Summer of Science

Final Report

June 10, 2020

Introduction to Quantum computing

Mithil Vakde

Mentor: Atri Dutta

1 Introduction

1.1 Landauer's limit

Deleting information is a dissipative process and is irreversible. However, increasing information is not. Landauer's principle states that if an observer loses information about a physical system, they lose the ability to extract work from that system. He set a limit on the minimum amount of energy required to erase one bit of information: $kT \ln 2$

If a computer works on many-valued logic (like a ternary computer), erasure of one unit of information is $kT \ln N$ where N is the dimension of the information vector. However, erasure of one bit of information from any type of system has the same limit $kT \ln 2$ (since a bit holds the same amount of information in any system)

1.2 Maxwell's Demon

Maxwell's thought experiment consisted of a gas-filled box with a partition containing a shutter operated by a demon. The demon lets fast molecules move from A to B and slow molecules to move from B to A by operating the shutter at negligible work expense. Heat flows from a cold place to a hot place at almost no cost, a violation of the second law of Thermodynamics. This problem is resolved when we consider that the demon must store information about the molecules. Using Landauer's principle, we associate some entropy with the recorded information and thus "exorcise" the demon.

1.3 Universal Gates

Currently, we can express all possible truth tables using functionally complete gates such as NAND or NOR (FANOUT is implied). These can be used to construct any boolean function. However, these gates are irreversible. We cannot recover the input from its output. This means that while using these gates, information is lost and hence, energy is used up.

NOT and XOR are examples of 1-input and 2-input reversible gates respectively. However, these 2 together don't constitute a universal set. This changed with the discovery of the Toffoli Gate: a 3 input reversible and functionally complete (in classical computing) gate. Other examples of gates in this category include Fredkin Gate.

1.4 Church-Turing Thesis

Turing developed an abstract notion equivalent to a programmable computer called Turing Machine, that could carry out calculations from inputs by manipulating symbols on a tape. A function is called Turing computable if it can

be computed by some Turing Machine. Church developed another model for computing called λ-calculus. Both of these models were proven to be equivalent. Turing also introduced the idea of a Universal Turing Machine: one that can simulate any arbitrary Turing machine on any arbitrary input. Church and Turing together came up with the Church-Turing Thesis, which states that any algorithmic process can be simulated using a Turing machine.

A modification of the thesis stated that any algorithmic process can be simulated efficiently using a probabilistic Turing machine. Here efficiently means that the problem can be solved in polynomial time. However, there exist problems that have (as of now) no known efficient algorithms on such classical computers. Shor found an algorithm that has an efficient solution to finding the prime factors of an integer. This algorithm works in polynomial time (polynomial in the log of on a quantum computer. No efficient solutions have been found on classical computers. This indicates that this form of the thesis is wrong and that Quantum computers are more powerful. However, it hasn't been conclusively disproven yet.

1.5 Error Correction

Shannon established that there is a maximum rate of error-free data that can theoretically be transferred over a channel if it is subject to random errors (noisy-channel coding theorem). He showed that error-correcting codes can be applied to protect information in the presence of noise, but placed an upper bound on this protection. Similarly, quantum error-correcting codes have been discovered. However, no quantum analogue to the noisy-channel coding theorem has been discovered as of now.

1.6 Cryptography

Currently, the most widely used public-key encryption system is the RSA encryption. Using the public key to break the encryption is very hard to do on a classical computer. Currently, it would take very long to do so. This problem is closely related to the factoring problem and an efficient factoring algorithm would mean an efficient RSA encryption breaker. Shor's algorithm can be carried out in quantum computers and hence can be used to break RSA. Quantum Cryptography can be used to make unbreakable encryption

2 Quantum Information

2.1 Quantum Info

Quantum information differs widely from classical information. The most basic unit of quantum information is the qubit. Unlike its classical analogue, a qubit is continuously valued. Quantum information is limited by the no-cloning theorem which states that it is impossible to create an identical copy of an arbitrarily

unknown quantum state. There is no place for true randomness in classical and deterministic systems but there are quantum phenomena that are truly random.Quantum mechanics does not specify the outcomes of such experiments, we can only talk about probabilities. The uncertainty principle also puts another limit. We cannot measure any property of one system without disturbing its other properties (non-commuting observables).

2.2 Qubit

The qubit is a vector in a two-dimensional complex vector space with inner product, spanned by the basis states $|0\rangle$ and $|1\rangle$ and is normalized.

$$|\psi\rangle \quad = \quad a\,|0\rangle \quad + \quad b\,|1\rangle \qquad\qquad |a|^2 + |b|^2 = 1$$

When a qubit is measured, it collapses into either of the computational basis states. And it continues to be in this state post-measurement (before any other disturbance occurs). Which of the states it collapses into is completely and truly random. This behaviour is a fundamental property of Quantum mechanics. $|a|^2$ represents the probability that measurement of $|\psi\rangle$ obtains the result $|0\rangle$ and $|b|^2$ represents the probability of the measurement yielding $|1\rangle$. It is not possible to find out the values of a and b from a single qubit. We would need an infinite number of identically prepared qubits to determine this.

We can also write the qubit as follows:

$$|\psi\rangle \quad = \quad e^{i\gamma}\left(\cos\left(\frac{\theta}{2}\right)|0\rangle \quad + \quad e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle\right)$$

We can ignore the factor $e^{i\gamma}$ since it corresponds to a global phase factor which has no observable physical effects. Multiplying the qubit by any global phase doesn't change the qubit. Geometrically, qubits are represented using Bloch spheres. Each qubit can be considered as a point on the Bloch Sphere

2.3 Multiple Qubits

If we have N qubits, the quantum state is a vector space of 2^N dimensions. The state vector is a complex linear combination of 2^N vectors with the same normalisation condition. The coefficients of each basis are called amplitudes. This combined quantum state is obtained by taking the tensor product of that many qubits. This is just a statement of a general postulate of quantum mechanics for composite systems (w.r.t qubits). The general statement is that: *The state space of a composite physical system is the tensor product of the state spaces of the component physical systems.* For example, a two qubit state is represented this way:

$$|\psi_1\psi_2\rangle \quad = \quad |\psi_1\rangle \otimes |\psi_2\rangle \quad = \sum_{i,j\epsilon\{0,1\}} \alpha_{ij}\,|ij\rangle \qquad\qquad \sum |\alpha_{ij}|^2 = 1$$

A 2 qubit state space is fundamentally different from having 2 independent qubits. If we consider 2 qubits prepared independently, the information stored in them is less than that if we prepare them together. For example, the vector

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

(called Bell state) cannot be obtained by (tensor) multiplying 2 phase vectors from independent single qubit systems. The fundamental difference here is entanglement. If we measure the outcome of the first bit, the output of the second bit will also be the same. The measurement outcomes are correlated. Bell proved that this correlation is stronger than ever possible in classical systems.

2.4 Quantum gates

Quantum gates can be represented as matrices that act on the state vectors. They are multiplied with the state vectors to obtain the output. For example, a quantum gate having matrix form $\mathbf{U}$ acts on the quantum vector $|\psi\rangle$ pre multiplication: $\mathbf{U}\,|\psi\rangle$ The only condition required for permitting a matrix to act as a quantum gate is that it should be unitary. ($\mathbf{U}\mathbf{U}^\dagger = \mathbf{I}$). This also means that all quantum gates are reversible since the inverse of a unitary matrix is a unitary matrix.

The number of qubits in the input and the output of gates must be equal. This means that FANOUT is not a valid quantum gate. The matrix form of a quantum gate acting on n qubits has an order $2^n \times 2^n$. These gates act upon quantum states that are vector in 2^n dimensions. Geometrically, a quantum gate acting on a single qubit can be considered to be a rotation of the Bloch sphere about an axis.

2.5 Bloch Sphere

Consider the complex number representation of a qubit:

$$|\psi\rangle \quad = \quad \cos\left(\frac{\theta}{2}\right)|0\rangle \quad + \quad e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle$$

The numbers θ and ϕ are considered as the co-ordinates of a point on the Bloch sphere. This representation is the same as the spherical co-ordinate system with a fixed $r = 1$. Thus a qubit has 2 degrees of freedom. Common vectors on this sphere are:

$+z$ axis $\equiv |0\rangle$ $\quad$ $+x$ axis $\equiv \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \equiv |+\rangle$ $\quad$ $+y$ axis $\equiv \frac{1}{\sqrt{2}}\left(|0\rangle + i\,|1\rangle\right)$

$-z$ axis $\equiv |1\rangle$ $\quad$ $-x$ axis $\equiv \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) \equiv |-\rangle$ $\quad$ $-y$ axis $\equiv \frac{1}{\sqrt{2}}\left(|0\rangle - i\,|1\rangle\right)$

3 Single qubit quantum logic gates

3.1 Pauli X

Denoted by σ_x or X, this gate is also known as the quantum NOT gate. It switches the amplitudes of $|0\rangle$ and $|1\rangle$. This gate corresponds to a rotation of the Bloch sphere around the x axis by π radians. Since applying this gate twice returns the original qubit, this matrix is hermitian. (As it is a unitary matrix and is equal to its inverse). Its matrix form is:

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

3.2 Pauli Y

Denoted by σ_y or Y, it switches the amplitudes of $|0\rangle$ and $|1\rangle$, multiplies them by i and negates the new amplitude of $|0\rangle$. This gate corresponds to a rotation of the Bloch sphere around the y axis by π radians. This matrix is hermitian too. Its matrix form is:

$$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

3.3 Pauli Z

Denoted by σ_z or Z, it doesn't affect $|0\rangle$ but flips the sign of $|1\rangle$. This gate corresponds to a rotation of the Bloch sphere around the z axis by π radians. This matrix is hermitian too. Its matrix form is:

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

3.4 Hadamard

This gate does the following

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

$$|1\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

It is also a hermitian matrix and hence applying this gate twice is the same as identity. On the Bloch sphere, this matrix rotates the sphere by $\frac{\pi}{2}$ radians about the y axis, followed by π radians about the x axis. Its matrix form is:

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

3.5 Phase Shift Gates

Family of gates that map $|1\rangle$ to $e^{i\phi}|1\rangle$, but leave $|0\rangle$ unchanged. This is called "phase shift" because it changes the *relative* phase between the two basis states (unrelated to global phase). Geometrically, this is represented by rotating the Bloch sphere around the z axis by ϕ radians. Its matrix form is:

$$\begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix}$$

Examples of this family of gates include:

- S gate: Also called the phase gate, $\phi = \frac{\pi}{2}$
- T gate: Also called the $\frac{\pi}{8}$ gate, although $\phi = \frac{\pi}{4}$
- Pauli Z gate

3.6 Other notes on single qubit gates

3.6.1 Serial Gates

When 2 or more gates are applied serially, we continue to multiply the matrix of a gate with the output vector of the previous gate. However, we can also simplify the action of multiple gates as the action of a single gate whose matrix is found by multiplying the matrix gates in the order of their action from right to left. For example, if on a circuit, we first apply gate **A** followed by **B**, then the equivalent matrix is **BA** and not **AB**

3.6.2 Pauli Matrices

Each Pauli matrix is Hermitian, and together with the identity matrix I they form a basis for the real vector space of 2 × 2 Hermitian matrices. Since they are unitary too, squaring them yields the identity matrix. Their determinants and traces are same and equal to -1 and 0 respectively. An important relation of the Pauli matrices is $\sigma_i^* = -\sigma_y \sigma_i \sigma_y$

3.6.3 Decomposing single qubit gates

All single qubit gates can be represented by 2x2 matrices. Every one of these gates can be decomposed into a rotation around the z axis, followed by a rotation around the x axis and followed by another rotation around the z axis together with a global phase shift. (Any 2 perpendicular axes can be used, need not be x or z). Another way of saying this is that every Unitary matrix can be decomposed into a constant multiplied by 3 matrix multiplications:

$$\mathbf{U} = e^{i\alpha} \begin{bmatrix} e^{-i\frac{\beta}{2}} & 0 \\ 0 & e^{i\frac{\beta}{2}} \end{bmatrix} \begin{bmatrix} \cos\frac{\gamma}{2} & -\sin\frac{\gamma}{2} \\ \sin\frac{\gamma}{2} & \cos\frac{\gamma}{2} \end{bmatrix} \begin{bmatrix} e^{-i\frac{\delta}{2}} & 0 \\ 0 & e^{i\frac{\delta}{2}} \end{bmatrix}$$

Where $\alpha, \beta, \gamma, \delta$ are all real numbers.

4 Multiple qubit quantum logic gates

These gates take in inputs that are the combined state of 2 or more qubits. 2 individual single gates acting in parallel is equivalent to a single gate acting on the quantum register of 2 qubits. The matrix form of the single larger gate is equal to the tensor product of the 2 parallel quantum gates (applied in a particular order). This can be extended to n gates acting in parallel. For example, $\mathbf{A}$ acting on $|\psi\rangle$ and $\mathbf{B}$ acting on $|\phi\rangle$ in parallel is the same as $(\mathbf{A} \otimes \mathbf{B})$ acting on $|\psi \otimes \phi\rangle$ in that particular order. One important point to note is that not all multiple qubit gates can be decomposed as a tensor product of single gates. For example, any 2-qubit controlled-U gate cannot be decomposed into $\mathbf{A} \otimes \mathbf{B}$ in general.

4.1 CNOT

This is also called the cX gate. It is a part of the family of Controlled Pauli gates (cX, cY, cZ). Of the two input qubits, one of them is named the control qubit and the other one is named the target qubit. The control qubit is unchanged by this gate, while the target qubit is mapped to an XOR between the 2 qubits. Another way of defining its action is that it applies NOT to the target qubit if the control qubit is 1. With respect to the bases, it is defined as $|10\rangle \leftrightarrow |11\rangle$ while $|00\rangle, |01\rangle$ are unchanged. All 2 qubit gates can be represented by 4x4 matrices (obta ined from tensor product of two matrices) with the same condition - the matrix must be unitary. The matrix form of CNOT is:

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

4.2 SWAP

This gate simply interchanges the 2 input qubits. With respect to the bases, $|10\rangle \leftrightarrow |01\rangle$ while $|00\rangle, |11\rangle$ are unchanged. Its matrix form is:

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

4.3 Toffoli

Also called the CCNOT gate, it is a three input quantum gate. It applies Pauli X on the third input if the first 2 inputs are 1 each. With respect to bases, it maps $|111\rangle$ to $|110\rangle$ and vice versa while leaving the other bases unchanged.

The matrix form is:

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

4.4 Fredkin

Also called the CSWAP gate, it performs a controlled swap. If the first qubit is 1 then the other two are swapped. It maps $|101\rangle$ to $|110\rangle$ and vice versa while leaving the others unchanged. Its matrix representation is:

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

4.5 Hadamard Transform

$\mathbf{H_n}$ is a hadamard transformation acting on a register on n qubits. It is equal to n Hadamard gates acting on n qubits in parallel.

$$\mathbf{H_n} \;=\; \mathbf{H} \otimes \mathbf{H} \otimes ... \otimes \mathbf{H}(n\ times) \;=\; \mathbf{H}^{\otimes n}$$

Its matrix representation is given by $(\mathbf{H}_n)_{ij} = 2^{-\frac{n}{2}}(-1)^{ij}$

5 Measurement

Measurement is irreversible is hence not a quantum gate. The postulate of Quantum Mechanics related to Measurements states that *Quantum measurements are a collection of operators* $\{\mathbf{M}_m\}$ *acting on the state space of the system.* The probability of result m being obtained when $|\phi\rangle$ is measured is $p(m) = \langle\phi| \mathbf{M}_m^\dagger \mathbf{M}_m |\phi\rangle$. Measurement changes the state of the system. In the previous example, post measurement the state of the system becomes

$$\frac{\mathbf{M}_m |\phi\rangle}{\sqrt{\langle\phi| \mathbf{M}_m^\dagger \mathbf{M}_m |\phi\rangle}} = \frac{\mathbf{M}_m |\phi\rangle}{\sqrt{p(m)}}$$

An important point to note is that non orthogonal quantum states cannot be distinguished

5.1 Projection Valued Measurement

This formalism of measurement is a special case of the general of discussed above. A projective measurement is described by an *observable*. An observable is a hermitian operator which corresponds to the action of performing a measurement. When it acts on a quantum state, it gives the output as an eigenvalue and the state changes to that eigenstate. These eigenvectors of the observable form an orthonormal basis for the quantum state. Each possible outcome corresponds to theat eigenvector. (Note that any orthonormal basis can be used) The observable has a spectral decomposition

$$\mathbf{M} = \sum_m m\mathbf{P}_m$$

Where $\mathbf{P}_m$ is the projector onto the eigenspace of $\mathbf{M}$ with eigenvalue m. The probability of getting the result as m on measuring $|\phi\rangle$ is

$$p(m) = \langle phi| \mathbf{P}_m |\phi\rangle$$

The new state of the system is

$$\frac{\mathbf{P}_m |\phi\rangle}{\sqrt{p(m)}}$$

The average value of the measurement is

$$E(\mathbf{M}) = \langle\phi| \mathbf{M} |\phi\rangle$$

5.2 Positive Operator Valued Measures

This formalism of measurement is a generalisation of PVMs. This is used mainly when we just want the probabilities of the respective outcomes when we just measure the system at the end of the experiment. We define a set of positive operators $\mathbf{E}_m = \mathbf{M}_m^\dagger \mathbf{M}_m$ with the properties

$$\sum_m \mathbf{E}_m = \mathbf{I} \quad , \quad p(m) = \langle\phi| \mathbf{E}_m |\phi\rangle$$

The set of operators $\mathbf{E}_m$ is called POVM. POVMs can be used to distinguish between 2 non orthogonal states without ever misidentifying the state. But the catch is that sometimes we will not gain any information about the state. For example, if the 2 states to be distinguished were $|1\rangle$ and $|+\rangle$ Then the POVM operators

$$\mathbf{E}_1 = \alpha |0\rangle \langle 0| \quad \mathbf{E}_2 = \alpha |-\rangle \langle -| \, and \mathbf{E}_3 = \mathbf{I} - \mathbf{E}_1 - \mathbf{E}_2$$

can be used. If the result is $\mathbf{E}_1$ then we know that the state measured was $|+\rangle$. If the result is $\mathbf{E}_2$ then we know that the state measured was $|1\rangle$. If the result is $\mathbf{E}_3$ then we have obtained no information about the state. measured.

6 Postulates of Quantum Mechanics

1. Associated with any isolated physical system is a Hilbert space known as *state space* of thee system. This system is completely described by *state vectors*, which are unit vectors in the state space

2. The evolution of a closed quantum system is described by a unitary transformation.

3. The time evolution of the state of a closed quantum system is described by the Schrodinger equation

$$i\hbar \frac{d\left|\psi\right\rangle}{dt} \;=\; H\left|\psi\right\rangle$$

   Where H is the Hamiltonian of the system.

4. The quantum measurement postulate that describes general measurements as stated above

5. The postulate describing composite systems as stated above

7 Entanglement and Applications

When a group of particles is such that their individual quantum states cannot be described independent of the other. (The overall state cannot be decomposed into tensor products of the states of individual particles)

7.1 Local Realism

7.1.1 EPR Paper

Einstein along with Podolsky and Rosen proposed a though experiment and argued that its result proves that Quantum Mechanics is an incomplete description of the Universe. The experiment is presented in a modified manner as follows: Suppose we measure the first qubit of the quantum state $\frac{|01\rangle+|10\rangle}{\sqrt{2}}$, and at the same time (relativistically) another person measures the second qubit of the same entangled pair of qubits (and they are located far away). Then we can predict with certainty the outcome of the other experiment (which is bound to be the opposite of what we measured). Einstein argued that there must be something in the quantum state that tells us the outcome of that experiment without measuring it. However, quantum mechanics tell us that nothing in the quantum state tells us that. Therefore there must be a more complete theory that can tell us that.

7.1.2 Bell's Inequality

Here we describe another thought experiment: 2 particles are prepared and sent to Alice and Bob. These particles have 4 measurable properties, P, Q, R and S, each having 2 possible values: {1,-1}. We assume these 4 properties are objective - their value is fixed for each particle once it is prepared and only revealed by measurement. Let A_p, A_q be the outcomes of Alice's measurement (she can only measure P, Q) and B_r, B_s are similarly defined. Once the particles are sent to Alice and Bob, they randomly choose to measure any one of the properties. This measurement is done at the same time and far away from each other (in a causally disconnected manner). Based on probability, we can easily derive the following inequality:

$$\mathbf{E}(A_p B_r) \; + \; \mathbf{E}(A_q B_s) \; + \; \mathbf{E}(A_q B_r) \; - \; \mathbf{E}(A_p B_s) \;\; \leq \;\; 2$$

Let us consider a version of this experiment with a quantum system of 2 qubits prepared in the entangled state

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

The first qubit is passed to Alice and the second is passed to Bob. Let the properties (here, observables) be

$$A_p = X_1 \qquad B_r = \frac{-X_2 - Y_2}{\sqrt{2}}$$

$$A_q = Y_1 \qquad B_s = \frac{+X_2 - Y_2}{\sqrt{2}}$$

By calculations we can show that

$$\langle A_p B_r \rangle = \frac{1}{\sqrt{2}} \qquad \langle A_q B_r \rangle = \frac{1}{\sqrt{2}} \qquad \langle A_q B_s \rangle = \frac{1}{\sqrt{2}} \qquad \langle A_p B_s \rangle = -\frac{1}{\sqrt{2}}$$

But, when we calculate the Bell inequality

$$\langle A_p B_r \rangle + \langle A_q B_r \rangle + \langle A_q B_s \rangle - \langle A_p B_s \rangle \quad = \quad 2\sqrt{2}$$

This is clearly a violation of the Bell Inequality. This means that if Quantum Mechanics is an accurate description of the universe, the Bell inequality is false and so is Einstein's argument. (Though it holds in the classical world). When this was experimentally carried out, Nature agreed with Quantum Mechanics.

7.1.3 Resolving the paradox

While deriving Bell's inequality, we made 2 assumptions. It is now known that either or both of these assumptions are wrong. The first assumption is that the physical properties are objective and have definite values independent of

measurement. This is called the assumption of *realism*. The other assumption, of *locality*, is that Alice's measurement doesn't influence Bob's. Although these assumptions are a given in the classical world, and they are intuitive, Bell's inequality clearly shows that atleast one of them is wrong. Therefore our Universe is not locally realistic.

7.2 Superdense coding

This application of quantum mechanics allows us to transmit 2 classical bits of information between 2 parties by the interaction of only 1 qubit. Suppose that Alice and Bob have the 1st and 2nd qubit of the entangled state $|-\rangle$. By applying elementary gates to her qubit, Alice can change the overall quantum state to one of the 4 Bell states. She then sends her qubit over to Bob, who can identify which state it is in. Alice applies the gates based on what information she wants to send. Since this information has 4 possible outcomes, the amount of it being sent is 2 bits. (Isomorphic map to 00,01,10,11)

7.3 Quantum teleportation

This application is used to transfer an unknown qubit between 2 people when they can only send classical information to one another. Since the qubit is unknown and only a single copy of it exists, we cannot find out its complete state by measurement. To send this qubit, we exploit entanglement. Suppose Alice and Bob share an EPR pair between each other with each of them having one qubit (let Alice have $|\beta_1\rangle$ and bob have $|\beta_2\rangle$ of an EPR state $|\beta\rangle$, and Alice wants to transfer another unknown qubit $|\phi\rangle$ to Bob using classical information only. The process occurs as follows: Alice applies CNOT, on $|\phi\rangle\,|\beta_1\rangle$, followed by Hadamard on $|\phi\rangle$. Alice measures her qubits and send the output(one of $\{00,01,10,11\}$) to Bob. Based on which of the 4 outputs Bob receives, he applies one of 4 transformations on his qubit to obtain $|\phi\rangle$

8 Quantum Search

Suppose we want to search for a particular element in a database of N elements, Classically the algorithm takes $O(N)$ operations to do so. There exists many quantum search algorithms that can offer a speed faster than this (atmost a quadratic speedup). Grover's algorithm, the fastest possible quantum search algorithm, takes $O(\sqrt{N})$ operations to complete.

8.1 The Oracle

We need a quantum oracle O that can recognize the correct solutions to the search problem. Let there be M solutions out of the N elements, with $N > M$. We take $n = log_2(N)$ qubits to specify all the possible indices x. Let the oracle

be described by

$$f(x) = 1 \quad \text{if x is a solution to the search}$$

$$f(x) = 0 \quad \text{if x is not one of the solutions}$$

This recognition of when index register $|x\rangle$ is a solution can be implemented by using an ancillary oracle qubit $|q\rangle$ which is flipped when $f(x) = 1$

$$|x\rangle |q\rangle \rightarrow |x\rangle |q \oplus f(x)\rangle \qquad (\oplus \text{ denotes addition modulo 2 / XOR})$$

$|q\rangle$ can be initialised to $|0\rangle$, but it is preferred to initialise it to $|-\rangle$. This is because $|- \oplus f(x)\rangle = (-1)^{f(x)} |-\rangle$. Therefore, the oracle works as: $|x\rangle |-\rangle \rightarrow (-1)^{f(x)} |x\rangle |-\rangle$. Which means we can effectively ignore the oracle qubit and consider $|x\rangle \rightarrow (-1)^{f(x)} |x\rangle$. The oracle doesn't know all the solutions to a problem, but it can recognize whether a particular element is a solution or not.

8.2 Procedure

The algorithm requires $n = log_2 N$ qubits, and the oracle will require a few more qubits (unspecified here).

- It begins with the n qubits initialised to a state $|0\rangle^{\otimes n}$.
- Then the Hadamard gate is applied to create an equal superposition of basis states. We denote this state as

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle$$

- Then we apply a quantum subroutine called the "Grover iteration" a $O(\sqrt{N})$ number of times (The specific number of times we apply it depends on M and N). This subroutine we denote by G.
- Then we finally measure the circuit

8.3 Grover Iteration

The Grover Iteration consists of the following steps:

1. Apply the oracle O
2. Apply the n-qubit Hadamard transform
3. We apply a conditional phase shift so that every n-bit computational basis other than $|0\rangle^{\otimes N}$ receives a factor of -1

$$|x\rangle \rightarrow -(-1)^{\delta_0} |x\rangle \qquad \text{where x is an n-qubit basis}$$

4. Finally, we apply the n-qubit Hadamard transform again.

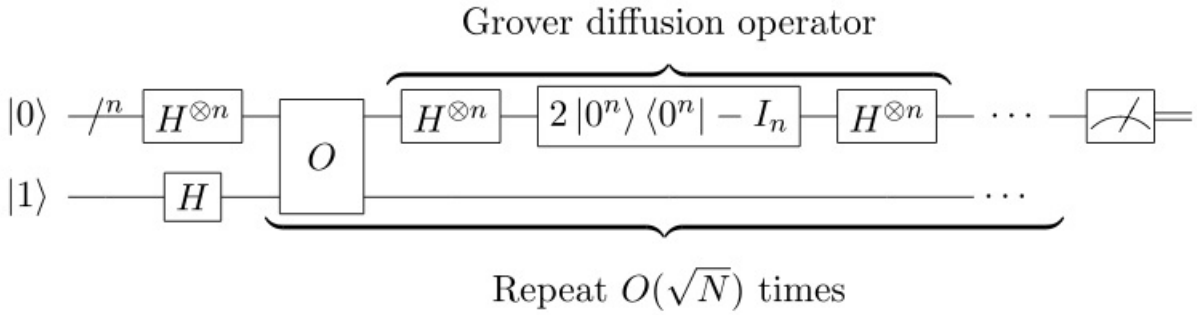

Figure 1: Circuit implementation of Grover's algorithm

The conditional phase shift operator corresponds to $(2\,|0^n\rangle\langle 0^n| - I)$. Its matrix form is:

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

When we consider the 2 Hadamard operations along with the phase shift, (step 2 and 4), we get the operator

$$D = H^{\otimes N}(2\,|0^n\rangle\langle 0^n| - I)H^{\otimes N} = (2\,|\psi\rangle\langle\psi| - I)$$

. This is called the Grover Diffusion operator. Its matrix form is:

$$\begin{bmatrix} \frac{2}{N}-1 & \frac{2}{N} & \frac{2}{N} & . & . & . & \frac{2}{N} \\ \frac{2}{N} & \frac{2}{N}-1 & \frac{2}{N} & . & . & . & \frac{2}{N} \\ \frac{2}{N} & \frac{2}{N} & \frac{2}{N}-1 & . & . & . & \frac{2}{N} \\ . & . & . & . & & & . \\ . & . & . & & . & & . \\ . & . & . & & & . & . \\ \frac{2}{N} & \frac{2}{N} & \frac{2}{N} & . & . & . & \frac{2}{N}-1 \end{bmatrix}$$

Therefore, the the Grover iteration $G = DO$.

8.4 Geometric Visualisation

Let $|\beta\rangle$ denote the normalized uniform superposition of all M solutions to the search problem and Let $|\alpha\rangle$ denote the normalized uniform superposition of $|x\rangle$ that are not solutions to the search problem in N. Let us consider the 2 dimensional vector space spanned by $|\alpha\rangle$ and $|\beta\rangle$. It is obvious that these 2 vectors form the orthonormal basis for this space. It is also obvious that $|\psi\rangle$ lies somewhere in this 2-dimensional space.

$$|\alpha\rangle = \frac{1}{\sqrt{N-M}} \sum_{x}^{\mathbf{M'}} |x\rangle$$

$$|\beta\rangle = \frac{1}{\sqrt{M}} \sum_{x}^{\mathbf{M}} |x\rangle$$

$$|\psi\rangle = \sqrt{\frac{N-M}{N}} |\alpha\rangle + \sqrt{\frac{M}{N}} |\beta\rangle$$

The oracle works by flipping the sign of any solution to the search. Therefore the action of O on any vector $|\phi\rangle = \phi_1 |\alpha\rangle + \phi_2 |\beta\rangle$ is

$$O|\phi\rangle = \phi_1 |\alpha\rangle - \phi_2 |\beta\rangle$$

Therefore, the vector gets reflected about $|\alpha\rangle$. The Diffusion operator $D = 2|\psi\rangle\langle\psi| - I$ acts similarly, by reflecting any vector about $|\psi\rangle$. We know that the product of 2 reflections is a rotation, hence the Grover Iteration rotates the any vector by a fixed angle. By comparing the oracle to the Diffusion matrix, we can say that the oracle can be expressed as

$$O = \mathbf{I} - 2|\beta\rangle\langle\beta|$$

Let us consider $\cos\frac{\theta}{2} = \sqrt{\frac{N-M}{N}}$ (the angle between $|\psi\rangle$ and $|\alpha\rangle$ being $\frac{\theta}{2}$), then the Grover iteration rotates any vector by an angle θ in an anticlockwise direction. We apply this iteration many times. Therefore after k iterations, we get

$$G^k |\psi\rangle = \cos\left(\frac{2k+1}{2}\theta\right) |\alpha\rangle + \sin\left(\frac{2k+1}{2}\theta\right) |\beta\rangle$$

. Hence the matrix form of G is

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Since this is a rotation, we always get a vector that lies in the same vector space. The objective of this algorithm is to get the vector as close as possible to $|\beta\rangle$ so that a measurement in the computational basis produces one of the outcomes in $|\beta\rangle$ with a high probability. Therefore, we have to minimise $\sin\left(\frac{2k+1}{2}\theta\right)|\beta\rangle$.

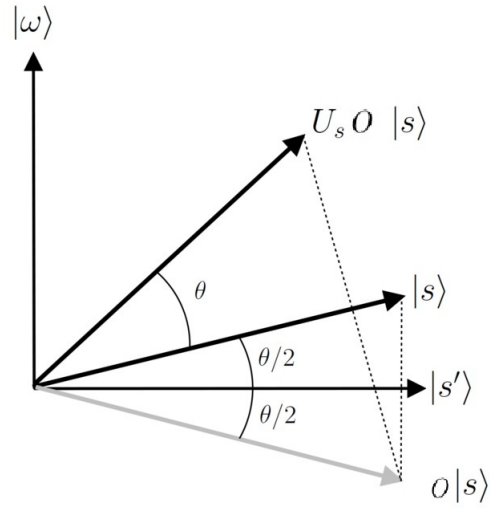

Figure 2: Visualisation of Grover's Algorithm

To do so, we apply the iteration $r = \dfrac{\arccos\left(\sqrt{M/N}\right)}{\theta}$ number of times (rounded to the nearest integer). Since $r < \lceil \pi/2\theta \rceil$, we can get an upper bound on the number of iterations required if we have a lower bound on θ. This upper bound is $r \leq \left\lceil \frac{\pi}{4}\sqrt{\frac{N}{M}} \right\rceil$ which is of the order $O\left(\sqrt{\frac{N}{M}}\right)$

8.5 Working of G

The oracle operator can be described as $I - 2\left|\beta\right\rangle\left\langle\beta\right|$. Its operation on the basis states are as follows:

$$O\left|\beta\right\rangle = \left(I - 2\left|\beta\right\rangle\left\langle\beta\right|\right)\left|\beta\right\rangle = \left|\beta\right\rangle - 2\left|\beta\right\rangle \cdot 1 = -\left|\beta\right\rangle$$

$$O\left|\alpha\right\rangle = \left(I - 2\left|\beta\right\rangle\left\langle\beta\right|\right)\left|\alpha\right\rangle = \left|\alpha\right\rangle - 2\left|\beta\right\rangle \cdot 0 = \left|\alpha\right\rangle$$

The operation of the Diffusion transform on the basis states is as follows:

$$D\left|\beta\right\rangle = \left(2\left|\psi\right\rangle\left\langle\psi\right| - I\right)\left|\beta\right\rangle$$

$$= \left(\sqrt{\frac{N-M}{N}}\left|\alpha\right\rangle + \sqrt{\frac{M}{N}}\left|\beta\right\rangle\right) \cdot 2\frac{M}{N} - \left|\beta\right\rangle$$

$$= 2\frac{M}{N}\left(\sqrt{\frac{N-M}{N}}\left|\alpha\right\rangle + \left(\sqrt{\frac{M}{N}} - \frac{N}{2M}\right)\left|\beta\right\rangle\right)$$

$$D\left|\alpha\right\rangle = \left(2\left|\psi\right\rangle\left\langle\psi\right| - I\right)\left|\alpha\right\rangle$$

$$= \left(\sqrt{\frac{N-M}{N}}\,|\alpha\rangle + \sqrt{\frac{M}{N}}\,|\beta\rangle\right) \cdot \frac{N-M}{N} - |\alpha\rangle$$

$$= 2\frac{(N-M)}{N}\left(\left(\sqrt{\frac{N-M}{N}} - \frac{N}{2(N-M)}\right)|\alpha\rangle + \sqrt{\frac{M}{N}}\,|\beta\rangle\right)$$

This is what happens in the first iteration of G:

$$G\,|\psi\rangle = DO(\sqrt{\frac{N-M}{N}}\,|\alpha\rangle + \sqrt{\frac{M}{N}}\,|\beta\rangle) = D(\sqrt{\frac{N-M}{N}}\,|\alpha\rangle - \sqrt{\frac{M}{N}}\,|\beta\rangle)$$

$$= \left(2\left(\frac{N-M}{N}\right)^2 + \left(2\left(\frac{M}{N}\right)^{3/2} - 1\right)\sqrt{\frac{N-M}{N}}\right)|\alpha\rangle$$

$$+ \left(\left(1 + \frac{N-M}{N}\right)\sqrt{\frac{M}{N}} - 2\left(\frac{M}{N}\right)^2\right)|\beta\rangle$$

9 Quantum simulation

To derive Grover's algorithm, we assume that the search has only 1 solution $|\omega\rangle$. We try to find a hamiltonian H such that a quantum system evolves from $|\psi\rangle$ (we take this to be a uniform superposition of all states) to $|\omega\rangle$ over some time. This hamiltonian satisfies the need:

$$H = |\psi\rangle\langle\psi| + |\omega\rangle\langle\omega|$$

After time t, the system will reach a state of $exp(-iHt)\,|\psi\rangle$. This lies in the space spanned by $|\psi\rangle$ and $|\omega\rangle$ and we decide the orthonormal basis to be $|\omega\rangle$ and some $|\gamma\rangle$ such that $|\psi\rangle = p\,|\omega\rangle + q\,|\gamma\rangle$ with $p^2 + q^2 = 1$. Here,

$$H = |\psi\rangle\langle\psi| + |\omega\rangle\langle\omega|$$

$$= \begin{bmatrix} p^2 & pq \\ pq & q^2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = I + p(qX + pZ)$$

After time t we get the quantum state

$$\cos\left(\frac{t}{\sqrt{N}}\right)|\psi\rangle - i\sin\left(\frac{t}{\sqrt{N}}\right)|x\rangle$$

After time $t = \pi\sqrt{N}/2$, the state reaches $|x\rangle$ and thus we reach our objective. We simulate the Hamiltonian by simulating its 2 terms individually for small amount of time Δt. To get a high probability of success of $O(1)$, We need to call the Oracle $O(N)$ times, which is the same as a classical algorithm. Suppose the simulation step is performed to an accuracy $O(\Delta t^r)$, we need $O(N^{r/2(r-1)})$ calls to the oracle that simulates H. Taking the limit of $r \to \infty$, we get the minimum number of calls possible to be $O(\sqrt{N})$, Hence Grover's algorithm is an optimal quantum search algorithm.

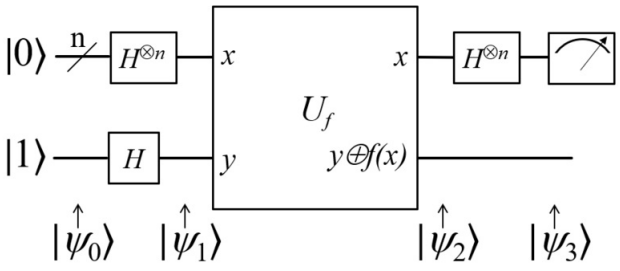

Figure 3: Circuit implementation of the Deutsch-Jozsa Algorithm

10 The Deutsch-Jozsa Algorithm

This algorithm shows that quantum algorithm can determine some global properties of a function faster than any classical algorithm can. The description of the Deutsch-Jozsa problem is: assume we have a function f such that it can be either balanced or constant (no other alternative allowed). Balanced means that $f(x)$ has 2 possible values with each outcome occurring equal number of times in the domain of this function while constant means that $f(x)$ has the same outcome for all x in the domain. We assume that the domain of the function has N discrete elements that can be simulated by $n = \log_2 N$ qubits. Classically, we would need $N/2$ calls to the oracle (that simulates $f(x)$). However, the Deutsch-Jozsa algorithm lets us accomplish this on a quantum computer with only 1 call to the quantum oracle

10.1 Procedure

Without loss of generality, we take take the function f to have a range $\{0, 1\}$. The quantum black box performs the following transformation: $|x\rangle |y\rangle \rightarrow |x\rangle |y \oplus f(x)\rangle$. We use $n + 1$ qubits for this algorithm. The n qubits correspond to the N elements in the domain and the 1 qubit is used to store the answer (The qubit is flipped if the function is constant and left alone otherwise). The register is initialised to $|0\rangle^{\otimes n} |1\rangle$. Then we apply a n-Hadamard transform on the n qubits and a single Hadamard on the answer qubit. The quantum state is now

$$|\psi_1\rangle = \left[\sum_{x\epsilon\{0,1\}^n} \frac{|x\rangle}{\sqrt{N}}\right] \left[\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right]$$

Performing XOR on the answer qubit just changes its sign ($\text{XOR}|-\rangle = \pm|-\rangle$). Therefore, we can consider the oracle operation to be

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

Which leaves us with the quantum state:

$$|\psi_2\rangle = \left[\sum_{x\epsilon\{0,1\}^n} \frac{(-1)^{f(x)}\,|x\rangle}{\sqrt{N}}\right]\left[\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right]$$

The last step is to apply the n-bit Hadamard again (ignoring the oracle qubit). To calculate the resulting state, we check the effect of the n-Hadamard on any vector $|x\rangle$

$$H^{\otimes n}\,|x\rangle = \sum_{z\epsilon\{0,1\}^n} \frac{(-1)^{x\cdot z}\,|z\rangle}{\sqrt{N}}$$

$$\implies H^{\otimes n}\left[\sum_{x\epsilon\{0,1\}^n} \frac{(-1)^{f(x)}\,|x\rangle}{\sqrt{N}}\right] = \sum_z \sum_x \frac{(-1)^{x\cdot z + f(x)}\,|z\rangle}{N}$$

Therefore, the quantum state of the register now becomes:

$$\implies |\psi_3\rangle = \left[\sum_z \sum_x \frac{(-1)^{x\cdot z + f(x)}\,|z\rangle}{N}\right]\left[\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right]$$

We are only concerned with the amplitude of the state $|000...0\rangle$ whose amplitude is

$$\sum_x \frac{(-1)^{f(x)}}{N}$$

. If the function is constant, the amplitude of x is either 1 or -1 and the amplitude of the rest of the states is 0. Hence, if the function is constant, the only possible outcome of measurement of the register is $|000...0\rangle$. If the function is balanced out, the amplitude of this basis becomes 0 and there must be some non zero amplitude in any of the other N-1 states (Since the vector has unit length). Therefore, if we measure $|0\rangle^{\otimes n}$, we know that the function is constant. If the outcome is anything else, the function is balanced.

11 The Quantum Fourier Transform

11.1 Discrete Fourier transform

It transforms a sequence of N complex numbers $\{x_n\} := x_0, x_1, ..., x_{N-1}$ to another sequence on N complex numbers $\{y_n\} := y_0, y_1, ..., y_{N-1}$ such that

$$y_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{\frac{-2i\pi}{N}kn}$$

Similarly, the quantum Fourier transform on an orthonormal basis is a linear operator acting on the basis states$|0\rangle, |1\rangle, ... |N-1\rangle$ in the following manner:

$$|k\rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} |n\rangle\, e^{\frac{-2i\pi}{N}kn}$$

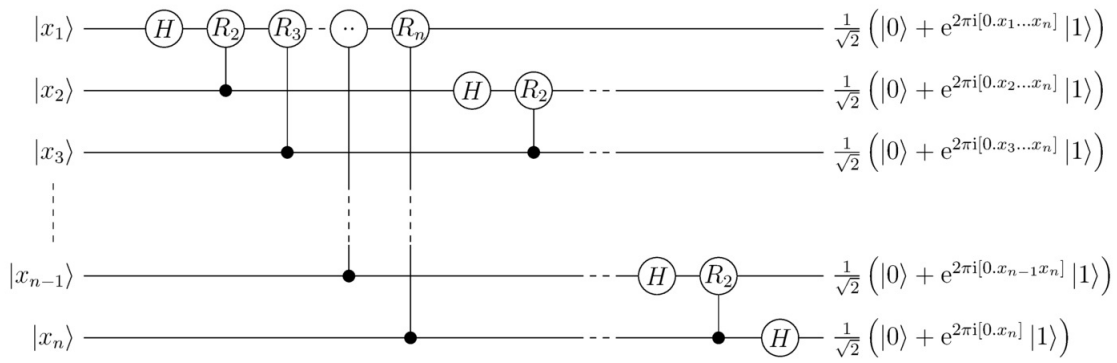

Figure 4: Circuit implementation of the quantum Fourier transform

. This is a unitary transformation and can be implemented in quantum circuits. The representation of these computational basis states becomes easier in binary notation. Instead of $|0\rangle, |1\rangle, ... |N-1\rangle$, we represent them as $|000...00\rangle, |000...01\rangle, ... |111...11\rangle$. We need $n = log_2 N$ qubits to represent these. The Fourier transform is easier to denote in this representation:

$$|j_1, j_2, ...j_n\rangle \rightarrow \frac{(|0\rangle + e^{2\pi i(0.j_n)}|1\rangle)(|0\rangle + e^{2\pi i(0.j_{n-1}j_n)}|1\rangle)...(|0\rangle + e^{2\pi i(0.j_1 j_2...j_n)}|1\rangle)}{\sqrt{2^n}}$$

Where $(0.j_k j_{k+1}...j_l)$ is a binary equivalent of fraction. We also need another unitary operator R_k whose matrix form is:

$$\begin{bmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{bmatrix}$$

11.2 Procedure

We apply the Hadamard gate followed by Controlled-R_k gates on each qubit. The circuit is as follows: We apply $R_n R_{n-1}...R_2 H$ on the first qubit, $R_{n-1}R_{n-2}...R_2 H$ on the second qubit and so on till the last qubit on which we apply only a hadamard gate. This gives us a quantum state of:

$$\frac{(|0\rangle + e^{2\pi i(0.x_1 x_2...x_n)}|1\rangle)...(|0\rangle + e^{2\pi i(0.x_{n-1}x_n)}|1\rangle)(|0\rangle + e^{2\pi i(0.x_n)}|1\rangle)}{\sqrt{2^n}}$$

Then we apply Swap gates to reverse the order of qubits to give us the desired quantum Fourier transform:

$$\frac{(|0\rangle + e^{2\pi i(0.x_n)}|1\rangle)(|0\rangle + e^{2\pi i(0.x_{n-1}x_n)}|1\rangle)...(|0\rangle + e^{2\pi i(0.x_1 x_2...x_n)}|1\rangle)}{\sqrt{2^n}}$$

This circuit proves that the quantum Fourier transform is unitary. The number of gates required pre-swap is $n(n+1)/2$ while the swap can be accomplished

using atmost $3n/2$ C-NOT gates. Therefore this is a $\theta(n^2)$ algorithm which is exponentially faster than best classical algorithms that implement the Fourier transform.

12 Phase estimation

This algorithm is used to determine the phase ϕ in the eigenvalue $e^{2\pi i\phi}$ of an eigenvector $|u\rangle$ of a unitary operator U. This algorithm requires an oracle that can perform the operation (U^{2^j})

12.1 Procedure

We take 2 registers, 1st one having t qubits initialised to $|0\rangle$ (t is chosen based on the accuracy we want in determining ϕ) and the second register is initialised to $|u\rangle$ and has the minimum number of qubits required to store $|u\rangle$.We apply Hadamard gates to each of the t qubits, followed by a sequence of Controlled U^{2^x} gates to the second register with x increasing from 0 to $t-1$. The second register stays on the same state $|u\rangle$ and gives us the following state in the first register:

$$\frac{1}{\sqrt{2^t}}\sum_{k=0}^{2^t-1} e^{2\pi i\phi k}\,|k\rangle$$

Now we apply the inverse Fourier transform on this state (by reversing the QFT circuit)

$$QFT\left(\frac{1}{\sqrt{2^t}}\sum_{k=0}^{2^t-1} e^{2\pi i\phi k}\,|k\rangle\right) = \frac{1}{\sqrt{2^t}}\sum_{k=0}^{2^t-1} e^{2\pi i\phi k}\frac{1}{\sqrt{2^t}}\sum_{x=0}^{2^t-1} e^{(\frac{-2\pi ikx}{2^t})}\,|x\rangle$$

$$= \frac{1}{2^t}\sum_{k=0}^{2^t-1}\sum_{x=0}^{2^t-1} e^{2\pi i\phi k}e^{(\frac{-2\pi ikx}{2^t})}\,|x\rangle$$

$$= \frac{1}{2^t}\sum_{k=0}^{2^t-1}\sum_{x=0}^{2^t-1} e^{-(\frac{2\pi ik}{2^t})(x-2^t\phi)}\,|x\rangle$$

We round off $2^t\phi$ to the nearest integer a $2^t\phi = a + 2^t\delta$ where $0 \le 2^t\delta \le 0.5$. Therefore the 1st register can be written as:

$$\frac{1}{2^t}\sum_{k=0}^{2^t-1}\sum_{x=0}^{2^t-1} e^{-\frac{2\pi ik}{2^t}(x-a)}e^{2\pi ik\delta}\,|x\rangle$$

Measuring this gives us $|a\rangle$ with a probability higher than $\frac{4}{\pi^2}$. The probability increases when the value of δ decreases (when $2^t\phi$ is very close to an integer) but

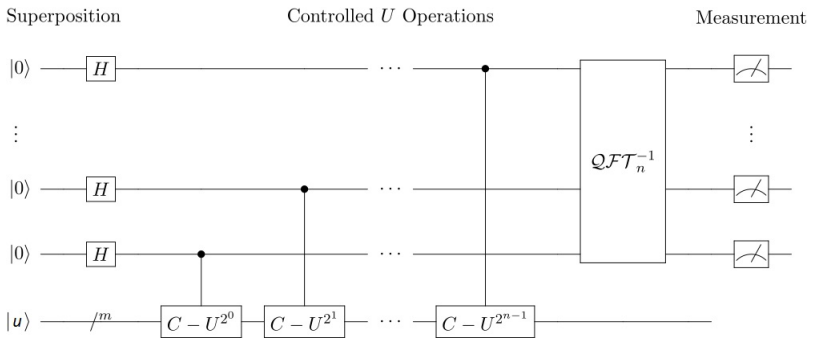

Figure 5: Circuit implementation of the phase estimation algorithm

decreases with the increase in the number of qubits used in the second register (t). However, a increase in t lets us determine the phase with a higher accuracy. Therefore there is an accuracy/probability trade off here. The probability of measuring $|a\rangle$ is:

$$P(|a\rangle) = \left| \langle a| \frac{1}{2^t} \sum_{k=0}^{2^t-1} \sum_{x=0}^{2^t-1} e^{-\frac{2\pi i k}{2^t}(x-a)} e^{2\pi i k\delta} |x\rangle \right|^2$$

$$P(|a\rangle) = \begin{cases} 1 & \delta = 0 \\ \frac{1}{2^{2t}} \left| \frac{\sin(2^t \pi\delta)}{\sin(\pi\delta)} \right|^2 & \delta \neq 0 \end{cases}$$

12.2 Order finding

Let r be the order of x modulo N. The quantum order finding algorithm is the same as the phase estimation algorithm applied to the unitary operator $U|y\rangle = |xy(modN)\rangle$. The eigenstates of this operator are:

$$|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{(\frac{2\pi i k s}{r})} \left|x^k modN\right\rangle$$

The eigenvalues of this operator are:

$$e^{(\frac{2\pi i s}{r})}$$

Hence when we apply the phase estimation algorithm using this operator, we get some value ϕ which is close to the value of $\frac{s}{r}$. It is an approximation of $\frac{s}{r}$ accurate upto $2L+1$ bits. From this value, we get the value of r. To do this, we use the continued fraction algorithm

12.2.1 Continued fraction expansion

This classical algorithm is based on the theorem:

$$\left|\frac{s}{r} - \phi\right| \leq \frac{1}{2r^2} \implies \frac{s}{r} \text{ is a convergent of the continued fraction of } \phi$$

We perform classical continued fraction expansion on ϕ to get approximations a/b till 2 conditions are satisfied:

$$b < N \text{ and } \left|\frac{a}{b} - \phi\right| \leq \frac{1}{2^{2L+1}} \qquad (\text{since } \frac{1}{2^{2L+1}} \leq \frac{1}{2r^2})$$

Assuming $\frac{a}{b}$ is irreducible, b is very likely to be either equal to r or a factor of it.

13 Shor's algorithm

Shor's algorithm is a quantum computing algorithm for factoring integers. The algorithm runs in polynomial time (in logN where N is the integer to be factorised) which is almost exponentially faster than the fastest classical algorithm (general number field sieve which works in sub exponential time)
Assuming N to be the product of 2 coprime integers greater than 2, The chinese remainder theorem tells us that there are atleast 4 distinct square roots of 1 modulo N. we tried to find such a square root (say α) (aside from the trivial 1 and -1) because:

$$\alpha^2 = 1(modN)$$

$$\implies (\alpha - 1)(\alpha + 1) = Np \qquad (\text{for some non-zero integer p})$$

This implies that $HCF(N, \alpha - 1)$ and $HCF(N, \alpha + 1)$ both are non trivial divisors of N.
Another important theorem that Shor's algorithm depends upon is:
Suppose $N = p_1^{q_1} p_2^{q_2} ....p_m^{q_m}$ (the prime factorization of an odd composite positive integer). Let x be chosen uniformly at random from Z_N, (the set of all positive integers less than N and co-prime to it) and let r be the order of x, (modulo N). Then

$$p(\text{r is even and } x^{r/2} \neq -1(modN)) \geq 1 - \frac{1}{2^m}$$

13.1 Reducing factoring to order-finding

We will assume that N is the product of 2 coprime integers greater than 2 and try to find these integers. We pick a random integer a less than N and coprime to it. (Other possible integers are quickly eliminated at the start of the algorithm). We know that all integers in this group have a finite order r. Therefore we know $a^r = 1(modN)$ exists for some r. We also know that 1 (modulo N) has atleast 4 square roots. Because of the second theorem, it is very likely that r is even and hence its arithmetic square root exists. The square root of a^r cannot be

equal to 1 (mod N) since by definition, r is the least integer that satisfies this property. We are left with the possibilities that its square root is -1 or atleast 2 other integers. If the square root turns out to be -1, then the algorithm fails for that particular a. If however, its square root is $\neq$ -1, then $(a^{r/2}+1)(a^{r/2}-1)$ is a multiple of N and atleast one of $HCF(N, a^{r/2}-1)$ and $HCF(N, a^{r/2}+1)$ is a non trivial factor of N.

13.2 Procedure

- Pick a random $a < N$
- If $HCF(a, N)$ (calculated using the euclidean algorithm) $\neq 1$, we found the non-trivial factor
- Else, find the order of a modulo N using the quantum order finding algorithm
- If r is odd or $a^{r/2} = -1(modN)$, then the algorithm fails for this particular a.
- Else atleast one of $HCF(N, a^{r/2}-1)$ and $HCF(N, a^{r/2}+1)$ is a non trivial factor of N.

14 References

- Quantum Computation and Quantum Information 10th Anniversary Edition by Michael A. Nielsen, Isaac L. Chuang
- Quantum Information Science 1 - MITx: 8.3.70 (Archived course, edX)
- Lecture Notes of Prof. John Preskill for Ph219/CS219, Caltech University - All circuit images are taken from their respective pages on Wikipedia (en.wikipedia.org)