):

$$|s\rangle \equiv |111...10\rangle$$
$$|t\rangle \equiv |111...11\rangle$$

- Therefore, we need to consider the corresponding 2x2 matrix, labelled by $V$:

$$V = |0\rangle\langle 0| + |1\rangle\langle 1| + ... + (\alpha|s\rangle + \beta|t\rangle)\langle s| + (\beta|s\rangle - \alpha|t\rangle)\langle t| \tag{151}$$

where

$$V|s\rangle \equiv \alpha|s\rangle + \beta|t\rangle$$
$$V|t\rangle \equiv \beta|s\rangle - \alpha|t\rangle$$

- If it acts on other states $|u\rangle$ outside the subspace, then $V|u\rangle = |u\rangle$.
- If we look at it closely, the operation $V$ is nothing but multiple controlled rotation that we covered before, e.g.

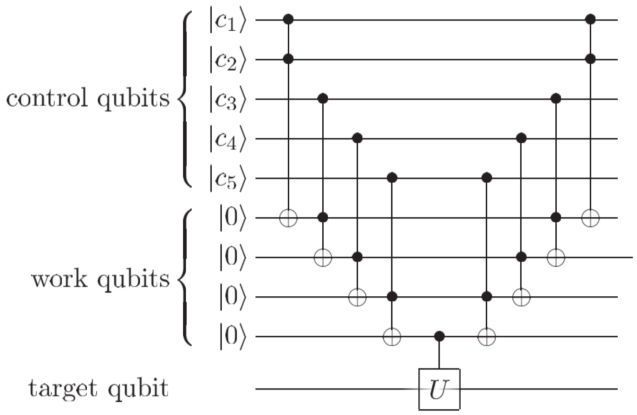
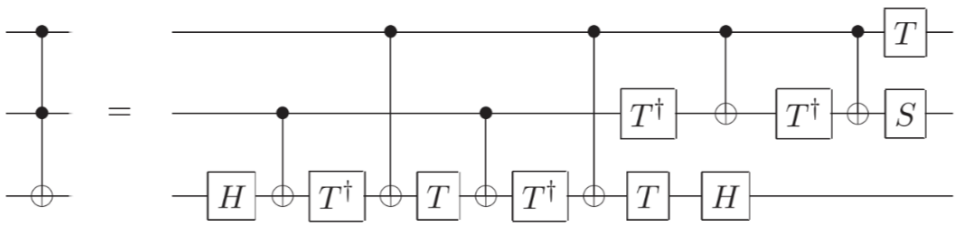



**Figure 11:** example of 6 qubits

- So, we have just sorted out one of the 2x2 unitaries, which can be simulated by the single-qubit gates together with CNOTs (of course, we also need ancilla qubits).
- There is a famous decomposition for the Toffoli gate




- Note that if we are able to have $T$ gates, then we can also have $T^\dagger = T^7$ (see below)

**General pairs (k,k+1) with Gray code**

- In general, the pairs $\{|k\rangle, |k+1\rangle\}$ may not have simple strcuture like the previous one.
- From the previous example, we know that the "trick" can be repeated, whenever the two bit-strings in the computational basis are different by a single bit.
- One way to fix this problem (meaning that putting any pair of bit strings to be different only by one bit) is the so-called Gray code.
- Let us consider an explicit example of an arbitrary pair of bit strings: $s = 101001$, $t = 110011$.
- The goal is to find a sequence of 1-bit operations, to make these two strings to be differed only by 1 bit.
- Explicitly,

$$
\begin{aligned}
1\,0\,1\,0\,0\,1 &= s && \textit{old value of } s \\
1\,0\,1\,0\,1\,1 &= s' && \\
1\,0\,0\,0\,1\,1 &= s'' && \textit{new value of } s \\
1\,1\,0\,0\,1\,1 &= t &&
\end{aligned}
$$

- The remaining question is that how to achieve the above transformation.
- Recall that each time, the gray code changes 1 bit only. So, we have apply the "trick" of applying multiple controlled NOT gate, over and over to obtain the result.
- Recall that these controlled operations are nothing but permutation.
- Explicitly, consider the 2x2 unitary gates that we **want** to realize (s,t=s+1) , labelled it by $V_{st}$,

$$V_{st} = |0\rangle\langle 0| + |1\rangle\langle 1| + ... + \boxed{(\alpha|s\rangle + \beta|t\rangle)\langle s| + (\beta|s\rangle - \alpha|t\rangle)\langle t|} + ... \tag{152}$$

- Suppose we construct multiple controlled gates such that

$$U'|101001\rangle = |101011\rangle = |s'\rangle$$
$$U''|101011\rangle = |101011\rangle = |s''\rangle$$

- Then, what we need to do, is to consider the following:

$$U'V_{st}U' = |0\rangle\langle 0| + |1\rangle\langle 1| + .. |s\rangle\langle s|.. + (\alpha|s'\rangle + \beta|t\rangle)\langle s'| + (\beta|s'\rangle - \alpha|t\rangle)\langle t| + ... \tag{153}$$

- Similarly,

$$U''U'V_{st}U'U'' = |0\rangle\langle 0| + |1\rangle\langle 1| + .. |s'\rangle\langle s'|.. + (\alpha|s''\rangle + \beta|t\rangle)\langle s''| + (\beta|s''\rangle - \alpha|t\rangle)\langle t| \tag{154}$$

- This gate can be realized by the "trick" mentioned before, labelled it as $V$ (recycle the symbol), which implies that

$$V_{st} = U'U''VU''U' \tag{155}$$

- With this I can already construct any 2x2 unitary matrices using single-qubit gates together with CNOTs. This finishes the second part of the argument.
- Therefore, we can also finish the whole proof of the universal gate set.

## 15. Lecture 15

**Bounding the errors of quantum cirucits with imperfect gates**

- Quantum gates are never perfect. Is it just a mathemtical game?
- (If I remember correctly), at some point in the past, people thought about using continuous variables instead of bits to build computers. They found that continuous computing models can outperform discrete models of computing.
- At the end, when they took into account the finite precision, the continuous model no longer has the power anymore.
- In the following, we will be interested in looking at the imperfection of the unitary gates.

- For example, suppose we want to realize a quantum circuit of $m$ quantum gates, using the set of the universal gates discussed previously, labelled as $U \equiv U_m U_{m-1} ... U_1$.
- But for some unknown reason, we realized instead $V \equiv V_m V_{m-1} ... V_1$. Then, the question is, how bad is it?
- Would it happen that one imperfect gate ruins the whole computation?
- Let us first define or quantify what we mean by error.

$$E(U,V) \equiv max_{|\psi\rangle} \; || \; (U-V) \; |\psi\rangle \; || \tag{156}$$

where we consider the Euclidean norm

$$|| \; |a\rangle \; || \; \equiv \sqrt{\langle a|a\rangle} \tag{157}$$

- Next, we discuss the **physical meaning** of such norm.
- We shall show that

$$\boxed{|P_U - P_V| \; \leq 2 \; E(U,V)} \tag{158}$$

where $P_U \equiv \langle\psi|U^\dagger M U|\psi\rangle$, $P_V \equiv \langle\psi|V^\dagger M V|\psi\rangle$ are the probabilities of some measurement
outcome $M$. For example, for single qubit, $M = |0\rangle\langle 0|$.

**Proof**

- We define the **unnormalized** vector $|\Delta\rangle \equiv (U-V)\,|\psi\rangle$ for any state $|\psi\rangle$.
- Let us look at the left hand side,

$$|P_U - P_V| = |\langle\psi|U^\dagger M U|\psi\rangle - \langle\psi|V^\dagger M V|\psi\rangle| \tag{159}$$

where we can write $\langle\psi|U^\dagger M U|\psi\rangle = \langle\psi|U^\dagger M|\Delta\rangle + \langle\psi|U^\dagger M V|\psi\rangle$, and similarly, we can also write

$$-\langle\psi|V^\dagger M V|\psi\rangle = \langle\psi|\left(U^\dagger - V^\dagger\right) M V|\psi\rangle - \langle\psi|U^\dagger M V|\psi\rangle = \langle\Delta|MV|\psi\rangle - \langle\psi|U^\dagger M V|\psi\rangle.$$

- Therefore, we have

$$|P_U - P_V| = |\langle\psi|U^\dagger M|\Delta\rangle + \langle\Delta|MV|\psi\rangle| \leq |\langle\psi|U^\dagger M|\Delta\rangle| + |\langle\Delta|MV|\psi\rangle| \tag{160}$$

- Next, we need to impose further bounds for the right hand side. Note that

$$|\langle\psi|U^\dagger M|\Delta\rangle| \equiv |\langle\psi_U|M|\Delta\rangle| \leq || \; |\Delta\rangle| | \tag{161}$$

- To see this, let us consider a single qubit situation, i.e., $|\Delta\rangle = a|0\rangle + b|1\rangle$. Then, $M|\Delta\rangle = a|0\rangle$. Therefore, $|\langle\psi_U|M|\Delta\rangle| = |a| \; |\langle\psi_U|0\rangle| \leq |a|$, which is obviously smaller than $|| \; |\Delta\rangle|| = \sqrt{|a|^2 + |b|^2}$.
- Apply the similar agrument for the other terms, we can then prove the final result.

- Practically, we want to have a bound that takes into account of the individual gate errors.
- Furtunately, we can further bound the overall error by individual errors, i.e., (homework)

$$\boxed{E(U,V) = E(U_m U_{m-1} \ldots U_1, V_m V_{m-1} \ldots V_1) \leq \sum_{j=1}^{m} E(U_j, V_j)} \tag{162}$$

- If we want to just bound the error in the probability, e.g. we want to make sure that $|P_U - P_V| \leq \delta$ for some constant value $\delta$. Then, we have to make

$$E(U_j, V_j) \leq \delta / (2m) \tag{163}$$

as

$$|P_U - P_V| \leq 2\, E(U,V) \leq 2 \sum_{j=1}^{m} E(U_j, V_j) \leq 2 \sum_{j=1}^{m} \frac{\delta}{2m} = \delta \tag{164}$$

- Am important message here is that if we want to run a quantum circuit with a long sequence of gates, then it is necessary that we can good single (and two) qubit gate fidelity.

---

**Supplementary material: Canonical decomposition of two-qubit gates**

- In general, any two-qubit gate can be expressed in the so-called canonical form (Kraus and Cirac 2001):

$$\boxed{U = A_L \otimes B_L \, e^{i(\alpha_{xx} XX + \alpha_{yy} YY + \alpha_{zz} ZZ)} \, A_R \otimes B_R} \tag{165}$$

where $A_L, A_R$ and $B_L, B_R$ are local unitary gates, and the $\alpha's$ are real numbers.

- To justify this, we need to introduce a couple of facts .

**Fact 1**:

- For any unitary matrix $V$, one can always perform the following SVD-like decomposition (see e.g. Tucci quant-ph/0507171):

$$V = O_L D_\phi O_R \tag{166}$$

where $O_L$, $O_R$ are orthogonal matrices (real square matrix whose columns and rows are orthonormal vectors), and $D_\phi$ is a diagonal matrix with complex phases $e^{i\phi}$ as the elements.

**Fact 2**:

- For any given state vector with real elements in the computational basis, i.e., $|\psi_{real}\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$, and the following "entangler" matrix:

$$M \equiv \sum\nolimits_k |m_k\rangle\langle k| \tag{167}$$

where the $|m\rangle's$ are called the magic basis

$$|m_{00}\rangle \equiv \left(|00\rangle + |11\rangle\right)/\sqrt{2} \tag{168}$$

$$|m_{01}\rangle \equiv i\left(|01\rangle + |10\rangle\right)/\sqrt{2} \tag{169}$$

$$|m_{10}\rangle \equiv i\left(|01\rangle - |10\rangle\right)/\sqrt{2} \tag{170}$$

$$|m_{11}\rangle \equiv \left(|00\rangle - |11\rangle\right)/\sqrt{2} \tag{171}$$

one **always** obtain a maximally entangled state by applying $M$ to $|\psi_{real}\rangle$, i.e.,

$$M|\psi_{real}\rangle = a|m_{00}\rangle + b|m_{01}\rangle + c|m_{10}\rangle + d|m_{11}\rangle \tag{172}$$

- This fact comes directly from the well-known result that the concurrence $C \equiv \langle\psi|\sigma_y \otimes \sigma_y|\psi^*\rangle$ (an entanglement measure) of any pure two-qubit state, expressed by the magic basis, can be calculated directly by $C = |a^2 + b^2 + c^2 + d^2|$; whenever the coefficients are all real, it is equal to $C = 1$ as required by the normalization condition.
- This is true as

**Fact 3:**

- The magic states are eigenstates of the middle part (see e.g. Kraus and Cirac 2001),

$$MD_\phi M^\dagger = \sum_k e^{i\phi_k}|m_k\rangle\langle m_k| = e^{i(\beta_{xx}XX + \beta_{yy}YY + \beta_{zz}ZZ)} \tag{173}$$

**Fact 4**:

- For any pair of maximally-entangled state in the magic basis with real elements, i.e., $|o_1\rangle = \sum_{ij} g_{ij}|m_{ij}\rangle$ and $|o_2\rangle = \sum_{ij} h_{ij}|m_{ij}\rangle$, where $\sum_{ij} g_{ij}^2 = \sum_{ij} h_{ij}^2 = 1$ and $\sum_{ij} g_{ij}h_{ij} = 1.$
- Consider the following combined state,

$$|\psi_+\rangle \equiv \left(|o_1\rangle + i|o_2\rangle\right)/\sqrt{2} \equiv |A_+\rangle \otimes |B_+\rangle \tag{174}$$

which is normalized and must be a product state, as $\sum_{ij}(g_{ij} + ih_{ij})^2 = 0.$

- Similarly, we can also define $|\psi_-\rangle \equiv \left(|o_1\rangle - i|o_2\rangle\right)/\sqrt{2} \equiv |A_-\rangle \otimes |B_-\rangle$,
- Suppose the two states are orthogonal, i.e., $\langle o_2|o_1\rangle = 0$, then we have $\langle\psi_+|\psi_-\rangle = 0$ and hence $\langle A_+|A_-\rangle\langle B_+|B_-\rangle = 0$.
- In fact, we can see $\langle A_+|A_-\rangle = 0$ and $\langle B_+|B_-\rangle = 0$. It is because the following is the schmidt decomposition (the reduced density matrix of any maximally entangled state must be an identity matrix)

$$|o_1\rangle = \left(|A_+\rangle \otimes |B_+\rangle + |A_-\rangle \otimes |B_-\rangle\right)/\sqrt{2} \tag{175}$$

$$|o_2\rangle = -i\left(|A_+\rangle \otimes |B_+\rangle - |A_-\rangle \otimes |B_-\rangle\right)/\sqrt{2} \tag{176}$$

**Proof:**

- Now, we are ready to show how to arrive at the canonical form of $U$: let us first consider

$$V \equiv M^\dagger U M \tag{177}$$

- From fact 1, we can perform a SVD-like decompsition for $V = O_L D_\phi O_R$, and get

$$\boxed{U = M O_L M^\dagger \cdot M D_\phi M^\dagger \cdot M O_R M^\dagger} \tag{178}$$

- Next, by applying $M$ to $O_L$,we have an operator mapping product states to maximally entangled states, i.e., $MO_L = \sum_k |o_k\rangle\langle k|$.
- From fact 4, the first two states are spanned by the orthonormal pairs $\left\{|A_+\rangle|B_+\rangle, |A_-\rangle|B_-\rangle\right\}$. Since the maximally entangled pairs $|o_3\rangle$ and $|o_4\rangle$ are also orthonormal to others, they must be spanned by $\left\{|A_+\rangle|B_-\rangle, |A_-\rangle|B_+\rangle\right\}$. Therefore, we can write the other states as

$$|o_3\rangle = e^{i\lambda_3}\left(e^{i\delta}|A_+\rangle|B_-\rangle + e^{-i\delta}|A_-\rangle|B_+\rangle\right) \tag{179}$$

$$|o_4\rangle = e^{i\lambda_4}\left(e^{i\delta}|A_+\rangle|B_-\rangle - e^{-i\delta}|A_-\rangle|B_+\rangle\right) \tag{180}$$

- Finally, we have

$$MO_L = A \otimes B \sum_k |m_k\rangle\langle k| e^{i\phi_k} \tag{181}$$

where $A = |0\rangle\langle A_+| + |1\rangle\langle A_-|e^{i\delta}$ and $B = |0\rangle\langle B_+| + |1\rangle\langle B_-|e^{-i\delta}$.

- After perform the same procedure for the right-hand side, we obtain what we want.

---

## 16. Lecture 16

**Univeral quantum gates (approximate)**

- Previously, we rely on the use of the set of single-qubit gates. For some technical reasons related to fault-tolerant quantum computation, we may want **to replace the set of single qubit gates with just Hadamard and T gates**:

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad , \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \tag{182}$$

- But the method presented below can be generalized to other pairs of gates.
- As a result, we may say the universal set of gates are just $H$, $T$ and CNOT for **approximately** simulating any given quantum circuit.
- In the following, we will basically just to figure out how to approximate single qubit rotation with the $H$ and $T$ gates.
- Recall that, in general, any single qubit gate can be decomposed by a pair of rotational gates like (generalizing the Z-Y-Z decomposition):

$$U = e^{i\alpha} R_n(\beta_1) R_m(\gamma_1) R_n(\beta_2) R_m(\gamma_2) ... \tag{183}$$

- The total number (finite and depending on the angle between the unit vectors $n$ and $m$) of this expression is left as a challenge for the student.
- Recall that any 2x2 matrix can be expanded by the Pauli basis: $A = a_0 I + a_x X + a_y Y + a_z Z$.
- We now **try** the following (check details in my lecture notes):

$$R_n(\theta_{unit}) \propto THTH = \cos\frac{\theta_{unit}}{2} I - i(n \cdot \sigma)\sin\frac{\theta_{unit}}{2} \tag{184}$$

  which fixes both $n = \left(\cos\frac{\pi}{8}, \sin\frac{\pi}{8}, \cos\frac{\pi}{8}\right)$ and $\theta_{unit}$.
- Now, we have to worry about creating the continuous values for $R_n$.
- Of course, if we consider $(THTH)^k$, then we have $R_n(k\theta_{unit})$.
- So, the question is, can we approximate any continuous value using discrete set of $R_n(\theta_{unit})$? i.e., find some $k$ such that the following holds

$$R_n(\beta) \sim R_n(k\theta_{unit}) = R_n(\theta_{unit})^k = (THTH)^k \tag{185}$$

- The point is that it is necessary to make sure (here assumed) that $\theta_{unit}$ is an irrational multiple of $2\pi$, i.e., $\theta_{unit}/2\pi \equiv \alpha$ is a irrational number.
- More precisely, for any $\epsilon > 0$ (input), there exists an integer $k$, such that

$$E(R_n(\beta), R_n(k\theta_{unit})) < \epsilon \tag{186}$$

**Denseness on the circle**

- (Fact 1): **If** $\alpha \in [0, 1)$ is a irrational number, then the points $n\alpha \ (mod\ 1)$ are all distinct.
  - Proof: let us consider the case where two points become the same, i.e., $n\alpha - m\alpha = k$, where $k$ is an integer. However, then we have $\alpha = k/(n-m)$, which is a rational number. This can lead to a contradtion, if we assumed $\alpha$ to be irrational.
- (Fact 2): The space centered around each point with width $\epsilon$ is called open interval. For a total of $N$ points, i.e., $n\alpha \ (mod\ 1)$ for $n = 1, 2, 3, ..., N$. Then, if $N\epsilon > 1$, then at least two of these intervals must intersect.
  - Proof: Think about it, for $N$ points, if there is no intersection at all, then we should have $N\epsilon \leq 1$.

- Now, we know that we can create the situation where two points can get close to $\epsilon$ on a ring with a unit perimeter.
- In other words, the positive number $r \equiv |n-m|$ satisfies

$$r\,\alpha\ (mod\ 1)\ <\ \epsilon \tag{187}$$

- Note the multiple of $r\,\alpha\ (mod\ 1)$ forms a equally spaced points on the unit circle with interval less than $\epsilon$.
- Return to the case of rotation operators (**homework**),

$$E(R_n(\theta_1), R_n(\theta_2)) \leq |e^{i\theta_1} - e^{i\theta_2}| \leq |\theta_1 - \theta_2| \tag{188}$$

- Similarly, we define the $R_m(\gamma)$ by

$$R_m(\gamma) \equiv HR_n(\gamma)H = HTHT \tag{189}$$

- Consequently, we can make the following claim:

$$E(U, R_n(k_1\theta_{unit})R_m(k_2\theta_{unit})R_n(k_3\theta_{unit})R_m(k_4\theta_{unit})\dots) < c\epsilon \tag{190}$$

where $c$ is the number of rotational terms.

- Now, what about the $e^{i\alpha}$ phase factor in the decomposition?

**Big-O notation (from wikipedia):**

The notation can also be used to describe the behavior of *f* near some real number *a* (often, *a* = 0): we say

$$f(x) = O(g(x)) \text{ as } x \to a$$

if and only if there exist positive numbers *δ* and *M* such that

$$|f(x)| \leq\ Mg(x) \text{ when } 0 < |x-a| < \delta.$$

**Product** [ edit ]

$$f_1 = O(g_1) \text{ and } f_2 = O(g_2) \Rightarrow f_1 f_2 = O(g_1 g_2)$$
$$f \cdot O(g) = O(fg)$$

**Sum** [ edit ]

$$f_1 = O(g_1) \text{ and } f_2 = O(g_2) \Rightarrow f_1 + f_2 = O(\max(g_1, g_2))$$

This implies $f_1 = O(g) \text{ and } f_2 = O(g) \Rightarrow f_1 + f_2 \in O(g)$, which means that $O(g)$ is a convex cone.

**Multiplication by a constant** [ edit ]

Let *k* be constant. Then:

$O(|k|g) = O(g)$ if *k* is nonzero.

$f = O(g) \Rightarrow kf = O(g).$

Also see:
https://www.freecodecamp.org/news/big-o-notation-why-it-matters-and-why-it-doesnt-

1674cfa8a23c/

# 17. Lecture 17

### Solovay-Kitaev algorithm

- Now, we have learned that using things like THTH sequences, we can approximate any single qubit gates to any accuracy.
- But, it seems that the performance would be quite bad.
- The Solovay-Kitaev algorithm is designed to speedup the process of approimating gates.
- The applicability of the SK algorithm is broader than we have discussed.
- Therefore, we have to learn the concept of $\epsilon$-net.
- What is $\epsilon$-net $V_\epsilon$? We already have one (in the last equation above).
- In simple language, given an arbitrary unitary gate $U$, if we say we have an epsilon-net, then we should be able to approximate $U$ by some gate $V \in V_\epsilon$ with an accuracy $\epsilon$, i.e.,

$$\boxed{|| U - V || \equiv \text{tr}\sqrt{(U-V)^\dagger (U-V)} \leq \epsilon} \tag{191}$$

which is equivalent to $|| U V^\dagger - I || \leq \epsilon$.

- **Additionally, the inverse should exist, i.e., if we have $V$, then we must have $V^\dagger$.**
- Let us look at the (trace) norm closedly for an arbitrary matrix : $||A|| = \text{tr}\sqrt{A^\dagger A}$
  - If $A$ is hermitian, i.e., $A^\dagger = A$, then $||A|| = |a_1| + |a_2| + |a_3| + ...$ is the sum of the absolute values of the eigenvalues.
  - For any unitary matrix $U$, $||U\,A|| = \text{tr}\sqrt{A^\dagger U^\dagger U A} = ||A||$.
- See also:

## Schatten norms [ edit ]

*Further information: Schatten norm*

The Schatten *p*-norms arise when applying the *p*-norm to the vector of singular values of a matrix. If the singular values of the $m \times n$ matrix $A$ are denoted by $\sigma_p$ then the Schatten *p*-norm is defined by

$$\|A\|_p = \left( \sum_{i=1}^{\min\{m,n\}} \sigma_i^p(A) \right)^{\frac{1}{p}}.$$

These norms again share the notation with the induced and entrywise *p*-norms, but they are different.

All Schatten norms are submultiplicative. They are also unitarily invariant, which means that $\|A\| = \|UAV\|$ for all matrices $A$ and all unitary matrices $U$ and $V$.

The most familiar cases are $p = 1, 2, \infty$. The case $p = 2$ yields the Frobenius norm, introduced before. The case $p = \infty$ yields the spectral norm, which is the operator norm induced by the vector 2-norm (see above). Finally, $p = 1$ yields the **nuclear norm** (also known as the *trace norm*, or the Ky Fan 'n'-norm[5]), defined as

$$\|A\|_* = \text{trace}\left(\sqrt{A^*A}\right) = \sum_{i=1}^{\min\{m,n\}} \sigma_i(A),$$

where $\sqrt{A^*A}$ denotes a positive semidefinite matrix $B$ such that $BB = A^*A$. More precisely, since $A^*A$ is a positive semidefinite matrix, its square root is well-defined. The nuclear norm $\|A\|_*$ is a convex envelope of the rank function $\text{rank}(A)$, so it is often used in mathematical optimization to search for low rank matrices.

- What we want to do, is to construct from the epsilon-net, a more accurate

approximation $\epsilon^{3/2} = \epsilon \cdot \epsilon^{1/2} < \epsilon$.

$$\boxed{|| \, U \, V^{\dagger} - W \, || = || \, U \, - WV \, || = O\left(\epsilon^{3/2}\right)} \tag{192}$$

where $W$ is a **product** of gates from the epsilon net.

- This is a kind of surprising result, as normally, we would think that the approximation should be $O(\epsilon)$.
  - If each step costs $O(\epsilon)$, n steps would be $O(n\epsilon)$.
  - Remember, here the error is **systematic**, which means that it is fixed.

**Proof:**

- First, let us define (you can always do it like that)

$$U \, V^{\dagger} \equiv e^{iA} \tag{193}$$

where we $||A|| = O(\epsilon)$, as $|| \, e^{iA} - I \, || = || \, iA|| = O(\epsilon)$. Equivalently, we may just write $A = O(\epsilon)$.

- Next, we want to find a product of gates, labelled by $W$, to minimize the norm: $|| \, U \, V^{\dagger} - W \, || = || \, e^{iA} - W \, ||$.
- To proceed, there are is an interesting relation: for any given $A$, one can always find a pair (they may not be unique) of matrices $B$ and $C$ such that

$$[B, C] = \, - iA \tag{194}$$

where $||B|| = ||C|| = O\left(\epsilon^{1/2}\right)$.

- For qubits, you may imagine we can rotate the reference frame, such that $A = O(\epsilon)Z$. Then, natrually, we may choose $B = O\left(\epsilon^{1/2}\right)X$ and $C = O\left(\epsilon^{1/2}\right)Y$.
- Remarks: although $A$ is kind of small, its mathematical expression is known exactly. Therefore, $B$ and $C$ are known as well.
- Now, let us try the following:

$$\boxed{W \equiv e^{iB'} e^{iC'} e^{-iB'} e^{-iC'} \, = I - [B', C'] + O\left(\epsilon^{3/2}\right)} \tag{195}$$

where the gate $e^{iB'}$ is an approximation of $e^{iB}$, and similarly $e^{iC'}$ is an approximation of $e^{iC}$ using the epsilon net.

- It means that $||e^{iB'} - e^{iB} \, || = O(\epsilon)$, $||e^{iC'} - e^{iC} \, || = O(\epsilon)$, which means $||B' - B \, || = O(\epsilon)$ and $||C' - C \, || = O(\epsilon)$.
- Let us do the **checking:**

$$e^{iB'} e^{iC'} e^{-iB'} e^{-iC'} \, = (I + iB' + O(\epsilon))(I + iC' + O(\epsilon))(I - iB' + O(\epsilon))(I - iC' + O(\epsilon)) \tag{196}$$

- If you just look at the first order terms, they should get cancelled.
- Why the second contains only $[B', C']$? (**homework**)
- Now, if we look closer to the commutator, we have

$$[B', C'] = [B + O(\epsilon), C + O(\epsilon)] = [B, C] + O\left(\epsilon^{3/2}\right) \tag{197}$$

- This means that you can just replce $[B', C']$ by $[B, C] = -iA$.
- In other words,

$$W = I + iA + O\left(\epsilon^{3/2}\right) = e^{iA} + O\left(\epsilon^{3/2}\right) \tag{198}$$

which completes the proof.

## 18. Lecture 18

- The point is that now from the $\epsilon$-net, we effectively are able to construct a $\epsilon^{3/2}$-net, at the cost of applying 5 total gates, namely $WV = e^{iB'}e^{iC'}e^{-iB'}e^{-iC'}V$, from the $\epsilon$-net.
- Now, if we replace the whole argument again, then we get a better net (if we don't care about the additional gates): ~$\left(\epsilon^{3/2}\right)^{3/2} = \epsilon^{(3/2)^2}$.
- More rigourously, let us quantify the new error by (upper bound)

$$\epsilon_1 = c\,\epsilon_0^{3/2} \quad or \quad c^2\epsilon_1 = \left(c^2\epsilon_0\right)^{3/2} \tag{199}$$

where $\epsilon_0$ is the original value in the epsilon-net, and $c$ is some constant.

- Moreover, if it takes (maximally) $L_0$ elementary gates for realizing $V$ in the $\epsilon$-net, then it takes $L_1 = 5L_0$ gates $\left\{e^{iB'}, e^{iC'}, e^{-iB'}, e^{-iC'}, V\right\}$ to construct the $\epsilon^{3/2}$-net.
- If we take one step further, then we have

$$\epsilon_2 = c\,\epsilon_1^{3/2} \quad or \quad c^2\epsilon_2 = \left(c^2\epsilon_1\right)^{3/2} = \left(c^2\epsilon_0\right)^{(3/2)^2} \tag{200}$$

- The next time (level 2), you would need to apply the set from the $\epsilon^{3/2}$-net to approximate the set of gates like $\left\{e^{iB''}, e^{iC''}, e^{-iB''}, e^{-iC''}, WV\right\}$. That would be 5($5L_0$) gates.
- In general, at level $k$, we have

$$c^2\epsilon_k = \left(c^2\epsilon_{k-1}\right)^{3/2} = \left(c^2\epsilon_0\right)^{(3/2)^k}. \tag{201}$$

and we need a total of $L_k = 5L_{k-1} = 5^k L_0$.

- Now, we have the situation that if we are willing to pay the cost of applying exponential number of gates (in terms of k), then we get a double-exponential effect in shrinking the errors.
- So, there may some advantage that we can gain (**homework**).

$$\frac{L_k}{L_0} = 5^k = \left(\frac{\log\left(1/c^2\epsilon_k\right)}{\log\left(1/c^2\epsilon_0\right)}\right)^{\log 5/\log(3/2)} . \tag{202}$$

- If we forget about the constants, then we have the more standard form:

$$L_k = O\left(\log(1/\epsilon_k)^{3.97}\right) \tag{203}$$

- This expression tells us that if we want to achieve an approximation of size $\epsilon_k$, then you expect that each gate would have an extra cost of $O\left(\log(1/\epsilon_k)^{3.97}\right)$, which would just impose an polynomial overhead.
- For example, in a certain quantum algorithm of $n$ qubits, we have a poly(n) circuit. To improve each gate to $\epsilon_k$, we still have a polynomial circuit.

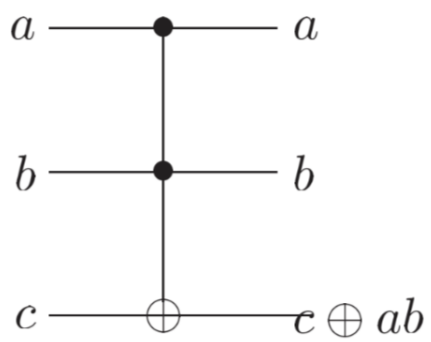
## 19. Lecture 19

- So far, our discussion is quite general; a framework that is applicable to most quantum algorithms.
- Now, it is time for us to look into details of some "real" quantum algorithms.
- But these algorithms may not be "practical", in the sense that we are just trying to demonstrate the idea of quantum computation.
- The first thing we need to sort out is how quantum computer can perform classical computation.
- Here, the important point is that the Toffoli gate is "universal" for classical computer.
- Imagine we have a classical circuit, with AND, OR, NOT etc. logic gates, the idea is to replace each logical gate with a Toffoli gate (together with some ancilla bits).
- There are two issues to figure out.
  - First, classical computation is irreversible (loss of information), but quantum computation is reversible (information is conserved).
  - Second, some classical computation invovles randomness.
- Both problems can be overcome with quantum computation, by keeping information and use quantum measurement e.g. on $|+\rangle$ state.

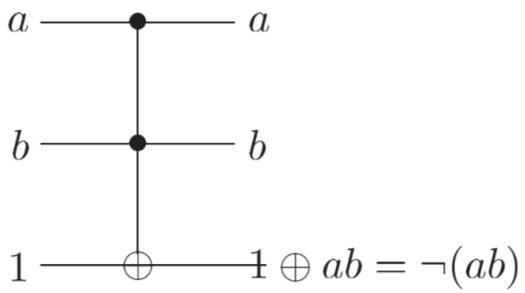
**Toffoli gate**

- Let us take a good at the logic table of the toffoli gate.

| Inputs | | | Outputs | | |
|---|---|---|---|---|---|
| $a$ | $b$ | $c$ | $a'$ | $b'$ | $c'$ |
| 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | 1 |
| 0 | 1 | 0 | 0 | 1 | 0 |
| 0 | 1 | 1 | 0 | 1 | 1 |
| 1 | 0 | 0 | 1 | 0 | 0 |
| 1 | 0 | 1 | 1 | 0 | 1 |
| 1 | 1 | 0 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 0 |




- [**Fact**] NAND gate (AND gate followed by the NOT gate) is universal for classical computation.
- It is sufficient to show that if Toffoli gate can simulate NAND gate, then Toffoli gate is also unviersal for classical computation.




- For checking, let us look at the NAND logic table:

| input a | input b | output |
|---|---|---|
| 0 | 0 | 1 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

- As Toffoli gate can be constructed using elementary quantum gates, classical computation can be archieved with a quantum computer.

- The caveat is that there will be additional overheads for quantum computers to perform classical computation.
- I recently heard from someone made a good analogy: it is true that quantum computer can simulate classical computation, but quantum computers are not likely to replace classical computers. Why? It is just like a plane can all be operated like a car, but we don't replace cars by planes.

**Quantum parallelism**

- When I was a graduate student, I was the only working on quantum computation in my group. In one group meeting, I was trying to explain quantum computation, then I wrote this ($x \in \{0,1\}^n$):

$$\boxed{|x\rangle|0\rangle \rightarrow |x\rangle|f(x)\rangle} \tag{204}$$

other people in the group is happy with it, but when I wrote it in a superposition form:

$$\boxed{\sum_x |x\rangle|0\rangle \rightarrow \sum_x |x\rangle|f(x)\rangle} \tag{205}$$

they didn't believe it was true.

- May be I should comment: can quantum computer do something like this: $|x\rangle \rightarrow |f(x)\rangle$? but this is not reversible in general.
- In reality, what we need, is also a bunch for ancilla qubits. Therefore, we actually need to perform something like:

$$|x\rangle|0\rangle|0\rangle_{anc} \rightarrow |x\rangle|f(x)\rangle|g(x)\rangle_{anc} \rightarrow |x\rangle|f(x)\rangle \tag{206}$$

which is still a simplified version. May be the best way is really go through some explicit example.

- For example, if the logical operation to perform is exactly the NAND gate.
- Single state:

$$|ab\rangle|1\rangle \ \rightarrow \ |ab\rangle|1+ab\rangle \tag{207}$$

- Superposition:

$$(|00\rangle + |01\rangle + |10\rangle + |11\rangle)|1\rangle \ \rightarrow \ |00\rangle|1\rangle + |01\rangle|1\rangle + |10\rangle|1\rangle + |11\rangle|0\rangle \tag{208}$$

- Now, imagine, if we need to further process the output qubit.

$$|ab\rangle|1\rangle \ \rightarrow \ |ab\rangle|1+ab\rangle|1\rangle \rightarrow \ |ab\rangle|1+ab\rangle|1+a(1+ab)\rangle \tag{209}$$

which means that the middle part would case problem for superposition:

$$(|00\rangle|1\rangle + |01\rangle|1\rangle + |10\rangle|1\rangle + |11\rangle|0\rangle)|1\rangle \tag{210}$$

$$\rightarrow |00\rangle|1\rangle|1\rangle + |01\rangle|1\rangle|1\rangle + |10\rangle|1\rangle|0\rangle + |11\rangle|0\rangle|1\rangle \tag{211}$$

$$|00\rangle|g(00)\rangle_{anc}|f(00)\rangle + |01\rangle|g(01)\rangle_{anc}|f(01)\rangle + |10\rangle|g(10)\rangle_{anc}|f(10)\rangle + |11\rangle|g(11)\rangle_{anc}|f(11)\rangle$$

- Here the function f is the application of two NAND gates.
- The point is that the ancilla qubits would be entangled with the input qubits.
- To get rid of it, we may just apply the Toffoli gates again.

**Quantum superposition**

- With a single qubit, we know that

$$H|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2} \tag{212}$$

- With two qubits:

$$H^{\otimes 2}|00\rangle = (|0\rangle + |1\rangle)(|0\rangle + |1\rangle)/2 = (|00\rangle + |01\rangle + |10\rangle + |11\rangle)/2 \tag{213}$$

- With multiple qubits,

$$H^{\otimes n}|0^n\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} |x\rangle \tag{214}$$

- Using $n$ Hadmard gates on $n$ qubits, one can create a quantum superposition with $2^n$ input terms.
- The exponential Hilbert space is a necessary condition (but not sufficient) for making it a hard problem for classical simulation.
- Combining it with the previous discussion, we can then claim that there exists a unitary circuit $U_f$, such that

$$\boxed{\sum_x a_x U_f|x\rangle|0\rangle = \sum_x a_x|x\rangle|f(x)\rangle} \tag{215}$$

where $f(x)$ is some function that can be represented as a classical cirucit.

## 20. Lecture 20

**Deutsch's algorithm (an oracle-based algorithm)**

- We should first understand the concept of an oracle, i.e., something regarded as a black box, for example somebody designed a classical circuit without explaining to you what it is, and you are not allowed to look into it.
- For practical purpose, we just mean that for the classical computation part,

$$\boxed{|x\rangle|0\rangle \to |x\rangle|f(x)\rangle} \tag{216}$$

we don't care how $f(x)$ is constructed.

- In addition, we can also perform the following operation (for Boolean function)

$$\boxed{|x\rangle \to (-1)^{f(x)}|x\rangle} \tag{217}$$

- We can achieve (sometimes called phase kick-back) it in the following way (adding an

extra ancilla qubit in a superposition state):

$$U_{CNOT}|f(x)\rangle(|0\rangle - |1\rangle) = |f(x)\rangle(|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle) = (-1)^{f(x)}|f(x)\rangle(|0\rangle - |1\rangle) \quad (218)$$

- In other words, we have to apply the sequence (ignoring some ancilla qubits):

$$|x\rangle \rightarrow |x\rangle|f(x)\rangle \rightarrow (-1)^{f(x)}|x\rangle|f(x)\rangle \rightarrow (-1)^{f(x)}|x\rangle \quad (219)$$

- Deutsch's algorithm solves the following problem:
  - Given an oracle Boolean function $f(x)$, which takes one bit as input and one bit as output.
  - There are two possibilities: either balance or constant,
  - Case 1 (constant): $f(0) = f(1)$
    - two possibilites, $f(0) = f(1) = 0$ or $f(0) = f(1) = 1$
  - Case 2 (balance): $f(0) \neq f(1)$
    - two possibilities: $\{f(0) = 1,\ f(1) = 0\}$, or $\{f(0) = 0,\ f(1) = 1\}$
  - Classically, one must evaluate the function twice.
    - It you just call the oracle once, you can only know either $f(0)$ or $f(1)$, but not both.
- Deutsch's algorithm allows us to solve the problem (determining it either case 1 or 2), by calling the oracle one time only.

**Step 1:**

- Let us first create a superposition of the input, i.e.,

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad (220)$$

**Step 2:**

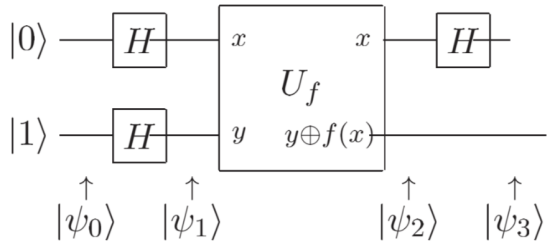
- We apply the oracle:

$$\frac{1}{\sqrt{2}}\left((-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle\right) \quad (221)$$

- Note that if it is case 1, then this is just $|+\rangle$ (up to a sign); if it is case 2, then we have $|-\rangle$ (up to a sign).

**Step 3:**

- Finally, we just need to distinguish between the $|+\rangle$ or $|-\rangle$ state, which can be acheived by applying the Hadamard gate: $H|+\rangle = |0\rangle$ and $H|-\rangle = |1\rangle$.




**Deutsch-Jozsa aglorithm**

- Later, Jozsa worked with Deutsch to extend the original algorithm to the multiple qubit cases.
- The problem DJ algorithm solving is as follows:
  - Given an oracle of n-bit Boolean function, $f(x) \rightarrow \{0,1\}$, for any $x \in \{0,1\}^n$.
  - (**promised problem**) Determine the two cases: constant or balanced
    - **constant:** $f(x) = 0$ or $f(x) = 1$ for all x's
    - **balanced:** number($f(x) = 0$) = number($f(x) = 1$)
- This is now a more artifical problem.
- Example: for two-bit cases:
  - **constant:** $f(00) = f(01) = f(10) = f(11) = 0$
  - **balanced:** $f(00) = f(01) = 0$ and $f(10) = f(11) = 1$
- Let us consider classical way to solve the problem.
- From the point of view of computer science, we often are interested in performance of algorithms in the **worst-case scenarios; if you try less than half of the $2^n$ combinations, you still cannot be sure that the function is constant of balanced.**
- Therefore, for this problem, classically the worst-case scenarios requires us to try more than $2^{n-1}$ combinations.
- Let us now consider the quantum algorithm:

**Step 1:**

- We also create a quantum superposition:

$$H^{\otimes n}|0^n\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle \tag{222}$$

**Step 2:**

- We apply the same trick of encoding the result of the Boolean function as a sign, i.e.,

$$\frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} U_f|x\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} (-1)^{f(x)}|x\rangle \tag{223}$$

- If it is constant, then we have the original state in the first step.

- If it is balanced, then we something more complicated, but we can make it simple.

**Step 3:**

- Recall that $H|0\rangle = \left(|0\rangle + |1\rangle\right)/\sqrt{2}$ and $H|1\rangle = \left(|0\rangle - |1\rangle\right)/\sqrt{2}$, the minus sign can only be next to the $|1\rangle$ state.

$$\frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} (-1)^{f(x)} H^{\otimes n}|x\rangle = \frac{1}{2^n}\left(\sum_{x\in\{0,1\}^n} (-1)^{f(x)}\right)|000...0\rangle + ... \tag{224}$$

where the part not showing here contains at least one '1' in the basis vector.

- For example: $H^{\otimes 2}|01\rangle = \left(|0\rangle + |1\rangle\right)\left(|0\rangle - |1\rangle\right)/2 = (1/2)|00\rangle + ...$
- Now, we can see quite clearly, that the amplitude for constant is 1, but the amplitude for balanced is 0.
- In conclusion, we can solve the problem by applying the oracle once, and check the final output to see if it is in the $|000...0\rangle$.

## 21. Lecture 21

**Quantum Simulation**

- When I was a student, the term "quantum simulation" means using classical computers to solve quantum problems, e.g. dynamics or ground state etc.
- I think now, the term now means solving quantum problems with quantum devices.
- What we would be interested, is to study the problem of simulating quantum unitary operation for some "local" Hamiltonains.
- Let us recall the Schrodinger's equation ($\hbar = 1$):

$$i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle \tag{225}$$

where $H$ is assumed to be a time independent Hamiltonian.

- All you need to know, is that one can re-write the above equation as follows:

$$|\psi(t)\rangle = e^{-iHt}\,|\psi(t=0)\rangle \tag{226}$$

- Our goal of quantum simulation is to prepare the time-dependent state $|\psi(t)\rangle$ using a set of (elementary) quantum gates (assuming that the initial state is easy to prepare, or given).

**Defining the errors**

- Given an operator (matrix) $X$, consider another operator $Y \equiv X + \Delta_X$. We write

$$X + O(\epsilon) \quad \leftrightarrow \quad ||\,\Delta_X\,|| = O(\epsilon) \tag{227}$$

- Here the matrix norm is not specified.

- Recall the definition of the exponential operator,

$$e^{-iHt} = I - iHt + \sum_{m=2}^{\infty} \frac{(-iHt)^m}{m!} . \tag{228}$$

- Recall that operator norms satisfy the following triangular inequality:

$$|| X + Y || \leq || X || + || Y || \tag{229}$$

- Therefore, we can bound the operator:

$$|| \sum_{m=2}^{\infty} \frac{(-iHt)^m}{m!} || \ \leq (||H||\ t)^2 \sum_{m=0}^{\infty} \frac{(||H||\ t)^m}{(m+2)!} \tag{230}$$

- Suppose we restrict the evolution time, such that

$$||H||\ t \ < 1 \tag{231}$$

- The last term can be bounded by,

$$\sum_{m=0}^{\infty} \frac{(||H||\ t)^m}{(m+2)!} \leq \sum_{m=0}^{\infty} \frac{1}{(m+2)!} = e - 1 - 1 < 1 \tag{232}$$

- Because we can consider $e^x = 1 + x + \sum_{m=2}^{\infty} \frac{x^m}{m!}$, and hence $e = 1 + 1 + \sum_{m=0}^{\infty} \frac{1}{(m+2)!}$.
- As a result, we can write

$$\boxed{e^{-iHt} = I - iHt + O\left(||H||^2 t^2\right)} \tag{233}$$

**k-local Hamiltonian**

- Let us consider a system of $n$ spin-1/2 particles (qubits), the total Hamiltonian is a sum of the interactions among the spins, i.e.,

$$H = \sum_{j=1}^{L} H_j \tag{234}$$

where $H_j$ acts only on at most $k < n$ spins.

- The important point is that you can change $n$, but keeping $k$ fixed.
- Also, here we consider more the mathematics, so physically locality is not imposed.
- So, $k$-local does not mean local in space.
- A technical issue to address is that we need to make sure that $L$ cannot be exponential.
- For example, Hamiltonians with terms containing exactly $k$-body intereactions, then

$$L \leq C_k^n = O\left(n^k\right) = poly(n) \tag{235}$$

**Commuting Case**

- As a special case, consider the situation where all of terms commute with each other, i.e., $[H_j, H_i] = 0$.
- This can happen, for example, for the Ising model: $H = \sum J_{ij} Z_i Z_j$.
- In this case, we can separate each term like this:

$$e^{-iHt} = e^{-iH_1 t} e^{-iH_2 t} \cdots e^{-iH_L t} \tag{236}$$

- To ensure the final error is bounded by $\epsilon$, one can arhieve this by requiring each oeprator to have an error bounded by $\epsilon / L$.
- If we consider the simulation of each elementary gate, using $H$ and $T$ gates, then we need to have $O\left(\log^c(L/\epsilon)\right)$ from the Solavay-Kitaev theorem.
- Because we have $L$ such terms, we need to consider $O\left(L \log^c(L/\epsilon)\right)$ gates.

**Non-Commuting Case**

- In the cases where $[H_j, H_i] \neq 0$, then our expansion no longer works, i.e.,

$$e^{-iHt} \neq e^{-iH_1 t} e^{-iH_2 t} \cdots e^{-iH_L t} \tag{237}$$

- Of course, when $t$ is small (let us replace it by $\Delta t$), we can approximately put

$$e^{-iH\Delta t} \approx e^{-iH_1 \Delta t} e^{-iH_2 \Delta t} \cdots e^{-iH_L \Delta t} \tag{238}$$

- The point is that we would like to make this approximation as rigorous as possible.
- To get started, we do have an equality like:

$$e^{-iHt} = \lim_{n \to \infty} \left(e^{-iH_1 t/n} e^{-iH_2 t/n} \cdots e^{-iH_L t/n}\right)^n \tag{239}$$

- But this equality is not so practical, as it talks about $n \to \infty$.
- We are interested in a finite number of terms.

---

**Theorem 3 (Lie-Trotter product formula)**

For a pair of Hermitian operatiors $A$ and $B$, where $||A|| < s < 1$ and $||B|| < s < 1$, we have

$$\boxed{e^{-iA} e^{-iB} = e^{-i(A+B)} + O\left(s^2\right)} \tag{240}$$

Proof.

- First of all, $e^{-iA} = I - iA + O\left(s^2\right)$ and $e^{-iB} = I - iB + O\left(s^2\right)$, now we can expand them both:

$$e^{-iA} e^{-iB} = \left(I - iA + O\left(s^2\right)\right)\left(I - iB + O\left(s^2\right)\right) = I - i(A+B) + O\left(s^2\right) \tag{241}$$

- On the other hand, we might also need to assume: $||A + B|| < ||A|| + ||B|| < 2s < 1$

$$e^{-i(A+B)} = I - i(A+B) + O\left(||(A+B)^2||\right) \tag{242}$$

where $O\left(||(A+B)^2||\right) = O\left(s^2\right)$.

---

- Now, to apply the Lie-Trotter product formula for the simulation problem, we need to bound the following (assuming for each term, we have $|| H_i || \Delta t < s < 1/L$ ):

$$|| H_1 \Delta t + H_2 \Delta t + ... + H_L \Delta t || < L\, s \ < \ 1 \tag{243}$$

- The strategy we will use is to combine the terms one by one, using the fact that $|| A || \ = || A\, U ||$.
- Let us now consider the product

$$e^{-iH_1\Delta t} e^{-iH_2\Delta t} \cdots e^{-iH_L\Delta t} = \left(e^{-i(H_1+H_2)\Delta t} + O\left(s^2\right)\right) e^{-iH_3\Delta t} ... e^{-iH_L\Delta t} \tag{244}$$

- Then, we have

$$e^{-i(H_1+H_2)\Delta t} e^{-iH_3\Delta t} ... e^{-iH_L\Delta t} + O\left(s^2\right) \tag{245}$$

where we used $e^{-i(H_1+H_2)\Delta t} = I - i(H_1+H_2)\Delta t + O\left(||(H_1+H_2)^2||\, \Delta t^2\right)$.

- Next, we just repeat the process, i.e.,

$$e^{-i(H_1+H_2)\Delta t} e^{-iH_3\Delta t} = e^{-i(H_1+H_2+H_3)\Delta t} + O\left(2^2 s^2\right) + O\left(s^2\right) \tag{246}$$

- The resulting expression looks like this:

$$\boxed{e^{-iH_1\Delta t} e^{-iH_2\Delta t} \cdots e^{-iH_L\Delta t} = e^{-i(H_1+H_2+...H_L)\Delta t} + O\left(L^3 s^2\right)} \tag{247}$$

- Note that $O\left(L^3 s^2\right) = O\left(s^2\right) + O\left(2^2 s^2\right) + O\left(3^2 s^2\right) + ... O\left((L-1)^2 s^2\right)$ , as

$$O\left(||(H_1 + H_2 + ... H_k)^2||\, \Delta t^2\right) = O\left(k^2 s^2\right) \tag{248}$$

- Finally, to make $\Delta t = t/N$ small, we can make a large $N$,

$$e^{-iHt} = \left(e^{-i(H_1+H_2+...H_L)t/N}\right)^N \tag{249}$$

- The question is, how large $N$ should be? It depends on the requirement of the final error.
- If we want a total error to be bounded by $O(\epsilon)$, then each term, $e^{-i(H_1+H_2+...H_L)t/N}$ is requried to be $O(\epsilon/N)$.
- It means that

$$O\left(L^3 s^2\right) = O(\epsilon/N) \tag{250}$$

- Let us asume all $||H_i|| < H_{\max}$ , then $|| H_i || \Delta t < H_{\max}\, t/N$.
- Therefore, we have $L^3 H_{\max}^2\, t^2/N^2 = \epsilon/N$, which means that we need to take

$$N = O\left(L^3 H_{\max}^2 t^2 / \epsilon\right) \tag{251}$$

- The cost of simulating each term like $e^{-iH_k \Delta t}$ would be similar to the commuting case.