Quantum Computation

Mohan Raj Gupta , Anusheel Bhushan*
Supervisor: Prof. Sandeep Sen
Department of Computer Science and Engineering,
IIT Delhi.

II^{nd} Semester 2001-2002

Independent Study, CS310S

*{csu99152,csu99140}@cse.iitd.ernet.in

Contents

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 algoritms 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 Probability Amplitudes

This section uses the Double slit experiment to bring about the concepts of Probability Amplitudes and Quantum Basis.

2.1 The double slit experiment

We are studying a property P of a Quantum System.
Accordingly we define -

Definition 1 *For a given experiment, an event is a set of initial and final conditions.*

Accordingly, we define an experiment as the transition of the system from initial conditions to the final conditions.

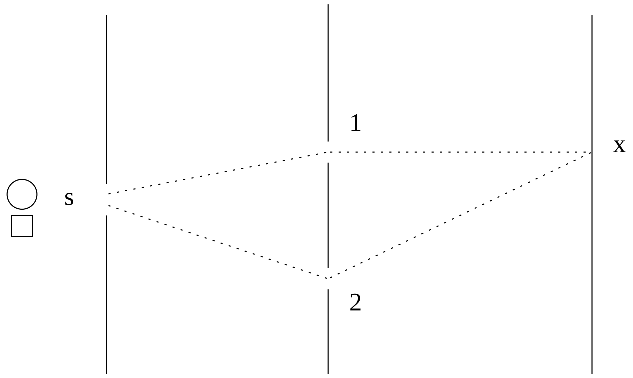


Figure 1: Double Slit Experiment

Definition 2 *The probability of an event is given by the square norm of a complex number α (known as the* probability amplitude*).*

$$p = \|\alpha\|^2$$

The probability amplitude of an event will be noted as follows :

$$\langle \text{final condition} \mid \text{initial condition} \rangle$$

For example, the event in the Double Slit Experiment is *a photon leaves the sorce s and arrives in the detector at position x*. This is noted as

$$\langle \text{photon detected at x} \mid \text{a photon leaves s} \rangle$$

or as

$$\langle x|s \rangle$$

The probability amplitude follows the laws of addition and multiplication of probabilities.
Thus,

$$\begin{aligned} \langle x|s \rangle &= \langle x|\text{wall}\rangle\langle \text{wall}|s\rangle \\ &= \langle x|1\rangle\langle 1|s\rangle + \langle x|2\rangle\langle 2|s\rangle \end{aligned}$$

also, we have

$$\begin{aligned}\langle 1|1\rangle &= \langle 2|2\rangle = 1\\ \langle 1|2\rangle &= \langle 2|1\rangle = 0\end{aligned}$$

Thus $\langle x|s\rangle$ is completely determined by amplitudes to transit from the holes. The holes thus act as natural elements for expressing events. Accordingly we define:

Definition 3 *The set $B = \{i$: i is the label of some condition$\}$ is a set of* basis states *if for all i,j $\in B$*

$$\langle i|j\rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Thus, we may view the basis as the set of all classically distinct alternatives of a given property P.

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

In Quantum Mechanics a *State*[1] is a *ray* in a *Hilbert Space*.

- **Hilbert Space** H

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

[1] A 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 $|x\rangle$. $\langle x||y\rangle$ denotes the inner product of $|x\rangle$ and $|y\rangle$ and is also denoted as $\langle x|y\rangle$.

3.2 Qubits

The indivisible unit of classical information is the bit also known as the *S*hannon 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 α , $\beta \,\epsilon\, \mathbf{C}$ and $\|\alpha\|^2 + \|\beta\|^2 \;=\; 1$ and $\{|\mathbf{0}\rangle, |\mathbf{1}\rangle\}$ form an orthonormal basis for the hilbert space.
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 $\mathbf{1}$[2]. 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*.

[2] we will also denote $|\mathbf{x}\rangle$ as $\mathbf{x}$

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

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

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 obsevation of a state vector

Let $|\psi\rangle$ be a vector in space $\mathbf{H}$ and let $O = \{E_1, \ldots, E_k\}$ 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. One of the E_i's will be selected with probability $\|\alpha_i\|^2$. $\|\alpha_i\|$'s are also known as probability amplitudes.

2. 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*.

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 ecoded in the polarisation of a photon. We can take $\{\uparrow, \rightarrow\}$ as the orthonormal basis of the hilbert space. We will denote the qubit as,

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

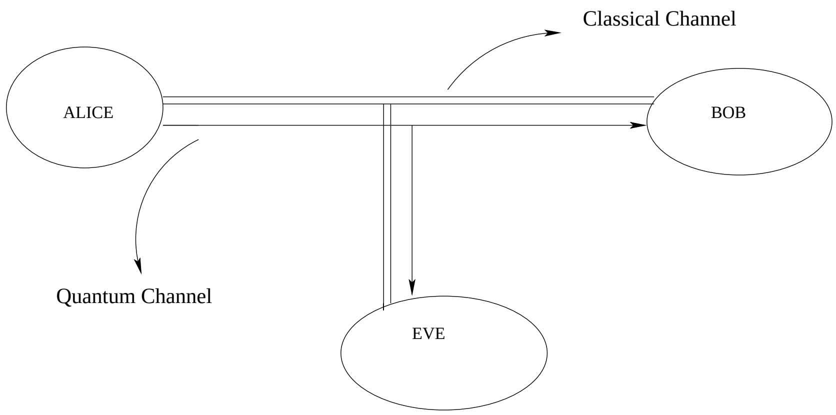
Suppose, we pass the photon through a filter F that is aligned to the horizontal (that is, it will let only horizontally polarised light pass through it). Note that this an act of measuring the qubit using the probe F. We can take the observable associated with F to be $O = \{\uparrow, \rightarrow\}$. 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 $O = \{\nearrow, \searrow\}$. 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 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 private 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, Alice and Bob exchange the basis they used for encoding and decoding of each qubit, over the open two-way channel. They identify all the bits for which the sending and receiving bases agree and construct the key. 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(obsevables), 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(Bob gets wrong value given that he uses the correct basis) = Pr(Eve uses wrong basis) $\times$ Pr(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 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

For computing with quantum states, some transformations need to be performaed. If $O = \{E_1, \ldots, E_k\}$, 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). The tranformations(U) are unitary(and linear) and are mappings of vectors in Hilbert Space. The transformations can also be represented as matrices. They preserve absolute values of the inner products.

$$|\langle\phi|\psi\rangle| \;=\; |\langle\phi'|\psi'\rangle|$$

where $U|\psi\rangle \;=\; |\psi'\rangle$.

$$U^*U \;=\; UU^* \;=\; I$$

The unitary transformations form a group -

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

$$(U_1.U_2)|\phi\rangle \;=\; U_1(U_2|\phi\rangle).$$

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

$$
\begin{array}{llc}
I: & |0\rangle \to |0\rangle \quad , \quad |1\rangle \to |1\rangle & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
X: & |0\rangle \to |1\rangle \quad , \quad |1\rangle \to |0\rangle & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
Z: & |0\rangle \to |0\rangle \quad , \quad |1\rangle \to -|1\rangle & \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
Y: & |0\rangle \to -|1\rangle \quad , \quad |1\rangle \to |0\rangle & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}
\end{array}
$$

Note that $Y = ZX$.

6 Multiple Qubits

As in classical computation, we have quantum registers composed of multiple qubits. Suppose we have n particles. Classically, they combine using the cartesian product. In Quantum Mechanics they combine using the *tensor product*. If

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

then

$$V \otimes W = span\{v_i \otimes w_i\}^3$$

If $dim(V) = n$ and $dim(W) = k$ then $dim(V \otimes W) = nk$. 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 transforamtion 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

[3] *Notation*: We write $(b_0 \otimes b_1 \otimes \cdots \otimes b_n)$ as $|b_0 b_1 \dots b_n\rangle$. *e.g.* $|0\rangle \otimes |1\rangle = |01\rangle$.

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.

6.2 Transformations

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

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

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

$$\begin{aligned} C_{not} \quad : \quad & |00\rangle \rightarrow |00\rangle \\ & |01\rangle \rightarrow |01\rangle \\ & |10\rangle \rightarrow |11\rangle \\ & |11\rangle \rightarrow |10\rangle \end{aligned}$$

It can also be thought of as an array-

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

6.2.1 Walsh-Hadamard Transformation

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

Therefore, H creates a superposition of states. 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,

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

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 $(|\psi\rangle \otimes \phi\rangle)$. But not all two qubit systems can be decomposed in this nice form, as a tensor product of one qubit states[4]. 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 probabilties 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 EPR[5] 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.

8 Dense Coding

8.1 Dense Coding

Dense Coding uses one quantum bit together with an EPR pair[6] $\psi_0 = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ to transmit 2 bits of classical information.

[4]We 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 $2^{n_1+n_2}$.

[5]Einstein,Podolsky, Rosen

[6]entangled particles

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} tranformation 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
$\psi_0 = \frac{1}{\sqrt{2}}(\|00\rangle + \|11\rangle)$	$\frac{1}{\sqrt{2}}(\|00\rangle + \|10\rangle)$	$\frac{1}{\sqrt{2}}(\|0\rangle + \|1\rangle)$	$\|0\rangle$
$\psi_1 = \frac{1}{\sqrt{2}}(\|10\rangle + \|01\rangle)$	$\frac{1}{\sqrt{2}}(\|11\rangle + \|01\rangle)$	$\frac{1}{\sqrt{2}}(\|1\rangle + \|0\rangle)$	$\|1\rangle$
$\psi_2 = \frac{1}{\sqrt{2}}(-\|10\rangle + \|01\rangle)$	$\frac{1}{\sqrt{2}}(-\|11\rangle + \|01\rangle)$	$\frac{1}{\sqrt{2}}(-\|1\rangle + \|0\rangle)$	$\|1\rangle$
$\psi_3 = \frac{1}{\sqrt{2}}(\|00\rangle - \|11\rangle)$	$\frac{1}{\sqrt{2}}(\|00\rangle - \|10\rangle)$	$\frac{1}{\sqrt{2}}(\|0\rangle - \|1\rangle)$	$\|0\rangle$

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 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 :

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

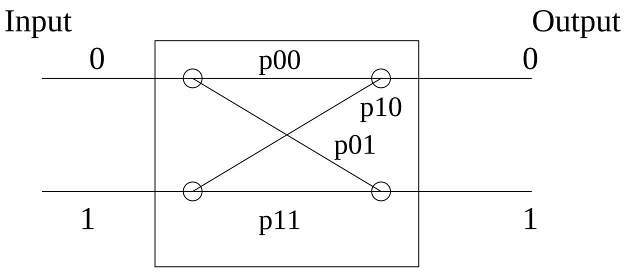
Clearly, we can measure the second qubit with probability 1, iff either
$a = d = 0$ or $b = c = 0$.
In either case the initial system is entangled in either one of the following states:-

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

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$

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 2) 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 $\sqrt{not}$ gate.

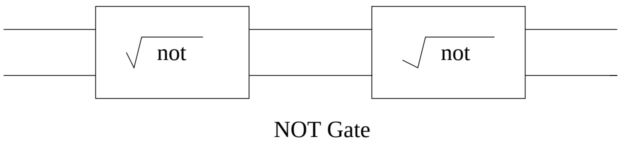


Figure 2: Concatenation of two $\sqrt{not}$ gates gives a not gate.

9.1 Empirical Verification

Consider Figure 3. L is a half-silvered mirror, i.e. a mirror which with probability 50% reflects a photon which impinges on it and with probability

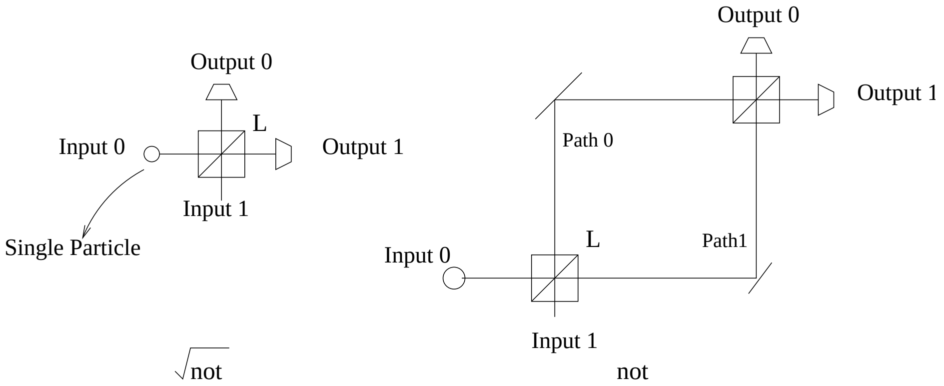


Figure 3: Experimental realisation of $\sqrt{not}$ gate and not gate.

50% allows it to pass through. The experimental realisation of the two concatenated $\sqrt{not}$ gates is known as *Single particle interfenece*. 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 $0 \to 0$ in the composite machine can happen via two mutually exclusive paths, $0 \to 0 \to 0$ and $0 \to 1 \to 0$. Thus using the axiom of addivity of probabilities, we have

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

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 $0 \to 0 \to 0$ and $0 \to 1 \to 0$ are mutually exclusive. Let us explain the process using quantum mechanics. We have already introduced the concept of *probability amplitudes* in Section 2.1. 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 serarately. Unlike probabilities, probability amplitudes can be both positive and negative and

they therefore can cancel out each other. Consider for example the $\sqrt{not}$ gate. It can be represented by the following unitary transformation:

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

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

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

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

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

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

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

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

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 4.

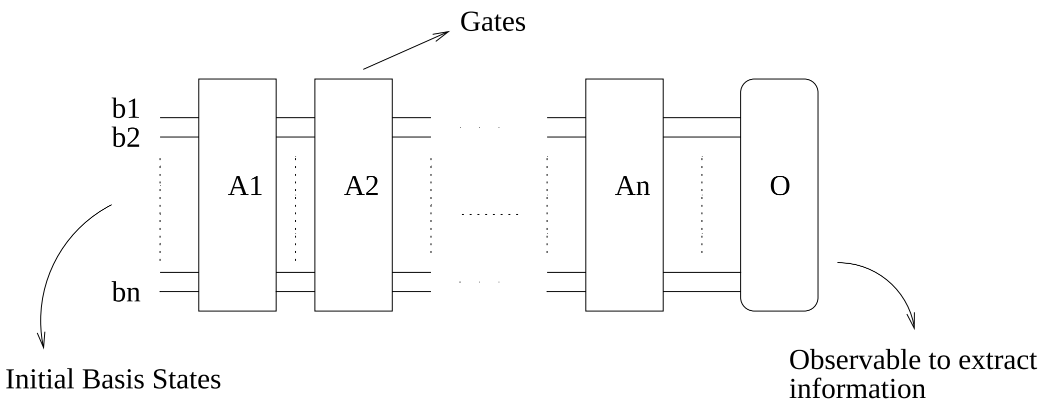


Figure 4: Quantum Gate Arrays

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".[7]

Definition 4 *A function $f : \{0,1\}^n \rightarrow \{0,1\}$ is* **non-balanced** *if one of the two values of f has a majority.*

Definition 5 *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 polymial time (with bounded error probability) on a 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.

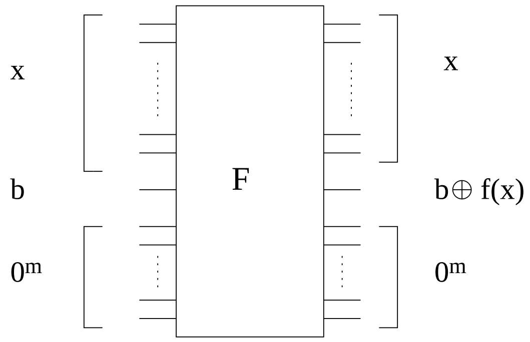
[7]The original Deutsch-Jozsa Problem dealt with strings rather than functions.

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 reversible [8], we have the following corollary:

Corollary 10.3 *A Turing-computable function f is always computable on a quantum gate array (with a neglible 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 in 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:



[8]recall that tranformations on a state are always unitary and thus reversible

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 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:

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

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

$$S_n \, |\underbrace{0 \ldots 0}_{n}\rangle \quad = \quad \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i\rangle$$

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:

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

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 **1**.

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

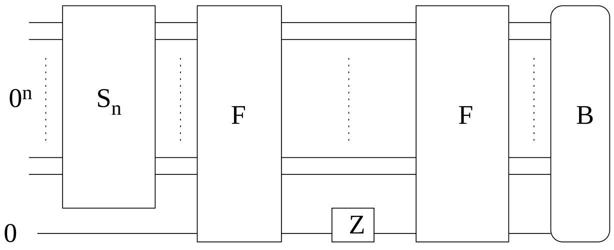


Figure 5: MDJP

In Figure 5 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 tranferred 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,

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

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 $|a_{l-1}a_{l-2}\ldots a_1a_0\rangle$. 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:

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

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 gates[9] for the quantum fourier transform.
The claim is that applying the above sequence of transformations will result in a quantum state,

$$|\phi\rangle \;=\; \frac{1}{\sqrt{2^l}} \sum_{b=0}^{2^l-1} exp(2\pi i a c/2^l)\, |b\rangle$$

where b is the bit reversal [10] 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 $|01\rangle$ $(l = 2)$:

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

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

[9]We 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.

[10]that is, if $|c\rangle = |c_{l-1}c_{l-2}\ldots c_0\rangle$ then $|b\rangle = |c_0 \ldots c_{l-2}c_{l-1}\rangle$

time, in order.

$$
\begin{aligned}
R_1\,|01\rangle &\rightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\otimes|1\rangle \\
&= \frac{1}{\sqrt{2}}(|01\rangle + |11\rangle) \\
S_{0,1}\big(\frac{1}{\sqrt{2}}(|01\rangle + |11\rangle)\big) &\rightarrow \frac{1}{\sqrt{2}}(S_{0,1}|01\rangle + S0,1|11\rangle) \\
&= \frac{1}{\sqrt{2}}(|01\rangle + e^{i\pi/2}|11\rangle) \\
R_0\big(\frac{1}{\sqrt{2}}(|01\rangle + e^{i\pi/2}|11\rangle)\big) &\rightarrow \frac{1}{\sqrt{2^2}}\Big(\big(|0\rangle\otimes(|0\rangle - |1\rangle)\big) \\
&\quad + \big(e^{i\pi/2}|1\rangle\otimes(|0\rangle - |1\rangle)\big)\Big) \\
&= \frac{1}{2}\big(|00\rangle - |01\rangle + e^{i\pi/2}|10\rangle - e^{i\pi/2}|11\rangle\big) \\
|\psi\rangle &= \frac{1}{2}\big(|00\rangle + e^{i\pi}|01\rangle + e^{i\pi/2}|10\rangle + e^{3i\pi/2}|11\rangle\big) \\
reverse-bits(|\psi\rangle) &\rightarrow \frac{1}{2}\big(|00\rangle + e^{i\pi/2}\,|01\rangle + e^{i\pi}\,|10\rangle + e^{3i\pi/2}|11\rangle\big)
\end{aligned}
$$

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 $|a\rangle = |a_{l-1}\ldots a_0\rangle$ to $|b\rangle = |b_{l-1}\ldots b_0\rangle$ where b is the reverse of c[11]. 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 1. 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 6, where a phase factor of $exp(i\pi/2)$ is introduced with $|11\rangle$.

[11]Note 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

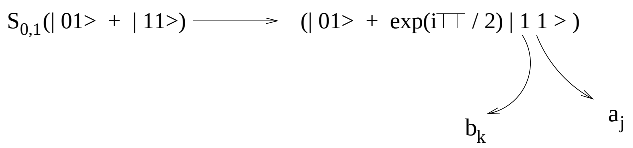


Figure 6: Example

When we apply $S_{j,k}$ for $j < k$, then we have converted all the bits $(a_{l-1} \ldots a_{j+1})$ to $(b_{l-1} \ldots b_{j+1})$. 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:

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

which can also be written as

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

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

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

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

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

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

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

---

calculating the corresponding amplitude associated with it.

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

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

and so we can write,

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

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

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 albhabetized 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 interfence as we have already seen.

12.1 The Problem

Let a system have $N = 2^n$ states which are labelled $S_1, S_2, \ldots, S_N$. 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 $x_i \in \{0, 1, \ldots, 2^n - 1\}$.

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 $X_0, X_1, \ldots, X_N$ 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 $X_i' = 2A - X_i$.

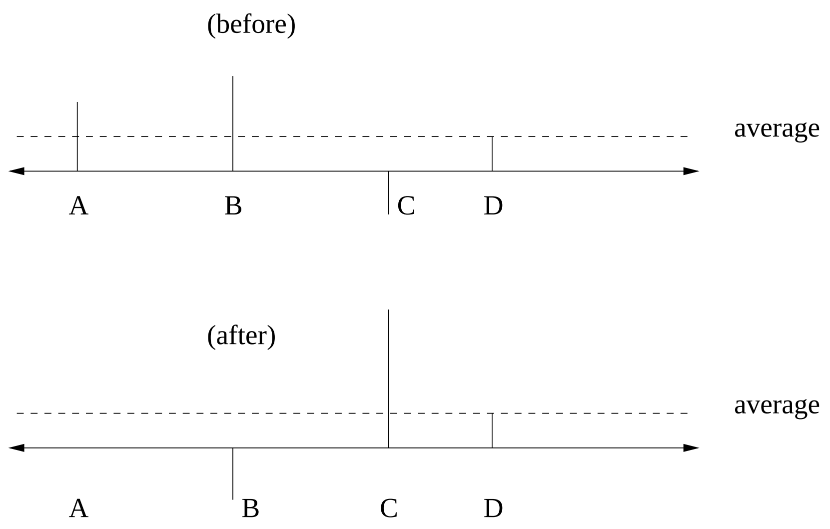


Figure 7: Inversion about average operation

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 9 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 8.

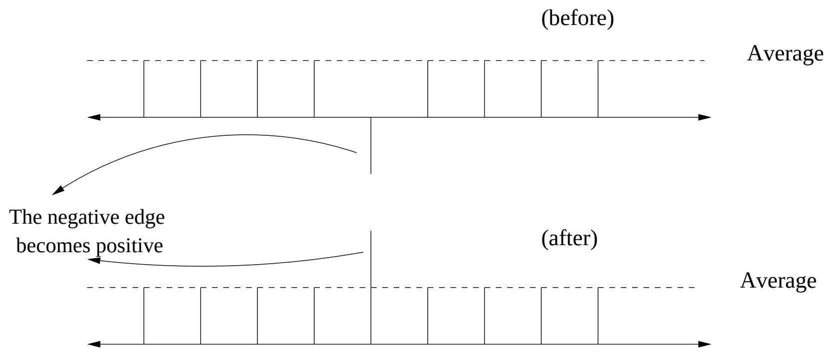


Figure 8: Inversion about average operation is applied to a distribution in which all but one are $O(\frac{1}{\sqrt{N}})$ and the one component is negative

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, $WRW = W(R' - 1)W = WR'W - I$. 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 $\left|\frac{k}{l}\right| < \sqrt{N}$. 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 satifies $P(S) = 1$, 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,

$$k_1 \;=\; -k \; and \; l_1 \;=\; l$$

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 atleast $1/\sqrt{2}$).

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 $x^r \equiv 1(mod\, \text{n})$. As with the factoring problem no algorithm is known for solving this problem.

Program (for Factorisation):
Input : an input n *odd integer*
x $\leftarrow$ random$\{0, \ldots, n\}$
r $\leftarrow$ use the oracle to find the order of x (mod n)
Output: if r is odd or $x^{r/2} \equiv -1\ (mod\ \ n)$ then fail
else return $\gcd(x^{r/2} - 1\,, n)$

13.1 Finding Order of an element

We now describe Shor's algorithm to find the order r of an element x mod 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 $|a, x^a mod\ n\rangle$.
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,

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

We measure the part of the superposition that contains the value $x^j mod\ n$. Let us observe the value y. The function $f(a) = x^a mod\ n$ is periodic with period r since $x^r \equiv 1(mod\, n)$. 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,

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

Now,

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

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

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

where,

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

Therefore,

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

Now measuring the state ϕ returns a value $c = \lambda q/r$ for some $\lambda \in \{0, \ldots r-1\}$. 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,

$$co - primes(r) \ \geq \ r/log\ r$$

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

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

Let $|f(z)\rangle$ be the result of measuring $|f(x)\rangle$. Now there are two x st

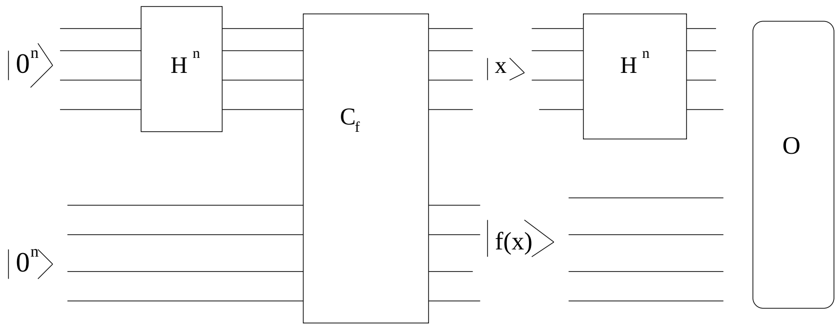


Figure 9: Simon's algorithm

$f(x) = f(z)$: 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:

$$H_n\left(\frac{1}{\sqrt{2}}|z\rangle + \frac{1}{\sqrt{2}}|z+u\rangle\right)$$

$$= \frac{1}{\sqrt{2}}\sum_{y=0}^{2^n-1}\frac{(-1)^{z.y}}{2^{n/2}}|y\rangle + \frac{1}{\sqrt{2}}\sum_{y=0}^{2^n-1}\frac{(-1)^{(z+u).y}}{2^{n/2}}|y\rangle$$

$$= \frac{1}{2^{n/2}\sqrt{2}}\sum_{y=0}^{y=2^n-1}(-1)^{x.y}\left(1+(-1)^{u.y}\right)|y\rangle.$$

Now, if $u.y = 1$, the amplitude of $|y\rangle$ is 0. All $|y\rangle$ such that $u.y = 0$ have the same non-zero amplitudes. Hence the output of the measurement is a random y such that $y.u = 0$. Hence, if we get $n-1$ linearly independent

vectors y st $u.y = 0$, 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.

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-Eigth 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.