ANALYSIS OF QUANTUM ALGORITHMS

Paresha Farastu

Rafiduzzaman Sonnet

Saddat Hasan

Sandipon Paul

Supervisor: Professor Mahbub Alam Majumdar

A thesis submitted in partial fulfillment for the degree of Bachelor of Science in the Department of Computer Science and Engineering, BRAC University

August 2017

DECLARATION

We, hereby, declare that the thesis titled “ANALYSIS OF QUANTUM ALGORITHMS ” is submitted to the Department of Computer Science and Engineering of BRAC University in partial fulfilment of the Bachelor of Science degree. This is our original work and was not submitted elsewhere for the award of any other degree or any other publication.

Date:

Dr. Mahbub Majumdar

Thesis Supervisor

_ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _

Paresha Farastu(EEE)

ID: 13321067

_ _ _ _ _ _ _ _ _ _ _ _ _

Rafiduzzaman Sonnet(CS)

ID: 17141002

_ _ _ _ _ _ _ _ _ _ _ _ _

Saddat Hasan(CS)

ID: 12201112

_ _ _ _ _ _ _ _ _ _ _ _ _

Sandipon Paul(CS)

ID: 17141012

_ _ _ _ _ _ _ _ _ _ _ _ _

ACKNOWLEDGEMENTS

We would like to convey our utmost gratitude to our supervisor, Dr. Mahbub Majumdar for his constant encouragement and support. His unwavering belief in our work catapulted our dedication and his expert advice proved to be the perfect guideline to tackle this difficult endeavour.

Furthermore, we would like to mention a few books which played a significant role in developing the core skills to write the thesis report. Quantum Computation and Quantum Information by Michael A. Nielsen and Isaac L. Chuang , Introduction to Quantum Mechanics by David J. Griffiths, Quantum Information by Stephen M. Barnett.

Contents

Chapter 1

Formalism of Quantum Mechanics

1.1 Linear Vector space

Definition: A *linear vector space* V is a collection of objects ψ_a, ψ_b, ..., called *vectors*, which satisfy the following postulates:

1. **Closure**: If $\psi_a, \psi_b \in V$, there is a unique vector $\psi_c = \psi_a + \psi_b$ in V. In other words, a vector space is closed under addition.

2. **Associativity of vector addition**:

$$\psi_a + \psi_b = \psi_b + \psi_a$$

3. **Commutativity of vector addition**:

$$\psi_a + (\psi_b + \psi_c) = (\psi_a + \psi_b) + \psi_c$$

4. **Additive identity**: There is a vector in V called the null vector and denoted by ϕ satisfying

$$\psi_a + \phi = \phi + \psi_a$$

for every ψ_a in V.

5. **Inverse**: For every vector ψ_a in V there is another vector ψ'_a in V such that

$$\psi_a + \psi'_a = \phi$$

We denote ψ'_a as $-\psi_a$.
Note: We use the notation $\psi_a - \psi_b$ to mean $\psi_a + (-\psi_b)$

6. **Associativity of scalar multiplication**:

$$(\lambda\mu)\psi_a = \lambda(\mu\psi_a)$$

7. **Distributivity of scalar multiplication with respect to vector addition**:

$$\lambda(\psi_a + \psi_b) = \lambda\psi_a + \lambda\psi_b$$

8. **Distributivity of scalar multiplication with respect to scalar addition**:

$$(\lambda + \mu)\psi_a = \lambda\psi_a + \mu\psi_a$$

9. Multiplication by scalars 0 and 1 (**multiplicative identity**) are defined by

$$0\psi_a = \phi$$

$$1\psi_a = \psi_a$$

for any ψ_a in V.

1.2 Inner-product space or a unitary vector space

For a general linear complex vector space, product of vectors (i.e., multiplication of two vectors) need not be defined. However, we will restrict ourselves to spaces in which a *scalar product* or an *inner product* is defined.

A linear vector space is called *unitary* if a scalar product is defined in it. To every pair of vectors ψ_a and ψ_b in V there corresponds a unique scalar (in general complex), called the scalar product.

$$(\psi_a, \psi_b) = (\psi_a^*)\psi_b$$

The scalar product is defined to have the following properties:

(a) $(\psi_a, \psi_b) = (\psi_b, \psi_a)^*$

(b) $(\psi_a, \lambda\psi_b) = \lambda(\psi_a, \psi_b)$

(c) $(\lambda\psi_a, \psi_b) = \lambda^*(\psi_a, \psi_b)$

(d) $(\psi_a, \psi_b + \psi_c) = (\psi_a, \psi_b) + (\psi_a, \psi_c)$

(e) $(\psi_a + \psi_b, \psi_c) = (\psi_a, \psi_c) + (\psi_b, \psi_c)$

(f) $(\psi_a, \psi_a) \geq 0$; the equality holds only if ψ_a is the null vector

1.3 Norm of a vector

If a scalar product is defined in a vector space, then the scalar product gives us the concept of the 'magnitude' or 'length' of a vector. In a general vector space the 'magnitude' or 'length' of a vector is called the norm of the vector. We simply define the norm of a vector ψ_a as

$$\|\psi_a\| \stackrel{def}{=} \sqrt{(\psi_a, \psi_a)}$$

The norm has the following properties:

(a) $\|\psi_a\| \geq 0$, the equality holds only if the vector is null

(b) $\|\psi_a + \psi_b\| \geq \|\psi_a\| + \|\psi_b\|$, this is called the triangle inequality

(a) $\|\psi_a - \psi_b\| = \|\psi_b - \psi_a\|$

A vector whose norm is unity is called a unit vector. For any given non-null vector, a unit vector can be formed by dividing the vector by its norm. Thus

$$u_a = \frac{\psi_a}{\|\psi_a\|}$$

is normalized.

1.4 Independence, Basis and Dimension

Linear Independence

The set of vectors $\psi_a, \psi_2, ..., \psi_N$ are *linearly independent* if none of them can be expressed as a linear combination of the others. Mathematically this means that the equation

$$\sum_{j=1}^{N} c_j \psi_j = 0$$

cannot be satisfied except by $c_j = 0$ for all j, where c_j is a constant.

Vectors that Span a Subspace

If all the vectors (v_i) in a subspace W can be represented by a linear combination of a set of vectors $S = \psi_a, \psi_2, ..., \psi_N$, we then say that the space W is *spanned* by S[8].

$$v_i = \sum_{j=1}^{N} c_j \psi_j$$

For example, the set $\{(1,0,0),(0,1,0),(1,1,0)\}$ spans the space $\mathbb{R}^2$ (a subspace of $\mathbb{R}^3$). The set of vectors is not necessarily linearly independent.

A basis for a vector space

A *basis* for a vector space V is a set of vectors that are linearly independent and they span the space V. It means that every vector in V is a linear combination of the basis vectors and that combination is unique for every vector because the basis vectors are linearly independent.

For example, the set of vectors $\{(1,0,0),(0,1,0),(0,0,1)\}$ span the whole $\mathbb{R}^3$ and are linearly independent and thus called a basis for the space $\mathbb{R}^3$.

Dimension of a vector space

The dimension of a vector space is the number of vectors in every basis. A vector space is said to be n-dimensional if there exists n linearly independent vectors (basis vectors).

1.5 Orthogonality and Orthonormality

Two vectors ψ_a and ψ_b are *orthogonal* if their inner product is zero, i.e., if

$$(\psi_a, \psi_b) = 0$$

The unit vectors u_1, u_2, ..., u_N form an *orthonormal* set if they are mutually orthogonal, i.e., if

$$(u_i, u_j) = \delta_{ij}$$

1.5.1 Orthonormality and linear independence

A set of mutually orthogonal non-zero vectors (not necessarily normalized) are necessarily linearly independent. The converse is not true, however. That is, a set of linearly independent vectors may not be mutually orthogonal. For example, the vectors (1,0) and (1,1) are linearly independent but their inner product is non-zero proving that they are not orthogonal.

It is always possible to orthonormalize a set of linearly independent vectors. By this we mean that from a given set of N linearly independent vectors, it is possible to form a set of N orthonormal vectors. This procedure is called Gram-Schmidt orthonormalization method[4].

1.5.2 Gram-Schmidt orthonormalization method

Suppose $\psi_1, \psi_2, ..., \psi_N$ is a set of linearly independent vectors.

Let $$u_1 = \frac{\psi_1}{\|\psi_1\|} \qquad \text{eq 1}$$

Then $$(u_1, u_1) = 1 \qquad \text{i.e., } u_1 \text{ is normalized.}$$

Next construct the vector ψ_2' as follows:

$$\psi_2' = \psi_2 - u_1(u_1, \psi_2) \qquad \text{eq 2}$$

i.e., to obtain ψ_2' we have subtracted the 'component' of ψ_2 along the u_1 "direction". Then it follows that

$$\begin{aligned}(u_1, \psi_2') &= (u_1, \psi_2) - (u_1, u_1)(u_1, \psi_2) \\ &= (u_1, \psi_2) - (u_1, \psi_2) \\ &= 0\end{aligned}$$

i.e., ψ_2' is orthogonal to u_1. We then normalize ψ_2', i.e.,

$$u_2 = \frac{\psi_2'}{\|\psi_2'\|} \qquad \text{eq 3}$$

We can continue the process until we exhaust all the vectors. For example, in the next step we can write

$$\psi_3' = \psi_3 - u_1(u_1, \psi_3) - u_2(u_2, \psi_3) \qquad \text{eq 4}$$

We note immediately that ψ_3' is orthogonal to both u_1 and u_2, i.e.,

$$(u_1, \psi_3') = (u_2, \psi_3') = 0$$

We normalize ψ_3' to get u_3, i.e.,

$$u_3 = \frac{\psi_3'}{\|\psi_3'\|} \qquad \text{eq 5}$$

Finally, in the N^{th} step, we write

$$\psi_N' = \psi_N - u_1(u_1, \psi_N) - u_2(u_2, \psi_N) - .... - u_{N-1}(u_{N-1}, \psi_N)$$

ψ_N' is orthogonal to $u_1, u_2, ..., u_{N-1}$, i.e.,

$$(u_1, \psi_N') = (u_2, \psi_N') = ... = (u_{N-1}, \psi_N') = 0$$

Normalizing ψ_N' we get

$$u_N = \frac{\psi_N'}{\|\psi_N'\|} \qquad \text{eq 6}$$

Thus, the set $\{u_1, u_2, ..., u_N\}$ is an orthonormal set of vectors.

1.6 Hilbert Space

1.6.1 Complete vector space

A sequence of vectors ψ_n in the vector space V is called a Cauchy sequence if for every $\epsilon > 0$ there exists an integer N such that

$$\|\psi_n - \psi_m\| < \epsilon$$

if $n, m > N$. In other words, the vectors in the sequence come 'closer' if the index increases.

In particular

$$\|\psi_n - \psi_m\| \to 0 \text{ as } n, m \to \infty$$

A linear vector space is said to be *complete* if any Cauchy sequence converges to a vector in the space.

1.6.2 Hilbert Space

A Hilbert space is a vector space H with an inner product (f, g) such that the norm defined by

$$|f| = \sqrt{(f, f)}$$

turns H into a complete metric space. If the metric defined by the norm is not complete, then H is instead known as an inner product space.

1.7 Dirac Notation

Vectors

A vector v of a vector space is called a *ket*. It is denoted as

$$|v\rangle = \begin{pmatrix} v_0 \\ v_1 \\ . \\ . \\ . \\ v_n \end{pmatrix}$$

Scalar product

With each pair of kets $|\phi\rangle$ and $|\psi\rangle$, taken in this order, we associate a complex number, which is their scalar product $(|\phi\rangle, |\psi\rangle)$ and which satisfies various properties discussed earlier in 1.2.

Dual vector space

Linear functional:

A linear functional x is a linear operation on the kets such that x operating on a ket $|\psi\rangle$ gives a complex scalar:

$$x\,|\psi\rangle \rightarrow \text{scalar, where } |\psi\rangle \in V$$

and

$$x(\lambda_1\,|\psi_1\rangle + \lambda_2\,|\psi_2\rangle) = \lambda_1 x\,|\psi_1\rangle + \lambda_2 x\,|\psi_2\rangle\,.$$

The set of all linear functionals defined on the kets of a vector space V themselves form a linear vector space called the dual space of V and symbolized by V^*.

Bra notation for the vectors of V^*

Any element, or vector, of the space V^* is called a "bra vector", or, more simply, a bra. It is denoted as

$$\langle v| = |v\rangle^*$$

Correspondence between kets and bras

The existence of the scalar product in V will now enable us to show that we can associate with every ket $|\phi\rangle \in V$ an element of V^*, that is a bra, which will be denoted by $\langle\phi|$.

The ket $|\phi\rangle$ does indeed enable us to define a linear functional, the one which associates with each $|\psi\rangle \in V$ a complex number which is equal to the scalar product $(|\phi\rangle, |\psi\rangle)$[4]. Let $\langle\phi|$ be this linear functional. It is thus defined by the relation

$$\langle\phi|\psi\rangle = (|\phi\rangle, |\psi\rangle).$$

The correspondence is antilinear

Let $\lambda_1 |\phi_1\rangle + \lambda_2 |\phi_2\rangle$ be a ket. Then

$$\begin{aligned}(\lambda_1 |\phi_1\rangle + \lambda_2 |\phi_2\rangle, |\psi\rangle) &= \lambda_1^*(|\phi_1\rangle, |\psi\rangle) + \lambda_2^*(|\phi_1\rangle, |\psi\rangle) \\ &= \lambda_1^* \langle\phi_1|\psi\rangle + \lambda_2^* \langle\phi_2|\psi\rangle \\ &= (\lambda_1^* \langle\phi_1| + \lambda_2^* \langle\phi_2|) |\psi\rangle\end{aligned}$$

Thus

$$\lambda_1 |\phi_1\rangle + \lambda_2 |\phi_2\rangle \overset{dc}{\to} \lambda_1^* \langle\phi_1| + \lambda_2^* \langle\phi_2|$$

where "dc" is short for dual correspondence.

Dirac notation for the scalar product

The scalar product of $|\psi\rangle$ by $|\phi\rangle$ is represented in three ways, namely, $(|\phi\rangle, |\psi\rangle)$, $\langle\phi|\psi\rangle$ and $(\langle\phi|)(|\psi\rangle)$, $\langle\phi|$ being the bra associated with the ket $|\phi\rangle$.

We shall mostly use the Dirac notation $\langle\phi|\psi\rangle$. In the table below we summarize, in Dirac notation, the properties of the scalar product:

1. $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$
2. $\langle\phi|\lambda_1\psi_1 + \lambda_2\psi_2\rangle = \lambda_1 \langle\phi|\psi_1\rangle + \lambda_2 \langle\phi|\psi_2\rangle$
3. $\langle\lambda_1\phi_1 + \lambda_2\phi_2|\psi\rangle = \lambda_1^* \langle\phi_1|\psi\rangle + \lambda_2^* \langle\phi_2|\psi\rangle$
4. $\langle\psi|\psi\rangle$ is real, positive; zero if and only if $|\psi\rangle = \phi$ (null).

1.8 Operators in a Hilbert space

An operator is a linear transformation by which every vector ψ_a in a Hilbert space H is associated with another vector ψ_b in the space:

$$\hat{A} : \psi_a \to \psi_b$$

for $\psi_a, \psi_b \in H$. We usually employ the notation

$$\psi_b = \hat{A}\psi_a$$

In Dirac notation, we write

$$|b\rangle = \hat{A} |a\rangle$$

where both $|a\rangle$ and $|b\rangle$ belong in the ket space.

An operator can also act on a bra vector (bra-space is also a Hilbert space; it is dual to the ket space) changing it to another bra-vector. The notation we employ is

$$\langle\psi| = \langle\phi| \hat{A}$$

Here the operator $\hat{A}$ acts on the bra-vector $\langle\phi|$ to produce the bra vector $\langle\psi|$. We place the bra-vector on which the operator acts on the left of the operator.

1.8.1 Linear operators

An operator $\hat{A}$ is said to be a linear operator if it has the following property: For any vectors $|a\rangle$ and $|b\rangle$ and any complex number λ_1 and λ_2, we have

$$\hat{A}(\lambda_1 |a\rangle + \lambda_2 |b\rangle) = \lambda_1 \hat{A} |a\rangle + \lambda_2 \hat{A} |b\rangle$$

A linear ooperator can act on a bra vector also.

$$(\lambda_1 \langle a| + \lambda_2 \langle b|)\hat{A} = \lambda_1 \langle a| \hat{A} + \lambda_2 \langle b| \hat{A}$$

The operator $\hat{A}$ is antilinear if

$$\hat{A}(\lambda_1 |a\rangle + \lambda_2 |b\rangle) = \lambda_1^* \hat{A} |a\rangle + \lambda_2^* \hat{A} |b\rangle$$

Two operators $\hat{A}$ and $\hat{B}$ are equal if

$$\hat{A} |\psi\rangle = \hat{B} |\psi\rangle$$

for all $|\psi\rangle$ in the vector space.

Sum of two operators $\hat{A}$ and $\hat{B}$ is defined as

$$(\hat{A} + \hat{B}) |\psi\rangle = \hat{A}| |\psi\rangle + \hat{B} |\psi\rangle$$

Product of two operators $\hat{A}$ and $\hat{B}$ is defined as

$$(\hat{A}\hat{B}) |\psi\rangle = \hat{A}(\hat{B} |\psi\rangle)$$

This equation says that the operator $\hat{A}\hat{B}$ acting on $|\psi\rangle$ produces the same vector which would be obtained if we first let $\hat{B}$ on $|\psi\rangle$ and then $\hat{A}$ acts on the result of the previous operation. In general $\hat{A}\hat{B} \neq \hat{B}\hat{A}$, although in exceptional cases we may have $\hat{A}\hat{B} = \hat{B}\hat{A}$.

1.8.2 Commutator of two operators

The commutator of two operators $\hat{A}$ and $\hat{B}$ is defined as

$$[\hat{A}, \hat{B}] \stackrel{def}{\equiv} \hat{A}\hat{B} - \hat{B}\hat{A}$$

In general $[\hat{A}, \hat{B}] \neq 0$ (null operator). If $[\hat{A}, \hat{B}] = 0$, we say that $\hat{A}$ and $\hat{B}$ commute with each other.

Some properties of commutators

$$\begin{aligned}
[\hat{A}, \lambda\hat{B}] &= \lambda[\hat{A}, \hat{B}] \\
[\lambda\hat{A}, \hat{B}] &= \lambda[\hat{A}, \hat{B}] \\
[\hat{A}, \hat{B} + \hat{C}] &= [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}] \\
[\hat{A}, \hat{B}] &= -[\hat{B}, \hat{A}] \\
[\hat{A}, \hat{B}\hat{C}] &= \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C} \\
[\hat{A}\hat{B}, \hat{C}] &= \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}
\end{aligned}$$

1.8.3 Anticommutator of two operators

The anticommutator of two operators is defined by

$$\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$$

Chapter 2

Quantum Mechanics

Quantum mechanics is not just the physics of small things, it is the complete overhaul of Newtonian mechanics. In general it is a rethink of how the world works.

2.1 Superposition

The principle of quantum superposition is simple enough to explain, but once understood it wreaks havoc in one's mind. Much like waves in classical physics, superimposing(adding) any two or more quantum states produces another quantum state. But quantum states are strictly binary. there are exactly two of them and if superposition creates another state then we are trapped within our own logic, unless both the quantum states exist at the same time and only takes form as any one state when observed. It is a clever way to keep the quantum logic intact while throwing the classical logic out of the canonical cat box.
This phenomenon asks an intriguing and vital question, what do we even mean by observing? Why can only conscious minds cause collapse? It seems like when we are not looking everything is possible. But what would it even mean for one object to have two different speeds at the same time? The superposition principle clearly states that while we are not looking at an object, in general it acts like it is in many states at once. No matter how ridiculous this may sound, once it is coupled with the rule of measurement it starts making some sense. The measurement rule states, whenever we measure a certain property of an object it must choose one state to be in. Understanding the explanation of superposition is never satisfactory as it poses more questions than it answers, but in any case, this is the current Quantum Mechanical world-view. [Refer to 2.10.1]

2.2 Double Slit Experiment

Thomas Young[1], well before the advent of Quantum Mechanics, in 1801, came up with an experiment which made the wave theory of light acceptable. A refinement to Young's experimental model lead scientists to the Double-slit experiment, which established the unwavering demonstration that matter can display the characteristics of both classical waves and particles. When electrons are passed through two slits from a single source, it is observed at the screen to have produced a pattern which corresponds to the pattern that occurs when two waves have been passed through the slits. Electrons behave as if they have been spread out everywhere To ensure

that the particles have not miraculously ceased being particles, when we check them at the screen, they seem to be intact particles.
So naturally we decide to track the particles and look at it from start to end during the experiment to see if it is engaging in any funny business. Lo and behold! Our particles do not produce the pattern that justified our intrusive suspicion.
Therefore, the electrons doing all the crazy stuff while we were not looking actually made a big difference to what happened when we were.

2.2.1 Entanglement

The world of quantum physics is populated with a rich culture of concepts and according to Bohrs Copenhagen interpretation, our experience of a well-defined material universe only has meaning at the moment of measurement.

Einstein disproved of Bohrs idea of abandoning the assumption of realism and locality. Locality is the idea that each bit of universe only acts in its immediate surroundings, which is fundamental to Einsteins relativity, therefore fixing the speed of light as a cosmic limit.

Along with Podolsky and Rosen, Einstein wrote the EPR paper which introduced one of the more mysterious ideas of Quantum Mechanics, Quantum Entanglement.

When two particles interact briefly and as they influence each other, their properties become somehow connected. Hence, we describe the particle pair with 1 wave function that encompasses all possible states of both particles, making an entangled pair and according to the Copenhagen interpretation, any measurement of the particle automatically collapses the entire entangled wave function and also affects the measurement of the other particles.

This phenomenon holds even when the particles of the entangled pair are separated by any theoretical distance thus violating location and causality.

2.3 The Wave Function

From the double slit experiment we learned about the wave-particle duality but we never figured out what the particles are doing behind our backs. Quantum Mechanics deals with this by basically side stepping this issue. Instead it only talks about the wave function associated with that particle. Before we can understand what the wave function is, we should talk about what it does. Recalling the superposition principle, it tells us how the wave function looks like in each of its many possible states.

$$\Psi(x,t)$$

Therefore, from the experiment, the wave function of the particle is the superposition of the wave functions of going through slit 1 and slit 2. Using the wave function, we can predict the outcome of any experiment done so far, and it's the reason that we get all the quantum phenomenon like entanglement, the Heisenberg's uncertainty principle etc. To assume that the particle is doing all possible things at once is an overstatement as we do not have any idea as to what the particle is actually doing, all we know is its wave function. But the wave function has no physical interpretation. What we are concerned with is the square of the magnitude of the wave function, evaluated at a certain place and at a certain time is proportional to finding the particle there at that time.

As probabilities must be a positive real quantity, we multiply the complex wave function with its conjugate.

$$|\Psi(x,t)|^2 = \Psi(x,t)^*\Psi(x,t) = \rho(x,t)$$

The war waged between the probabilistic and the deterministic views of reality goes on, and it dares to ask whether we should even worry about the interpretations of scientific theory. Should it be enough that the theory makes the right predictions even though we do not understand why it works?

2.4 Stationary states

Although the behaviour of all subatomic particles is inherently probabilistic, Schrodinger's equation[2] does not itself contain any probabilities. The probability of every possible observation is determined by the wave function, but prior to the observation, the wave function changes and evolves in a completely deterministic manner. This deterministic way in which a wave function behaves without any observer present is what Schrodinger's Equation describes.
Charged particles exert forces on one another. We can view this by saying that at different points in space a charged particle has a different amount of potential energy(V).
Schrodinger equation becomes:

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi$$

where

$$\begin{aligned}
\hbar &= \text{Planck's constant}/2\pi \\
m &= \text{mass of the particle} \\
\psi &= \text{wave function of the particle at each point in space \& time} \\
V(x,t) &= \text{specified Potential Energy}
\end{aligned}$$

We shall consider V to independent of time, t.

$$\begin{aligned}
\psi(x,t) &= \psi(x)\phi(t) \\
\frac{\partial\psi}{\partial t} &= \psi\frac{\partial\phi}{dt} \\
\frac{\partial^2\psi}{\partial x^2} &= \frac{d^2\psi}{dx^2}\phi
\end{aligned}$$

Schrodinger's equation now becomes

$$\begin{aligned}
i\hbar\psi\frac{d\phi}{dt} &= -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}\phi + V\psi\phi \\
i\hbar\frac{1}{\phi}\frac{d\phi}{dt} &= -\frac{\hbar^2}{2m}\frac{1}{\psi}\frac{d^2\psi}{dx^2} + V \\
i\hbar\frac{1}{\phi}\frac{d\phi}{dt} &= E \qquad \text{eq 1}
\end{aligned}$$

or,
$$\frac{d\phi}{dt} = -\frac{iE}{\hbar}\phi$$

and
$$-\frac{\hbar^2}{2m}\frac{1}{\psi}\frac{d^2\psi}{dx^2} + V = E$$

or,
$$-\frac{\hbar^2}{2m}\left(\frac{d^2\psi}{dx^2}\right) + V\psi = E\psi \qquad \text{eq 2}$$

The first equation dissolved down to:
$$\phi(t) = e^{-iEt/\hbar}$$

The second equation is the time-independent Schrodinger equation, which needs the potential $V(x)$ to be specified to proceed further.

The vitality of separable solutions can be expressed in 2 physical and 1 mathematical ways[2].

(i) The wave function depends on t but the probability density does not. The time dependence also cancels out while calculating the expectation value of any dynamic variable. Therefore the separable solutions to the time-independent Schrodinger's equation are quantum states which are stationary states.

Stationary states are called so, because with the progression of time the quantum system remains in the same state. Thus, for a single-particle Hamiltonian this means that the particle has a constant probability distribution for its velocity, spin and position. We have assumed that the particle's environment is static.

This stationary state is much like the approximate stationary state of an orbital of a one-electron system (i.e. atomic orbital or molecular orbital).

$$\Psi(x,t) = \psi(x)e^{-Et/\hbar} \qquad \rightarrow \text{wave function dependent on t}$$

$$\begin{aligned} |\Psi(x,t)|^2 &= \Psi^*\Psi \\ &= \Psi^* e^{+iEt/\hbar}\ \Psi e^{-iEt/\hbar} \\ &= |\psi(x)|^2 \qquad \rightarrow \text{probability density independent of t} \end{aligned}$$

something something 1.36 momentum and 1.33

(ii) Through the concept of stationary states we have established that the Hamiltonian is unchanging in time. So they are states of definite total energy. The total energy of a classical system is the sum of kinetic and potential energies, and this is called the Hamiltonian:

$$H(x,p) = \frac{p^2}{2m} + V(x)$$
$$P \rightarrow \left(\frac{\hbar}{i}\right)\left(\frac{\partial}{\partial x}\right)$$
$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$$

Using equation (ii) the TISE becomes

$$\hat{H}\left|\psi\right\rangle = E\psi\left|\psi\right\rangle$$
$$\hat{H}\psi = E\psi$$

where $|\psi\rangle$ is an eigenvector of $\hat{H}$ and $E\psi$ is its eigenvalue.
The expectation value of the total energy is:

$$\langle H\rangle = \int \psi^* \hat{H}\psi dx = E\int |\psi|^2 dx = E\int |\psi|^2 dx = E$$
$$\hat{H}^2\psi = \hat{H}(\hat{H}\psi) = \hat{H}(E\psi) = E(\hat{H}\psi) = E^2\psi$$
$$\langle H^2\rangle = \int \psi^* \hat{H}^2\psi dx = E^2\int |\psi|^2 dx = E^2$$

This suggests that the variance of H is 0.

$$\sigma^2 H = 0$$

Ass all measurements of the total energy will return the value E, therefore this can be regarded as a property of separable solutions[3].

(iii) The T.I.S.E produces an infinite collection of solutions and each of these solutions has its own value of the separation constant (E). Hence, the linear combination of the separable solutions produces a general solution and every different wave function is responsible for each allowed energy.

$$\Psi_1(x,t) = \psi_1(x)e^{-iE_1 t/\hbar}$$
$$\Psi_2(x,t) = \psi_2(x)e^{-iE_2 t/\hbar}$$

General solution: $$\Psi(x,t) = \sum_{n=1}^{\infty} C_n\psi_n x(x)e^{-iE_n t/\hbar} = \sum_{n=1}^{\infty} C_n\Psi_n(x,t)$$

Separable solution: $$\Psi(x,t) = \psi_n(x)e^{-iE_n t/\hbar}$$

Thus solving the time independent Schrondinger's equation essentially will g .....

2.5 The Infinite Square Well

The infinite square well demonstrates a particle bound at two ends, but is completely free in between those ends. An infinite force will prevent the particle from escaping. The probability of finding the particle outside the well, where the potential is infinite, is zero. But inside the well where the potential is zero, the TISE[2] becomes:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi$$

Classical simple harmonic equation:

$$\frac{d^2\psi}{dx^2} = -k^2\psi, \qquad \text{where } k = \frac{\sqrt{2mE}}{\hbar}$$

We have also assumed that $E \geq 0$.
General solution:

$$\psi(x) = A\sin kx + B\cos kx$$

Here the constants A and B are determined by the boundary conditions.

$$V(x) = \begin{cases} 0, & \text{if } 0 \leq x \leq a \\ X, & \text{otherwise} \end{cases}$$

$$\begin{aligned} \psi(0) &= \psi(a) = 0 \\ \psi(0) &= A\sin 0 + B\cos\theta = B \\ B &= 0 \end{aligned}$$

$$\begin{aligned} \psi(x) &= A\sin kx \\ \psi(a) &= A\sin ka \\ ka &= 0, \pm\pi, \pm 2\pi... \end{aligned}$$

$k = 0$ implies that $\psi(x) = 0$.
So distinct solutions are:

$$\begin{aligned} k_n &= \frac{n\pi}{a} \qquad n = 1, 2, 3 \\ E_n &= \frac{\hbar^2 {k_n}^2}{2m} = \frac{n^2\pi^2\hbar^2}{2ma^2} \end{aligned}$$

For a quantum particle in the infinite square well we get an infinite set of solutions of the TISE. The first of those solutions are similar to standing waves on a string. The solution with the lowest energy is the ground state and with the proportional increase in energies to n^2, the excited states are produced.

$$\Psi_n(x) = \sqrt{\frac{2}{a}}\ \sin\left(\frac{n\pi x}{a}\right)$$

Some of the properties[3] of $\psi_n(x)$ are:

1. With respect tot he centre of the well, the solutions are alternately even and odd, starting with ψ_1 as even.

2. A node is a zero crossing except for end points. Starting with ψ_1 at none, with the increase of energy, each successive state has 1 more node.

3. The states are mutually orthogonal.

$$\int \psi_m(x)^* \psi_n(x) dx = S_{mn}$$

S_{mn} is the kronecker delta,

$$S_{mn} = \begin{cases} 0, & \text{if } m \neq n \\ 1, & \text{if } m = n \end{cases}$$

4. The stationary states can be used in a linear combination to explain a function, hence they are complete.

Therefore, the most general solution to the time dependent Schrodinger equation is a linear combination of stationary states

$$\Psi(x,t) = \sum_{n=1}^{\infty} C_n \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) e^{-i(n^2\pi^2\hbar/2ma^2)t}$$
$$C_n = \sqrt{\frac{2}{a}} \int_0^a \sin\left(\frac{n\pi}{a}x\right) \psi(x,0) dx$$

2.6 The Harmonic Oscillator

The quantum harmonic oscillator severs a very vital purpose in QM by being analogous to the classical harmonic oscillator and by having an exact, analytical solution.

Looking at the classical harmonic oscillator; we can begin with Hooke's Law

$$F = -kx = m\frac{d^2x}{dt^2}$$
$$x(t) = A\sin(\omega t) + B\cos(\omega t)$$
$$\omega = \sqrt{\frac{k}{m}} = \text{Angular frequency}$$
$$V(x) = \frac{1}{2}kx^2 = \text{Potential Energy}$$

For the quantum problem[2]

$$V(x) = \frac{1}{2}m\omega^2 x^2$$

TISE:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2 x^2 \psi = E\psi$$

The above equation can now be written as:

$$\frac{1}{2m}[p^2 + (m\omega x)^2]\psi = E\psi$$

The Hamiltonian of the particle is:

$$H = \frac{1}{2m}[p^2 + (m\omega x)^2]$$

The momentum operator is:

$$p = \left(\frac{\hbar}{p}\right)\frac{d}{dx}$$

Here, p and x are operators and they do not commute.
We have,

$$a\pm = \frac{1}{\sqrt{2\hbar m\omega}}(\mp ip + m\omega x)$$

as ladder operators.
The canonical commutation relations is established as

$$[x, p] = i\hbar$$

The Hamiltonian[3] can be expressed as;

$$H = \hbar\omega(a_- a_+ - 1/2)$$

or

$$H = \hbar\omega(a_+ a_- + 1/2)$$

and the TISE for the harmonic oscillator becomes (with respect to $a_\pm$)

$$\hbar\omega(a_\pm a_\mp \pm 1/2)\psi = E\psi$$

So;

$$H(a_+\psi) = (E + \hbar\omega)(a_+\psi)$$

or,

$$H(a_-\psi) = (E - \hbar\omega)(a_-\psi)$$

Therefore using the ladder operators (a_- lower operator and a_+ raising operator), we can start with only one solution.
We use $a_-\psi_0 = 0$ to determine $\psi_0(x)$ as:

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega}{2\hbar}x^2}; \qquad A^2 = \sqrt{\frac{m\omega}{\pi\hbar}}$$

where A is the normalization constant.
Hence, plugging $\psi_0(x)$ into the Schrodinger equation,

$$\hbar\omega(a_+ a_- + 1/2)\psi_0 = E_0\psi_0$$

we can get the energy of this state

$$E_0 = \frac{1}{2}\hbar\omega$$

Therefore we can get all the stationary states of the harmonic oscillator from,

$$\psi_n(x) = A_n(a_+)^n\psi_0(x); \qquad E_n = \left(n + \frac{1}{2}\right)\hbar\omega$$

as well as the allowed energies.

2.7 The Free Particle

For the infinite square well, $V(x)$ wsa 0 only within the well ($x \leq x \leq a$) but a free particle[2] not bound by an external force which equivalently means the $V(x)$ doesn't vary in the region of the free particle; $V(x) = 0$ everywhere.
TISE:

$$\begin{aligned} -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} &= E\psi \\ \frac{d^2\psi}{dx^2} &= -k^2\psi \qquad \text{where } k = \frac{\sqrt{2mE}}{\hbar} \\ \psi(x) &= Ae^{ikx} + Be^{-ikx} \qquad \text{[general equation]} \end{aligned}$$

The above equations are the same as the equations as inside the infinite square well where $V(x) = 0$, except the general equation is now represented in the exponential form. And there are no boundary condition on the free particle, there it can carry any positive energy.
The standard time dependence ($e^{-iEt/\hbar}$), hence gives,

$$\psi(x,t) = Ae^{ik(x-\frac{\hbar k}{2m}t)} + Be^{ik(x+\frac{\hbar k}{2m}t)}$$

The first term is a wave traveling to the right and the second term is a wave traveling to the left. Both terms are of the same energy.

$$\begin{aligned} \psi_k(x,t) &= Ae^{kn-\frac{\hbar k^2}{2m}t} \\ k &= \pm\frac{\sqrt{2mE}}{\hbar} \qquad\qquad k > 0 \implies \text{traveling to right} \\ & \qquad\qquad\qquad\qquad k < 0 \implies \text{traveling to left} \end{aligned}$$

So the stationary states of the free particle represent propagating waves with wavelength $\lambda = \frac{2\pi}{|k|}$
The moment now stands;

$$p = \hbar k = \sqrt{2mE} \qquad \text{(de Broglie formula)}$$

For pure kinetic energy $E(V = 0)$, the classical speed of a free particle is

$$\begin{aligned} E &= \frac{1}{2}mv^2 \\ V_{classical} &= \sqrt{\frac{2E}{m}} = 2V_{quantum} \end{aligned}$$

A paradox arises here, which states that the quantum mechanical wave function travels at half the speed of the particle. So for a free particle, the separable solutions do not represent any physically realizable states as the wave function is not normalizable, therefore a free particle has no definite energy.
But this poses no concern as we can still use the separable solutions, as the time dependent Schrodinger equation is still a linear combination of them just with an integral over the continuous variable k.

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\alpha}^{+\alpha} \phi(k) e^{i(kx - \frac{\hbar k^2}{2m} t)} dk$$
$$\Psi(x,0) = \frac{1}{\sqrt{2\pi}} \int_{-\alpha}^{+\alpha} \phi(k) e^{ikx} dk$$
$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\alpha}^{+\alpha} \Psi(x,0) e^{-ikx} dx$$

2.8 Bound States and Scattering States

Comparing potential $V(x)$ and total Energy E[2]:

1. $V(x) > E$ creates a bound state where the particle is confined with a potential well between the turning points.

2. $V(x) < E$ creates a scattering states where the particle comes in from infinity and goes back to infinity.

Bound states are only admitted by some potentials (like) infinite square well and harmonic oscillator and scattering states are admitted by others such as the free particle potential.

$$E < [V(-\alpha) \text{ and } V(+\alpha)] \implies \text{bound state}$$
$$E > [V(-\alpha) \text{ and } V(+\alpha)] \implies \text{scattering state}$$

Before we move on to work on some potentials we look at the Dirac Delta function

$$\delta(x) = \begin{Bmatrix} 0, & \text{if } x \neq 0 \\ \alpha, & \text{if } x = 0 \end{Bmatrix} \quad \text{with} \int_{-\alpha}^{+\alpha} \phi(x) dx = 1$$

$$f(x)\delta(x-a) = f(a)\delta(x-a)$$
$$\int_{-\alpha}^{+\alpha} \delta(x-a)dx = f(a) \int_{-\alpha}^{+\alpha} \delta(x-a)dx = f(a)$$

Now we take an artificial potential and work with it to determine the bound and scattering states.

$$V(x) = \alpha\delta(x)$$

Schrodinger Equation:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} - \alpha S(x)\psi = E\psi$$

for Bound States, consider $(E < 0)$

$$x < 0\ ; \qquad \psi(x) = Be^{kx}$$
$$x > 0\ ; \qquad \psi(x) = Ee^{-kx}$$

$$k = \frac{\sqrt{-2mE}}{\hbar} \qquad \text{and} \qquad \frac{d^2\psi}{dx^2} = -\frac{2mE}{\hbar^2}\psi = k^2\psi$$

Keeping in mind that, ψ is always continuous; $\frac{d\psi}{dx}$ is continuous except where the potential is infinite, we find that the delta-function well has exactly 1 bound state:

$$\psi(x) = \frac{\sqrt{m\alpha}}{\hbar}e^{-m\alpha|x|/\hbar^2}\ ; \quad E = \frac{m\alpha^2}{2h^2}$$

For scattering states we use $E > 0$

$$x < 0$$

Schrodinger equation:

$$\frac{d^2\psi}{dx^2} = -\frac{2mE}{\hbar^2}\psi = -k^2\psi$$
$$k = \frac{\sqrt{2mE}}{\hbar}$$

$$x < 0: \qquad \psi(x) = Ae^{ikx} + Be^{-ikx}$$
$$x > 0: \qquad \psi(x) = Fe^{ikx} + Ge^{-ikx}$$

Continuity of $\psi(x)$ at $x = 0$ yields

$$F + G = A + B$$

Continuity of $d\psi/dx$ except at boundaries where $V(x) = \alpha$

$$F - G = A(1 + 2i\beta) - B(1 - 2i\beta);\ \ \beta = \frac{m\alpha}{\hbar^2 k}$$

Solving the above 2 equations we get:

$$B = \frac{i\beta}{1 - i\beta}A \qquad F = \frac{1}{1 - i\beta}A$$

Reflection coefficient: $$R = \frac{|B|^2}{|A|^2} = \frac{\beta^2}{1 + \beta^2}$$

Transmission coefficient: $$T = \frac{|F|^2}{|A|^2} = \frac{1}{1 + \beta^2}$$

$$R + T = 1$$

$$R = \frac{1}{1 + (2\hbar^2 E/m\alpha^2)} \quad ; \quad T = \frac{1}{1 + (m\alpha^2/2\hbar^2 E)}$$

These scattering wave functions are not normalizable without involving a range of energies, R and T are to be taken as approximate reflection & transmission probabilities for particles in the region of E.
By imposing proper boundary conditions on the functions we determined the probability that a particle would either bounce off or pass through, the potential.
Therefore, by taking the linear combinations of states spread over all space, wave functions that are concentrated about a (moving) point were constructed.

2.9 The Uncertainty Principle

The Heisenberg's Uncertainty Principle[3 is perhaps the most misunderstood and misused ideas in physics. To put it in laymen's terms, the particles have a lot of positions and momentums, but if the particle is "mostly" in a small range of positions then it is "mostly" in a large range of momentum. But the act of measurement will collapse the state of the particle and we will observe it at a single classical position. Moreover we cannot get the other positions that the particle may have been at before measuring after our initial observation.
Although this principle may seem like it is about measurement, the particle follows the principle whether we measure the state of the particle or not. The uncertainty does not arise due to the disturbance to the particle while measuring. Uncertainty is an unfortunate term to describe this phenomenon, it seems to tell us that there is a position but we are uncertain about it.

A general proof of the uncertainty principle:

Stationary states are determinate states[2] of the Hamiltonian; which suggests that a measurement of the total energy, on a particle in the stationary state Ψ_n, will yield the corresponding "allowed" energy E_n.

$$\begin{aligned}
\sigma^2 &= \langle (\hat{Q} - \langle Q \rangle)^2 \rangle \\
&= \left\langle \psi \middle| (\hat{Q} - q)^2 \psi \right\rangle \\
&= \left\langle (\hat{Q} - q)\psi \middle| (\hat{Q} - q)\psi \right\rangle \\
&= 0
\end{aligned}$$

So, for observable A,

$$\sigma_A^2 = \left\langle (\hat{A} - \langle A \rangle)\psi \middle| (\hat{A} - \langle A \rangle)\psi \right\rangle = \langle f | f \rangle$$

where $f \equiv (\hat{A} - \langle A \rangle)$.

Similarly for any other observables like B,

$$\sigma_B^2 = \langle g|g\rangle$$

where $g \equiv (\hat{B} - \langle B\rangle)\psi$.
From Schwarz inequality,

$$\left|\int_a^b f(x)^* g(x)dx\right| \leq \sqrt{\int_a^b |f(x)|^2 dx \int_a^b |g(x)|^2 dx}$$

we have

$$\sigma_A^2\sigma_B^2 = \langle f|f\rangle \langle g|g\rangle \geq |\langle f|g\rangle|^2$$

$$|z|^2 = [\mathrm{Re}(z)]^2 + [\mathrm{Im}(z)]^2 \geq [\mathrm{Im}(z)]^2 = \left[\frac{1}{2i}(z - z^*)\right]^2$$

Let $z = \langle f|g\rangle$

$$\sigma_A^2\sigma_B^2 \geq \left(\frac{1}{2i}[\langle f|g\rangle - \langle g|f\rangle]\right)^2$$

But,

$$\begin{aligned}
\langle f|g\rangle &= \left\langle (\hat{A} - \langle A\rangle)\psi \middle| (\hat{B} - \langle B\rangle)\psi \right\rangle \\
&= \left\langle \psi \middle| (\hat{A} - \langle A\rangle)(\hat{B} - \langle B\rangle)\psi \right\rangle \\
&= \left\langle \psi \middle| (\hat{A}\hat{B} - \hat{A}\langle B\rangle - \hat{B}\langle A\rangle + \langle A\rangle\langle B\rangle)\psi \right\rangle \\
&= \left\langle \psi \middle| \hat{A}\hat{B}\psi \right\rangle - \langle B\rangle \left\langle \psi \middle| \hat{A}\psi \right\rangle - \langle A\rangle \left\langle \psi \middle| \hat{B}\psi \right\rangle + \langle A\rangle\langle B\rangle\langle\psi|\psi\rangle \\
&= \left\langle \hat{A}\hat{B} \right\rangle - \langle B\rangle\langle A\rangle - \langle A\rangle\langle B\rangle + \langle A\rangle\langle B\rangle \\
&= \left\langle \hat{A}\hat{B} \right\rangle - \langle A\rangle\langle B\rangle
\end{aligned}$$

Therefore,

$$\langle g|f\rangle = \left\langle \hat{B}\hat{A} \right\rangle - \langle A\rangle\langle B\rangle$$

so,

$$\langle f|g\rangle - \langle g|f\rangle = \left\langle \hat{A}\hat{B} \right\rangle - \left\langle \hat{B}\hat{A} \right\rangle = \left\langle [\hat{A}, \hat{B}] \right\rangle$$

where $[\hat{A}, \hat{B}]$ is the commutator as defined in 1.8.2.
Finally,

$$\sigma_A^2\sigma_B^2 \geq \left(\frac{1}{2i}\left\langle [\hat{A}, \hat{B}] \right\rangle\right)^2$$

Now suppose the first observable is position and the second observable is momentum.

$$\hat{A} = x$$
$$\hat{B} = \frac{\hbar}{i}\frac{d}{dx}$$

Commutator: $[\hat{x}, \hat{p}] = i\hbar$

so, $\sigma_x^2 \sigma_p^2 \geq \left(\frac{1}{2i} i\hbar\right) = \left(\frac{\hbar}{2}\right)^2$

$\sigma_x \sigma_p \geq \frac{\hbar}{2}$ (Heisenberg Uncertainty principle)

This is only one application of a much more general theorem.

2.10 The postulates of quantum mechanics

2.10.1 Postulate 1

There exists a corresponding Hilbert space to every physical system, and each possible state of the system is given by a vector in the Hilbert space.
Discussions:
This postulate implies the principle of superposition[4], which states that if $|\psi_1\rangle$ and $|\psi_2\rangle$ are two vectors, in the Hilbert space representing two possible states of the system, then a linear combination of them,

$$|\psi\rangle = c_1 |\psi\rangle + c_2 |\psi\rangle$$

is also a vector in the Hilbert space and therefore, represents a possible state of the system. Furthermore, we have assumed that if a vector corresponding to a state of the system is multiplied by any complex number(effect of superposing a vector with itself), except zero, the resulting vector will correspond to the same state. Thus, a state is specified by the 'direction', of a vector in the Hilbert space, and any assigned magnitude of the vector is irrelevant. In case of a classical system, when one superposes a state of a vibrating string with itself, the result is a different state with double the amplitude of the original state. Also, there is a classical state with zero amplitude of oscillation everywhere, but such a state of rest does not exist for a quantum system.

2.10.2 Postulate 2

To every physical quantity there corresponds a hermitian operator acting in the Hilbert space of the system[4]. For instance, the operators $\hat{x}$ (coordinate) and $\hat{p}$ (momentum), satisfy the commutator relation.

$$[\hat{x}, \hat{p}] = i\hbar \mathbb{1}$$

Discussions: Any classical physical quantity may be expressed as a function of coordinate and momentum, $Q = Q(x, p)$ thus replacing $x \to \hat{x}$ and $p \to \hat{p}$ yields the operator $\hat{Q} = Q(\hat{x}, \hat{p})$, hence establishing one to one correspondence between operators and observables. However, spin operators cannot be obtained through such substitution as they are purely quantum operators. The Hamiltonian(operator), for the classical hamiltonian function $H(p, x)$.

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + V(\hat{x}) \quad \text{[For a conservative system]}$$

The recipe,

$$\hat{Q} = Q(x \to \hat{x}, p \to \hat{p}) \quad \text{[ambiguous]}$$

So if $Q = xp$

then $\hat{Q} = \hat{x}\hat{p}$

As there are no general method to solve such ambiguities, we will use the symmetric sum as long as Q does not involve products of two or more powers of $\hat{x}$ with two or more powers of $\hat{p}$

$$\hat{Q} = \frac{1}{2}(\hat{x}\hat{p} + \hat{p}\hat{x})$$

2.10.3 Postulate 3

The only possible result of the measurement of a physical quantity A is one of the eigenvalues of the corresponding hermitian operator $\hat{A}$. The eigenvectors of a hermitian operator $\hat{A}$ representing an observable, form a complete orthonormal set of vectors and therefore form a basis of the Hilbert space[4].

Discussions:
Discrete eigenvalue spectrum,

orthonormality: $$\langle a_n, r | a_n{}' \rangle = S_{nn'} S_{rr'}$$

completeness: $$\sum_n \sum_{r=1}^{g_n} |a_n, r\rangle \langle a_n, r| = \hat{\mathbb{1}}$$

The eigenspace of eigenvalue a_n is of dimension g_n (eigenvectors) Eigenvectors belonging to different eigenvalue are automatically orthogonal, since $\hat{A}$ is a hermitian operator. Similarly, the orthonormality and completeness conditions of the eigenvectors of can be written for eigen spectrums that are purely continuous or partly discrete and partly continuous.

2.10.4 Postulate 4

When the physical quantity A is measured on a system in the normalized state ψ , the probability, $P_\psi(a_n)$, of obtaining the eigenvalue of the corresponding hermitian operator $\hat{A}$ is[4],

$$P_\psi(a_n) = |\langle a_n | \psi \rangle|^2$$

If the eigenvalue*, is g_n-fold degenerate, then

$$P_\psi(a_n) \geq \sum_{i=1} g_n |\langle a_n, i | \psi \rangle|^2$$

where, $\{|a_n, i\rangle \, ; i = 1, 2, ..., g_n\}$ are the orthonormalised eigenvectors of $\hat{A}$ all belonging to the same eigenvalue a_n. For continuously distributed eigenvalues,the probability $dP_\psi(a)$ of obtaining a result between a and $a + da$ is

$$dP_\psi(a) = |\langle a | \psi \rangle|^2$$

or, if there is any degeneracy,

$$dP_\psi(a_n) = \sum_{i=1}^{g_a} |\langle a, i|\psi\rangle|^2 da$$

Discussions: The expectation value of the observable A in the normalized state $|\psi\rangle$ is,

$$\left\langle \hat{A} \right\rangle = \sum_n a_n P_\psi(a_n)$$

Normalisation of the state vector $|\psi\rangle$ to unity ensures that the sum of the probabilities of obtaining the various eigenvalues of $\hat{A}$ is unity,

$$\begin{aligned}
\langle\psi|\psi\rangle &= \sum_n \langle\psi|a_n\rangle \langle a_n|\psi\rangle \\
&= \sum_n |\langle a_n|\psi\rangle|^2 \\
&= \sum_n P_\psi(a_n) \\
&= 1
\end{aligned}$$

Alternative form of the expectation value,

$$\begin{aligned}
\left\langle \hat{A} \right\rangle &= \sum_n a_n P_\psi(a_n) \\
&= \sum_n a_n |\langle a_n|\psi\rangle|^2 \\
&= \sum_n a_n \langle\psi|a_n\rangle \langle a_n|\psi\rangle \\
&= \langle\psi| \left(\sum_n a_n |a_n\rangle \langle a_n|\right) |\psi\rangle \\
&= \langle\psi| \hat{A} |\psi\rangle
\end{aligned}$$

Pointers: If the particle is in an eigenstate of $\hat{A}$, then $\left\langle \hat{A} \right\rangle = a$. We only need the state vector and the operator to calculate $\left\langle \hat{A} \right\rangle$. We average over an ensemble of a large number of particles each in the same state ψ

$$P_\psi(a_n) = \frac{N_n}{N} \;\; ; \;\; n = 1, 2, 3...$$

Therefore, this postulate asserts that,

$$P_\psi(a_n) = |\langle a_n|\psi\rangle|^2$$

The quantum mechanical ensemble average is then,

$$\left\langle \hat{A} \right\rangle = \sum_n a_n P_\psi(a_n)$$
$$= \sum_n a_n \frac{N_n}{N}$$

Ensemble average,

$$\left\langle \hat{A} \right\rangle = \langle \psi | \, \hat{A} \, | \psi \rangle$$

Chapter 3

Qubits and Quantum Circuits

3.1 A quantum bit

In classical information, the smallest unit of information is called a *bit* which can take one of the two values {0, 1}. In quantum information, the similar concept is called a quantum bit or *qubit* which describes a state in the 2-state quantum system. The state of a qubit at any given instant is represented by a vector in $\mathbb{C}^2$ with orthonormal basis vectors $\{|0\rangle, |1\rangle\}$[14]

A quantum bit is represented as a two-element column vector because it is not wholly 0 or wholly 1 at any given instant. A qubit can be found to be in one of the two states when measured. Whenever it is in uncollapsed (not yet measured) state, it exists in a superposition of the states as described in the first postulate(2.10.1):

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle$$

where a and b are complex numbers and are called the amplitude of the $|0\rangle$ and $|1\rangle$ components respectively. The amplitudes can be thought of as the probability that a given state will be observed when the superposition is collapsed.

A physical qubit can be thought of as the two directions of rotation of a circularly polarized photon. There can be other physical embodiment of the quantum system, for example, two discrete energy levels of an electron orbiting an atom or the two directions of spin of an electron[14]. All of them would work equally well, but for our purposes we will think of a qubit as an abstract particle.

There is a constraint on the vectors:

$$|a|^2 + |b|^2 = 1$$

This is done to properly normalize the qubit. It guarantees that when a qubit is measured it will either be in state $|0\rangle$ with probability $|a|^2$ or in state $|1\rangle$ with probability $|b|^2$. The measurement is done in the standard computational basis states $\{|0\rangle, |1\rangle\}$

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad\qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Since the states are orthonormal they satisfy the equation $|a|^2 + |b|^2 = 1$, and therefore a qubit can be described by a 2-dimensional Hilbert space.

3.2 Bloch sphere representation of a qubit

The states of a single qubit can be visualized in a unit three-dimensional sphere called the Bloch sphere. A qubit is any point on the the surface of the sphere. There are infinite points on the unit-sphere. By that logic, one might argue that theoretically one can store the entire Wikipedia in the infinite expansion of θ. However, when a qubit is observed we get only either 0 or 1. The measurement changes the superposition state of the qubit to a discrete state and it differs from the previously held state. Further measurements will yield the same value. As discussed in previous chapter, this is one of the postulates of quantum mechanics. Therefore, we can only get one bit of information from a qubit.

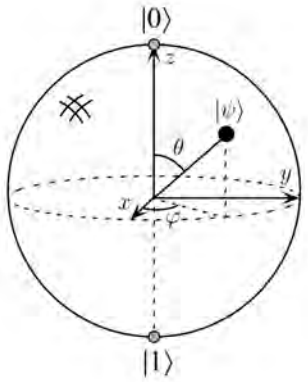

Figure 3.1: A bloch sphere[3]

We may rewrite the state of a single qubit $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$ in terms of azimuth and elevation angles as:

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

where θ, ϕ and γ are real numbers. One thing that can be confusing with this picture is that the vectors $|0\rangle$ and $|1\rangle$ are 180° apart even though they are orthogonal. The way the Bloch sphere is constructed results in orthogonal states to be the antipodal points on the sphere. This is why $|0\rangle$ is at the North Pole and $|1\rangle$ is at the South Pole.

3.3 Classical logic gates

3.3.1 Irreversible gates

Logic gates are the heart of computing. All computation can be decomposed into a sequence of logic gates act on only a few bits at a time. The most common logic gates are AND, OR and NOT. Figure 3.2 and 3.3 show the truth table for AND and OR gates.

It is possible to create any boolean circuit with a combination of AND, OR and NOT gates but the circuits become complex quite quickly. There is a special class of logic gates, called *universal* gates. These gates alone can be used to create all other gates necessary whereas you need a combination of AND/NOT/OR to create a complex boolean circuit. Examples of these gates are

a	b	$a \wedge b$
0	0	0
0	1	0
1	0	0
1	1	1

Figure 3.2: Truth table of AND[16]

a	b	$a \vee b$
0	0	0
0	1	1
1	0	1
1	1	1

Figure 3.3: Truth table of OR[16]

NOT and NAND. The advantages of using such universal gates are that the chip designers need only focus on miniaturizing a circuit based on a single type of gate and the circuits are much simpler. One way of implementing NOT using NAND is as follows:

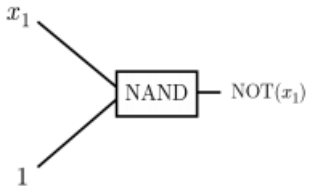

Figure 3.4: creating a NOT gate using NAND

a	a	$a \mid a$	$\neg a$
0	0	1	1
1	1	0	0

Figure 3.5: truth table for NOT[16]

The AND, OR and NAND gates are logically *irreversible*, which means that you cannot determine unique inputs for all outputs. The NAND gate takes as input two bits and produces one single bit as output. Some information is *erased* when an irreversible logic gate operates because the input cannot be recovered from the output. For example, if the output of a NAND gate is 1, it cannot be determined for which input - 00, 01 or 10 - we got the output 1. The information erased is dissipated as energy. This energy consumption can be a performance bottleneck.

3.3.2 Reversible gates

NOT gate is a perfect example of a *reversible* gate. If the output is 1, we can determine the input for this particular output, which is 0. In reversible computation, no information is erased because the input can always be recovered from the output. Some of these gates are NOT, SWAP and CNOT. A SWAP does exactly what its name suggests. It takes 2 bits as inputs, swaps them and produces 2 bits as output. A CNOT gate is also a 2-bit input/output gate. It flips the second bit *iff* the first bit is 1.
Just like classical irreversible computing, there are universal gates for classical reversible computing too. Two of the well-known examples are the FREDKIN (controlled-SWAP) gate and the TOFFOLI (controlled-CNOT) gate. These are the smallest universal reversible gates that take 3 bits as input and produces 3 bits as output. The FREDKIN gate swaps the second and third input bits *iff* the first input bit is set to 1 and the TOFFOLI gate flips the third bit *iff* the first two bits are set to 1.
Moore's Law has been true for decades but now it's getting difficult to increase the number of transistors in the chip without dissipating too much heat. Moore's Law is coming to an end and chip designers are trying out different ways to increase the efficiency and reduce the power

a	b	a'	b'
0	0	0	0
0	1	1	0
1	0	0	1
1	1	1	1

Figure 3.6: Truth table of SWAP[16]

a	b	a'	b'
0	0	0	0
0	1	0	1
1	0	1	1
1	1	1	0

Figure 3.7: Truth table of CNOT[16]

dissipation. Landauer showed that energy is dissipated whenever information is erased[15]. So whenever a bit is erased in irreversible computation a minimum amount of $k_B T \ln 2$ Joules is dissipated as heat. Reversible computing would greatly decrease the power dissipation and theoretically would require no energy to compute but creating gates that are both logically and physically reversible is very difficult. Reversible computing plays an important role in quantum computing and we will see later why.

3.4 Quantum logic gates

3.4.1 1-qubit gates

Pauli-X or NOT gate

The Pauli X matrix is similar to the classical NOT gate. It maps $|0\rangle$ to $|1\rangle$ and $|1\rangle$ to $|0\rangle$.

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

However, it is not a true quantum NOT gate. It does not negate every state vector of the Bloch sphere to its antinodal states[16].

Pauli-Y gate

The Pauli-Y gate maps $|0\rangle$ to $i\,|1\rangle$ and $|1\rangle$ to $-i\,|0\rangle$. It's a rotation around the Y-axis of the Bloch sphere by π radians.

$$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Pauli-Z gate

It does a rotation around the Z-axis of the Bloch sphere by π radians. It maps $|1\rangle$ to $-|1\rangle$ but leaves $|0\rangle$ unchanged. It is sometimes called the phase-flip gate.

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

$\sqrt{\text{NOT}}$ gate

The first truly non-classical gate we see is the $\sqrt{\text{NOT}}$ gate:

$$\sqrt{\text{NOT}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{\frac{1}{2}} = \begin{pmatrix} \frac{1}{2}+\frac{i}{2} & \frac{1}{2}-\frac{i}{2} \\ \frac{1}{2}-\frac{i}{2} & \frac{1}{2}+\frac{i}{2} \end{pmatrix}$$

The application of a $\sqrt{\text{NOT}}$ gate corresponds neither to the classical bit 0 nor 1. It results in a superposition of states.

Hadamard gate

One of the most important and useful quantum gates is the Hadamard gate.

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

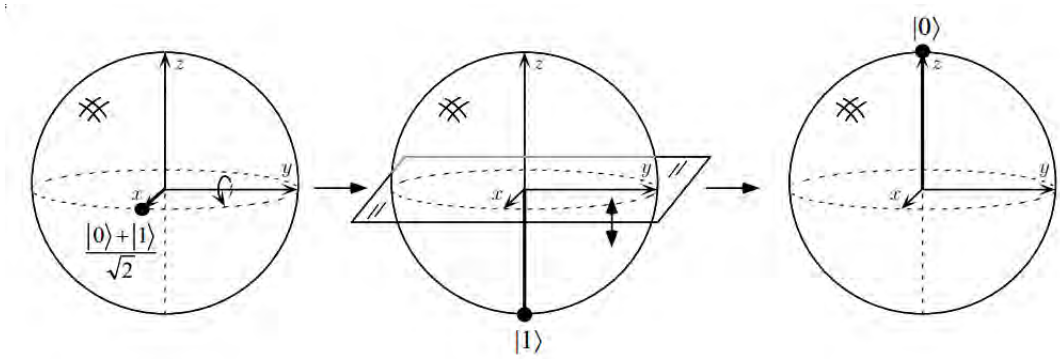

Figure 3.8: Hadamard gate acting on the input state $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ [3]

As seen in the figure 3.8, the Hadamard operation is just a rotation about the $\hat{y}$ axis by 90°, followed by a rotation about the $\hat{x}$ axis by 180°. It is sometimes referred to as the "square-root of NOT gate" because it transforms $|0\rangle$ to $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $|1\rangle$ to $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$. Note, however, H^2 is not a NOT gate.

$$
\begin{aligned}
H^2 &= (\frac{|0\rangle + |1\rangle}{\sqrt{2}}).(\frac{|0\rangle - |1\rangle}{\sqrt{2}}) \\
&= \frac{|0\rangle \,.\, |0\rangle - |0\rangle \,.\, |1\rangle + |1\rangle \,.\, |0\rangle - |1\rangle \,.\, |1\rangle}{2} \\
&= \frac{1+1}{2} \\
&= 1
\end{aligned}
\tag{3.1}
$$

Applying H twice to a state does nothing to it.

3.4.2 Other gates

CNOT

The quantum CNOT gate logically works exactly the same way the classical CNOT gate we saw earlier. CNOT acts on 2 qubits. It performs the NOT operation on the second qubit only when the first qubit is $|1\rangle$, and otherwise leaves it unchanged.

$$
CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}
$$

Chapter 4

Quantum Algorithms

4.1 A brief intro on complexities

Computational Complexity is needed to measure how much work is required to solve different problems. It helps to classify tool for tackling discrete deterministic problems. We use it to determine, if a problem is hard or easy to solve for a computer [12]. The Computational Complexity Theory helps us to sort the problems into classes on the basis of the problem's property. In order to understand the computational power of a quantum computer, first we need to understand the complexity of classical computation.
A complexity class is a set of languages, which decides within a time or space complexity bound $t(n)$ or $s(n)$ with respect to some fixed method of solution or computation to a problem. $t(n)$ is called the time constructible, i.e., if there exist a Turing machine which takes input of a string of n 1s halts after exactly $t(n)$ steps. Similarly $s(n)$ is called space constructible, i.e., if there exists a Turing machine which takes input of a string of n 1s halts after visiting exactly $s(n)$ tape cells[11]. In classical computation we are interested in doing deterministic computation. As $t(n)$ and $s(n)$ are said to be time and space constructible functions, we can define the classes $TIME(t(n))$ and $S(s(n))$ as[11],

$$\begin{aligned}\mathbf{TIME}(t(n)) &= \{X \subseteq \{0,1\}^* \ : \ \exists T \in \mathfrak{T} \forall n(time_T(n) \leq t(n)) \text{ and } T \text{ decides } X\} \\ \mathbf{SPACE}(s(n)) &= \{X \subseteq \{0,1\}^* \ : \ \exists T \in \mathfrak{T} \forall n(space_T(n) \leq s(n)) \text{ and } T \text{ decides } X\}\end{aligned}$$

Since the polynomials in variable n are of order $\mathcal{O}(n^k)$ for some k, we can define the classes polynomial time and polynomial space as

$$\mathbf{P} = \bigcup_{k \in \mathbb{N}} \mathbf{TIME}(n^k)$$

and,

$$\mathbf{PSPACE} = \bigcup_{k \in \mathbb{N}} \mathbf{SPACE}(n^k)$$

Now that we have defined the Polynomial classes, we can also define the Exponential classes,

$$\mathbf{EXP} = \bigcup_{k \in \mathbb{N}} \mathbf{TIME}(2^{n^k}) \quad \text{(exponential time)}$$

and,

$$\mathbf{L} = \mathbf{SPACE}(\log(n)) \quad \text{(logarithmic space)}$$

Now that we have defined the deterministic polynomial complexity classes, we can also define its counterpart the analogous non-deterministic complexity classes as,

$$\mathbf{NTIME}(t(n)) = \{X \subseteq \{0,1\}^* \ : \ \exists N \in \mathfrak{N} \forall n(time_N(n) \leq t(n)) \text{ and } N \text{ decides } X\}$$
$$\mathbf{NSPACE}(s(n)) = \{X \subseteq \{0,1\}^* \ : \ \exists N \in \mathfrak{N} \forall n(space_N(n) \leq s(n)) \text{ and } N \text{ decides } X\}$$

The classes NP (non-deterministic polynomial time) and NSPACE (non-deterministic polynomial space), NEXP (non-deterministic exponential time) and NL (non-deterministic logarithmic space) are defined analogously to **P**, **NP**, **EXP**, and **L** as, $\mathbf{NP} = \bigcup_{k \in \mathbb{N}} \mathbf{NTIME}(n^k)$
We can say the relationship between the complexity classes are,

$$\mathbf{L} \subseteq \mathbf{NL} \subseteq \mathbf{P} \subseteq \mathbf{NP} \subseteq \mathbf{PSPACE} \subseteq \mathbf{EXP} \subseteq \mathbf{NEXP} \subseteq \mathbf{EXPSPACE}$$

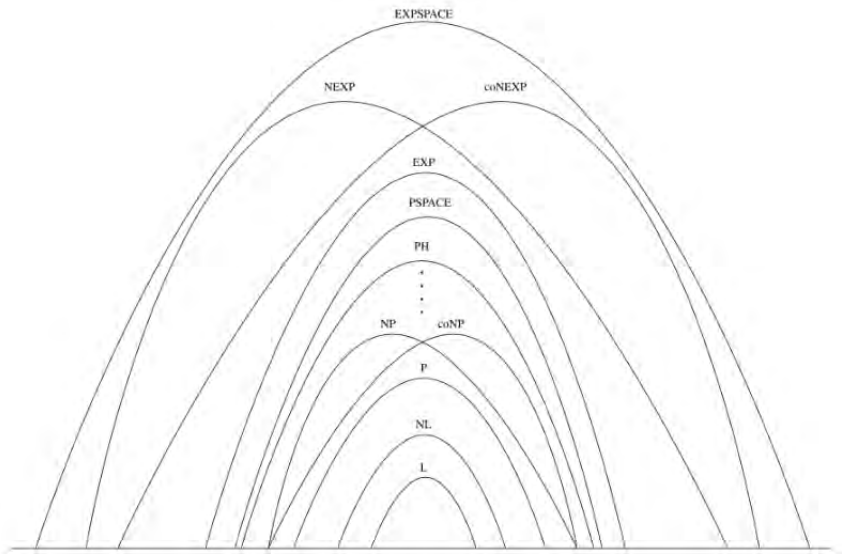

Figure 4.1: Inclusion Relationship Among Major Complexity Classes[11]

Decision problems can be solved in polynomial time with the deterministic Turing Machines, which are in P-Class. On the other hand, the problems those requires a nondeterministic Turing machine in order to be solved efficiently are in NP-Class. All the hardest problems in NP-Class are considered NP-Complete problems. Every problems in NP-Class can be reduced to NP-Complete Class. Now, if we can map a single NP-Complete problem to P-Class, both classes will collapse into each other i.e. we will be able to solve all NP-Complete Problems in Polynomial time, which will give us conclusive proof of P=NP.
Our current hierarchy model of the complexity classes range from P to EXP which is the most robust benchmarks of computational difficulty now available. Beyond the hierarchy there are a wide range of additional complexity classes are available which demarcate additional structure either inside P or between P and NP.

The BPP or Bounded-error Probabilistic Polynomial Time class can be defined using the Probabilistic Turing Machine model. BPP addresses those decision problems which can be solved in polynomial time with probabilistic turing machine, with a chance that the solution can be wrong. Probabilistic turing machine addresses those problems with random decision with directly access to some truly random inputs. In BPP, the error of the solution is bounded in a parameter, that the probability of the answer being correct must be at least two-thirds.
We can define BPP as, with problem X such that, there exist a probabilistic turing machine, $C \in \mathfrak{C}$ and a boundary condition $\frac{1}{2} < p \leq 1$ with the following properties:

a) C runs in polynomial time for all inputs;

b) for all inputs xX, at least fraction p of the possible computations of C on x accept;

c) for all inputs xX, at least fraction p of the possible computations of C on x reject.

The definition describes if a problem is decidable or not, with a probabilistic algorithm as one for which there exists a decision procedure which can make undetermined choices during its computation but still solves the problem correctly in a majority of cases (i.e. with probability p bounded away from $\frac{1}{2}$).
Now, we introduce the BQP, the analogous class of BPP. BQP or Bounded Error Quantum Polynomial Time is defined using a Quantum Computation Model instead of Classical Model. Such a model can be defined as device that use the Quantum-Mechanical Phenomena (like: entanglement or interference) to perform computation on data represented by sequences qubits i.e. quantum superpositions of vectors of 0s and 1s.
BQP can be described as, it include those problems which can be solved in polynomial time P(x), where x is the number of instructions in time t. A Bounded error quantum algorithm x has properties of entanglement and superposition.
Here we are answering all the decisions in yes or no with a boundary condition of error $\frac{1}{3} \leq x \leq \frac{1}{4}$, where x is a solvable in polynomial time. Further we can say $BPQ^{BPQ} = BPQ$, that means a polynomial time algorithm calls a polynomial subroutine until it finishes the computation.

$$\mathrm{BQP}^{n_1^{n_2^{\cdot^{\cdot^{n_i}}}}} = \mathrm{BQP},\ 1 \leq i \geq p, n_i = \mathrm{BQP}$$

The expansion depends on the number of qubits. A deep subroutine needs exponential qubits $[n_1, [n_2, [n_3, [...], n_i]]]$. A nested subroutine needs asymptotic $\mathcal{O}(n^2)$ qubits, which is the violation to quantum polynomial phenomena.

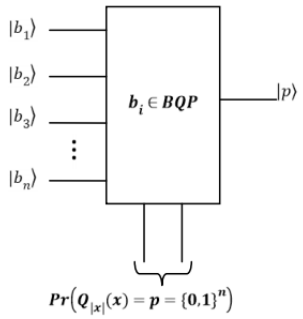

Figure 4.2: Quantum Turing machine solve all classical unsolvable problems into BQP time[9]

A quantum turing machine can solve all BQP problem with $|b_n\rangle$ qubits and computes results as

$$Pr(Q_{|x|}(x) = p = \{0,1\}^n)$$

The relation of BQP with different classes can be shown by

$$\mathbf{P} \subseteq \mathbf{BPP} \subseteq \mathbf{BQP} \subseteq \mathbf{PSPACE} \subseteq \mathbf{EXP}$$

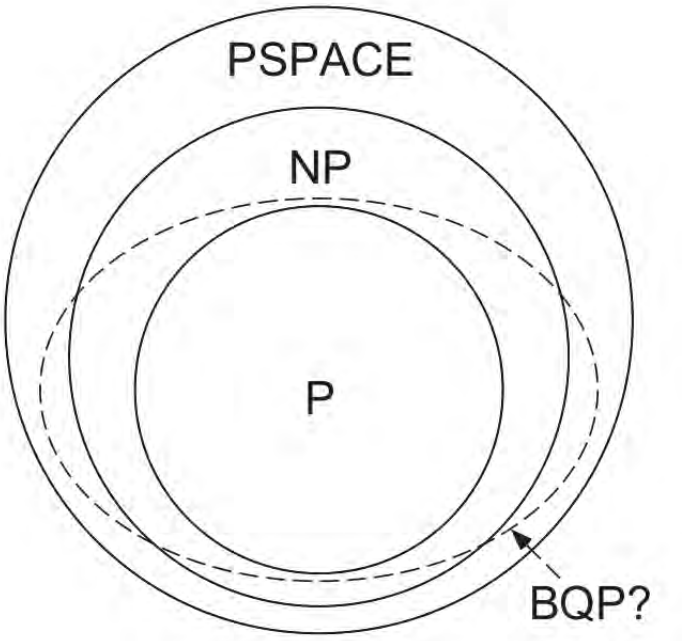

Figure 4.3: Relationship Between Classical and Quantum Complexity Classes

Algorithms have been developed which use such device and run faster than the best known classical algorithms for the same problem. We will now see how some quantum algorithms run order of magnitude times faster than their classical counterparts.

4.2 Simon's Algorithm

4.2.1 Simon's problem

We are given a function

$$f : \{0,1\} \to \{0,1\}^n$$

that maps n bit strings to n bit strings and that there exists a string $s \in \{0,1\}^n$ such that

$$f(x) = f(y) \longleftrightarrow x \oplus y \in \{0^n, s\}$$

for all $x, y \in \{0,1\}^n$ and $x \neq y$. The problem is to find the string s. In other words, there exists such a string s such that the function f results in the same value for two different inputs and the bitwise mod 2 addition of that two inputs give us the string s.

4.2.2 The classical approach

The function outputs random strings. There is no additional constraints on f that would help us find s. So, to find the string s we have to randomly guess the input strings x and y and determine if $f(x) = f(y)$ and then calculate $x \oplus y$. We will have to query $\Omega(2^{n/2})$ inputs before finding a pair that satisfy the function f[17]. The problem is hard classically, since the probability of finding the two input strings x and y after $\Omega(2^{n/2})$ queries is still less than $2^{-N/2}$.

4.2.3 The quantum approach

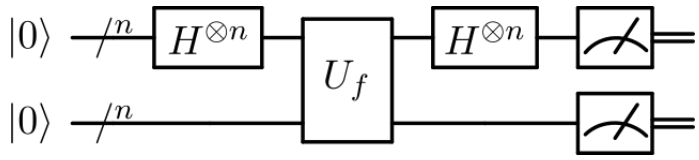

Figure 4.4: Circuit diagram for one iteration of Simon's algorithm[13]

Start with two n-bit registers to 0:

$$|0\rangle^{\otimes n} |0\rangle^{\otimes n}$$

Then apply Hadamard transform to the first register to get an equal superposition of states:

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

Querying the oracle yields the state vector:

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

Now measure the second register. The second register now holds the value of $f(x)$. Further calculations are now based on the first register. After the measurement of the first register, the state of the second register is reduced to:

$$\frac{1}{\sqrt{2}}(|x\rangle + |x \oplus s\rangle)$$

To isolate the information about s, we apply the Hadamard transform to each of the n remaining bits which gives us

$$\begin{aligned} & H^{\otimes n} \left[\frac{1}{\sqrt{2}}(|x\rangle + |x \oplus s\rangle) \right] \\ &= \frac{1}{\sqrt{2^{n+1}}} \sum_{y \in \{0,1\}^n} \left((-1)^{x.y} + (-1)^{(x \oplus s).y} \right) |y\rangle \\ &= \frac{1}{\sqrt{2^{n-1}}} \sum_{y \in \{0,1\}^n} (-1)^{x.y} |y\rangle \end{aligned}$$

After measuring the first register, we will find a string $y \in \{0,1\}^n$ such that $s.y = 0$. If we query the oracle $n-1$ times, then we end up with $n-1$ linear equations $\{y_i.s = 0\}$. We can then solve the set of linear equations to find s. Therefore, to sum up the we only have to query $\mathcal{O}(n)$ times to find s[17].

4.2.4 A worked example

Let $n = 3$, $s = 110$ and $f(x)$ defined by the following table[10]:

x	$f(x)$
000	101
001	010
010	000
011	110
100	000
101	110
110	101
111	010

Initialize two 3-bit registers to $|0\rangle^{\otimes 3}$:

$$|000\rangle\,|000\rangle$$

Now apply the Hadamard transform to obtain an equal superposition of the states:

$$\frac{1}{2\sqrt{2}} \sum_{x \in \{0,1\}^3} |x\rangle\,|f(x)\rangle$$

Measuring the second register collapses the first register to:

$$\frac{1}{\sqrt{2}}(|x\rangle + |x \oplus 110\rangle)$$

Then we apply a second Hadamard transform to the first register:

$$\begin{aligned} &H^{\otimes 3}\left[\frac{1}{\sqrt{2}}(|x\rangle + |x \oplus 110\rangle)\right] \\ &= \frac{1}{4} \sum_{y \in \{0,1\}^n} \left((-1)^{x.y} + (-1)^{(x \oplus 110).y}\right)|y\rangle \end{aligned}$$

Now we make $n-1$, in this case 2, observations such that $s.y = 0$ giving us two linearly independent equations for observations $y_1 = 001$ and $y_2 = 000$:

$$0(s_0) + 0(s_1) + 1(s_2) = 0$$

$$0(s_0) + 0(s_1) + 0(s_2) = 0$$

Solving the equations would result in $s = 110$, which is rightly associated with function f defined above.

4.3 Grover's algorithm

In computer science, searching algorithms are among the most important tools as they make the most mundane of tasks such as looking through phone books possible as well as help the grilling task of breaking cryptographic codes. This kind of algorithm is prevalent in this discipline. Therefore, any way of speeding up the task has a huge impact.

A standard search algorithm takes linear time which is a period of time that is roughly proportional to the number of elements in the search. This time complexity is applicable for the worst case scenario, where the algorithm has to search through all the elements to find the target. Some other classical search algorithms and their time complexities:

Linear Search $\mathcal{O}(N)$

Breadth First Search $\mathcal{O}(V + E)$

Depth First Search $\mathcal{O}(V + E)$

Binary Search $\mathcal{O}(\log N)$

A* Search (Heuristic) $\mathcal{O}(N)$

In 1996, Grover[6] used the key idea of superposition from Quantum mechanics, to formulate an algorithm which could only be implemented using a quantum computer. This was an ambitious idea , but by 1998, physicists demonstrated the first primitive quantum computer, and showed the execution of Grover's algorithm in that same year. But it was a very limited form of a quantum computer which worked only on a few qubits with no hope of scaling up to larger computations, as quantum decoherence presented a challenge for the practical realization of quantum computers with a large number of qubits.
But now, some 20 years later, the dawn of the scalable quantum computations has finally arrived, as researchers at the University of Maryland[5] have already executed Grover's algorithm on a scalable quantum computer for the first time.
Grover's algorithm uses the quantum superposition of states, to demonstrate how the qualities of quantum systems can be used to improve upon the lower runtime bounds of classical algorithms. Grover also exploits the qualities of quantum amplitude, which has no classical analogous, by applying amplitude amplification. Then the selective shifting of the phase of the target state of a quantum system, at each iteration is the crucial point of this algorithm

4.3.1 The mechanics of the Grover's Algorithm

Step 1: Initialization

We begin with a search space, $2^n = N$, where n is the number of qubits and everything is initialized to $|0\rangle$,

$$|0\rangle^{\otimes n} = |0\rangle$$

Step 2: Hadamard Transform

The first step is to apply a special type of a quantum gate called the Hadamard gate $H^{\otimes n}$, which requires $\Theta \lg N$ operations (refer to chap 3), to all the qubits. This will put the system into an equal superposition of states[3]:

$$|\psi\rangle = H^{\otimes n} |0\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle$$

Step 3: Grover's Iteration: Quantum Subroutine

This step provides the bulk of the experiment. The amplitude amplification is performed here with $\frac{\pi}{4}\sqrt{2^n}$ iterations. This repetition number has been set by Grover, to achieve optimal probability of getting the desired output, therefore the overall rotation of the phase is $\pi/4$.

4.3.2 Quantum Oracle

The quantum oracle is a quantum black box, which has a special property of observing and modifying the system without causing it to collapse into a classical state. If the oracle encounters the correct state then it will rotate the phase by π radians, otherwise it will do nothing. Hence at the beginning of Grover's iteration the oracle is called. In this implementation the oracle need not use any extra scratch qubit.

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

$$f(x) = \begin{cases} 1, & x = \text{ correct state} \\ 0, & \text{otherwise} \end{cases}$$

Calling the oracle is an elementary operation.

4.3.3 Diffusion Transform

After the states have undergone phase inversion, the next step is the diffusion transform, which performs the inversion about the mean. This will transform the amplitude of each of the states by putting it as far above the mean as it was as far below the mean before being transformed, and vice-versa. The diffusion transform process comprises of an application of the hadamard transform, then an application of the conditional phase shift which performs a (-1) phase shift to all the states except $|0\rangle$, and finally another hadamard transform[6].
Representation of the conditional phase shift using the unitary operator $2 |0\rangle \langle 0| - I$:

$$[2 |0\rangle \langle 0| - I] |0\rangle = 2 |0\rangle \langle 0|0\rangle - I = |0\rangle$$
$$[2 |0\rangle \langle 0| - I] |x\rangle = 2 |0\rangle \langle 0|x\rangle - I = - |x\rangle$$

Diffusion Transform becomes,

$$H^{\otimes n}[2 |0\rangle \langle 0| - I]H^{\otimes n} = 2H^{\otimes n} |0\rangle \langle 0| H^{\otimes n} - I = 2 |\psi\rangle \langle \psi| - I$$

Grover Iteration now stands to be,

$$[2 |0\rangle \langle 0| - I]\mathcal{O}$$

4.3.4 Computational Complexity Analysis

The precise runtime of the oracle is dependent of the specific problem and how we implemented it, therefore we take it as having one elementary operation. The cost of the two hadamard gates are $\mathcal{O}(n)$ and the cost of the conditional phase shift is $\mathcal{O}(n)$. Therefore the combined cost of the Grover iteration is $\Theta(2n)$. Hence, the total runtime of the complete Grover's algorithm is $\mathcal{O}(2^{n/2})$, considering that $\mathcal{O}(\sqrt{n}) = \mathcal{O}(\sqrt{2^n}) = \mathcal{O}(2^{n/2})$ iterations with a runtime of $\mathcal{O}(n)$ have been performed[6].

4.3.5 Visualizing the algorithm

If we have a problem which requires us to find one unique state for a given set of states,

$$f : \{0, ..., N-1\} \rightarrow \{0, 1\}$$

such that $f(x) = 1$, for exactly one x.
1) We perform phase inversion over all the elements of the unordered database, resulting in the inversion of only the target element, with the same amplitude, as shown in Fig. 4.5.

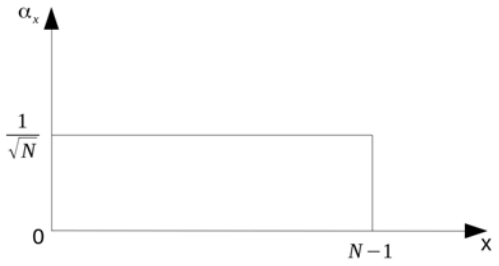

Figure 4.5

2) Then we perform the inversion about the mean for all the elements. The mean is slightly lower than the original amplitude of $1/\sqrt{N}$, as shown in Fig. 4.6.

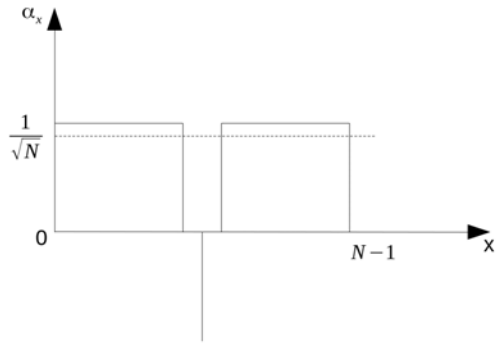

Figure 4.6

3) Finally, we get a result where the target element has an amplitude which is three times that of the original. Thus, amplitude amplification takes place, as shown in Fig. 4.7. If the entire set of activities(phase inversion and inversion about mean), is repeated enough times(roughly for $\mathcal{O}(\sqrt{N})$ times), then we end up with an amplitude of $1/\sqrt{2}$ for the target element, and the rest of the elements have an estimated amplitude of $1/\sqrt{2N}$.

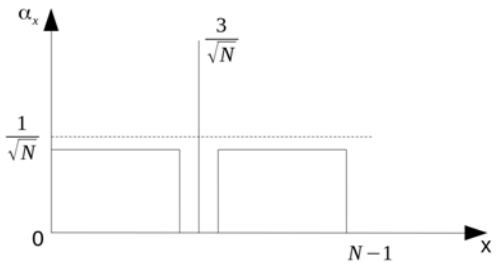

Figure 4.7

4.3.6 Four-Phase Improvement of Grover's Algorithm

When Grover's algorithm searches for a target in an unordered database, it achieves quadratic acceleration. But with an increase in the quantity of targets, the probability of obtaining correct results decreases. To combat this limitation, physicists have been proposing several different amendments along the years. In 2017, researchers at the University of Science and Technology of China, Hefei[7], proposed a four-phase improvement of Glover's algorithm that states that when the proportion of target is 1/3, the probability of success is greater than 97.82%, with only one iteration, using two different phases of 1.3789 and 1.8025. The probability of success further increases up to 99.63% when an optimal phase of 6.0215 is used, for the computational complexity of $\mathcal{O}(\sqrt{M/N})$.

References

[1] Feynman, Richard P.; Robert B. Leighton; Matthew Sands (1965). The Feynman Lectures on Physics, Vol. 3. US: Addison-Wesley.

[2] Griffiths, D. J. (2007). Introduction to quantum mechanics. Cambridge: Cambridge University Press.

[3] Nielsen, M. A., & Chuang, I. L. (2010). *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 10th ed. Cambridge: Cambridge University Press.

[4] Kabir,K.A (2017). *Lecture notes on Quantum Mechanics*. Course handout. Quantum Mechanics. Dept. of Theoretical Physics, Dhaka University.

[5] C. Figgatt, D. Maslov, K. A. Landsman, N. M. Linke, S. Debnath, & C. Monroe(2017). *Complete 3-Qubit Grover Search on a Programmable Quantum Computer*. arXiv:1703.10535v1 [quant-ph]

[6] L. K. Grover, A fast quantum mechanical algorithm for database search, in Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing. New York, New York, USA: ACM Press.

[7] Ma, B., Bao, W., Li, T., Li, F., Zhang, S., & Fu, X. (2017). A Four-Phase Improvement of Grovers Algorithm. *Chinese Physics Letters,34*(7), 070305th ser. doi:10.1088/0256-307X/34/7/070305

[8] Strang, G. (2009). *Introduction to Linear Algebra, 4th Edition*. Wellesley, MA, USA: Wellesley-Cambridge Pr.

[9] Sihare, S. R., & Nath, D. V. (2017). Analysis of Quantum Algorithms with Classical Systems Counterpart. *Modern Education and Computer Science Press,2*, 20-26. doi:10.5815/ijieeb.2017.02.03

[10] Strubell, E. (2011). An introduction to quantum algorithms. *COS498 Chawathe Spring, 13*, 19.

[11] Dean, W. (2015, July 27). Computational Complexity Theory. Retrieved August 10, 2017, from https://plato.stanford.edu/entries/computational-complexity/

[12] Tovey, C. A. (n.d.). Tutorial on Computational Complexity. *Interfaces, INFORMS,32*(3), 30-61. Retrieved August 15, 2017, from https://pdfs.semanticscholar.org/3d33/be88d84d1da3e06e104a60a279d3f71568af.pdf.

[13] – –. (2017) Quantum subroutine in Simons algorithm. Wikimedia, Inc. (Accessed 8 July, 2017). [Online]. Available: https://en.wikipedia.org/wiki/File:Simons_algorithm.svg

[14] Preskill, J. (1998). Lecture notes for physics 229: Quantum information and computation. *California Institute of Technology, 16*.

[15] Landauer, R. (1961). *Irreversibility and heat generation in the computing process. IBM journal of research and development, 5*(3), 183-191.

[16] Williams, C. P. (2010). *Explorations in quantum computing*. Springer Science & Business Media.

[17] Simon, D. R. (1997). On the power of quantum computation. *SIAM journal on computing, 26*(5), 1474-1483.