Introduction to quantum computation: Quantum Gates and circuits

Lecture Notes for Quantum Computing

Lecturer: Dorit Aharonov
Scribe: Ahud Mordechai and Oren Mizrahi

March 11-16, 2001

Summary of the second week of semester 2 (11.3.2001 - 16.3.2001)

Abstract

This week we learn about elementary quantum operations, which we call quantum gates. These must be unitary; We connect the unitarity restriction to Schrodinger's equation. We then define tensor product and quantum systems of n qubits, and define quantum circuits. We prove that the function computed by a circuit has error probability ϵ if each gate is approximated to within ϵ over the number of gates in the circuit. This shows that the sensitivity of a quantum circuit to errors is additive. We discuss universal quantum gates, and state the Solovay Kitaev theorem which shows that universal sets of gates approximate any gate on a constant number of qubits exponentially fast. We use this theorem to show that quantum complexity does not depend on the universal set of gates which is used, up to polylogarithmic factors.

1 Quantum Operations

We have seen how to describe the state of a qubit or two, and how to observe it. In order to compute quantumly, we need also to be able to manipulate the state, or to perform quantum operations. What kind of quantum operators are allowed? Quantum mechanics tells us that all quantum states can only go through *linear* operators. When we represent the state of quantum bits as a unit vector in a Hilbert space, a quantum operation can thus be represented as a matrix operating on this space. Moreover, not any matrix is allowed. Quantum mechanics tells us that all matrices should be unitary. We will explain this in a little bit, but let us first define and discuss unitary matrices.

1.1 Unitary matrices

If U is a linear operator, we use the notation $U^\dagger$ to represent the *adjoint* of U. The adjoint is the unique operator that satisfies

$$(Uv, w) = (v, U^\dagger w)$$

for all vectors v, w in the Hilbert space. When U is represented by a matrix, $U^\dagger$ is represented by the complex conjugate transpose of the matrix.
A matrix U is unitary iff: $U^\dagger = U^{-1}$ i.e. $UU^\dagger = U^\dagger U = I$

1.2 properties of unitary matrices

1. Preserve the inner product of the underlying Hilbert space

$$(V, W) = (UV, UW)$$

   Proof: $(UV, UW) = (V, U^\dagger UW) = (V, IW) = (V, W)$

2. The eigen values are of the form $\lambda = e^{i\theta}$

3. Similar to a diagonal matrix

$$\begin{pmatrix} e^{i\lambda_1} & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & e^{i\lambda_d} \end{pmatrix}$$

   Which relates to an orthogonal basis.

4. The rows of U form an orthonormal basis.

5. The columns of U form an orthonormal basis.

1.3 Why Unitary?

Shrȯdinger's equation is the equation of motion which describes the evolution in time of the quantum state.

$$i\hbar\frac{d|\Psi\rangle}{dt} = H|\Psi\rangle$$

$\hbar$ is a constant (called Planck's constant) and H is a linear *Hemiltonian* which is Hermitian, i.e. it has an orthonormal set of eigenstates, all with real eigenvalues.

For those of you who are familiar with Shrȯdinger's equation, the unitarity restriction on quantum gates is simply the time-discrete version of the restriction that the Hamiltonian is Hermitian.

We will now prove that if the system satisfies Shrȯdinger's equation, then its evolution in discrete time is described by Unitary operations. (We will assume that H is time independent.)
let $|\psi_0\rangle$ be the initial state of $|\psi\rangle$, we can choose a representation of ψ using the eigen vectors of H :

$$|\psi\rangle = \Sigma a_i|\phi_i\rangle$$

$$\Downarrow$$

$$i\hbar\frac{d\Sigma a_i|\phi_i\rangle}{dt} = H\Sigma a_i|\phi_i\rangle = \Sigma a_i\lambda_i|\phi_i\rangle$$

$$\Downarrow$$

$$i\hbar\frac{da_i}{dt} = \lambda_i a_i$$

$$\Downarrow$$

$$a_i(t) = e^{i\frac{\lambda_i}{\hbar}t}a_i(0)$$

$$\Downarrow$$

$$|\psi(t)\rangle = e^{-i\frac{\lambda_i}{\hbar}t}a_i(0)|\phi_i\rangle$$

We get that the change after a discrete time difference is unitary :

$$|\psi(t)\rangle = \begin{pmatrix} e^{-i\frac{\lambda_i}{\hbar}t} & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & e^{-i\frac{\lambda_d}{\hbar}t} \end{pmatrix} \begin{pmatrix} a_0 \\ \cdot \\ \cdot \\ a_d \end{pmatrix} = U|\psi(0)\rangle$$

2 Quantum Gates

In classical computation, complex computations are build of many elementary simple steps, called gates, like AND and NOT. These are easy to realize (physically) and can be combined together. In the same way, we would like to define quantum gates, as the elementary quantum operations which will be combined together to construct complicated quantum circuits. We will refer to operations on one and two qubits and their representations as quantum gates. Here are a few examples of quantum gates on one and two qubits.

2.1 NOT

The NOT gate operates on one qubit and can be represented by the matrix

$$\mathcal{U} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

So that, $|0\rangle \to |1\rangle$ and $|1\rangle \to |0\rangle$ Where $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

2.2 Rotation

A rotation or phase shift through an angle θ can be represented by the matrix

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

2.3 Hadamard transform

Perhaps the most useful quantum gate is the Hadamard gate. This can be represented by the matrix

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

$|0\rangle \to \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $|1\rangle \to \frac{|0\rangle-|1\rangle}{\sqrt{2}}$

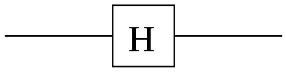


Figure 1: Hadamrd

2.4 Phase Flip

Another one qubit gate which will be useful is the so called phase shift:

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

This gate puts a minus phase in front of the basis state $|1\rangle$ leaving the other basis state unchanged. Thus, it flips the relative phase between the two basis states, which is what gives it its name.

2.5 Controlled NOT

Our first example of a two qubit gate is the controlled NOT gate, which can be represented by the matrix

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

where the basis elements are (in order) $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$. If inputs a and b are basis states, then the outputs are a and $a \oplus b$.

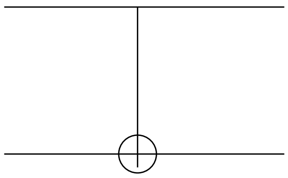


Figure 2: CNOT

3 No Cloning theorem

The CNOT gate copies the first bit to the second bit; If the input is $a, 0$ the output is a, a. However, this copying process works only for classical input bits. Can we construct a gate which will copy any state of one qubit onto a second qubit, i.e. it will take

$$|\alpha\rangle \otimes |0\rangle \longmapsto |\alpha\rangle \otimes |\alpha\rangle$$

for any $|\alpha\rangle$? A fundamental theorem in quantum mechanics says that the answer is no.

Theorem 1 No Cloning Theorem *No unitary transformation takes both* $|\Psi\rangle \otimes |s\rangle \to |\Psi\rangle \otimes |\Psi\rangle$ *and* $|\Phi\rangle \otimes |s\rangle \to |\Phi\rangle \otimes |\Phi\rangle$ *unless* $|\Psi\rangle$ *and* $|\Phi\rangle$ *are either orthogonal or parallel.*

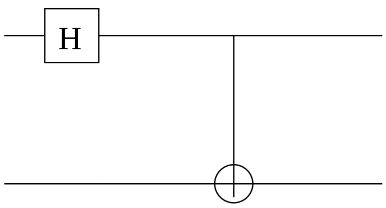
proof: We assume there is such a unitary transformation U.

$\langle\Psi|\Phi\rangle = \langle\Psi|\Phi\rangle\langle S|S\rangle = \langle\Psi \otimes S|\Phi \otimes S\rangle = \langle U(\Psi \otimes S)|U(\Phi \otimes S)\rangle = \langle\Psi|\Phi\rangle\langle\Psi|\Phi\rangle$. Taking absolute values on both sides, we get $= |\langle\Psi|\Phi\rangle| = |\langle\Psi|\Phi\rangle|^2$ which implies that $|\langle\Psi|\Phi\rangle|$ is equal to 1 or to 0 $\square$.

This means that it is possible to clone a state if you know that it is one out of an orthonormal set of vectors; But it is impossible to clone an unknown state, or a general state.

4 Bell states

An example of a quantum circuit :



$$|11\rangle \to^H \left(\frac{|0\rangle - 1}{\sqrt{2}}\right) \otimes |1\rangle = \frac{|0\rangle \otimes |1\rangle - |1\rangle \otimes |1\rangle}{\sqrt{2}} \to^{CNOT} \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

This is one of four Bell states :

$$\frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

$$\frac{|01\rangle + |10\rangle}{\sqrt{2}}$$

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

$$\frac{|00\rangle - |11\rangle}{\sqrt{2}}$$

These states are all created by the above circuit, when the input is one of the standard basis states.

The first Bell state, as we have seen last week in the Bell inequality game, has very strong correlation properties, stronger than any classical model can explain. It turns out that all the above states share this property of strong quantum correlations. Another name for these correlations is *Entanglement*. We will say that a state of two qubits is *entangled*, or that the two qubits are entangled, if the state cannot be written as a tensor product of the state of the first qubit and the state of the second qubit. The Bell states are the most entangled two qubit states, in the sense that the correlations between the two qubits are the strongest; We will get to definitions and quantifying this peculiar notion of entanglement only at the end of the course, if time permits. Next, we see another peculiar phenomena which can be done using the entangled Bell state.

5 Quantum Teleportation

Alice has a qubit $\Psi = \alpha|0\rangle + \beta|1\rangle$ that she wishes to send to Bob. However, they can only talk on the phone, and Alice cannot simply send Bob the qubit. Moreover, Alice cannot even tell Bob the coefficients α and β since she doesn't know them, and she cannot measure them since she only has one copy of her qubit. This, of course, cannot be done.

Surprisingly, if in the past Alice and Bob had met, created an EPR pair, and separated each with one of the qubits of this pair, then they can use this state to perform the task of sending Ψ from Alice to Bob, merely by talking on a (classical) phone. Note that they can do it even though the quantum state they share, the EPR, is completely independent of the coefficients α, β and might have been prepared long before they even dreamt that they will even need to send Ψ. This process is called *Teleportation.* A protocol for teleportation was found by Bennett Brassard Crepeau Josza Peres and Wootters in 1993, and has since been given many applications in seemingly unrelated areas like quantum error correction.

Teleportation Protocol

1. The state of the system in the beginning is
$$\frac{1}{\sqrt{2}}[\alpha|0\rangle + \beta|1\rangle \otimes (|00\rangle + |11\rangle).$$

2. Alice performs a Controlled-Not operation on her two qubits, where ψ is the control, and her half of the EPR pair is the target.
$$\frac{1}{\sqrt{2}}[\alpha|0\rangle \otimes (|00\rangle + |11\rangle) + \beta|1\rangle \otimes (|00\rangle + |11\rangle)]$$
$$\rightarrow^{CNOT} \frac{1}{\sqrt{2}}[\alpha|0\rangle \otimes (|00\rangle + |11\rangle) + \beta|1\rangle \otimes (|10\rangle + |01\rangle)]$$

3. Alice then performs a Hadamard transformation on ψ. This takes the state to
$$\frac{1}{\sqrt{2}}\left[\alpha\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) \otimes (|00\rangle + |11\rangle) + \beta\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \otimes |10\rangle + |01\rangle\right] =$$
$$\frac{1}{2}[|0\rangle \otimes |0\rangle \otimes (\alpha|0\rangle + \beta|1\rangle)]$$
$$+\frac{1}{2}[|0\rangle \otimes |1\rangle \otimes (\alpha|1\rangle + \beta|0\rangle)]$$
$$+\frac{1}{2}[|1\rangle \otimes |0\rangle \otimes (\alpha|0\rangle - \beta|1\rangle)]$$
$$+\frac{1}{2}[|1\rangle \otimes |1\rangle \otimes (\alpha|1\rangle - \beta|0\rangle)]$$

The circuit that describes the operations so far is given in the figure.

Note that in the above expression, the four states on Bob's qubit are very reminiscent of ψ. For example, the first state entangled to $|0\rangle \otimes |0\rangle$ is exactly

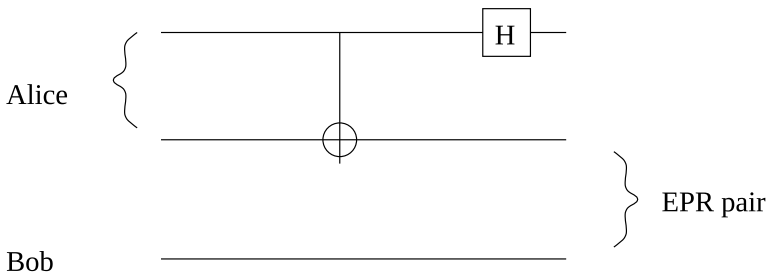


Figure 3: Teleportation Circuit

ψ. The other three states seem to be very close; If Bob only knew which state he has, he could apply a unitary transformation on his qubit to get the correct state. For example, if Bob's state is the second state, $\alpha|1\rangle + \beta|0\rangle$ Bob can apply NOT to get the correct state. On the third state, he need to apply the matrix Z, which we met before; and on the last state he needs to apply both Z and NOT.

So if Bob knew in which of the states he is, we are done. If Alice measures in the standard basis, Bob's qubit will collapse to one of these states, and moreover, Alice can tell Bob using only 2 classical bits which of the Basis states she received, or in which of the four possible states his qubit is. Using this information , Bob can determine which unitary transformation is necessary to recover Alice's original state ψ.

Remark: Note that at the end of the teleportation protocol, nor Alice neither Bob know the state; they both are completely ignorant about α and β. Note also, that at the end of the process, when Bob has the original state in his hand, Alice's state is lost, and has no connection with the original state. This is of course in agreement with the no cloning theorem.

6 Generalizing to n Qubits

So far, we have discussed only systems with one or two qubits. Now, we would like to proceed to larger systems on which we can perform general computations. Qubits are combined by tensor products, as was discussed in the previous week. We repeat here the definition of tensor products and show how matrices and gates operate on tensor product spaces.

6.1 Tensor Products

Definition 1 *The tensor product of two Hilbert spaces A, B of dimensions m, n is the vector space of dimension mn which is donated by $A \otimes B$, together with a function $f : A, B \to A \otimes B$ such that:*

$$\alpha, \beta \to \alpha \otimes \beta = f(\alpha, \beta)$$

which satifies

$$\alpha \otimes \beta + \alpha \otimes \gamma = \alpha \otimes (\beta + \gamma)$$

$$\alpha \otimes \beta + \gamma \otimes \beta = (\alpha + \gamma) \otimes \beta$$

$$a(\alpha \otimes \beta) = (a\alpha) \otimes \beta = \alpha \otimes (a\beta)$$

$A \otimes B$ is a Hilbert space with the inner products induced from A, B, i.e.

$$(\alpha \otimes \beta, \gamma \otimes \lambda)_{A \otimes B} = (\alpha, \gamma)_A \otimes (\beta, \lambda)_B$$

6.2 n Qubits

The Hilbert space of n qubits is thus: $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes ... \otimes \mathbb{C}^2$, the tensor product of n Hilbert spaces of one qubit.

The dimension of the n qubits space $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes ... \otimes \mathbb{C}^2$ (n times) is 2^n.

This result is intuitive: it's the number of configuration of n classical bits.

6.3 Applying a quantum gate

Let U be a transformation over a qubit. Suppose U operates on the k'th qubit in a system of n qubits. Then the state

$$\alpha = \sum_i a_i |i_1 > |i_2 > ...|i_n >$$

will go to the state:

$$\sum_i a_i |i_1 > |i_2 > ..|i_{k-1} > U|i_k > |i_{k+1}....|i_n >$$

i.e. U operates on the k'th coordinate and does nothing on the rest, i.e. it applies the identity on the rest of the qubits. The operation described is a linear operation on any state of n qubits; Thus, it can be written as a matrix operating on the Hilbert space of n qubits. We say that this operation is the tensor product of U on the $k'th$ qubit with the identity on the rest of the qubits. We denote this by

$$U\tilde{(k)} = I_{1,...,k-1,k+1,...n} \otimes U_k.$$

The tilde notation means extending U which operates on a small number of qubits to a matrix operating on all the qubits by applying the identity on the rest of the qubits.

What will the matrix $U\tilde{(k)}$ look like in matrix representation? This is easy to see by checking how $U\tilde{(k)}$ operates on any basis state. We know that

$$U\tilde{(k)}|i_1....i_{k-1} > |i_k > |i_{k+1}, ...i_n >= |i_1....i_{k-1} > (U|i_k >)|i_{k+1}, ...i_n >$$

which is a combination of two basis states, both of them have the original values on the rest of the qubits. this means that $U\tilde{(k)}$ is a block matrix:

$$U\tilde{(k)}_{i_1...i_n,j_1...j_n} = U_{i_k,j_k}$$

if all indices except the k'th index, $i_1 = j_1, ...i_n = j_n$. Otherwise,

$$U\tilde{(k)}_{i_1...i_n,j_1...j_n} = 0$$

.

A shorter way to write this is:

$$U\tilde{(k)}_{i_1...i_n,j_1...j_n} = U_{i_k,j_k} I_{i_1..,\hat{i_k},..,i_n,j_1..,\hat{j_k},..j_n}$$

It is worthwhile to try how a simple one qubit gate extended to three qubits looks like.

The definition of $U\tilde{(k)}$ is in fact the *tensor product* of U and the identity matrix; it is a special case of tensor product of matrices which is defined next.

6.4 Tensor products of matrices

Consider two matrices M_A, M_B on A, B. We want to define the matrix operating on $A \otimes B$ which corresponds to applying M_A on A and M_B on B.

Definition 2 *$M_A \otimes M_B$ is the matrix that satisfies*

$$(M_A \otimes M_B)(\alpha \otimes \beta) = M_A\alpha \otimes M_B\beta$$

for any $\alpha \in A, \beta \in B$.

It can be shown that the matrix which satisfies this is of the form :

$$(M_A \otimes M_B)_{i,j,k,l} = (M_{A(i,k)} M_{B(j,l)})$$

$$\Downarrow$$

$$M_A \otimes M_B = \begin{pmatrix} \left(M_{A(1,1)} M_B\right) & \left(M_{A(1,2)} M_B\right) & .. \\ . & . & .. \\ . & . & .. \\ M_{A(n,1)} M_B & . & .. \end{pmatrix}$$

6.5 Example:

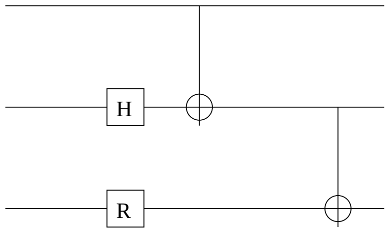
U_{2x2} on A of dimension 2. tensor product with the identity matrix on B of dimension 2 :

$$I \otimes U = \begin{pmatrix} U_{11} & U_{12} & 0 & 0 \\ U_{21} & U_{22} & 0 & 0 \\ 0 & 0 & U_{11} & U_{12} \\ 0 & 0 & U_{21} & U_{22} \end{pmatrix}$$

7 Quantum Circuits

We can now define a quantum circuit:

Definition 3 *A quantum circuit over a set of gates G is represented by a directed acyclic graph which satisfies: For each node the in degree equals to the out degree. Some of the inputs are denoted by constants: 0 or 1, and some of the outputs are denoted by 'm' (the "interesting" part - the part that will be measured). Each node is assigned a gate $g \in G$.*



7.1 example

The unitary matrix which the circuit operates is:

$U = \widetilde{CNOT}_{12}\widetilde{CNOT}_{23}\widetilde{H}_2\widetilde{R}_{\theta_1}$

The order of reading is according to the order of topological sorting. (i.e. a gate g_2 must appear after a gate g_1 if there is a path from g_1 to g_2. There can be many topological orders to a circuit; However, since matrices that operate on different sets of qubits commute, the topological order we choose does not matter.

Definition 4 *The function computed by the circuit is defined as follows: given an input i, apply the circuit on $|i\rangle$, measure the qubits denoted by m in the standard basis, and output the result.*

8 Universal Set of Gates

Definition 5 *A set of quantum gates, G, is called universal if for any $\epsilon > 0$ and any unitary matrix U on any number of bits, U can be approximated to within ϵ by a sequence of gates from G, i.e. $g_l ... g_1 = U‘$ and $\|U - U‘\| < \epsilon$. In other words, the subgroup generated by G is dense in the group of unitary operators $U(n)$, for all n.*

Reminder : $\|U - U'\| = max_{\|\alpha\|=1} |U(\alpha) - U'(\alpha)|$

8.1 Examples

- All the gates that operates on one qubit with CNOT
- $\begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$, CNOT, H

Why need many universal sets of gates? Universal sets of gates are our programming languages; WE will need to translate between different sets, for different purposes. A very basic example is this: whereas a certain set is easy to construct algorithms with, it might not be the set of gates which is easy to realize physically in the laboratory. We know that we can approximate any gate by a sequence of gates from a universal set of gates, so we know that this translation can be done. The question is what is the price we pay, or in other words, if we want to translate one set to the other, how many gates are needed to approximate the gates to within accuracy which does not harm the correctness of the computation too much? For this, we first have to find out what is the required accuracy. This is the question we deal with in the following section.

9 Sensitivity of Quantum Circuits

We will see that there are some problems for which quantum computers are more efficient than classical computers. But where does the power of quantum computation come from? So far when describing unitary transformations and quantum circuits we have assumed that the circuits can be implemented with arbitrary precision. Are we in fact exploiting the power of an analog computer? In 1979, Shamir showed that an analog computer that could perform arithmetic operations with infinite precision could be used to factor a number a composite number n in time $O(\log n)$. Do quantum computations require such infinite precision?

Suppose we can create in a laboratory $g_1, ..., g_l$ gates, and we want to create a CNOT gate. We'll simulate it. Every function we calculate using this new approximated gate has an error due to the error in the gate. It is important to predict the magnitude of this error. The next theorem deals with this problem

We shall see that because unitary transformations are linear, small errors in the gates only produce small errors in the output.

Theorem 2 *Given $\epsilon > 0$ and unitary transformations $\mathcal{U}_1, \ldots, \mathcal{U}_m, \mathcal{V}_1, \ldots, \mathcal{V}_m$ with*

$$\|\mathcal{U}_i - \mathcal{V}_i\| < \epsilon/m.$$

Let

$$\begin{aligned} \psi &= \mathcal{U}_1\mathcal{U}_2\cdots\mathcal{U}_m|\phi_0\rangle \\ \psi' &= \mathcal{V}_1\mathcal{V}_2\cdots\mathcal{V}_m|\phi_0\rangle. \end{aligned}$$

Let $\mathcal{D}(\psi)$ be the distribution of ψ. Then

$$\|\mathcal{D}(\psi) - \mathcal{D}(\psi)\|_1 < \epsilon.$$

Proof: Denote :

$$\psi_k = \mathcal{V}_1\mathcal{V}_2\cdots\mathcal{V}_{k-1}\mathcal{U}_k\mathcal{U}_{k+1}\cdots\mathcal{U}_m|\phi\rangle$$

for $k = 0, \ldots, m$.
And denote $A = \mathcal{V}_1, .., \mathcal{V}_{k-1}$ and $|\alpha\rangle = \mathcal{U}_{k+1}, .., \mathcal{U}_m|\phi\rangle$

$$\|\psi_k - \psi_{k+1}\| = \|A\mathcal{U}_k|\alpha\rangle - A\mathcal{V}_k|\alpha\rangle\| =$$

All of the $\mathcal{V}_i$s are unitary, then $\prod \mathcal{V}_i$ is also unitary (and unitary matrices preserve inner products) .

$$= \||\mathcal{U}_k|\alpha\rangle - \mathcal{V}_k|\alpha\rangle\| \leq \||\mathcal{U}_k - \mathcal{V}_k\| \cdot \||\alpha\rangle\| =$$

because $\||\alpha\rangle\| = 1$

$$= \frac{\epsilon}{m}$$

Then by the triangle inequality we have that

$$\|\psi_0 - \psi_m\| < \epsilon.$$

The rest of the proof was given as an exercise (ex 3 in problem set 1.)
And so we conclude that the error is adaptive, and we can afford an error of order $\frac{1}{m}$ in each gate.

9.1 Approximation rate

We now know that we only need to approximate each gate to within $\frac{\epsilon}{m}$, to get a reasonable error. But what is the price? How many gates from a universal set are needed to do this? A very pleasing answer is given by the following theorem:

Theorem 3 Solovay-Kitaev theorem: *Let $G \subset SU(d)$, G is closed under reciprocal ($g \in G \Leftrightarrow g^{-1} \in G$), and G is dense in $Su(d)$. Then any matrix $U \in SU(d)$ can be approximated to within ϵ by a product of l matrices from G, where $l \leq log^2(\frac{1}{\epsilon})$.*

$SU(d)$ is the set of $n \times n$ unitary matrices with determinant 1.
We won't prove this theorem- The proof uses Lie Algebra techniques, and is quite involved.

We conclude, that in order to approximate a gate using a set of universal gates without loosing accuracy of more than $\frac{1}{m}$, we can use words over the original set with length no more than l. Thus, the theorem shows that in order to approximate a gate on a constant number of qubits to within ϵ/m only a logarithmic(m/ϵ) number of gates is needed; This means that changing between universal sets of gates inserts only polylogarithmic factors, and so quantum complexity is almost not sensitive to which set of gates we use.