

Quantum Computation

Nipun Kwatra, Dhruv Kumar Mahajan, Sumit Jain, Vinay Gupta*

Supervisor: Prof. Sandeep Sen

Department of Computer Science and Engineering,
IIT Delhi.

IInd Semester 2002-03

Independent Study, CS310S

*{csu00122,csu00108,csu00136,csu00140}@cse.iitd.ernet.in

Contents

1	Introduction	1
2	Photon Polarisation	1
2.1	The Experiment	1
2.2	The Explanation	3
3	Basics of Quantum Computation	4
3.1	States in Quantum Mechanics	4
3.2	Qubits	6
3.3	Observables	6
3.3.1	Effect of an observation of a state vector	7
4	Quantum Cryptography	8
5	Transformations on a single Qubit	10
5.1	Some Important Properties	10
5.2	Some Basic Transformations	12
6	Multiple Qubits	12
6.1	Measurement	13
6.2	Transformations	14
6.2.1	Walsh-Hadamard Transformation	15
6.2.2	No Cloning	15
7	EPR Paradox	16
8	Dense Coding and Teleportation	17
8.1	Dense Coding	17
8.2	Need for entangled system for Dense Coding	18
8.3	Teleportation	19
9	Quantum Interference	21
9.1	Empirical Verification	22
9.2	Explanation	22

10 Quantum Gates	24
10.1 Quantum Gate Arrays	24
10.2 Modified Deutsch-Jozsa Problem (MDJP)	25
11 Quantum Fourier Transforms	30
12 Grover's Search Algorithm	35
12.1 The Problem	36
12.2 Algorithm	36
12.2.1 Inversion about average	37
12.3 Proofs	39
13 Shor's Factoring Algorithm	42
13.1 Finding Order of an element	42
14 Simon's Problem.	44
15 Quantum Circuits - some consruction issues	46
15.1 Controlled Operations	46
15.2 Universal Quantum Gates	49
15.2.1 Two-level unitary gates are universal	51
15.2.2 Single qubit and CNOT gates are universal	53
16 Acknowledgements	55
A Properties of tensor products	56

1 Introduction

In the early 80's it was observed by Richard Feynman that some Quantum mechanical effects could not be simulated on a classical computer. This led many scientists to think whether *quantum* effects could be harnessed for computation.

Classically, the computation time can be decreased by using parallel processors. An exponential decrease in time requires an exponential increase in the amount of resources. But, time improves exponentially with the size of the system in a quantum system.

The world of quantum computation received a boost with Peter Shor's polynomial time quantum algorithm for factoring integers. Our aim in this Independent Study will be to study and understand this and other algorithms that could run on a Quantum Computer.

In Section 2 we describe some of the experiments on the basis of which which can build on the ideas in quantum computation. In Section 3 we give some mathematical constructs for quantum computation.

2 Photon Polarisation

Photons are the only particles that we can directly observe. The following experiment demonstrates some of the principles of quantum mechanics through photons and their polarisation.

2.1 The Experiment

A beam of light shines on the projection screen. Filters A , B , and C are polarised horizontally, at 45° , and vertically, respectively, and can be placed so as to intersect the beam of light.

First, insert a filter A . Assuming that incoming light is randomly polarised, the intensity of the output will have half the intensity of the incoming

light. The outgoing photons are now all horizontally polarised.

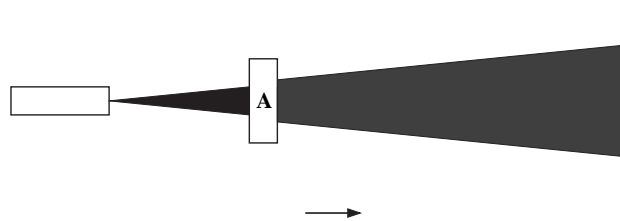


Figure 1: Filter *A* inserted

The function of filter *A* cannot be explained as a "seive" that only lets those photons pass that happen to be already horizontally polarised. If that were the case, few of the randomly polarised incoming photons would be horizontally polarised, so we could expect a much larger attenuation of the light as it passes through the filter.

Next, when filter *C* is inserted, the intensity of the output drops to zero. None of the horizontally polarised photons can pass through the vertical filter. A Seive model could explain this behavior.

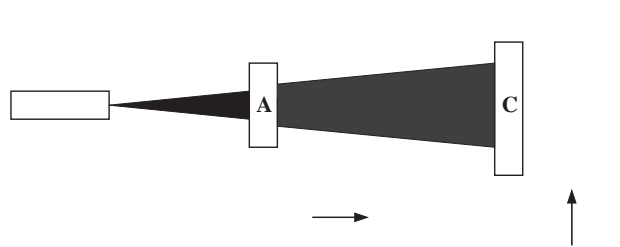
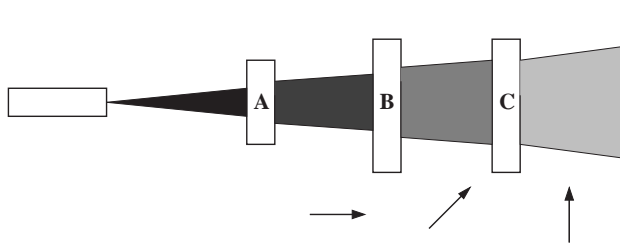


Figure 2: Filters *A* and *C* inserted

Finally, after filter *B* is inserted between *A* and *C*, a small amount of light will be visible on the screen, exactly one eighth of the original amount of light.

Here we have a non-intuitive effect. Classical experience suggests that adding a filter should only be able to decrease the number of photons getting

Figure 3: Filters A , B and C inserted

through. How can it increase it?

2.2 The Explanation

A photon's polarization state can be modelled by unit vector pointing in the appropriate direction. Any arbitrary polarization can be expressed as a linear combination of the horizontal and vertical polarizations (basis vectors).

The measurement postulate of quantum mechanics states that any device measuring a 2-dimensional system has an associated orthonormal basis with respect to which the quantum measurement takes place. Measurement of a state transforms the state into one of the measuring device's associated basis vectors. The probability that the state is measured as one of the basis vectors is the square norm of the amplitude of the component of the state in the direction of that basis vector (termed as the probability amplitude).

Quantum mechanics can explain the polarization experiment as follows. A polaroid measures the quantum state of photons with respect to the basis consisting of the vector corresponding to its polarization together with a vector orthogonal to its polarization. The photons which, after being measured by the filter, match the filter's polarization are let through. Therefore Filter A measures the photon polarization with respect to horizontal polarization basis vector. The photons that pass through filter A all have horizontal polarization.

Assuming random polarization, filter A measures 50% of all photons as horizontally polarized. Filter C will measure these photons with respect to vertical polarization basis vector. Hence state will be projected onto it with probability 0 (as component along this vector is 0), and no photons will pass through filter C .

Finally B measures these horizontally polarized photons with respect to basis vectors at 45° above the horizontal and other orthogonal to it. It is trivial to see that horizontal vectors have equal component along these vectors so photons passing through A will be measured by filter B with probability $1/2$. Therefore only 50% of the photons pass through B in state 45° to horizontal. As before, these photons will be measured by the filter C with probability $1/2$. Thus only one-eighth of the original photons manage to pass through the sequence of filters A , B and C .

3 Basics of Quantum Computation

This section defines some of the mathematical constructs and notations that are required in Quantum computation.

3.1 States in Quantum Mechanics

The *State*¹ space of a quantum system, consisting of positions, momentums, polarizations, spins, etc. of the various particles, can be modelled by a Hilbert space of wave functions. For quantum computing we need only deal with finite quantum systems, and it suffices to consider finite dimensional Hilbert Space. In Quantum Mechanics a *State* is a *ray* in a *Hilbert Space*.

- Hilbert Space H

1. It is a Vector Space over the set of Complex Numbers \mathbf{C}
2. It has an inner product $\langle \psi | \phi \rangle$ that maps an ordered pair of vectors $\psi, \phi \in H$ to \mathbf{C} .
3. It is complete in the norm $\| \psi \| = \langle \psi | \psi \rangle^{1/2}$

¹A state is a complete description of a physical system

- **Ray**

It is an equivalence class of vectors that differ by multiplication by a non-zero complex scalar. We generally choose a vector $|\psi\rangle$ (*Dirac Notation*, see below), having the unit norm as the representative of this class. Note that $|\psi\rangle$ and $\lambda|\psi\rangle$ represent the same physical state, where λ is a scalar.

- **Dirac Notation**

Quantum state spaces and the transformations acting on them can be described in terms of vectors and matrices or in the more compact bra/ket notation invented by Dirac. Kets like $|x\rangle$ denote column vectors and are used to denote quantum states. The matching bra, $\langle x|$, denotes the conjugate transpose of $|x\rangle$. For example, let the orthonormal basis of a 2-D Hilbert Space be $\{|0\rangle, |1\rangle\}$, then any vector in the space of the form $a|0\rangle + b|1\rangle$, can be written as $(a, b)^T$. Then the corresponding bra is (a^*, b^*) .

$\langle x|y\rangle$ denotes the inner product of $|x\rangle$ and $|y\rangle$ and is also denoted as $\langle x|y\rangle$. For example for the orthonormal bases $\{|0\rangle, |1\rangle\}$, above $\langle 0|1\rangle = 0$ and $\langle 1|1\rangle = 1$.

The notation $|x\rangle\langle y|$ is the outer product of $|x\rangle$ and $\langle y|$, and represents a transformation from \mathbf{H} to \mathbf{H} , where \mathbf{H} is the corresponding Hilbert Space. For example, $|0\rangle\langle 1|$ is the transformation that maps $|1\rangle$ to $|0\rangle$ and $|0\rangle$ to $(0, 0)^T$ since

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

$$|0\rangle\langle 1||0\rangle = |0\rangle\langle 1|0\rangle = |0\rangle 0 = (0, 0)^T$$

Equivalently $|0\rangle\langle 1|$ can be written in matrix form where $|0\rangle = (1, 0)^T$, $\langle 0| = (1, 0)$, and $\langle 1| = (0, 1)$. So

$$|0\rangle\langle 1| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (0, 1) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} (0, 1)$$

Thus this notation can be used to specify transformations neatly in terms of the basis vectors.

3.2 Qubits

The indivisible unit of classical information is the bit also known as the Shannon bit which can take either the value 1 or the value 0. The corresponding unit in Quantum Mechanics is called the *Qubit*.

Definition:

A *qubit* is a quantum state in a 2-dimensional Hilbert Space \mathbf{C}^2 , of the form,

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

where $\alpha, \beta \in \mathbf{C}$ and $\{|\mathbf{0}\rangle, |\mathbf{1}\rangle\}$ form an orthonormal basis for the Hilbert space. The normalized vector² i.e. if $\|\alpha\|^2 + \|\beta\|^2 = 1$, has a special property which is discussed when we come to measurements.

The qubit can be encoded in the polarization of the photon or the spin of an atom.

The difference between classical and quantum bit is that qubit can take any (uncountable) quantum superposition of $\mathbf{0}$ and $\mathbf{1}$ ³. Thus, an infinite amount of information can be encoded in a single qubit by appropriately varying (defining) α and β . A greater problem lies in how to *extract* information out of this qubit. This is done through *observables*.

3.3 Observables

Observables are a mathematical device used for measuring qubits. An observable is a property of a physical system that in principle can be measured. Suppose we have a probe (or a measuring device) P and a property of quantum state $|\psi\rangle$ to be measured. This property can be anything - *eg.* position or direction. An *observable* is a mathematical representation of the probe P.

Definition:

Let \mathbf{H} be the Hilbert space used to represent the state vectors of a quantum

²Remember the state is represented by a ray, i.e. the vectors are only upto a scale

³we will also denote $|\mathbf{x}\rangle$ as \mathbf{x}

system. An *observable* O is a set of subspaces $E_1, E_2, \dots, E_k \subseteq \mathbf{H}$ such that these subspaces completely partition \mathbf{H} , that is,

$$E_1 \times E_2 \times \dots \times E_k = \mathbf{H}$$

that is \mathbf{H} is isomorphic to the cartesian product of the various subspaces. Also for all i, j such that if $i \neq j$ then $E_i \perp E_j$. That is, the subspaces are orthogonal.

3.3.1 Effect of an observation of a state vector

Let $|\psi\rangle$ be a vector in space \mathbf{H} and let $O = \{E_1, \dots, E_k\}$ be an observable. Since O partitions \mathbf{H} into orthogonal subspaces, $|\psi\rangle$ can be written as a superposition,

$$|\psi\rangle = \sum_{i=1}^k \alpha_i |\psi_{E_i}\rangle$$

where, $|\psi_{E_i}\rangle$ lies in E_i .

Now, observing state $|\psi\rangle$ with O will cause the following-

1. $|\psi\rangle$ will be observed to be in one of the subspaces. The probability that it is selected in E_i is given by $\|\alpha_i\|^2$, if $|\psi\rangle$ is a **normalised vector**. $\|\alpha_i\|$'s (for a normalised vector) are also known as probability amplitudes.
2. If $|\psi\rangle$ is observed to be in E_i , then the state $|\psi\rangle$ collapses to $|\psi_{E_i}\rangle$.
3. The only classical information given by O is which subspace E_i is chosen. All information not in $|\psi_{E_i}\rangle$ is *lost*.

These rules are the *Quantum Postulates of Measurement*.

To each possible output value of the probe corresponds a subspace in the observable. Also, any observable is allowed for observing a quantum state. The standard observable for qubit is taken as $B = \{E_0, E_1\}$, where E_0 is spanned by $|0\rangle$ and E_1 is spanned by $|1\rangle$. We could as well have an observable $O = \{E'_0, E'_1\}$ where E'_0 and E'_1 are spanned by,

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

respectively.

Let us take an example. Let us take a *qubit* that has been encoded in the polarisation of a photon. We can take $\{|\uparrow\rangle, |\rightarrow\rangle\}$ as the orthonormal basis of the Hilbert space. We will denote the qubit as,

$$|\phi\rangle = \alpha|\uparrow\rangle + \beta|\rightarrow\rangle$$

Suppose, we pass the photon through a filter F that is aligned to the horizontal (that is, it will let only horizontally polarised light pass through it). Note that this is an act of measuring the qubit using the probe F . We can take the observable associated with F to be $O = \{|\uparrow\rangle, |\rightarrow\rangle\}$. Therefore, the qubit (photon) will *collapse* to $|\uparrow\rangle$ with probability $\|\alpha\|^2$ and to $|\rightarrow\rangle$ with probability $\|\beta\|^2$. Suppose we turn the filter F by 45° . The observable associated with the new probe can be taken to be $O = \{|\nearrow\rangle, |\searrow\rangle\}$. We can write the qubit as

$$|\phi\rangle = \frac{\alpha}{\sqrt{2}}(|\nearrow\rangle - |\searrow\rangle) + \frac{\beta}{\sqrt{2}}(|\nearrow\rangle + |\searrow\rangle)$$

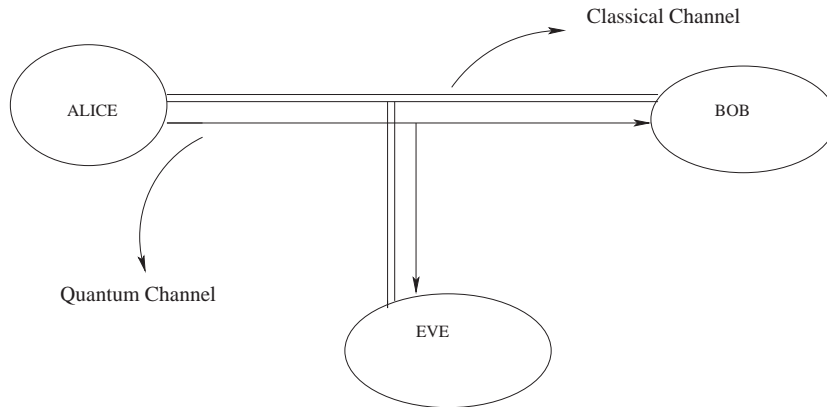
Thus, it will *collapse* to $|\nearrow\rangle$ with probability $\|\frac{\alpha+\beta}{\sqrt{2}}\|^2$ and to $|\searrow\rangle$ with probability $\|\frac{\beta-\alpha}{\sqrt{2}}\|^2$.

4 Quantum Cryptography

In this section we will take a look at how sequences of single *qubits* can be used to transmit private keys over an insecure channel. Let $|0\rangle, |1\rangle$ be the vertically and horizontally polarised photons and $|0'\rangle, |1'\rangle$ be the diagonally polarised photons. Then B and O (as defined above) are a horizontally polarised filter and a filter set at 45° to the horizontal respectively.

Alice and Bob want to exchange a secret key K . And Eve intends to eavesdrop on the transmission and learn the secret key. They have a two way open *classical* channel and a one way *quantum* channel. They agree that 0 will be encoded either in state $|0\rangle$ or $|0'\rangle$ and 1 will be sent as $|1\rangle$ or $|1'\rangle$ (based on a coin flip). Bob will read (measure) the photons with B or O with equal probability.

Suppose Alice wants to send a 0 and she chooses to encode it using $|0\rangle$. If Bob chooses to measure the qubit using B then he will get a 0 outcome



with probability 1. But if he uses O then

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

and therefore Bob has 50% probability of getting a $0'$ and 50% probability of getting a $1'$. After the bits have been transmitted, Bob communicates to Alice the basis he used for encoding and decoding of each qubit, over the open two-way channel. Then Alice sends all the measurements to Bob for which the sending and receiving bases agree. Alice and Bob then delete all bits for which they used incompatible quantum alphabet to produce their **raw keys**. Alice and Bob will agree 50% of the times.

Suppose Eve is eavesdropping on both the channels. She intercepts the qubit being sent and measures it using B or O with equal probability. She encodes the value observed using the same basis she used for decoding and sends the qubit to Bob. In 50% of the cases she chooses the wrong basis(observable), in which case she sends the bit encoded with the wrong basis. That is, in the above example 50% of the times Eve chooses to measure using O . She will encode whatever she observes using $|0'\rangle$ or $|1'\rangle$ and send it to Bob. Now even if Bob uses the correct basis(observable) which is B then he will observe the wrong value with 50% probability. Thus, $\Pr(\text{Bob gets wrong value given that he uses the correct basis}) = \Pr(\text{Eve uses wrong basis}) \times \Pr(\text{Bob observes the wrong value}) = 50\% \times 50\% = 25\%$. Thus, with high probability Bob observes the wrong value even when using the correct observable. Thus, the wrong values will be accepted in the key. Alice and Bob can overcome this problem by exchanging a sufficient number of parity

Alice										
	1	0	0	1	1	0	0	1	0	1
Eve										
	1	0	1	1	1	1	0	1	0	0
Bob										
	1	0	1	1	1	1	1	0	0	0

bits. Thus, if they find any discrepancy they will know whether somebody is eavesdropping or not.

Eve cannot make a copy of the intercepted bit and measure the copy due to the *no cloning* principle explained later.

5 Transformations on a single Qubit

So far we have looked at static quantum systems which change only when measured. The dynamics of a quantum system, when they are not being measured are called transformations.

In quantum system transformations are governed by Schrodinger's equation:

Definition 1 *Transformations must take states to states in a way that preserves orthogonality.*

5.1 Some Important Properties

Property 1: For a complex vector space, linear transformations that preserve orthogonality are unitary transformations.

Proof: Let M be a quantum transformation and let $B = \{b_1, \dots, b_n\}$ be the

orthogonal basis of complex vector space.

Under M

$$\begin{aligned}\langle b_i | &\rightarrow \langle Mb_i | \\ |b_i\rangle &\rightarrow |Mb_i\rangle \\ \langle b_i | b_i \rangle &\rightarrow \langle Mb_i | Mb_i \rangle .\end{aligned}$$

Since linear transformations preserve orthogonality. So

$$\begin{aligned}\langle Mb_i | Mb_i \rangle &= 1 \\ b_i^* M^* M b_i &= 1 \\ M^* M &= I.\end{aligned}$$

So M is a unitary transformation.

Property 2: Linear Transformations are reversible.

Proof: From property 1, since quantum transformations are unitary they are reversible.

Property 3: If $O = \{E_1, \dots, E_k\}$, then probability of observing E_k is $\|\alpha_k\|^2$ where $|\psi\rangle = \sum_i \alpha_i |\psi_{E_i}\rangle$.

The transformations should leave these probabilities unchanged (that is, when we apply it to both $|\psi\rangle$ and the device).

Property 4: Linear Transformations preserve absolute values of the inner products.

Proof: Follows directly from property 1 since each state can be represented as linear combination of orthogonal basis vectors.

Property 5: The unitary transformations form a group .

Proof:

1. They have an identity element (I).
2. They can be inverted. (By definition, $M^* = M^{-1}$.)
3. They are closed under multiplication

$$(M_1.M_2)|\phi\rangle = M_1(M_2|\phi\rangle).$$

Multiplication of two transformations can be thought of as their *composition*.

5.2 Some Basic Transformations

$$\begin{aligned}
 I : |0\rangle &\rightarrow |0\rangle, |1\rangle \rightarrow |1\rangle && \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
 X : |0\rangle &\rightarrow |1\rangle, |1\rangle \rightarrow |0\rangle && \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
 Z : |0\rangle &\rightarrow |0\rangle, |1\rangle \rightarrow -|1\rangle && \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
 Y : |0\rangle &\rightarrow -|1\rangle, |1\rangle \rightarrow |0\rangle && \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}
 \end{aligned}$$

I is the identity transformation, X is negation, Z is a phase shift operation, and $Y = ZX$ is a combination of both.

Note that matrix representation for each transformation is given. This representation assumes that basis vectors used to represent the state and the transformed state are same. The matrix will change if basis chosen are different.

6 Multiple Qubits

Consider a system of n particles. In classical physics, the state of the entire system can be expressed by specifying the states of the individual particles, but this is not the case for Quantum systems, i.e. the individual states of the n particles do not define the state of the entire system.

In classical systems the individual state spaces of n particles combine through the cartesian product, but in quantum system they do so by *Tensor Product*. i.e. the entire state of the n particle can be modelled as a vector in the tensor product space of the individual state spaces.

Now we define the tensor product of 2 vector spaces.

If

$$V = \text{span}\{v_i\} \text{ and } W = \text{span}\{w_i\}$$

then

$$V \otimes W = \text{span}\{v_i \otimes w_i\}^4$$

⁴Notation: We write $(b_0 \otimes b_1 \otimes \dots \otimes b_n)$ as $|b_0 b_1 \dots b_n\rangle$. e.g. $|0\rangle \otimes |1\rangle = |01\rangle$.

This definition is slightly non-intuitive, in the sense that in general vector space is defined by the set of vectors V in the vector space, but here we are defining the tensor product space by its set of basis vectors.

If $\dim(V) = n$ and $\dim(W) = k$ then $\dim(V \otimes W) = nk$.⁵ If A is a transformation on a n -dimensional vector space and B is a transformation on a k -dimensional vector space then $V \otimes W$ is a transformation on a nk -dimensional space. See *appendix* for properties of the tensor product.

So the state space for two qubits, each of which has the basis $\{|0\rangle, |1\rangle\}$ has the basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. If we have n qubits then the state space will have a dimension 2^n .

6.1 Measurement

We have seen that when we measure single qubits then it projects to one of the basis state. In multiple qubits also if we measure all the qubits then with a certain probability the state will collapse to one of the basis states. Consider, *e.g.* a two qubit state $a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$, where a, b, c, d are complex numbers such that $\|a\|^2 + \|b\|^2 + \|c\|^2 + \|d\|^2 = 1$. Suppose we want to measure the first qubit with the observable $B = \{|0\rangle, |1\rangle\}$. We can write,

$$\begin{aligned} a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle &= |0\rangle \otimes (a|0\rangle + b|1\rangle) + |1\rangle \otimes (c|0\rangle + d|1\rangle) \\ &= u|0\rangle \otimes (a/u|0\rangle + b/u|1\rangle) + v|1\rangle \otimes (c/v|0\rangle + d/v|1\rangle) \end{aligned}$$

where, $u = \sqrt{\|a\|^2 + \|b\|^2}$ and $v = \sqrt{\|c\|^2 + \|d\|^2}$.

If we measure the first qubit then with probability $\|u\|^2$ we will observe a $|0\rangle$ and the state will collapse to $|0\rangle \otimes (a/u|0\rangle + b/u|1\rangle)$. With probability $\|v\|^2$ we will observe a $|1\rangle$ and the state will collapse to $|1\rangle \otimes (c/v|0\rangle + d/v|1\rangle)$. Now we can measure the second qubit from the new state. We can treat multiple qubit measurements as a series of single qubit measurements.

⁵It is actually this multiplication in the state space that helps achieve exponential increase in complexities, by linear increase in the no. of qubits or particles.

In general measurement of k qubits of an n qubit system splits the 2^n -dimensional state space of \mathbf{H} into a cartesian product of orthogonal subspaces S_1, \dots, S_{2^k} with $\mathbf{H} = S_1, \dots, S_{2^k}$, where S_i is the subspace which has all the vectors with the first k qubits representing i in binary. The probability of the vector being observed in state S_i is given by the square of the amplitude of the component (normalised) of $|\psi\rangle$ in S_i .

6.2 Transformations

Definition 1 *If A is a transformation on $|\psi\rangle$ and B is a transformation on $|\phi\rangle$ then $A \otimes B$ is a transformation on $|\psi\rangle \otimes |\phi\rangle$.*

$$(A \otimes B)(|\psi\rangle \otimes |\phi\rangle) = (A|\psi\rangle \otimes B|\phi\rangle)$$

Let us prove that $(A \otimes B)$ is a valid linear transformation.

$$(A \otimes B)^*(A \otimes B) = (A^* \otimes B^*)(A \otimes B) \quad (1)$$

$$= (A^*A \otimes B^*B) \quad (2)$$

$$= (I \otimes I) \quad (3)$$

$$= I. \quad (4)$$

So $(A \otimes B)$ is a unitary transformation and hence is valid.

Consider a transformation on two qubit system, *controlled not* C_{not} that inverts the second qubit if the first qubit is a $\mathbf{1}$.

$$C_{not} : |00\rangle \rightarrow |00\rangle$$

$$|01\rangle \rightarrow |01\rangle$$

$$|10\rangle \rightarrow |11\rangle$$

$$|11\rangle \rightarrow |10\rangle$$

It can also be thought of as an array-

$$C_{not} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Here to represent C_{not} as a matrix we associate $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ to the standard 4-tuple basis $(1, 0, 0, 0)^T, (0, 1, 0, 0)^T, (0, 0, 1, 0)^T, (0, 0, 0, 1)^T$, in that order.

6.2.1 Walsh-Hadamard Transformation

For single bit Hadamard transformation is defined as

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

Therefore, H creates a superposition of states.

The transformation that applies H to n bits is called the Walsh, or Walsh-Hadamard transformation. It can be defined as a recursive decomposition of the form

$$W_1 = H, W_{n+1} = (H \otimes W_n)$$

When applied to n bits it acts on each of them and creates a superposition of 2^n states

6.2.2 No Cloning

The *No cloning* principle states that we cannot copy or clone an unknown qubit. This claim is proved by contradiction. Let U be the transformation that clones, *i.e.* $U(|a0\rangle) = |aa\rangle$. Let $|a\rangle$ and $|b\rangle$ be two orthonormal states. Consider $|c\rangle = \frac{1}{\sqrt{2}}(|a\rangle + |b\rangle)$. We have by linearity,

$$\begin{aligned} U(|c0\rangle) &= U\left(\frac{1}{\sqrt{2}}(|a0\rangle + |b0\rangle)\right) \\ &= \frac{1}{\sqrt{2}}(|aa\rangle + |bb\rangle) \\ &= |cc\rangle \end{aligned}$$

But,

$$\begin{aligned} (|cc\rangle) &= |c\rangle \otimes |c\rangle \\ &= \frac{1}{2}(|aa\rangle + |ab\rangle + |ba\rangle + |bb\rangle) \end{aligned}$$

which is not equal to what was obtained above.

7 EPR Paradox

We saw in the previous section that if $|\psi\rangle$ and $|\phi\rangle$ are two qubits then we can represent the combined two qubit system as $(|\psi\rangle|\phi\rangle)$. But not all two qubit systems can be decomposed in this nice form, as a tensor product of one qubit states⁶. Consider for example, $|\phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ which cannot be written as a tensor product. The states corresponding to an n particle system that cannot be written as the tensor product of n states (of their component particles) are called *entangled* states and the particles with that state are called *entangled particles*.

Let us see what happens if we measure the above state. If we measure the first particle using the standard observable then we will observe $|0\rangle$ with probability $\frac{1}{2}$ and the state will collapse to $|00\rangle$. Now if we measure the second particle then we will measure it to be $|0\rangle$ with probability 1. Thus, we see that measuring the first particle has affected the probabilities associated with the measurement of second particle. Thus for entangled particles, measurement of one affects the measurement of the other. We can extend above for more than two particles.

This leads us to the EPR⁷ paradox. Consider the quantum state $|\phi\rangle$ as defined above. Let one particle be sent to Alice and other to Bob. They are arbitrarily far apart. Suppose Alice measures her particle and observes a $|0\rangle$. The combined state will now be $|00\rangle$ and if now Bob measures his particle then he will definitely observe a $|0\rangle$. Thus, it appears that they can communicate faster than the speed of light.

⁶We can never write the state of a m qubit system as a tensor product of states of $n(> m)$ particles, because taking the tensor product of a state of n_1 particles and a state of n_2 particles gives a state corresponding to $n_1 + n_2$ particles, *i.e.*, the dimension of the state space becomes $2^{n_1+n_2}$.

⁷Einstein, Podolsky, Rosen

8 Dense Coding and Teleportation

Dense Coding uses one quantum bit together with an EPR pair⁸ $\psi_0 = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ to transmit 2 bits of classical information.

Teleportation is the opposite of dense coding, in that it uses two classical bits to transmit a single qubit. This is surprising in the light of no cloning principle of quantum mechanics, in that it enables the transmission of an unknown quantum state.

8.1 Dense Coding

Alice receives two bits 0 to 3. One particle each (of the entangled pair) is sent to Alice and Bob. Depending on this number she performs one of the $\{I, X, Y, Z\}$ transformation to her part. Note that she cannot perform any transformation on Bob's particle. So whenever she applies a transformation to her particle and identity transformation acts on Bob's particle.

Value	Transformation	New State
0	$\psi_0 = (I \otimes I)\psi_0$	$\frac{1}{\sqrt{2}}(00\rangle + 11\rangle)$
1	$\psi_1 = (X \otimes I)\psi_1$	$\frac{1}{\sqrt{2}}(10\rangle + 01\rangle)$
2	$\psi_2 = (Y \otimes I)\psi_2$	$\frac{1}{\sqrt{2}}(- 10\rangle + 01\rangle)$
3	$\psi_3 = (Z \otimes I)\psi_3$	$\frac{1}{\sqrt{2}}(00\rangle - 11\rangle)$

Alice sends her qubit to Bob.

Bob applies C_{not} to the two qubits of the entangled pair.

Initial State	Final State	First qubit	Second qubit
$\psi_0 = \frac{1}{\sqrt{2}}(00\rangle + 11\rangle)$	$\frac{1}{\sqrt{2}}(00\rangle + 10\rangle)$	$\frac{1}{\sqrt{2}}(0\rangle + 1\rangle)$	0⟩
$\psi_1 = \frac{1}{\sqrt{2}}(10\rangle + 01\rangle)$	$\frac{1}{\sqrt{2}}(11\rangle + 01\rangle)$	$\frac{1}{\sqrt{2}}(1\rangle + 0\rangle)$	1⟩
$\psi_2 = \frac{1}{\sqrt{2}}(- 10\rangle + 01\rangle)$	$\frac{1}{\sqrt{2}}(- 11\rangle + 01\rangle)$	$\frac{1}{\sqrt{2}}(- 1\rangle + 0\rangle)$	1⟩
$\psi_3 = \frac{1}{\sqrt{2}}(00\rangle - 11\rangle)$	$\frac{1}{\sqrt{2}}(00\rangle - 10\rangle)$	$\frac{1}{\sqrt{2}}(0\rangle - 1\rangle)$	0⟩

Bob measures the second qubit, if it is **1** then he knows that it is a 1 or 2, else the number was 0 or 3. Now he applies the transformation H to the

⁸entangled particles

first qubit (*i.e.* $H \otimes I$ on the two particles). Then he observes the first qubit.

First qubit	Second qubit	Number
$ 0\rangle$	$ 0\rangle$	0
$ 0\rangle$	$ 1\rangle$	1
$ 1\rangle$	$ 0\rangle$	3
$ 1\rangle$	$ 1\rangle$	2

Thus, 2 bits worth of classical information can be transferred.

8.2 Need for entangled system for Dense Coding

Lets start from a general state :

$$\psi = (a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle)$$

This is the most general state in which the system starts. Alice encodes using I,X,Y,Z the 2 bit information. The resulting states in the 4 cases are :

$$\begin{aligned} |\psi_I\rangle &= (a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle) \\ |\psi_X\rangle &= (a|10\rangle + b|11\rangle + c|00\rangle + d|01\rangle) \\ |\psi_Y\rangle &= (-a|10\rangle - b|11\rangle + c|00\rangle + d|01\rangle) \\ |\psi_Z\rangle &= (a|00\rangle + b|01\rangle - c|10\rangle - d|11\rangle) \end{aligned}$$

This is most general state in which Bob recieves the 2 qubits. Now his job is to use some tranformations to extract the 2 bit encoded information in the system.

He applies C_{not} to get :

$$\begin{aligned}
|\psi_I\rangle &= (a|00\rangle + b|01\rangle + c|11\rangle + d|10\rangle) \\
&= (a|0\rangle + d|1\rangle) \otimes |0\rangle + (b|0\rangle + c|1\rangle) \otimes |1\rangle \\
|\psi_X\rangle &= (a|11\rangle + b|10\rangle + c|00\rangle + d|01\rangle) \\
&= (a|1\rangle + d|0\rangle) \otimes |1\rangle + (b|1\rangle + c|0\rangle) \otimes |0\rangle \\
|\psi_Y\rangle &= (-a|11\rangle - b|10\rangle + c|00\rangle + d|01\rangle) \\
&= (-a|1\rangle + d|0\rangle) \otimes |1\rangle + (-b|1\rangle + c|0\rangle) \otimes |0\rangle \\
|\psi_Z\rangle &= (a|00\rangle + b|01\rangle - c|11\rangle - d|10\rangle) \\
&= (a|0\rangle - d|1\rangle) \otimes |0\rangle + (b|0\rangle - c|1\rangle) \otimes |1\rangle
\end{aligned}$$

Clearly, we can measure the second qubit with probability 1, iff either $a = d = 0$ or $b = c = 0$.

In either case the initial system is entangled in either one of the following states:-

$$\begin{aligned}
|\psi\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\
|\psi\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)
\end{aligned}$$

8.3 Teleportation

The objective is to transmit the quantum state of a particle using classical bits and reconstruct the exact quantum state at the receiver. Since quantum state cannot be copied, the quantum state of a given particle will necessarily be destroyed.

Alice has a qubit whose state is unknown. Let the qubit be

$$\phi = a|0\rangle + b|1\rangle$$

She wants to send the state of this qubit to Bob through classical channels. As with dense coding, Alice and Bob each possess one qubit of an entangled pair

$$\psi_0 = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Alice combines the states of the two particles she have. So the starting state is

$$\begin{aligned}\phi \otimes \psi_0 &= \frac{1}{\sqrt{2}}(a|0\rangle \otimes (|00\rangle + |11\rangle) + b|1\rangle \otimes (|00\rangle + |11\rangle)) \\ &= \frac{1}{\sqrt{2}}(a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle)\end{aligned}$$

of which Alice controls the first two bits and Bob controls the last one. Alice now applies $C_{not} \otimes I$ and $H \otimes I \otimes I$ to this state:

$$\begin{aligned}(H \otimes I \otimes I)(C_{not} \otimes I)(\phi \otimes \psi_0) \\ &= (H \otimes I \otimes I)(C_{not} \otimes I)\frac{1}{\sqrt{2}}(a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle) \\ &= (H \otimes I \otimes I)\frac{1}{\sqrt{2}}(a|000\rangle + a|011\rangle + b|110\rangle + b|101\rangle) \\ &= \frac{1}{2}(a(|000\rangle + |011\rangle + |100\rangle + |111\rangle) + b(|010\rangle + |001\rangle - |110\rangle - |101\rangle)) \\ &= \frac{1}{2}(|00\rangle \otimes (a|0\rangle + b|1\rangle) + |01\rangle \otimes (a|1\rangle + b|0\rangle) + |10\rangle \otimes (a|0\rangle - b|1\rangle) + |11\rangle \otimes (a|1\rangle - b|0\rangle))\end{aligned}$$

Alice measures the first two qubits to get one of $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ with equal probability. Depending on the result of the measurement, the quantum state of Bob's qubit is projected to $a|0\rangle + b|1\rangle, a|1\rangle + b|0\rangle, a|0\rangle - b|1\rangle, a|1\rangle - b|0\rangle$ respectively. Alice sends the result of her measurement as two classical bits to Bob.

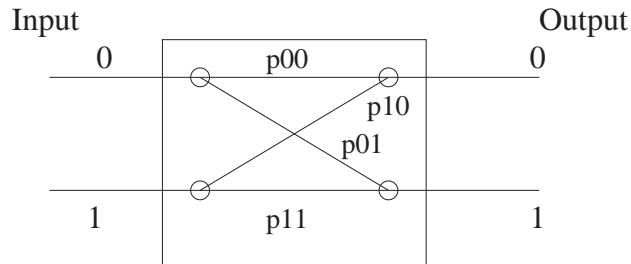
When Bob receives the two classical bits from Alice he knows how the state of his half of the entangled pair compares to the original state of Alice's qubit.

Bits received	State	Decoding
00	$a 0\rangle + b 1\rangle$	I
01	$a 1\rangle + b 0\rangle$	X
10	$a 0\rangle - b 1\rangle$	Z
11	$a 1\rangle - b 0\rangle$	Y

Bob can reconstruct the original state of Alice's qubit, ϕ , by applying the appropriate decoding transformation to his part of the entangled pair.

9 Quantum Interference

Consider a machine M , shown in the figure which given a 0 or 1 as input returns 0 or 1 as output.



p_{ij} is the probability of obtaining output j when input i is given. Note that $\sum_j p_{ij} = 1$. Therefore, if $p_{01} = 0$ and $p_{10} = 0$, then M is the identity machine. Similarly we can have the *not* machine. Consider $p_{00} = p_{01} = p_{10} = p_{11} = \frac{1}{2}$. This is a random machine which given an input i returns 0 or 1 with equal probability. When this particular machine is concatenated with an identical machine (Figure 4) the output is always the negation of the input. This a very counter-intuitive claim - each machine alone outputs 0 or 1 with equal probability and independent of the input, but the two machines one after the other, acting independently implement the logical operation **not**. The individual machines for this reason are called the $\sqrt{\text{not}}$ gate.

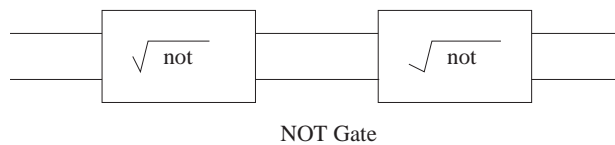


Figure 4: Concatenation of two $\sqrt{\text{not}}$ gates gives a not gate.

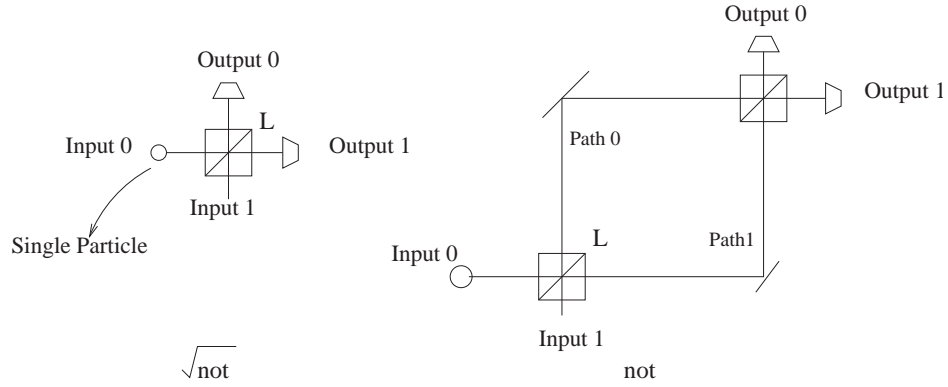


Figure 5: Experimental realisation of $\sqrt{\text{not}}$ gate and not gate.

9.1 Empirical Verification

Consider Figure 5. L is a half-silvered mirror, i.e. a mirror which with probability 50% reflects a photon which impinges on it and with probability 50% allows it to pass through. The experimental realisation of the two concatenated $\sqrt{\text{not}}$ gates is known as *Single particle interference*. A photon which enters the apparatus (known as *interferometer*) via Input 0 always strikes a detector at Output 1 and never a detector at Output 0.

Classically, we may argue that the transition $0 \rightarrow 0$ in the composite machine can happen via two mutually exclusive paths, $0 \rightarrow 0 \rightarrow 0$ and $0 \rightarrow 1 \rightarrow 0$. Thus using the axiom of additivity of probabilities, we have

$$p_{0 \rightarrow 0} = p_{00}p_{00} + p_{01}p_{10}$$

9.2 Explanation

Any (classical) explanation which assumes that the photon takes exactly one path through the interferometer leads to the conclusion that the two detectors should on an average each fire on half the occasions when the experiment is performed. But the experiment shows otherwise.

One thing that is wrong is the assumption that the processes $0 \rightarrow 0 \rightarrow 0$ and $0 \rightarrow 1 \rightarrow 0$ are mutually exclusive. Let us explain the process using quantum mechanics. We have already introduced the concept of *probability amplitudes*

in Section 2.2. The overall probability amplitude for the transition is the sum not of probabilities but of the probability amplitudes of each of the constituent transitions considered separately. Unlike probabilities, probability amplitudes can be both positive and negative and they therefore can cancel out each other. Consider for example the $\sqrt{\text{not}}$ gate. It can be represented by the following unitary transformation:

$$\sqrt{\text{not}} = \frac{1}{2} \begin{pmatrix} 1 - i & 1 + i \\ 1 + i & 1 - i \end{pmatrix}$$

Let q_{ij} denote the probability amplitude for the transition input i , output j .

$$(\sqrt{\text{not}})|0\rangle = \sqrt{\text{not}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}((1 - i)|0\rangle + (1 + i)|1\rangle)$$

$$(\sqrt{\text{not}})|1\rangle = \sqrt{\text{not}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2}((1 + i)|0\rangle + (1 - i)|1\rangle)$$

$$q_{00} = \frac{1}{2}(1 - i) \quad q_{01} = \frac{1}{2}(1 + i)$$

$$q_{10} = \frac{1}{2}(1 + i) \quad q_{11} = \frac{1}{2}(1 - i)$$

It can be easily verified that, $q_{0 \rightarrow 0} = q_{00}q_{00} + q_{01}q_{10} = 0$ and therefore, $p_{0 \rightarrow 0} = \|q_{0 \rightarrow 0}\|^2 = 0$.

If a machine starts in a particular initial configuration (input) then the probability that after its evolution via a sequence of intermediate configurations it ends up in a specific final configuration (output) is the squared modulus of the sum of all the probability amplitudes of the computational paths that connect the input with the output. The amplitudes are complex numbers and may cancel each other, which is called *destructive interference*, or enhance each other, referred to as *constructive interference*. The basic idea of quantum computation is to use quantum interference to amplify the correct outcomes and to suppress the incorrect outcomes of the computations. This will be illustrated in the Deutsch-Jozsa Problem in Section 10.2.

10 Quantum Gates

Just like gates in classical computers we have gates in Quantum Computers. The gates in quantum computation correspond to a transformation on the quantum state. Therefore, the computation performed by the gates must be reversible as the transformations have to be unitary. While the classical NOT gate is reversible, AND, OR and NAND gates are not. Thus it seems, due to this constraint, that it may not be possible to carry out all classical computations using quantum transformations(gates). But this is not the case. Consider a *Toffoli gate*, T that acts on 3 qubits and inverts the third if the first two are 1.

$$\mathbf{T} = |0\rangle\langle 0| \otimes I \otimes I + |1\rangle\langle 1| \otimes C_{not}$$

where, C_{not} is as defined in Section 6.2. We can realise the classical gates using T as follows,

$$\begin{aligned} Not(x) &= T(|1, 1, x\rangle) = |1, 1, \neg x\rangle \\ And(x, y) &= T(|x, y, 0\rangle) = |x, y, x \wedge y\rangle \end{aligned}$$

Deutsch has shown [6] that it is possible to construct reversible quantum gates for any classically computable function.

10.1 Quantum Gate Arrays

Classically, Turing machines and Uniform circuit families are used to model computation. The quantum counterpart of classical uniform circuit families are Quantum Gate Arrays, Figure 6.

In the figure, time flows from left to right. One might think of the particles composing the register as travelling through the different gates. At the right end is the observable that extracts information from the register after it has gone through all the gates. The sequence of A_i 's with observable O is what constitutes a quantum program.

We use Modified Deutsch-Jozsa Problem to illustrate the programming of Quantum Gate Arrays.

10.2 Modified Deutsch-Jozsa Problem (MDJP)

Input : a computable function $f : \{0, 1\}^n \rightarrow \{0, 1\}$

Problem : to answer if f is “non-balanced” or “non-constant”.⁹

Definition 2 A function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ is **non-balanced** if one of the two values of f has a majority.

Definition 3 A function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ is **non-constant** if there exist $x, y \in \{0, 1\}^n$ such that $f(x) \neq f(y)$.

Classically, we will need to potentially evaluate function f on all possible inputs to answer the problem. But we can solve this problem exponentially faster on a quantum computer than on a Turing machine.

Theorem 10.1 (Simon) *There exists an oracle relative to which there is a problem solvable in polynomial time (with bounded error probability) on a*

⁹The original Deutsch-Jozsa Problem dealt with strings rather than functions.

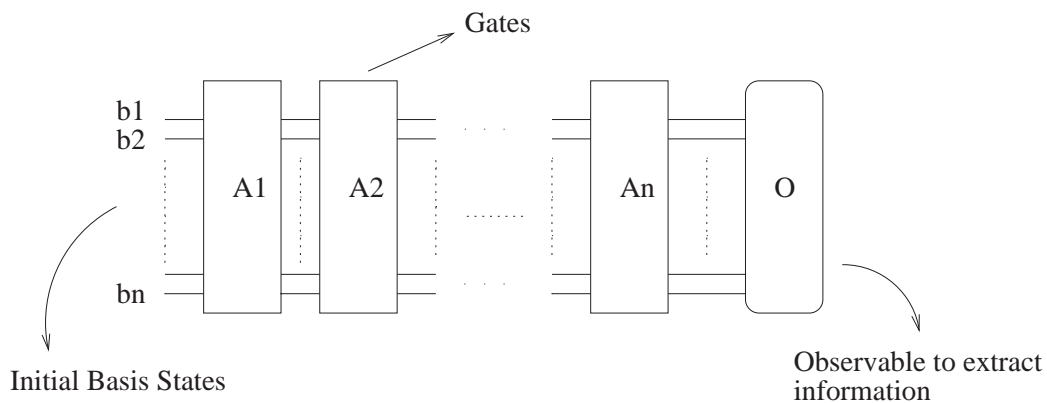


Figure 6: Quantum Gate Arrays

quantum computer, but any probabilistic Turing machine with bounded error probability claiming to solve this problem (using this oracle) will require exponential time on infinitely many inputs.

Simon's theorem is the strongest argument in favour of the superiority of quantum computers over Turing machines.

Theorem 10.2 (Lecerf-Bennett) *For any Turing machine T computing a function f there exists a reversible Turing machine T' computing $\langle x, f(x) \rangle$ on input x and whose running time is within a constant factor of the running time of T . The cost in space is also polynomial in $|x|$, but all the tape cells used in the process of computing $\langle x, f(x) \rangle$ will reset back to zero (reversibly). These tape cells are referred collectively as the workspace.*

Benioff and Deutsch have shown that quantum Turing machines can directly simulate reversible Turing machines. Since quantum Turing machines (and also quantum gate arrays) are reversible¹⁰, we have the following corollary:

Corollary 10.3 *A Turing-computable function f is always computable on a quantum gate array (with a negligible increase in the time complexity).*

Consider the MDJP. The input function is computable so by the Lecerf-Bennett theorem, there exists a reversible Turing machine that computes $\langle x, f(x) \rangle$ on input x . By definition of the problem, x is an n -bit value and $f(x)$ is a single bit. By corollary 10.3, this implies the existence of a unitary matrix F that computes f on n -bit values in the following sense. Consider the quantum gate corresponding to F :

The gate works on an $(n+1+m)$ -qubit register; the top n qubits encode the input $x \in \{0, 1\}^n$. Those qubits must have the same value before and after the gate: if they are changed during the computation itself, they must be returned to their initial value. The next qubit, initially set to b will

¹⁰recall that transformations on a state are always unitary and thus reversible

have the value of $b \oplus f(x)$ at the output of the gate. If we have 0 as the initial value then $f(x)$ appears as the output. The reason for this is that we cannot just overwrite the initial qubit value because we have to do the computations reversibly, that is we should be able to get the initial value from the final value. The last m qubits are the “workspace” that comes about in theorem 10.2. We donot specify the exact circuitry of the elementary gates that are used to make the F gate, but by theorem 10.2 and corollary 10.3 we are certain that it can be done in accordance with the quantum principles. For clarity, we do not usually display the qubits used as workspace since they serve no purpose outside of the gates themselves. The action of the gate F will be denoted as:

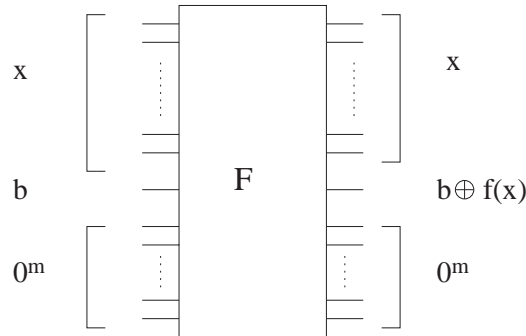
$$F|x, b\rangle = |x, b \oplus f(x)\rangle$$

Consider $S_n = \underbrace{H \otimes H \otimes \dots \otimes H}_{ntimes}$ where H is the *Walsh- Hadamard* transformation. We have,

$$S_n |\underbrace{0 \dots 0}_n\rangle = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i\rangle$$

That is, if the initial state of the register is all zero, S_n tranforms it in a superposition of all 2^n values of the first n qubits. Therefore, we can compute in one application of the gate all possible values for f in quantum superposition (using linearity of the quantum operations) as shown below:

$$F(S_n \otimes I) |\underbrace{0 \dots 0}_n, 0\rangle = F\left(\frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i, 0\rangle\right) = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i, f(i)\rangle$$



Now if we use an observable B to measure the state of the system then we will observe a single pair $\langle x, f(x) \rangle$ where x is chosen uniformly from the 2^n values. Thus we will need on an average an exponential number of runs to obtain all values of f . To get benefit from the superpositions we need to do something more.

Consider the phase inversion gate that multiplies the amplitude by -1 iff the qubit is set to $\mathbf{1}$.

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

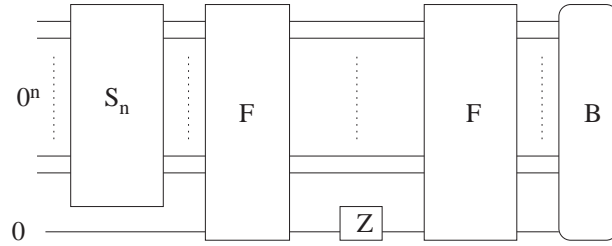


Figure 7: MDJP

In Figure 7 we know from our gate definitions that the value of the register just after Z gate is:

$$|\psi\rangle = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} (-1)^{f(i)} |i, f(i)\rangle$$

When this state goes through the final F gate the values of f are again computed and since $f(i) \oplus f(i) = 0$ for all $i \in \{0, 1\}^n$ we have the final state before observation as,

$$|\phi\rangle = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} (-1)^{f(i)} |i, 0\rangle$$

Thus, we have transferred the values of f into the amplitudes of each of the basis states. The power of quantum computation lies in the interference of these amplitudes and the observable used to read the quantum states.

Define the observable $B = \{E_a, E_b\}$ where E_a is the subspace spanned by:

$$|\psi_a\rangle = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i, 0\rangle$$

and $E_b = (E_a)^\perp = \text{span}\{\psi_b\}$ is the orthogonal complement of E_a . Let α and β be the projections of $|\phi\rangle$ along E_a and E_b then,

$$|\phi\rangle = \alpha|\psi_a\rangle + \beta|\psi_b\rangle$$

$$\begin{aligned} \alpha &= \langle \psi_a | \phi \rangle \\ &= \left(\frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} \langle i, 0 | \right) \left(\frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} (-1)^{f(j)} |j, 0\rangle \right) \\ &= \frac{1}{2^n} \sum_{i=0}^{2^n-1} \sum_{j=0}^{2^n-1} (-1)^{f(j)} \langle i, 0 | j, 0 \rangle \end{aligned}$$

But since $\langle i, 0 | j, 0 \rangle = 1$ iff $i = j$ and zero otherwise, the above expression simplifies to

$$\alpha = \frac{1}{2^n} \sum_{i=0}^{2^n-1} (-1)^{f(i)}$$

If f is a balanced function then $\alpha = 0$, which implies if we observe state ψ_a we can be sure that the function f is ‘**non-balanced**’. If f is a constant function then $\alpha = -1$ or 1 , that is $\|\alpha\|^2 = 1$ and $\|\beta\|^2 = 1 - \|\alpha\|^2 = 0$. In this case $\beta = 0$ and therefore if we observe state ψ_b then we can be sure that the function is ‘**non-constant**’. Therefore we can solve the problem using only two evaluations of the function f , an exponential speed up from the straightforward classical algorithm.

11 Quantum Fourier Transforms

The Quantum Fourier Transform is defined as follows. Consider a number a , $0 \leq a < q$ for some q ,

$$U_{QFT}|a\rangle \rightarrow \frac{1}{q^{1/2}} \sum_{c=0}^{q-1} e^{\frac{2\pi iac}{q}} |c\rangle \quad (5)$$

As all quantum transformations need to be unitary we need to show that U_{QFT} can be written as a unitary matrix. Imagine the transformation matrix such that the rows are indexed by input states and the columns by the output states, eg

$$X = \begin{array}{c|cccc|} & |a\rangle & |b\rangle & |c\rangle & |d\rangle & \\ |a\rangle & \alpha & \beta & \gamma & \delta & \\ |b\rangle & \cdots & \cdots & \cdots & \cdots & \\ |c\rangle & \cdots & \cdots & \cdots & \cdots & \\ |d\rangle & \cdots & \cdots & \cdots & \cdots & \end{array}$$

where a, b, c, d are the basis states. Thus, $X|a\rangle = \alpha|a\rangle + \beta|b\rangle + \gamma|c\rangle + \delta|d\rangle$. Therefore, for Quantum Fourier Transform we want to find a matrix A_q such that the entry $A_q(a,c)$ should be $\frac{1}{q^{1/2}} e^{\frac{2\pi iac}{q}}$. We assume that q is a power of 2. Let $q = 2^l$ and we represent a state $|a\rangle$ (having l bits) as $|a_{l-1}a_{l-2} \dots a_1a_0\rangle$. For quantum fourier transform we need to use only two types of gates. These gates are R_j that operates on the j^{th} bit of a bit vector,

$$R_j = \begin{array}{c|cc|} & |0\rangle & |1\rangle & \\ |0\rangle & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \\ |1\rangle & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \end{array}$$

and $S_{j,k}$, which operates on two bits in the positions j and k with $j < k$,

$$S_{j,k} = \begin{array}{c|cccc|} & |00\rangle & |01\rangle & |10\rangle & |11\rangle & \\ |00\rangle & 1 & 0 & 0 & 0 & \\ |01\rangle & 0 & 1 & 0 & 0 & \\ |10\rangle & 0 & 0 & 1 & 0 & \\ |11\rangle & 0 & 0 & 0 & e^{i\theta_{k-j}} & \end{array}$$

where, $\theta_{k-j} = \pi/2^{k-j}$. We can see that R_j is essentially applying the *Walsh-Hadamard Transformation* on the j^{th} bit. That is it takes it to a

superposition of $|0\rangle$ and $|1\rangle$. The transformation $S_{j,k}$ is just a phase change transformation, it cannot change the bits on which it is being applied. Now consider the following:

$$R_{l-1} S_{l-2,l-1} R_{l-2} S_{l-3,l-1} S_{l-3,l-2} R_{l-3} \dots R_1 S_{0,l-1} S_{0,l-2} \dots S_{0,2} S_{0,1} R_0$$

that is, we apply the gates R_j in reverse order from R_{l-1} to R_0 , and between R_{j+1} and R_j we apply the gates $S_{j,k}$ where $k > j$. For example, on 3 bits, the matrices would be applied in the order $R_2 S_{1,2} R_1 S_{0,2} S_{0,1} R_0$.

A total of l R_j gates are used and $l(l-1)/2$ $S_{j,k}$ gates are used. Thus, we need $l(l+1)/2$ quantum gates¹¹ for the quantum fourier transform.

The claim is that applying the above sequence of transformations will result in a quantum state,

$$|\phi\rangle = \frac{1}{\sqrt{2^l}} \sum_{b=0}^{2^l-1} \exp(2\pi iac/2^l) |b\rangle$$

where b is the bit reversal¹² of c , that is, the binary number obtained by reading the bits of c from right to left. Thus to calculate the actual quantum fourier transform we need to do a further computation to reverse the bits of $|b\rangle$ to obtain $|c\rangle$ (which takes atmost $O(l)$ time) or we can leave these bits in place and read them in reverse order whenever required.

Let us take an example. Consider the QFT of $|01\rangle$ ($l = 2$):

$$\begin{aligned} |01\rangle &= \frac{1}{\sqrt{2^2}} (\exp(2\pi i(1.0)/2^2) |00\rangle + \exp(2\pi i(1.1)/2^2) |01\rangle \\ &\quad + \exp(2\pi i(1.2)/2^2) |10\rangle + \exp(2\pi i(1.3)/2^2) |11\rangle) \\ &= \frac{1}{2} (|00\rangle + e^{i\pi/2} |01\rangle + e^{i\pi} |10\rangle + e^{3i\pi/2} |11\rangle) \end{aligned}$$

The corresponding transformation A_q is $R_1 S_{0,1} R_0$. Consider them one

¹¹We know that each of the these gates take a fixed number of bits as input and therefore take some fixed number of steps to compute the transformation.

¹²that is, if $|c\rangle = |c_{l-1}c_{l-2} \dots c_0\rangle$ then $|b\rangle = |c_0 \dots c_{l-2}c_{l-1}\rangle$

at a time, in order.

$$\begin{aligned}
R_1 |01\rangle &\rightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |1\rangle \\
&= \frac{1}{\sqrt{2}}(|01\rangle + |11\rangle) \\
S_{0,1}\left(\frac{1}{\sqrt{2}}(|01\rangle + |11\rangle)\right) &\rightarrow \frac{1}{\sqrt{2}}(S_{0,1}|01\rangle + S_{0,1}|11\rangle) \\
&= \frac{1}{\sqrt{2}}(|01\rangle + e^{i\pi/2}|11\rangle) \\
R_0\left(\frac{1}{\sqrt{2}}(|01\rangle + e^{i\pi/2}|11\rangle)\right) &\rightarrow \frac{1}{\sqrt{2^2}}\left((|0\rangle \otimes (|0\rangle - |1\rangle))\right. \\
&\quad \left.+ (e^{i\pi/2}|1\rangle \otimes (|0\rangle - |1\rangle))\right) \\
&= \frac{1}{2}(|00\rangle - |01\rangle + e^{i\pi/2}|10\rangle - e^{i\pi/2}|11\rangle) \\
|\psi\rangle &= \frac{1}{2}(|00\rangle + e^{i\pi}|01\rangle + e^{i\pi/2}|10\rangle + e^{3i\pi/2}|11\rangle) \\
\text{reverse - bits}(|\psi\rangle) &\rightarrow \frac{1}{2}(|00\rangle + e^{i\pi/2}|01\rangle + e^{i\pi}|10\rangle + e^{3i\pi/2}|11\rangle)
\end{aligned}$$

which is the required QFT of $|01\rangle$.

To show that A_q performs the required transform consider the entry $A_q(a, c)$ which gives the amplitude of going from $|a\rangle = |a_{l-1} \dots a_0\rangle$ to $|b\rangle = |b_{l-1} \dots b_0\rangle$ where b is the reverse of c ¹³. First, all the $1/\sqrt{2}$ multiply in the R matrices to produce a factor of $1/\sqrt{2^l}$, overall. Thus we need to worry only about the $\exp(2\pi iac/2^l)$ phase factor in the expression 5. The matrices $S_{j,k}$ do not change the values of any of the bits, but merely change the phases. There is thus only one way to switch the j^{th} bit from a_j to b_j , and that is to use the appropriate entry in R_j , e.g. if we want to switch j^{th} bit from $|0\rangle$ to $|1\rangle$ then we will look at entry (0,1) of R_j . This adds a phase of π if both a_j and b_j are 1 and leaves it unchanged otherwise. Further, the matrix $S_{j,k}$ adds $\pi/2^{k-j}$ to the phase if a_j and b_k are both 1 and leaves it unchanged otherwise. See Figure 8, where a phase factor of $\exp(i\pi/2)$ is introduced with $|11\rangle$.

¹³Note that applying A_q on a quantum state will return a superposition as we saw in the example. We are concentrating here on one of the vectors of the superposition and calculating the corresponding amplitude associated with it.

$$S_{0,1}(|01\rangle + |11\rangle) \longrightarrow (|01\rangle + \exp(i\pi/2)|11\rangle)$$

Figure 8: Example

When we apply $S_{j,k}$ for $j < k$, then we have converted all the bits $(a_{l-1} \dots a_{j+1})$ to $(b_{l-1} \dots b_{j+1})$. Further applications of $S_{j,k}$ do not modify this bits. Thus, b_k should be 1 for this phase to be introduced. So the total phase introduced on the path from $|a\rangle$ to $|b\rangle$ is:

$$\phi = \phi_R + \phi_S = \sum_{0 \leq j < l} \pi a_j b_j + \sum_{0 \leq j < k < l} \frac{\pi}{2^{k-j}} a_j b_k$$

which can also be written as

$$\phi = \sum_{0 \leq j \leq k < l} \frac{\pi}{2^{k-j}} a_j b_k$$

Since c is the bit reversal of b , i.e., $b_k = c_{l-1-k}$, this can be further written as

$$\phi = \sum_{0 \leq j \leq k < l} \frac{\pi}{2^{k-j}} a_j c_{l-k-1}$$

Making the substitution $l - p - 1$ for k in the sum, we have

$$\phi = \sum_{0 \leq j+k < l} 2\pi \frac{2^j 2^k}{2^l} a_j c_k$$

Now if we also all such j, k which have $j+k \geq l$ then we end up adding multiples of 2π to the phase ϕ and therefore, the overall phase remains unaltered. Thus,

$$\phi = \sum_{0 \leq j, k=0}^{l-1} 2\pi \frac{2^j 2^k}{2^l} a_j c_k = \frac{2\pi}{2^l} \sum_{j=0}^{l-1} 2^j a_j \sum_{k=0}^{l-1} 2^k c_k$$

where the last equality follows from the distributive law of multiplication.

$$a = \sum_{j=0}^{l-1} 2^j a_j, \quad c = \sum_{k=0}^{l-1} 2^k c_k$$

and so we can write,

$$\phi = 2\pi ac/2^l,$$

which is the phase for the amplitude of the transformation $|a\rangle \rightarrow |b\rangle$.

12 Grover's Search Algorithm

A large class of problems can be specified as search problems of the form “find some x in a set of possible solutions such that statement $P(x)$ is true”. Such problems range from database search to sorting. A sorting problem can be viewed as a search for a permutation for which the statement “the permutation x takes the initial state to the desired sorted state” is true.

An *unstructured* search problem is one where nothing is known (or no assumptions are used) about the structure of the solution space and the statement P . For example, determining $P(x_0)$ provides no information about the possible value of the possible value of $P(x_1)$ for $x_0 \neq x_1$. A *structured* search problem is one where information about the search space and statement P can be exploited. For instance, searching an alphabetized list is a structured search problem and the structure can be exploited to construct efficient algorithms. Consider the problem of searching for a phone number in an unsorted directory. This is an example of unstructured search problem. In order to find someone's phone number with a probability of $\frac{1}{2}$, any classical algorithm (whether deterministic or probabilistic) will need to look at minimum of $\frac{N}{2}$ names. In the general case of an unstructured problem, randomly testing the truth of statements $P(x_i)$ one by one is the best that can be done classically. For a search space of size N , the general unstructured search problem requires $O(N)$ evaluations of P (classically). On a quantum computer, however, Grover showed that the unstructured search problem can be solved with bounded probability within $O(\sqrt{N})$ evaluations of P . Grover's search algorithm [4] has been proved to be more efficient than any algorithm that could run on a classical computer.

A good starting point to think of quantum mechanical algorithms is probabilistic algorithms. In these algorithms, instead of having the system in a specified state, it is in a distribution over various states with a certain probability of being in each state. At each step, there is a certain probability of being in each state. Quantum mechanical algorithms work with a probability distribution over various states. However, unlike classical systems, the probability vector does not completely describe the system. As we have seen in Section 9 (Quantum Interference), in order to completely describe the system we need the *amplitude* in each state which is a complex number. We can have a state in quantum mechanics as $|\psi\rangle = \alpha|x\rangle + \beta|y\rangle$. The first thing

that comes to mind is that this is just like a classical state, which exists in $|x\rangle$ or $|y\rangle$ with certain probability. But this is not the case. The difference is that both $|x\rangle$ and $|y\rangle$ occur simultaneously which can lead to interference as we have already seen.

12.1 The Problem

Let a system have $N = 2^n$ states which are labelled S_1, S_2, \dots, S_N . These 2^n states are represented as n bit strings. Let there be a unique state, say S_v , that satisfies the condition $P(S_v) = 1$, whereas for all other states, S , $P(S) = 0$ (assume that for any state S , the condition $P(S)$ can be evaluated in unit time). The problem is to identify the state S_v .

12.2 Algorithm

The algorithm involves the following steps-

1. Prepare a register containing a superposition of all possible values of $x_i \in \{0, 1, \dots, 2^n - 1\}$.
2. Compute $P(x_i)$ for the register. We have a gate array U_P such that $U_P |x, 0\rangle \rightarrow |x, P(x)\rangle$.
3. Now out of the 2^n possible states for only one state will $P(x_i)$ be 1 and we want that state. Apply transformations on this superposition so that the probability of observing the required state increases. This involves two operations, which are repeated for some number of steps:
 - (a) Applying the Z transform which inverts the amplitude if the state is a 1 (done earlier).
 - (b) Inversion about the average
4. Observe the final state. If we have $|x, 1\rangle$ then x is the required element else repeat.

12.2.1 Inversion about average

Suppose there are N numbers X_0, X_1, \dots, X_N and their average is A . Then inversion about average of X_i means we want an X'_i such that if $X_i = A + x$ then $X'_i = A - x$ that is $X'_i = 2A - X_i$.

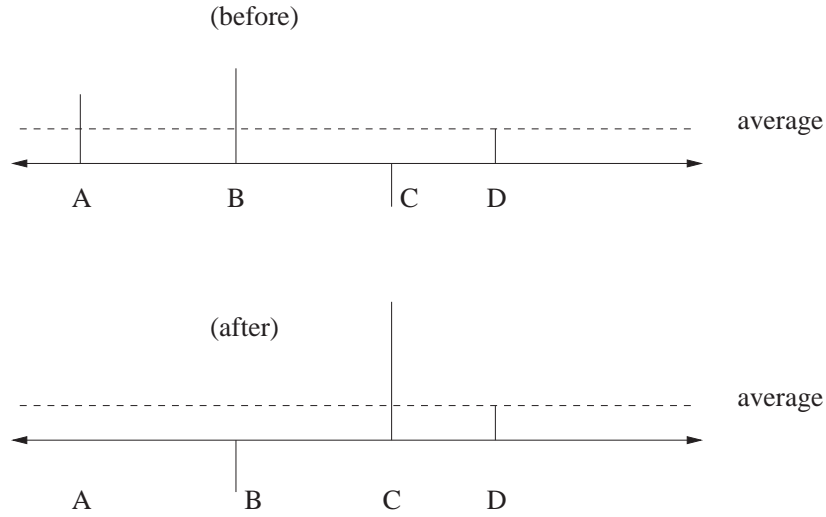


Figure 9: Inversion about average operation

The diffusion transform, D , is defined as follows:

$$D_{ij} = \begin{cases} \frac{2}{N} & \text{if } i \neq j \\ -1 + \frac{2}{N} & \text{if } i = j \end{cases}$$

Next we prove that D is the *inversion about average* as shown in figure 11 and that D is *unitary*.

Observe that D can be written as $D = -I + 2P$ where I is the identity matrix and P is a matrix with $P_{ij} = \frac{1}{N}$ for all i, j . The following two properties of P can be easily verified-

1. $P^2 = P$
2. P acting on any vector \vec{v} gives a vector each of whose components is equal to the average of all the components.

Using the fact that $P^2 = P$, it follows immediately from the representation $D = -I + 2P$ that $D^2 = I$ and hence D is unitary. In order to see that D is the *inversion about average*, consider what happens when D acts on an arbitrary vector \vec{v} . Expressing D as $-I + 2P$, it follows that

$$D\vec{v} = (-I + 2P)\vec{v} = -\vec{v} + 2P\vec{v}.$$

Now, each component of the vector $P\vec{v}$ is A where A is the average of all the components of vector \vec{v} . Therefore, the i^{th} component of the vector $D\vec{v}$ is given by $(-v_i + 2A)$ which is the inversion about average. Consider what happens when the *inversion about average* operation is applied to a vector where each of the components, except one, are equal to a value, say C , which is approximately equal to $\frac{1}{\sqrt{N}}$. The one component that is different is negative (due to the application of Z). The average A is approximately equal to C . Since each of the $(N - 1)$ components are equal to the average, it does not change significantly as a result of inversion about average. The one component that was negative to start out, now becomes positive and its magnitude increases by approximately $3C$, and becomes approximately $\frac{2}{\sqrt{N}}$. See Figure 10.

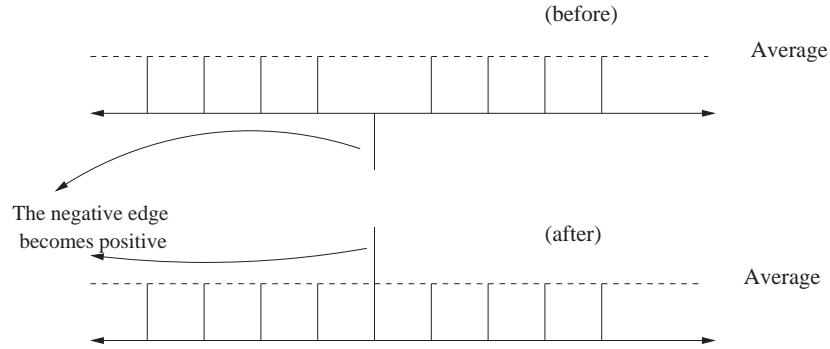


Figure 10: Inversion about average operation is applied to a distribution in which all but one are $O(\frac{1}{\sqrt{N}})$ and the one component is negative

In the loop of step of 3 first the amplitude of the required state is inverted(Z). Then the *inversion about the average* operation is carried out. This increases the amplitude of the required state in each iteration by $O(\frac{1}{\sqrt{N}})$.

12.3 Proofs

We showed in the previous section that D is unitary. Here we prove that D can be implemented as a sequence of three local quantum mechanical state transition matrices. It can be decomposed into $O(n) = O(\log(N))$ elementary quantum gates. We can write D as $D = WRW$ where W is the Walsh-Hadamard transform and

$$R = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & -1 \end{pmatrix}$$

To see that $D = WRW$, consider $R = R' - I$ where I is the identity and

$$R' = \begin{pmatrix} 2 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

Now, $WRW = W(R' - I)W = WR'W - I$. It is easily verified that ,

$$WR'W = \begin{pmatrix} \frac{2}{N} & \frac{2}{N} & \dots & \frac{2}{N} \\ \frac{2}{N} & \frac{2}{N} & \frac{2}{N} & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{2}{N} & \dots & \dots & \frac{2}{N} \end{pmatrix}$$

and thus, $WRW = WR'W - I = D$.

Theorem 12.1 *Let the state vector be as follows - for any one state the amplitude is k_1 , for each of the remaining $(N - 1)$ states the amplitude is l_1 . Then after applying the diffusion transform D , the amplitude in the one state is $k_2 = \left(\frac{2}{N} - 1\right)k_1 + 2\frac{(N-1)}{N}l_1$ and the amplitude in each of the remaining $(N - 1)$ states is $l_2 = \frac{2}{N}k_1 + \frac{(N-2)}{N}l_1$.*

Proof

Using the definition of the diffusion transform D , it follows that

$$\begin{aligned} k_2 &= \left(\frac{2}{N} - 1\right)k_1 + 2\frac{(N-1)}{N}l_1 \\ l_2 &= \left(\frac{2}{N} - 1\right)l_1 + \frac{2}{N}k_1 + \frac{2(N-2)}{N}l_1 \end{aligned}$$

Therefore:

$$l_2 = \frac{2}{N}k_1 + \frac{(N-2)}{N}l_1$$

Corollary 12.2 *Let the state vector be as follows - for any one state the amplitude is k , for each of the remaining $(N-1)$ states the amplitude is l . Let k and l be real numbers (in general the amplitudes can be complex). Let k be negative and l be positive and $|\frac{k}{l}| < \sqrt{N}$. Then after applying the diffusion transform both k_1 and l_1 are positive numbers.*

Corollary 12.3 *Let the state vector be as follows- for the state satisfies $P(S) = 1$, the amplitude is k , for each of the $(N-1)$ states the amplitude is l . Then if after applying the diffusion transformation D , the new amplitudes are respectively k_1 and l_1 as derived in theorem 12.1, then*

$$k_1^2 + (N-1)l_1^2 = k^2 + (N-1)l^2$$

Theorem 12.4 *Let the state vector before the first step of the iteration be as follows - for the one state that satisfies $P(S) = 1$, the amplitude is k , for each of the remaining $(N-1)$ states the amplitude is l such that $\left(0 < k < \frac{1}{\sqrt{2}}\right)$ and $l > 0$. The change in k (Δk) after the two steps of the iteration in step 3 of the algorithm is lower bounded by $\Delta k > \frac{1}{2\sqrt{N}}$. Also after these steps, $l > 0$.*

Proof- Denote the initial amplitudes by k and l , the amplitudes after the phase inversion by k_1 and l_1 and after the diffusion transform by k_2 and l_2 . Using Theorem 12.1, it follows that:

$$k_2 = \left(1 - \frac{2}{N}\right)k + 2\left(1 - \frac{1}{N}\right)l$$

where,

$$k_1 = -k \text{ and } l_1 = l$$

Therefore,

$$\Delta k = k_2 - k = -\frac{2k}{N} + 2\left(1 - \frac{1}{N}\right)l$$

Since $0 < k < \frac{1}{\sqrt{2}}$, it follows from corollary 12.3 that $|l| > \frac{1}{\sqrt{2N}}$ and since by the assumption in this theorem, l is positive, it follows that $l > \frac{1}{\sqrt{2N}}$. Thus,

$$\Delta k > -\frac{\sqrt{2}}{N} + \sqrt{2}\left(1 - \frac{1}{N}\right)\frac{1}{\sqrt{N}}$$

$$\Delta k > \frac{\sqrt{2}}{\sqrt{N}}\left(1 - \frac{1}{N} - \frac{1}{\sqrt{N}}\right)$$

Now, for $N > 9$, $\frac{1}{N} + \frac{1}{\sqrt{N}} < \frac{4}{9} < \frac{1}{2}$

Thus,

$$\Delta k > \frac{\sqrt{2}}{\sqrt{N}}\left(1 - \frac{1}{2}\right) = \frac{1}{\sqrt{2N}}$$

Thus, applying $O(\sqrt{N})$ iterations of step 3 we can get the required element with probability more than $1/2$ (because, after \sqrt{N} iterations the amplitude will increase by at least $1/\sqrt{2}$).

13 Shor's Factoring Algorithm

Every integer n has a unique decomposition into prime factors. However, finding this decomposition when n is large is a difficult problem. The faith in the hardness of this problem is so much that many cryptographic algorithms are based on it. It has been shown that one could solve the factoring problem if one could find the order of an element. Given x and n , find r (called the *order*) such that $x^r \equiv 1 \pmod{n}$. As with the factoring problem no algorithm is known for solving this problem.

Program (for Factorisation):

Input : an input n odd integer

$x \leftarrow \text{random}\{0, \dots, n\}$

$r \leftarrow$ use the oracle to find the order of $x \pmod{n}$

Output: if r is odd or $x^{r/2} \equiv -1 \pmod{n}$ then fail
else return $\text{gcd}(x^{r/2} - 1, n)$

13.1 Finding Order of an element

We now describe Shor's algorithm to find the order r of an element $x \pmod{n}$. We choose m qubits such that $n \leq q = 2^m \leq 2n^2$. We know that modular exponentiation can be done in polynomial time. So there exists a quantum gate, F , that efficiently implements this operation. Hence given an input $|a, \mathbf{0}\rangle$ the gate returns $|a, x^a \pmod{n}\rangle$.

The two steps in the algorithm are as follows:

1. First, we drive the system of m qubits into a superposition of all the possible states. We apply the gate F on the state. Thus we have the state,

$$\psi = \frac{1}{\sqrt{q}} \sum_{j=0}^{q-1} |j, x^j \pmod{n}\rangle$$

We measure the part of the superposition that contains the value $x^j \pmod{n}$. Let us observe the value y . The function $f(a) = x^a \pmod{n}$ is periodic with period r since $x^r \equiv 1 \pmod{n}$. Hence the system will collapse to a superposition of all those states j such that $y = f(j)$.

There will be q/r such values of j (Assuming q is a multiple of r). Thus we have the state,

$$\psi' = \sqrt{\frac{r}{q}} \sum_{j=0}^{\frac{q}{r}-1} |jr + l\rangle$$

where l is the smallest value such that $y = f(l)$.

2. Next we apply Quantum Fourier Transform on ψ' . For each of the basis states,

$$QFT(|jr + l\rangle) \rightarrow \frac{1}{\sqrt{q}} \sum_{c=0}^{q-1} e^{\frac{2\pi i(jr+l)c}{q}} |c\rangle$$

Hence we have,

$$\begin{aligned} \phi = QFT(\psi') &= \frac{\sqrt{r}}{q} \sum_{j=0}^{\frac{q}{r}-1} \sum_{c=0}^{q-1} e^{\frac{2\pi i(jr+l)c}{q}} |c\rangle \\ &= \frac{\sqrt{r}}{q} \sum_{c=0}^{q-1} e^{\frac{2\pi ilc}{q}} \sum_{j=0}^{\frac{q}{r}-1} e^{\frac{2\pi ijr c}{q}} |c\rangle \end{aligned}$$

Now,

$$\sum_{j=0}^{M-1} e^{\frac{2\pi ijc}{M}} = \begin{cases} M & \text{if } c \text{ is a multiple of } M \\ 0 & \text{otherwise} \end{cases}$$

where $M = q/r$. We have,

$$\phi = \sum_{c=0}^{q-1} \alpha_c |c\rangle$$

where,

$$\alpha_c = \begin{cases} \frac{1}{\sqrt{r}} e^{2\pi ilc/q} & \text{if } c \text{ is a multiple of } M \\ 0 & \text{otherwise} \end{cases}$$

Therefore,

$$\phi = \sum_{j=0}^{r-1} e^{2\pi ilc/q} |j \cdot \frac{q}{r}\rangle$$

Now measuring the state ϕ returns a value $c = \lambda q/r$ for some $\lambda \in \{0, \dots, r-1\}$. If $\gcd(\lambda, r)=1$ then we can determine r by cancelling c/q to an irreducible fraction. The number of primes less than or equal to N is $N/\log N$ for large values of N . Thus number of co-primes to r , less than or equal to r will be,

$$\text{co-primes}(r) \geq r/\log r$$

Hence, probability that $\gcd(\lambda, r) = 1$ is greater than $1/\log r$. Hence we may have to repeat the algorithm atmost $O(\log r) < O(\log N)$ times to successfully find the order r .

14 Simon's Problem.

We use the finite field Z_n^2 . Further, we can define, $x+y$ as the bitwise addition modulo 2 and $x.y$ the inner product.

Input: a function $f: Z_n^2 \rightarrow Z_n^2$ such that f is two-to-one and there exists a u such that $f(x) = f(x + u)$ for all $x \in Z_n^2$.

Problem: To find u .

The problem can be easily generalised to $r: 1$ functions. The *collision problem* is to find a collision in f under the promise that there is one.

This problem is of interest in cryptology, because some cryptographic protocols are based on the difficulty of finding collisions in hash functions.

Simon's algo uses the following circuit:

The basic idea is to take the initial system of n qubits to a superposition of all possible 2^n states. Then we apply the function f and measure the *2nd* qubit. Then when we measure the first qubit, we get a $y \in Z^n$ that is orthogonal to u .

After the C_f , the state is

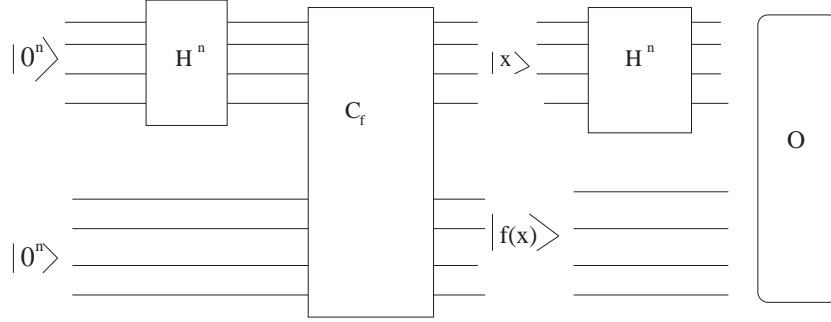


Figure 11: Simon's algorithm

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

Let $|f(z)\rangle$ be the result of measuring $|f(x)\rangle$. Now there are two x st $f(x) = f(z)$: one is z and another is $z + u$. Thus the state now is:

$$\frac{1}{\sqrt{2}} |z\rangle |f(z)\rangle + \frac{1}{\sqrt{2}} |z + u\rangle |f(z)\rangle.$$

Now if we apply H_n to the first n qubits, we get:

$$\begin{aligned} & H_n \left(\frac{1}{\sqrt{2}} |z\rangle + \frac{1}{\sqrt{2}} |z + u\rangle \right) \\ &= \frac{1}{\sqrt{2}} \sum_{y=0}^{2^n-1} \frac{(-1)^{z \cdot y}}{2^{n/2}} |y\rangle + \frac{1}{\sqrt{2}} \sum_{y=0}^{2^n-1} \frac{(-1)^{(z+u) \cdot y}}{2^{n/2}} |y\rangle \\ &= \frac{1}{2^{n/2} \sqrt{2}} \sum_{y=0}^{2^n-1} (-1)^{x \cdot y} (1 + (-1)^{u \cdot y}) |y\rangle. \end{aligned}$$

Now, if $u \cdot y = 1$, the amplitude of $|y\rangle$ is 0. All $|y\rangle$ such that $u \cdot y = 0$ have the same non-zero amplitudes. Hence the output of the measurement is a random y such that $y \cdot u = 0$. Hence, if we get $n - 1$ linearly independent vectors y st $u \cdot y = 0$, we can solve for u . Also the probability that we will get $n - 1$ independent vectors is very high.

Thus, if we solve for u , then after a few iterations, we will get the right u .

15 Quantum Circuits - some construction issues

15.1 Controlled Operations

The operations of the type "If A, then do B", are one of the most useful operations in both classical and quantum computations. In this section we see how to implement such operations from elementary operations.

First of all let us define the controlled operations more formally in the framework of quantum mechanics. The simplest and most common example is the controlled-NOT(CNOT) gate. It is a quantum gate with 2 input qubits, known as the *control* and *target* qubit respectively. In terms of the basis vectors, the action of CNOT is given by $|c\rangle|t\rangle \rightarrow |c\rangle|t \otimes c\rangle$, i.e. if(*c*) then invert *t*. In the computational basis $|control, target\rangle$ the matrix representations of CNOT is:

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

Diagrammatically we denote the CNOT operation as shown in fig 1.

More generally, we can define a controlled-U operations for any arbitrary

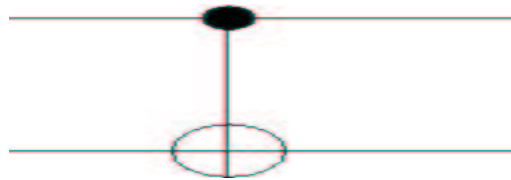


Figure 12: *Circuit representations of CNOT gate. The top qubit is the control qubit, while the bottom qubit is the target qubit.*

unitary operation U. The controlled-U operation is a two qubit operations which acts on a control qubit, and a target qubit. If the control qubit is

set, then U is applied to the target qubit, otherwise the target qubit remains unchanged, i.e. $|c\rangle|t\rangle \rightarrow |c\rangle U^c|t\rangle$. The circuit representation of controlled- U gate is shown in fig 2.

We now proceed to show how to implement any arbitrary controlled- U op-

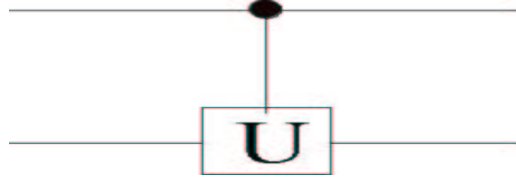


Figure 13: *Circuit representations of controlled- U gate. The top qubit is the control qubit, while the bottom qubit is the target qubit.*

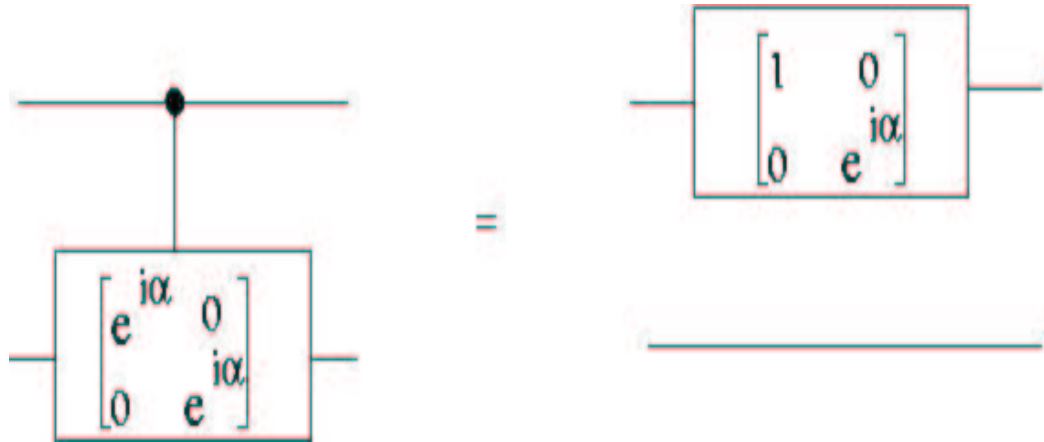
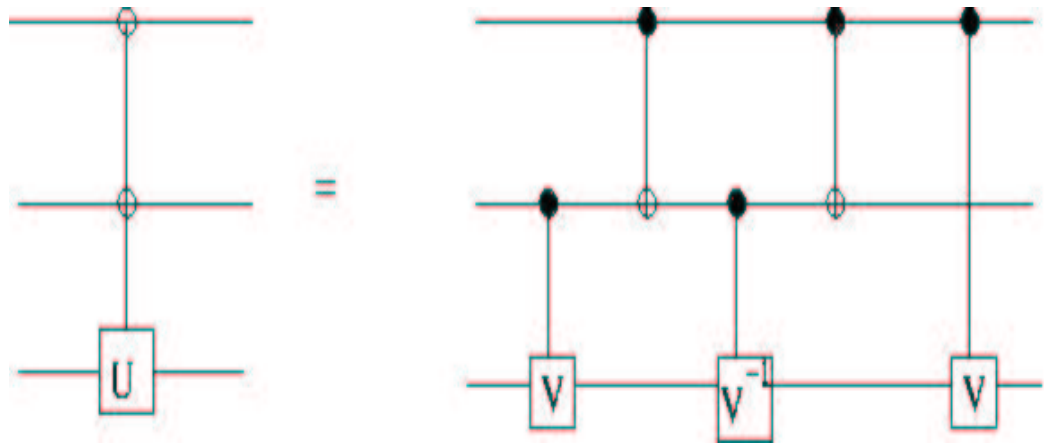
eration for any arbitrary single qubit operation, U , using only single qubit operations and the CNOT gate. We will use the theorem that any arbitrary single qubit operation can be decomposed as $U = e^{i\alpha} \mathbf{A} \mathbf{X} \mathbf{B} \mathbf{X} \mathbf{C}$, where \mathbf{A} , \mathbf{B} , and \mathbf{C} are single qubit operations, such that $\mathbf{A} \mathbf{B} \mathbf{C} = \mathbf{I}$, and \mathbf{X} is the normal *not* operation.

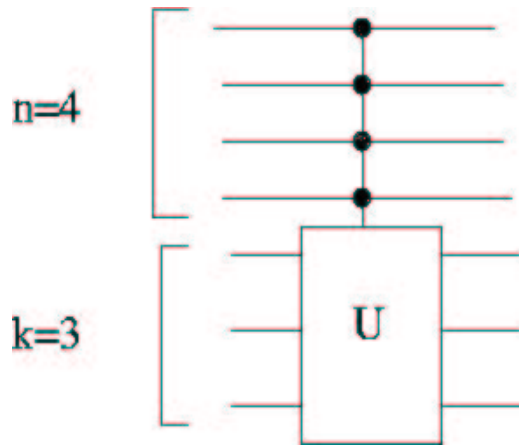
First of all we see how to imlement the phase shift operator $e^{i\alpha}$ on the targer qubit, controlled by the control qubit, using just a single qubit unitary gate. A circuit implementing this operation is shown in fig 3.

The complete construction of the controlled- U operation is shown in figure 4. It is easy to see that if the control qubit is set, then the operation $e^{i\alpha} \mathbf{A} \mathbf{X} \mathbf{B} \mathbf{X} \mathbf{C} = U$, is applied to the second qubit, otherwise the operation $\mathbf{A} \mathbf{B} \mathbf{C} = \mathbf{I}$, is applied to the second qubit, i.e. no change is made.

We now extend the above to conditioning on multiple qubits. Suppose that we have $n+k$ qubits, and U is a k , qubit unitary operator. Then the controlled operation $C^n|x_1x_2\dots x_n\rangle|\psi\rangle = |x_1x_2\dots x_n\rangle U^{x_1x_2\dots x_n}|\psi\rangle$, where x_1, x_2, \dots, x_n in the exponent of U means the product of the bits x_1, x_2, \dots, x_n . This means that if the last n qubits are all set, only then the operation U , is applied to the last k qubits, else nothing happens. The representation of such a operation is represented in fig 5. We here will deal with only the case for $k = 1$.

First of all we see how we can implement the *Toffoli Gate*. Consider a general single qubit operation U , such that there is a unitary operator V , so that $V^2 = U$. Then the operator C^2U may be implemented using the circuit shown in fig 6. The toffoli gate i.e. C^2X can be implemented as a special

Figure 14: *Implementation of the controlledPhaseShift Gate*Figure 15: *Circuit representations of CNOT gate*

Figure 16: *Circuit representations of CNOT gate*

case of the C^2U operation, where the corresponding $V = (1 - i)(I + iX)/2$.

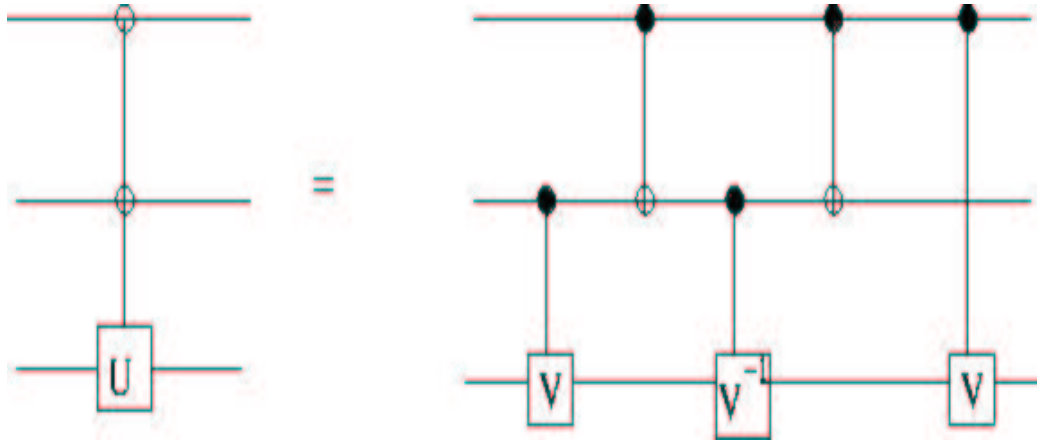
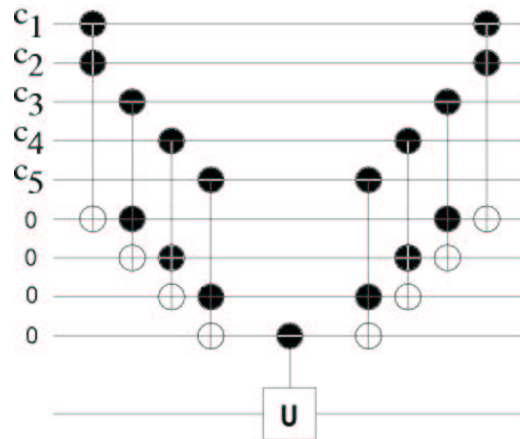
Now we can implement the C^nU gates now using the *Toffoli Gate* and the single qubit U . A particular case for $n = 5$, is shown in figure 7. The circuit can be divided into three stages:

1. The first stage reversible *ANDs*, the control bits c_1, \dots, C_n .
2. Next U , is applied to the target qubit, conditioned on the bit obtained by the *AND* of c_1, \dots, C_n in the previous step.
3. Finally the steps of the first stage are reversed, returning all the work qubits to their original state, $|0\rangle$.

The final result is therefore C^nU , as desired.

15.2 Universal Quantum Gates

We know that a small set of gates (AND, OR, NOT) are universal for classical computation. Infact since Toffoli gate can simulate all the universal classical gates, quantum circuits subsume classical circuits. A similar universality result is true for quantum computation. We will describe three universality constructions for quantum computation. These constructions build upon each other and culminate in a proof that any unitary operation

Figure 17: *Circuit representations of CNOT gate*Figure 18: *Circuit representations of CNOT gate*

can be approximated to arbitrary accuracy using Hadamard, phase, CNOT and $\pi/8$ gates.

The first construction shows that an arbitrary unitary operator may be expressed exactly as a product of unitary operators that each acts non-trivially on a subspace spanned by two basis states. The second construction combines the first construction with results of controlled operations to show that an arbitrary unitary operator may be expressed exactly using single qubit and CNOT gates. The third construction combines the second construction with a proof that single qubit operation may be approximated to arbitrary accuracy using Hadamard, phase and $\pi/8$ gates.

15.2.1 Two-level unitary gates are universal

Two-level unitary matrices

Unitary matrices which act non-trivially only on two or fewer vector components are called Two-level unitary matrices.

Let U be unitary matrix acting on a d -dimensional Hilbert space. We want to decompose it in to products of two-level unitary matrices. Let us try to understand the basic idea by considering the case when U is 3×3 . Suppose U is

$$\begin{pmatrix} a & d & g \\ b & e & h \\ c & f & j \end{pmatrix}$$

We will find two-level unitary matrices U_1, U_2, U_3 such that

$$U_3 U_2 U_1 U = I \tag{6}$$

It follows that

$$U = U_1^{-1} U_2^{-1} U_3^{-1} \tag{7}$$

Since U_1, U_2, U_3 are two-level unitary matrices so $U_1^{-1} U_2^{-1} U_3^{-1}$ are also two-level unitary matrices. Let us construct U_1, U_2, U_3 .

if $b = 0$ then set $U_1 = I$. Otherwise set U_1 as

$$U_1 = \begin{pmatrix} \frac{a^*}{\sqrt{|a|^2+|b|^2}} & \frac{b^*}{\sqrt{|a|^2+|b|^2}} & 0 \\ \frac{b}{\sqrt{|a|^2+|b|^2}} & \frac{-a}{\sqrt{|a|^2+|b|^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$U_1U = \begin{pmatrix} a' & d' & g' \\ 0 & e' & h' \\ c' & f' & j' \end{pmatrix}$$

Note that middle entry in left hand column is zero.

Now we apply similar procedure to find U_2 such that U_2U_1U has no entry in the bottom left corner. If $c' = 0$ we set

$$U_2 = \begin{pmatrix} a'^* & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Otherwise we set

$$U_2 = \begin{pmatrix} \frac{a'^*}{\sqrt{|a'|^2+|c'|^2}} & 0 & \frac{c'^*}{\sqrt{|a'|^2+|c'|^2}} \\ 0 & 1 & 0 \\ \frac{c'}{\sqrt{|a'|^2+|c'|^2}} & 0 & \frac{-a'}{\sqrt{|a'|^2+|c'|^2}} \end{pmatrix}$$

In either case

$$U_2U_1U = \begin{pmatrix} 1 & d'' & g'' \\ 0 & e'' & h'' \\ 0 & f'' & j'' \end{pmatrix}$$

Since U_2U_1U is also unitary, so $d'' = g'' = 0$. This is because first row of U_2U_1U should also have norm 1. Now set U_3 as

$$U_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e''^* & f''^* \\ 0 & h''^* & j''^* \end{pmatrix}$$

Note that $U_3U_2U_1U = I$ and thus $U = U_1^{-1}U_2^{-1}U_3^{-1}$ is the required decomposition of U in to two level unitaries.

Generalizing, suppose U acts on a d -dimensional space. Then in fashion similar to 3×3 case, we can find two-level unitary matrices U_1, \dots, U_{d-1} such that the matrix $U_1, \dots, U_{d-1}U$ has a one in the top left hand corner and all zeros elsewhere in first row and column. We then repeat the procedure for the $d-1$ by $d-1$ unitary submatrix in the lower right hand corner of $U_{d-1}U_{d-2} \dots U_1U$ and so on. The end result is that arbitrary $d \times d$ unitary matrix may be written

$$U = V_1 \dots V_k$$

where matrices V_i are two level matrices and $k \leq (d-1) + (d-2)\dots + 1 = d(d-1)/2$.

Thus arbitrary unitary matrix on a n qubit system can be written as product of atmost $2^{n-1}(2^n - 1)$ two level unitary matrices. Also the bound is tight.

15.2.2 Single qubit and CNOT gates are universal

We will shpw that single qubit and CNOT gates together can be used to implement an arbitrary two-level unitary operation on the state space of n qubits. Combining the results we see that single qubit and CNOT gates can be used to implement an arbitrary unitary operation on n qubits.

Let U is a two-level unitary matrix on n qubit system. Let U acts non trivially on the space spanned by states $|s\rangle$ and $|t\rangle$. Let \tilde{U} be the non-trivial 2×2 unitary submatrix of U ; \tilde{U} can be thought of as a unitary operator on a single qubit.

We want to construct a circuit implementing U . For this Gray Codes are needed. Suppose s and t are distinct binary numbers. A Gray Code connecting s and t is a sequence of binary numbers, starting with s and concluding with t such that adjacent members of the list differ in exactly one bit. For example with $s=101001$ and $t=110011$ we have the Gray code

$$\begin{array}{cccccc} 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{array}$$

The basic idea of quantum circuit implementing U is to perform a sequence of gates effecting the state changes $|g_1\rangle \rightarrow |g_2\rangle \rightarrow \dots \rightarrow |g_{m-1}\rangle$, then to perform a controlled- \tilde{U} operation, with the target bit located at single bit where $|g_{m-1}\rangle$ and $|g_m\rangle$ differ, and then to undo the first stage, transforming $|g_{m-1}\rangle \rightarrow |g_{m-2}\rangle \rightarrow \dots \rightarrow |g_1\rangle$.

The first step is to swap the states $|g_1\rangle$ and $|g_2\rangle$. Suppose g_1 and g_2 differ at i^{th} bit. Then we accomplish the swap by performing a controlled bit flip on the i^{th} qubit, conditional on the values of other qubits being identical to those in both $|g_1\rangle$ and $|g_2\rangle$. Next we use a controlled operation to swap g_2 and g_3 . We continue in this fashion until we swap $|g_{m-2}\rangle$ with $|g_{m-1}\rangle$. All other computational basis states are left unchanged by this sequence of operations. Next, suppose $|g_{m-1}\rangle$ and $|g_m\rangle$ differ in the j th bit. We apply a

controlled- \tilde{U} operation with the j th qubit as target, conditional on the other qubits having the same values as appear in both $|g_m\rangle$ and $|g_{m-1}\rangle$. Finally, we complete the U operation by undoing the swap operations: we swap $|g_{m-1}\rangle$ with $|g_{m-2}\rangle$, then $|g_{m-2}\rangle$ with $|g_{m-3}\rangle$ and so on, until we swap $|g_2\rangle$ with $|g_1\rangle$. Let us take a simple example. Suppose we wish to implement

$$U = \begin{pmatrix} a & 0 & 0 & 0 & 0 & 0 & 0 & c \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ b & 0 & 0 & 0 & 0 & 0 & 0 & d \end{pmatrix}$$

Here \tilde{U} is

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

U acts non-trivially on the states $|000\rangle$ and $|111\rangle$. Gray code connecting these two states are

A	B	C
0	0	0
0	0	1
0	1	1
1	1	1

The corresponding circuit is as shown in the figure. The first two gates shuffle the states so that $|000\rangle$ gets swapped with $|011\rangle$. Next, the operation \tilde{U} is applied to the first qubit of the states $|011\rangle$ and $|111\rangle$, conditional on second and third qubits being in the state $|11\rangle$. Finally, we unshuffle the states, ensuring that $|011\rangle$ gets swapped back with the state $|000\rangle$.

For general case, we see that implementing U requires at most $2(n-1)$ controlled operations to swap $|g_1\rangle$ with $|g_{m-1}\rangle$ and then back again. Each of these controlled operations can be realized using $O(n)$ single qubit and CNOT gates; the controlled- \tilde{U} operation also requires $O(n)$ gates. Thus implementing U requires $O(n^2)$ single qubit and CNOT gates. In previous section we showed that arbitrary unitary matrix on the 2^n -dimensional state space of n qubits may be written as product of $O(4^n)$ two-level unitary operations. Combining these results we see that an arbitrary unitary operation on n

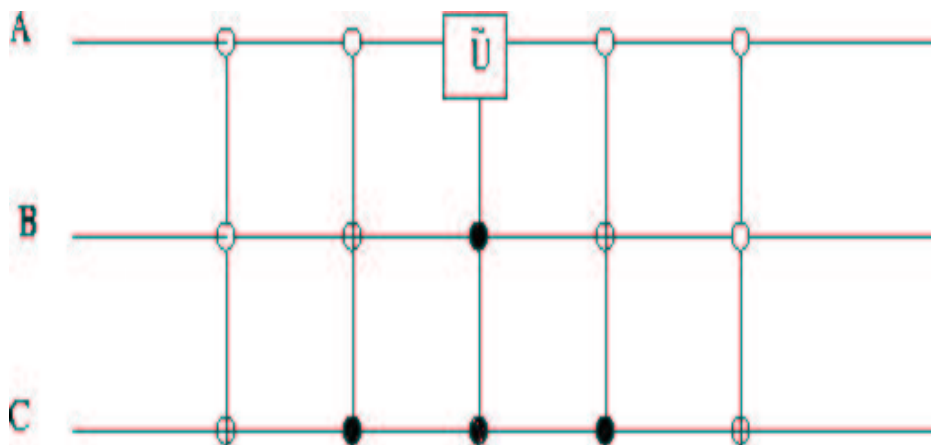


Figure 19: *Quantum Circuit for the example*

qubits can be implemented using a circuit containing $O(n^2 4^n)$ single qubit and CNOT gates.

16 Acknowledgements

We give special thanks to Mohan Raj Gupta and Anusheel Bhushan, whose previous independent study in quantum computing was extremely helpful in compiling this study.

A Properties of tensor products

For matrices A, B, C, D vectors x, y, u and scalars a, b we have,

$$\begin{aligned}
 (A \otimes B)(C \otimes D) &= (AC \otimes BD) \\
 (A \otimes B)(x \otimes y) &= (Ax \otimes By) \\
 (x + y) \otimes u &= (x \otimes u + y \otimes u) \\
 u \otimes (x \otimes y) &= u \otimes x + u \otimes y \\
 (ax \otimes by) &= ab(x \otimes y)
 \end{aligned}$$

For example,

$$(a|0\rangle + b|1\rangle) \otimes (c|0\rangle + d|1\rangle) = ab|00\rangle + bc|10\rangle + ad|01\rangle + bd|11\rangle$$

References

- [1] Eleanor Rieffel and Wolfgang Polak *An Introduction to Quantum Computation for Non-Physicists*, arXiv:quant-ph/9809016v2, 2000.
<http://xxx.lanl.gov/archive/quant-ph>.
- [2] André Berthiaume *Quantum Computation*, Centrum voor Wiskunde en Informatica, The Netherlands, 1996.
- [3] David Deutsch, Arthur Ekert and Rossella Lupacchini *Machines, Logic and Quantum Physics*, arXiv:math.HO/9911150v1 November 19, 1999.
<http://xxx.lanl.gov/archive/quant-ph>.
- [4] Lov K. Grover *A fast quantum mechanical algorithm for database search*, Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, 1996.
- [5] Lov K. Grover *Quantum Computing* The Sciences, July/August 1999, pp.24-30.
- [6] David Deutsch *Quantum theory, the Church-Turing principle and universal quantum computer*, Proceedings of the Royal Society of London Ser.A A400, 97-117, 1985.
- [7] Michael A. Nielsen and Isaac L.Chuang, “Quantum Computations and Quantum Information”