# Introduction to Quantum Computing

Dan Sehayek

January 2, 2018

# Contents

# 1 Introduction

Welcome to my introduction to quantum computing! Over the past year I've been taking the time to learn about the foundations and applications of quantum computing. This introduction is essentially a summary of my notes on quantum computing from a few different amazing resources that I was very fortunate to come across.

- John Watrous's Lecture Notes on Quantum Computing
- David McMahon's Book Notes on Quantum Computing Foundations
- Umesh Vazirani's edX Course on Quantum Mechanics and Quantum Computing

All of the people above did an incredible job of explaining some of the challenging concepts of quantum computing and I would like to thank these people for giving me the confidence and motivation to study the subject further. I would also encourage anybody reading this to go check them out!

As far as these notes go, they were originally just meant to summarize and review the content that I've learned from the resources above. But if you have also acquired a sudden interest in the subject and would like to learn more about it, please feel free to read this!

# 2 Quantum Computing Foundations

## 2.1 Qubits

Let us consider a system with $k$ distinguishable states. For example, an electron in a hydrogen atom may only occupy one of a discrete set of energy levels. The very first energy level is the ground state and the energy levels that follow are the first excited state followed by the second excited state and so on. If we establish that there are $k$ different states or energy levels then the electron will occupy the ground state or one of $(k-1)$ excited states. We may use the state of the system to store a number between 0 and $(k-1)$.

The **superposition principle** establishes that if a quantum system can be in one of two states then it can also be placed in a linear superposition of these two states. Let us represent the ground state of our $k$ system by $|0\rangle$ and the excited states by $|1\rangle \dots |k-1\rangle$. The electron has $k$ possible classical states. The quantum state of the electron is given by $\alpha_0 |0\rangle + \cdots + \alpha_{k-1} |k-1\rangle$ where $\alpha_0 \dots \alpha_{k-1}$ are complex numbers that are **normalized** such that $\sum_j \left|\alpha_j\right|^2 = 1$. Note that $\alpha_i$ is called the **amplitude** of the state $|i\rangle$. $|\alpha_i|^2$ gives the probability that the system will collapse to the state $|i\rangle$. If $k = 3$ then the state of the electron is written as:

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle + \alpha_2 |2\rangle$$

We can define the state vectors to be equal to the standard unit basis vectors as follows.

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad |2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

This allows us to rewrite our state $|\psi\rangle$ as a single vector.

$$|\psi\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix}$$

The normalization on the complex amplitudes means that the state of the system is a unit vector in a $k$ dimensional complex vector space. This complex vector space is called a **Hilbert space**.

A **qubit** is a 2 state system. The state of a qubit is typically written as follows.

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$$

If we are referring to the state of an electron in a hydrogen atom then $|0\rangle$ would represent the electron being in the ground state and $|1\rangle$ would represent the electron being in the first excited state. Once again we can set these vectors to be equal to the standard unit basis vectors.

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

Another set of basis vectors that is commonly used are the $|+\rangle$ and $|-\rangle$ states.

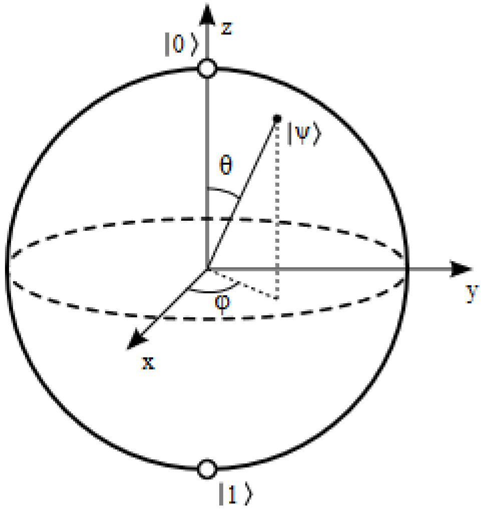
$$|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \qquad |-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$$

Notice that is different from our classical bit. The classical bit can exist only in state 0 or 1 while the qubit can exist in a superposition of the states 0 and 1.

Is there a way of giving our qubit a nice geometric representation? Let us rewrite our state $|\psi\rangle$ using the spherical coordinates $\theta$ and $\phi$.

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

We can provide a geometrical representation of the state $|\psi\rangle$ using the **Bloch sphere**.



## 2.2 Quantum Gates

Qubit transformations are represented using matrices. The general idea is that if $A$ is the transformation matrix and $|\psi\rangle$ is the initial state of the qubit then $|\psi'\rangle = A|\psi\rangle$ is the new state of the qubit after applying the transformation. Many of these matrices or **quantum gates** are analogous to the **logic gates** that you would find when dealing with classical bits.

We also have two qubit gates. Two qubit states are generally denoted $|ab\rangle$ where both $a$ and $b$ can be equal to 0 or 1. It should be noted that $|ab\rangle \equiv |a\rangle \otimes |b\rangle$ where $\otimes$ denotes the **tensor product**.

| | Matrix | Significance |
|---|---|---|
| X Pauli Gate | $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ | This gate essentially acts as a NOT gate since it transforms $\|0\rangle$ to $\|1\rangle$ and $\|1\rangle$ to $\|0\rangle$. For general cases this corresponds to a reflection about the $xy$ plane followed by a reflection about the $xz$ plane. |
| Rotation Gate | $R(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$ | This corresponds to a rotation relative to the $z$ axis by the angle $\gamma$ |
| Phase Flip Gate | $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ | Transforms $\alpha \|0\rangle + \beta \|1\rangle$ to $\alpha \|0\rangle - \beta \|1\rangle$.<br>This corresponds to $\phi \mapsto \phi + \pi$ |
| Phase Shift Gate | $P = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\gamma} \end{pmatrix}$ | $\phi \mapsto \phi + \gamma$ |
| S Gate | $S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$ | $\phi \mapsto \phi + \frac{\pi}{2}$ |
| T Gate | $T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$ | $\phi \mapsto \phi + \frac{\pi}{4}$ |
| Hadamard Gate | $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ | $\alpha \|0\rangle + \beta \|1\rangle \mapsto \alpha \|+\rangle + \beta \|-\rangle$<br>$\alpha \|+\rangle + \beta \|-\rangle \mapsto \alpha \|0\rangle + \beta \|1\rangle$<br>This gate essentially transforms $\|0\rangle$ to $\|+\rangle$ and $\|1\rangle$ to $\|-\rangle$ and vice versa. |

If $|\phi\rangle = \begin{pmatrix} a \\ b \end{pmatrix}$ and $|\chi\rangle = \begin{pmatrix} c \\ d \end{pmatrix}$ then:

$$|\phi\rangle \otimes |\chi\rangle = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}$$

One can verify that computing the tensor product for all possible combinations of $|0\rangle$ and $|1\rangle$ produces the 4 standard unit basis vectors in $\mathbb{R}^4$.

At this point we can introduce two new quantum gates that are classified as **controlled gates**. Controlled gates implement and if else type construct in which they basically rely on a **control qubit** and a **target qubit**.

Let us first consider the **CNOT gate** as shown below.

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

The CNOT gate flips the target qubit when the control bit is $|1\rangle$.

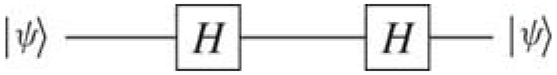
$$\begin{aligned} |00\rangle &\mapsto |00\rangle \\ |01\rangle &\mapsto |01\rangle \\ |10\rangle &\mapsto |11\rangle \\ |11\rangle &\mapsto |10\rangle \end{aligned}$$

Another controlled gate is the **Controlled Hadamard gate**.

$$CH = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

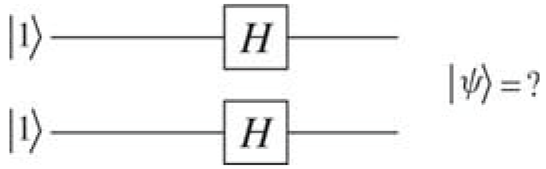
If the control qubit is $|0\rangle$ then nothing happens to the target qubit. If the control qubit is $|1\rangle$ then it applies a Hadamard gate to the target qubit.

Quantum gates act on quantum states in **quantum circuits**. If the state $|\psi\rangle$ was acted on by two Hadamard gates then the corresponding circuit would be drawn as follows.



We say that the two Hadamard gates act on the state $|\psi\rangle$ in **series**.

We can also have two Hadamard gates acting on two states in **parallel** as shown below.



To mathematically denote two quantum gates acting on two states in parallel we can use the tensor product. For the circuit above we would write:

$$(H \otimes H)\,|1\rangle\,|1\rangle = (H\,|1\rangle)(H\,|1\rangle) = \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) = \frac{1}{2}\,(|00\rangle - |01\rangle - |10\rangle + |11\rangle)$$

When $n$ Hadamard gates act in parallel on $n$ qubits it is called a **Hadamard transform**. For the case above this would be written as follows.

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

For the general case of $n$ Hadamard gates acting on $n$ qubits we would write:

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

When a quantum gate acts on a quantum state to change the probabilities of obtaining $|0\rangle$ and $|1\rangle$ then we say that **quantum interference** has occurred. **Positive interference** occurs when probability amplitudes add constructively and **negative interference** occurs when probability amplitudes add destructively. If we apply the Hadamard gate to $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ then we will get:

$$H\,|\psi\rangle = \left(\frac{\alpha + \beta}{\sqrt{2}}\right)|0\rangle + \left(\frac{\alpha - \beta}{\sqrt{2}}\right)|1\rangle$$

Here we would say the positive interference occurred for $|0\rangle$ since there is a higher probability of finding that state and negative interference occurred for $|1\rangle$ since there is a lower probability of finding that state.

It should be noted that all matrices representing quantum gates or transformations to quantum states must be **unitary**. This is necessary if we are to preserve the Euclidean length of the state which ensures that the sum of the probabilities is still equal to one. A matrix $U$ is unitary if and only if $U^{\dagger}U = I$.

## 2.3 Density Operator

Before we describe the density operator it will useful to introduce some new notation.

- For any nonempty set $\Sigma$ let $\mathbb{C}(\Sigma)$ denote the vector space of all column vectors indexed by $\Sigma$. As before column vectors are denoted by **kets** $|\psi\rangle$.
- It will be typical and often useful to assign scripted letters to spaces of the form $\mathbb{C}(\Sigma)$. For example one might write $\mathcal{X} = \mathbb{C}(\{0,1\}^n)$ to indicate that the space $\mathcal{X}$ is indexed by the set $\{0,1\}^n$.
- $\mathcal{X}^\dagger$ or $\mathcal{X}^*$ will refer to the corresponding space of row vectors denoted by $\langle\psi|$.
- When we consider a particular quantum system we assume that it has some finite set $\Sigma$ of associated classical states. We will typically use the term **register** to refer to abstract physical devices such as qubits or collections of qubits. Associated with any register having classical state set $\Sigma$ is the vector space $\mathbb{C}(\Sigma)$.

Suppose that a register X having the classical state set $\Sigma$ is in a quantum state $|\psi\rangle \in \mathcal{X}$ for $\mathcal{X} = \mathbb{C}(\Sigma)$. The new way of representing this state will be:

$$|\psi\rangle\langle\psi|$$

For example suppose that $\Sigma = \{0,1\}$ and $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Then

$$|\psi\rangle\langle\psi| = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}\begin{pmatrix}\bar{\alpha} & \bar{\beta}\end{pmatrix} = \begin{pmatrix}\alpha\bar{\alpha} & \alpha\bar{\beta}\\ \alpha\bar{\beta} & \beta\bar{\beta}\end{pmatrix} = \begin{pmatrix}|\alpha|^2 & \alpha\bar{\beta}\\ \alpha\bar{\beta} & |\beta|^2\end{pmatrix}$$

This is called a **density matrix** or **density operator**. This density matrix describes what is called a **pure state**.

Suppose we have a probability distribution $(p_1...p_k)$ as well as unit vectors $|\psi_1\rangle ... |\psi_k\rangle \in \mathcal{X}$ where $\mathcal{X} = \mathbb{C}(\Sigma)$ is the space corresponding to some register X. Someone randomly chooses $j \in \{1...k\}$ according to the probability distribution $(p_1...p_k)$ and prepares the register X in the state $\left|\psi_j\right\rangle$ for the chosen $j$. They then hand you X without telling you the value of $j$. The collection $\{(p_1, |\psi_1\rangle)...(p_k, |\psi_k\rangle)\}$ that describes the different possible states $\left|\psi_j\right\rangle$ with their associated probabilities is called a **mixture**. The density matrix corresponding the this mixture is:

$$\sum_{j=1}^{k} p_k \left|\psi_j\right\rangle\left\langle\psi_j\right|$$

It is essentially a weighted average of the pure state $\left|\psi_j\right\rangle\left\langle\psi_j\right|$. This density matrix is said to correspond to a **mixed state**.

Let us consider an example. Suppose that Alice has a qubit A. She flips a coin. If the results if HEADS then she prepares A in the state $|0\rangle$. If the result is TAILS then she prepares A in the state $|1\rangle$. She gives Bob the qubit without revealing the result of the coin flip. Bob's knowledge of the qubit is described by the density matrix

$$\frac{1}{2}\,|0\rangle\langle 0| + \frac{1}{2}\,|1\rangle\langle 1| = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

The density operator is normally denoted using the symbol $\rho$. It satisfies the following properties:

- The trace of the density matrix is 1. $\mathrm{Tr}(\rho) = 1$
- Every density matrix is positive semidefinite. This means that $\rho$ is Hermitian ($\rho = \rho^\dagger$) and that all eigenvalues of $\rho$ are nonnegative real numbers or that $\langle\phi|\,\rho\,|\phi\rangle \geq 0$

Any operation $\Phi$ on the density matrix $\rho$ that can be written as

$$\Phi(\rho) = \sum_{j=1}^{k} A_j \rho A_j^\dagger$$

where $\sum_{j=1}^{k} A_j^\dagger A_j = I$ represents an operation that can in an idealized sense be physically implemented. Such operations are called **admissible operations** or **completely positive trace preserving** operations. If $\rho$ is a density matrix and $\Phi$ is admissible then $\Phi(\rho)$ represents the new transformed density matrix.

Suppose X and Y are registers with corresponding spaces $\mathcal{X}$ and $\mathcal{Y}$. Let $\mathrm{D}(\mathcal{X})$ denote the set of all density matrices of $\mathcal{X}$. A mixed state of these two registers is represented by some element of $\mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$. If $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ is a density matrix representing the state of (XY) and Y is discarded then the resulting state of X is denoted $\mathrm{Tr}_{\mathcal{Y}}(\rho)$. We say that this state is the **reduced state** of X and we call the admissible operation $\mathrm{Tr}_{\mathcal{Y}}$ the **partial trace**. We also refer to the action corresponding to this operation as **tracing**. If we let $\Sigma$ denote the set of classical states of Y then:

$$\mathrm{Tr}_{\mathcal{Y}}(\rho) = \sum_{a \in \Sigma} (I \otimes \langle a|)\,\rho\,(I \otimes |a\rangle)$$

Suppose that the pair of qubits (XY) is in the state

$$\left|\phi^+\right\rangle = \frac{1}{\sqrt{2}}\,|00\rangle + \frac{1}{\sqrt{2}}\,|11\rangle$$

The corresponding density matrix is therefore $\left|\phi^+\right\rangle\langle\phi_+|$. The effect of discarding the second qubit is computed as follows:

$$\mathrm{Tr}_{\mathcal{Y}}\left|\phi^+\right\rangle\left\langle\phi^+\right| = (I \otimes \langle 0|)\left|\phi^+\right\rangle\left\langle\phi^+\right|(I \otimes |0\rangle) + (I \otimes \langle 1|)\left|\phi^+\right\rangle\left\langle\phi^+\right|(I \otimes |1\rangle)$$

In order to simplify this note that:

$$(I \otimes \langle 0|) \left|\phi^+\right\rangle = (I \otimes \langle 0|) \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}} |0\rangle$$

$$(I \otimes \langle 1|) \left|\phi^+\right\rangle = (I \otimes \langle 1|) \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}} |1\rangle$$

Thus:

$$\mathrm{Tr}_{\mathcal{Y}} \left|\phi^+\right\rangle \left\langle\phi^+\right| = \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1|$$

This specific density matrix is sometimes referred to as the **totally mixed state**.

# 3 Applications of Entanglement

## 3.1 Superdense Coding

Consider a scenario in which Alice would like to send Bob two classical bits of information using a single qubit. Alice and Bob begin by sharing an entangled pair of particles. The system begins in the state $|\beta_{00}\rangle$.

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

Alice has the first qubit and Bob has the second qubit. Alice acts on her qubit with a quantum gate of her choosing depending on what bit string $xy$ she wants to send to Bob. If she wants to send Bob the classical bit string 00 then she leaves her qubit alone. If she wants to send Bob the classical bit string 01 then she applies the $X$ gate. The similar idea applies for the 10 and 11 classical bit strings.

$$(X \otimes I)\,|\psi\rangle = \frac{|10\rangle + |01\rangle}{\sqrt{2}} = |\beta_{01}\rangle$$

$$(Z \otimes I)\,|\psi\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}} = |\beta_{10}\rangle$$

$$(iY \otimes I)\,|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} = |\beta_{11}\rangle$$

Alice then sends her qubit to Bob. Once Bob receives this qubit, he applies a CNOT gate to the pair and then applies a Hadamard gate to Alice's qubit. Finally he measures both qubits. The output will correspond to the classical bit string $xy$ with certainty.

| xy | Alice applies her gate | Bob applies the CNOT gate | Bob applies the Hadamard gate |
|---|---|---|---|
| 00 | $\frac{\lvert 00\rangle+\lvert 11\rangle}{\sqrt{2}}$ | $\left(\frac{1}{\sqrt{2}}\lvert 0\rangle + \frac{1}{\sqrt{2}}\lvert 1\rangle\right)\lvert 0\rangle$ | $\lvert 00\rangle$ |
| 01 | $\frac{\lvert 10\rangle+\lvert 01\rangle}{\sqrt{2}}$ | $\left(\frac{1}{\sqrt{2}}\lvert 1\rangle + \frac{1}{\sqrt{2}}\lvert 0\rangle\right)\lvert 1\rangle$ | $\lvert 01\rangle$ |
| 10 | $\frac{\lvert 00\rangle-\lvert 11\rangle}{\sqrt{2}}$ | $\left(\frac{1}{\sqrt{2}}\lvert 0\rangle - \frac{1}{\sqrt{2}}\lvert 1\rangle\right)\lvert 0\rangle$ | $\lvert 10\rangle$ |
| 11 | $\frac{\lvert 01\rangle-\lvert 10\rangle}{\sqrt{2}}$ | $\left(\frac{1}{\sqrt{2}}\lvert 0\rangle - \frac{1}{\sqrt{2}}\lvert 1\rangle\right)\lvert 1\rangle$ | $\lvert 11\rangle$ |

### 3.2 Quantum Teleportation

Suppose that Alice has a qubit that she wants to send to Bob. The state of the qubit is $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$.

Let us suppose that Alice and Bob create the entangled state

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

The first member of the pair belongs to Alice and the second member belongs to Bob. Thus the starting state is:

$$(|\psi\rangle = \alpha|0\rangle + \beta|1\rangle)\left(|\beta_{00}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle)$$

Alice applies a CNOT gate using the first qubit as the control bit and the second qubit as the target qubit. This transforms the state to:

$$\frac{1}{\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|110\rangle + \beta|101\rangle)$$

Alice then applies a Hadamard gate to the first qubit. This transforms the state to:

$$\frac{1}{2}|00\rangle(\alpha|0\rangle + \beta|1\rangle) + \frac{1}{2}|01\rangle(\alpha|1\rangle + \beta|0\rangle) + \frac{1}{2}|10\rangle(\alpha|0\rangle - \beta|1\rangle) + \frac{1}{2}|11\rangle(\alpha|1\rangle - \beta|0\rangle)$$

Alice then measures her pair. At this point there are four different possible cases. Notice that each of these possible cases occur with probability 1/4.

- **Case 1**: Alice measures 00. The state of the three qubits then becomes $|00\rangle(\alpha|0\rangle + \beta|1\rangle)$. Alice then transmits the classical bits 00 to Bob. Because is he sent the classical bits 00, he does not perform any operation on his qubit and his qubit remains in the state $\alpha|0\rangle + \beta|1\rangle$.

- **Case 2**: Alice measures 01. The state of the three qubits then becomes $|01\rangle(\alpha|1\rangle + \beta|0\rangle)$. Alice then transmits the classical bits 01 to Bob. Because is he sent the classical bits 01, he performs the NOT operation on his qubit. Thus the state of his qubit becomes $\alpha|0\rangle + \beta|1\rangle$.

- **Case 3**: Alice measures 10. The state of the three qubits then becomes $|10\rangle(\alpha|0\rangle - \beta|1\rangle)$. Alice then transmits the classical bits 10 to Bob. Because is he sent the classical bits 10, he performs the $\sigma_z$ gate on his qubit. Thus the state of his qubit becomes $\alpha|0\rangle + \beta|1\rangle$.

- **Case 4**: Alice measures 11. The state of the three qubits then becomes $|11\rangle(\alpha|1\rangle - \beta|0\rangle)$. Alice then transmits the classical bits 11 to Bob. Because is he sent the classical bits 11, he performs the NOT operation on his qubit followed by the $\sigma_z$ gate. Thus the state of his qubit becomes $\alpha|0\rangle + \beta|1\rangle$.

# 4 Quantum Algorithms

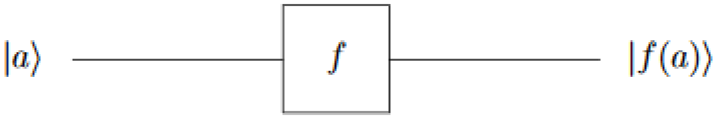
## 4.1 Deutsch's Algorithm

Suppose that we have a device that computes some function $f : \{0, 1\} \to \{0, 1\}$. It is useful for the purposes of the present investigation to think of this device as a black box. There are four possible functions:

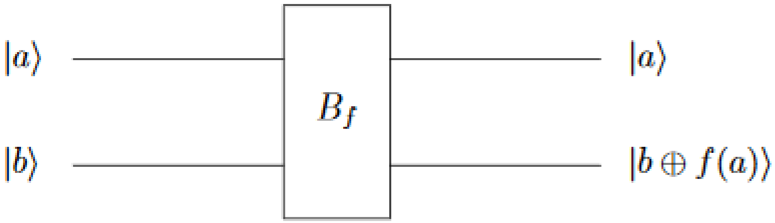
| | $f_0$ | $f_1$ | $f_2$ | $f_3$ |
|---|---|---|---|---|
| 0 | 0 | 0 | 1 | 1 |
| 1 | 0 | 1 | 0 | 1 |

Suppose we are interested in determining whether $f$ is **constant** or **balanced**. Balanced refers to having each output appear the same number of times. $f_0$ and $f_3$ are constant while $f_1$ and $f4$ are balanced. Obviously two evaluations of the function are required to determine whether $f$ is constant or balanced.

Now let us consider the same question in the content of quantum information. Here we cannot consider the black box as a one qubit gate as its corresponding matrix is not unitary.

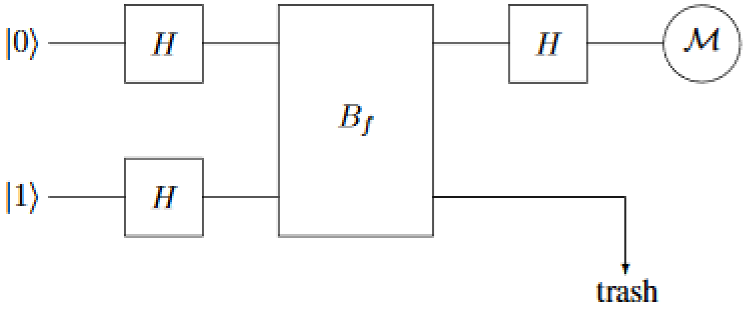


Instead we need to consider it as a two qubit quantum gate $B_f$.



Mathematically the quantum transformation $B_f$ can be written as $B_f |x\rangle |y\rangle = |x\rangle |y \oplus f(x)\rangle$ where $\oplus$ denotes XOR.

Classically we would still require two evaluations of $B_f$ to answer the question. However using a quantum algorithm only one application of $B_f$ is necessary to solve the problem. The quantum circuit diagram below explains this procedure.



This procedure is known as **Deutsch's Algorithm** and the problem of determining whether a one bit function is constant or balanced is sometimes called **Deutsch's Problem**. The output of the measurement is a single bit and the interpretation is that the value 0 indicates that the function was constant and the value 1 indicates that the function was balanced.

Let us analyze this algorithm to determine that it works correctly. If the initial state is $|0\rangle$ and $|1\rangle$ then the state after the first two Hadamard transforms is:

$$\left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right)\left(\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle\right)$$

We can partially expand this state as:

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

Performing the $B_f$ operation transforms this state to:

$$\frac{1}{2}|0\rangle(|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle) + \frac{1}{2}|1\rangle(|0 \oplus f(1)\rangle - |1 \oplus f(1)\rangle)$$

If we use the fact that $|0 \oplus a\rangle - |1 \oplus a\rangle = (-1)^a(|0\rangle - |1\rangle)$ then we can rewrite this as:

$$\frac{1}{2}(-1)^{f(0)}|0\rangle(|0\rangle - |1\rangle) + \frac{1}{2}(-1)^{f(1)}|1\rangle(|0\rangle - |1\rangle)$$

Finally we can write this as:

$$\left(\frac{1}{\sqrt{2}}(-1)^{f(0)}|0\rangle + \frac{1}{\sqrt{2}}(-1)^{f(1)}|1\rangle\right)\left(\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle\right)$$

Notice that the $B_f$ transformation has not changed the state of the second qubit. The important effect however is that the factors $(-1)^{f(0)}$ and $(-1)^{f(1)}$ have appeared in the state of the first qubit. This phenomenon is sometimes referred to as **phase kickback**. The state of the first qubit can be rewritten as:

$$(-1)^{f(0)}\left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}(-1)^{f(0)\oplus f(1)}|1\rangle\right)$$

We can use the observation that

$$H\left(\frac{1}{\sqrt{2}\,|0\rangle} + \frac{1}{\sqrt{2}}(-1)^a\,|1\rangle\right) = |a\rangle$$

This can easily be verified by considering cases $a = 0$ and $a = 1$. We can then see that applying the final Hadamard transform takes this state to:

$$(-1)^{f(0)}\,|f(0) \oplus f(1)\rangle$$

The measurement therefore results in the value $f(0) \oplus f(1)$ with certainty. This value is 0 if $f$ is constant and 1 if $f$ is balanced.

## 4.2 Deutsch Jozsa Algorithm

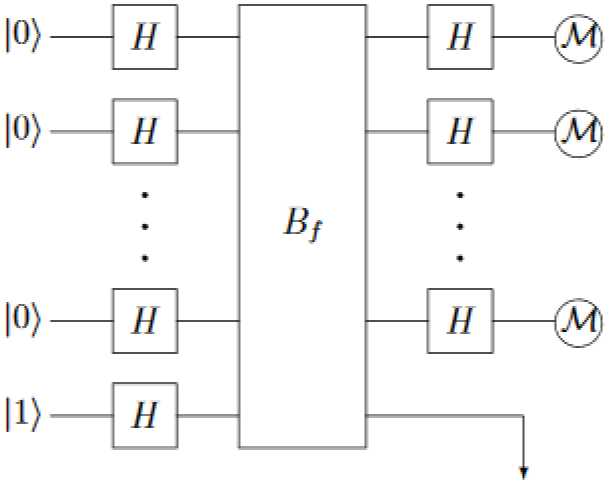
The **Deutsch Jozsa Algorithm** is a generalization of Deutsch's Algorithm. This time we assume that we are given a function $\{0,1\}^n \to \{0,1\}$. We are promised that two possibilities hold:

1. $f$ is **constant**. This means that either $f(x) = 0$ for all $x \in \{0,1\}^n$ or $f(x) = 1$ for all $x \in \{0,1\}^n$

2. $f$ is **balanced**. This means that the number of inputs $x \in \{0,1\}^n$ for which the function takes values 0 and 1 are the same.

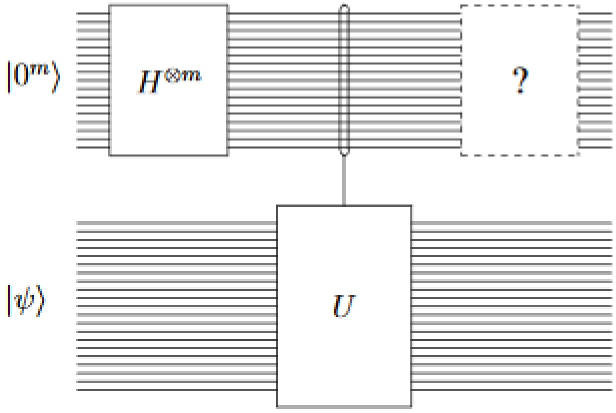
The goal is to determine which of the two possibilities holds. Classically this problem is pretty easy given a small number of queries if we allow randomness and accept that there may be a small probability of error. However in the quantum case, one query will be sufficient to determine with certainty whether the function is constant or balanced. The procedure responsible for doing this is the Deutsch Jozsa Algorithm. Its quantum circuit is shown below.

There are $n$ bits resulting from the measurements. If all $n$ measurement results are 0 then we conclude that the function was constant. Otherwise we conclude that the function was balanced.



## 4.3 Phase Estimation

Suppose we have a quantum circuit $Q$ acting on $n$ qubits as shown below.



Because $U$ is unitary we know from linear algebra that it has a complete orthonormal collection of eigenvectors.

$$|\psi_1\rangle \ldots |\psi_N\rangle$$

It also has associated eigenvalues of the form

$$e^{2\pi i\theta_1} \ldots e^{2\pi i\theta_N}$$

where $\theta_i \in [0, 1)$. This of course means that $U\left|\psi_j\right\rangle = e^{2\pi i\theta_j}\left|\psi_j\right\rangle$. The reason why each of the eigenvalues has the form $e^{2\pi i\theta}$ (which is equivalent to saying that these eigenvalues exist on the complex unit circle) is that $U$ is unitary and therefore preserves Euclidean length. The problem is as follows: A quantum circuit $Q$ performs a unitary operation $U$ on a quantum state $|\psi\rangle$ that is promised to be an eigenvector of $U$ such that $U|\psi\rangle = e^{2\pi i\theta}|\psi\rangle$. We want to determine an approximation to $\theta \in [0, 1)$. This is known as the **phase estimation problem**.

We will now describe the phase estimation procedure. Suppose that $U$ is the unitary transformation acting on $n$ qubits and suppose that $m$ is any positive integer. Then we let $\Lambda_m(U)$ denote the unique unitary transformation on $m + n$ qubits that satisfies:

$$\Lambda_m(U)|k\rangle|\phi\rangle = |k\rangle(U^k|\phi\rangle)$$

In other words the first $m$ qubits specify the number of times that $U$ is applied to the remaining $n$ qubits. From our quantum circuit schematic we see that the initial state is $|0^m\rangle|\psi\rangle$ and after the Hadamard transforms are performed the state becomes:

$$\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)^{\otimes m}|\psi\rangle = \frac{1}{2^{m/2}}\sum_{k=0}^{2^m-1}|k\rangle|\psi\rangle$$

We then apply the $\Lambda_m(U)$ transformation. This transforms the state to:

$$\frac{1}{2^{m/2}}\sum_{k=0}^{2^m-1}|k\rangle(U^k|\psi\rangle)$$

Notice that $U^k|\psi\rangle = (e^{2\pi i\theta})^k|\psi\rangle = e^{2\pi ik\theta}|\psi\rangle$. Thus we can rewrite our state as:

$$\frac{1}{2^{m/2}}\sum_{k=0}^{2^m-1}|k\rangle(e^{2\pi ik\theta}|\psi\rangle) = \frac{1}{2^{m/2}}\sum_{k=0}^{2^m-1}e^{2\pi ik\theta}|k\rangle|\psi\rangle$$

Given that they are in a product tensor state we can rewrite this state as follows:

$$\left(\frac{1}{2^{m/2}}\sum_{k=0}^{2^m-1}e^{2\pi ik\theta}|k\rangle\right)|\psi\rangle$$

So if we discard the last $n$ qubits we are left with the state:

$$\frac{1}{2^{m/2}}\sum_{k=0}^{2^m-1}e^{2\pi ik\theta}|k\rangle$$

Let us consider a simple case in which we suppose that $\theta = j/2^m$ for $j \in \{0...2^m - 1\}$. If we define $\omega \equiv e^{2\pi i}/2^m$ then our state can be rewritten as:

$$\frac{1}{2^{m/2}}\sum_{k=0}^{2^m-1}\omega^{jk}|k\rangle$$

Let us define:

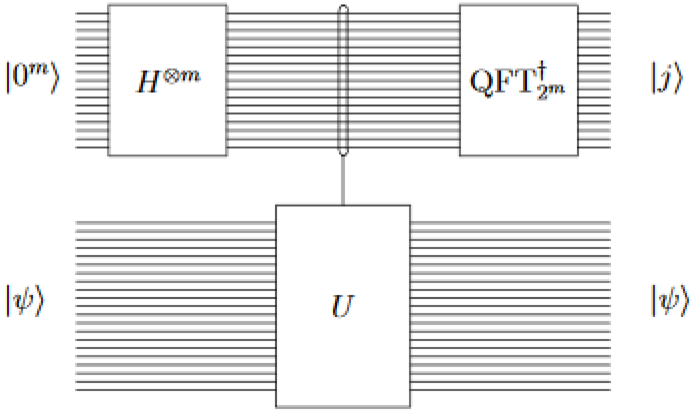
$$\left|\phi_j\right\rangle = \frac{1}{2^{m/2}} \sum_{k=0}^{2^m-1} \omega^{jk} |k\rangle$$

for each $j \in \{0...2^m - 1\}$. We know that the state of the first $m$ qubits of our circuit is one of the state $\left\{\left|\phi_j\right\rangle \,:\, j = 0...2^m - 1\right\}$ and the goal is to determine which one.

It can be shown that the set $\{|\phi_0\rangle \ldots |\phi_{2^m-1}\rangle\}$ is indeed orthonormal. Therefore there exists a unitary transformation $F$ that satisfies $F|j\rangle = \left|\phi_j\right\rangle$. We can describe this matrix by allowing the vectors $\left|\phi_j\right\rangle$ to determine the columns of $F$.

$$F = \frac{1}{\sqrt{2^m}}\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{2^m-1} \\ 1 & \omega & \omega^4 & \cdots & \omega^{2(2^m-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{2^m-1} & \omega^{2(2^m-1)} & \cdots & \omega^{(2^m-1)^2} \end{pmatrix}$$

This matrix describes the **discrete Fourier transform**. In the context of quantum computing we will often refer to this as the **quantum Fourier transform**. It is sometimes denoted $\mathrm{QFT}_{2^m}$. Plugging the inverse of this transformation into our circuit from before gives:



Thus measuring the first $m$ qubits and dividing by $2^m$ tells us precisely the value $\theta$.

What about dealing with arbitrary values of $\theta$? Recall that the state of the circuit immediately before the inverse of the QFT is applied is

$$\frac{1}{2^{m/2}} \sum_{k=0}^{2^m-1} e^{2\pi i k\theta} |k\rangle |\psi\rangle$$

Once again we know that the second collection of qubits is uncorrelated with the first $m$ qubits and thus can be ignored. After applying the transformation $\mathrm{QFT}^{\dagger}_{2m}$ to the first $m$ qubits results in the state:

$$\sum_{j=0}^{2^m-1}\left(\frac{1}{2^m}\sum_{k=0}^{2^m-1} e^{2\pi i k(\theta - j/2^m)}\right)|j\rangle$$

The probability that the measurement results in the outcome $j$ is therefore:

$$p_j = \left|\frac{1}{2^m}\sum_{k=0}^{2^m-1} e^{2\pi i k(\theta - j/2^m)}\right|^2$$

We see that if $\theta = j/2m$ then $p_j = 1$. Recall the formula for the sum of a geometric series:

$$\sum_{k=0}^{n-1} x^k = \frac{x^n - 1}{x - 1}$$

We may then simplify $p_j$ to get:

$$p_j = \frac{1}{2^{2m}}\left|\frac{e^{2\pi i(2^m\theta - j)} - 1}{e^{2\pi i(\theta - j/2^m)} - 1}\right|^2$$

It can be shown that the probability $p_j$ is large for values of $j$ that satisfy $j/2m \approx \theta$ and small otherwise. We can run the phase estimation procedure several times and use the value of $j$ that occurs the most frequently to give an approximation of $\theta$.

### 4.4 Shor's Algorithm

It is time to apply the phase estimation technique to an interesting computational problem. But first we should review some notation. If $a$ $b$ and $N$ are integers with $N \geq 1$ then we write

$$a \equiv b(\mod N)$$

to mean that $N$ divides $a - b$. We let $\mathbb{Z}_N$ denote the set $\mathbb{Z}_N = \{0...N-1\}$. We write $\mathbb{Z}_N^*$ to denote the following set:

$$\mathbb{Z}_N^* = \{a \in \mathbb{Z}_N : \gcd(a, N) = 1\}$$

If $a \in \mathbb{Z}_N^*$ then the **order** of $a$ in $\mathbb{Z}_N^*$ is the smallest positive integer $r$ such that

$$a^r \equiv 1(\mod N)$$

As an example let us consider $a = 4$ and $N = 35$. Because $\gcd(4, 35) = 1$ we have $4 \in \mathbb{Z}_{35}^*$. Computing powers of 4 modulo 35 gives:

$$4^1 \equiv 4 \quad 4^2 \equiv 16 \quad 4^3 \equiv 29 \quad 4^4 \equiv 11 \quad 4^5 \equiv 9 \quad 4^6 \equiv 1$$

Thus the order of 4 modulo 35 is 6. Overall the problem of **order finding** can be summarized as follows:

- Input: A positive integer $N \geq 2$ and an element $a \in \mathbb{Z}_N^*$
- Output: The order of $a$ in $\mathbb{Z}_N^*$

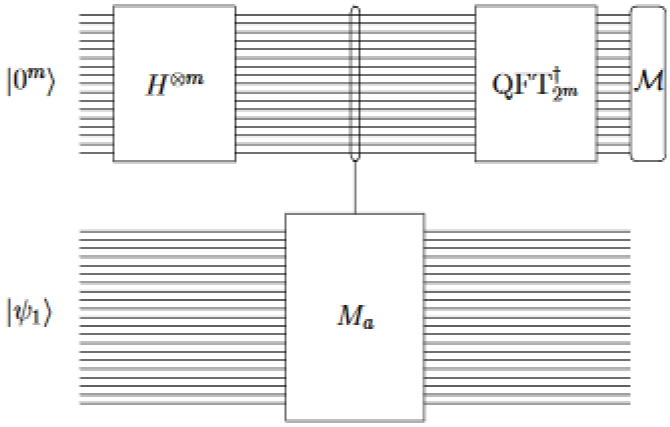
We may now look at the algorithm itself. We begin by defining the unitary transformation $M_a$ as

$$M_a |x\rangle = |ax \text{ ( mod } N)\rangle$$

If we define $\omega_r \equiv e^{2\pi i/r}$ then it can be shown that the eigenstates of $M_a$ are given by:

$$\left|\psi_j\right\rangle = \frac{1}{\sqrt{r}}\left(|1\rangle + \omega_r^{-j}|a\rangle + \omega_r^{-2j}\left|a^2\right\rangle + \cdots + \omega_r^{-j(r-1)}\left|a^{r-1}\right\rangle\right)$$

Our corresponding eigenvalue equation is $M_a\left|\psi_j\right\rangle = \omega_r^j\left|\psi_j\right\rangle$. Consider the phase estimation procedure for the state $|\psi_1\rangle$.



The eigenvalue associated with $|\psi_1\rangle$ is $\omega_r = e^{2\pi i(1/r)}$. If we performed the procedure several times and took the most common result, we would have an approximation $j/2^m$ that with very high probability is within distance $1/2^{m+1}$ of $1/r$. Of course we would compute the reciprocal of our approximation to obtain $r$. Since this is an approximation we would also need to round to the nearest integer. It can be shown that if we take $m = 2n$ then the resulting accuracy will be sufficient to find $r$.

Unfortunately there is no known way to prepare $|\psi_1\rangle$ without knowing $r$. Thus a different approach is required. Instead we consider running the phase estimation procedure on the state

$$|1\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} |\psi_k\rangle$$

Using this state would produce a result that would be indistinguishable from running the phase estimation procedure on an eigenvector $\psi_k$ for a random value of $k \in \{0...r-1\}$ chosen uniformly. In other words the phase estimation procedure will give an approximation

$$\frac{j}{2^m} \approx \frac{k}{r}$$

for $k \in \{0...r-1\}$ chosen uniformly. The question is how do find $r$ for such an approximation? It is important for us to keep in mind that $k$ will not be known explicitly when we attempt to find $r$.

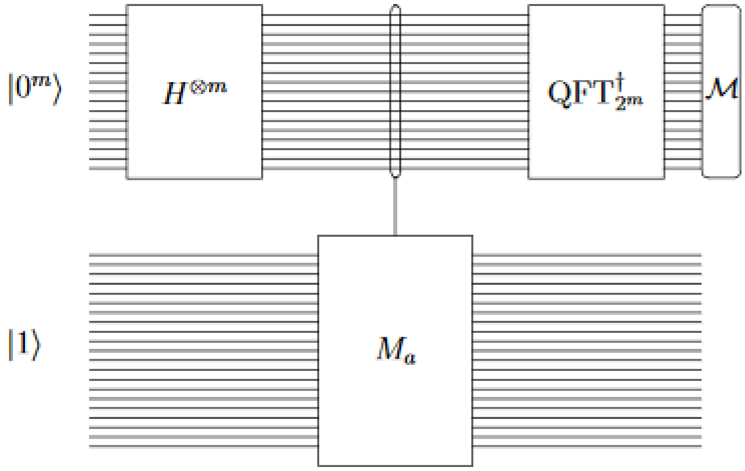
The answer is that you cannot be guaranteed to find $r$ given just one sample. But if you repeat the process several times (each time for a different value of $k$) you can find $r$ with high probability. Plugging the measurement outcome $j/2^m$ along with $N$ into the **continued fraction algorithm** will produce the values $x$ and $y$ such that

$$\frac{x}{y} = \frac{k}{r}$$

In other words we can find $k/r$ in lowest terms. Our value of $y$ will not return $r$ if $\gcd(k, r) > 1$ (if $k/r$ can be simplified) but at least we will know that $y$ divides $r$. Repeating several times (each time for a different value of $k$) and taking the least common multiple/largest of the resulting $y$ values gives $r$ with high probability. An example is shown below for $r = 6$.

| | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ | $k = 5$ |
|---|---|---|---|---|---|
| $r = 6$ | $\frac{1}{6}$ | $\frac{2}{6}$ | $\frac{3}{6}$ | $\frac{4}{6}$ | $\frac{5}{6}$ |
| | | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{3}$ | |
| | 6 | 3 | 2 | 3 | 6 |

Our algorithm can be summarized as follows. First we choose $m = 2n$ and run the phase estimation procedure given by the following quantum circuit diagram.



Plug the measurement outcome $j/2^m$ along with $N$ into the continued fraction algorithm to obtain a fraction $x/y$. Repeat this entire process several times and take the least common multiple of the $y$ values. The result will be $r$ with high probability. This entire process is known as **Shor's algorithm for order finding**.

### 4.5 Grover's Algorithm

Suppose that we have a function $f : \{0,1\}^n \to \{0,1\}$. We want to find a string $x \in \{0,1\}^n$ such that $f(x) = 1$ or conclude that no such $x$ exists if $f$ is identically 0. It is important to note that this problem is **unstructured** as there are no promises on the function $f$. Thus it is not possible to use any fast searching method to efficiently solve this problem classically. The purpose of **Grover's Algorithm** is to solve this problem. To describe this algorithm we will need to define the following two unitary transformations on $n$ qubits:

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

$$Z_0 |x\rangle = \begin{cases} -|x\rangle & \text{if } x = 0^n \\ |x\rangle & \text{if } x \neq 0^n \end{cases}$$

Grover's algorithm can then be stated as follows:

1. Let X be an $n$ qubit quantum register (a collection of $n$ qubits to which we assign the name X). Let the starting state of X be $|0^n\rangle$ and perform $H^{\otimes n}$ on X.
2. Apply to the register X the transformation $G = -H^{\otimes n} Z_0 H^{\otimes n} Z_f$ $k$ times.
3. Measure X and output the result.

Let us define the sets of strings $A$ and $B$ as follows:

$$A = \{x \in \{0,1\}^n : f(x) = 1\}$$
$$B = \{x \in \{0,1\}^n : f(x) = 0\}$$

We will think of $A$ as the good set of strings. The goal of the algorithm is to find one of these strings. The set $B$ contains the bad set of strings that do not satisfy the search criterion. We will let $a = |A|$ and $b = |B|$. Next we will define:

$$|A\rangle = \frac{1}{\sqrt{a}} \sum_{x \in A} |x\rangle$$
$$|B\rangle = \frac{1}{\sqrt{b}} \sum_{x \in B} |x\rangle$$

The state of the register X immediately after step 1 is:

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

We can rewrite this as:

$$|h\rangle = \sqrt{\frac{a}{N}} |A\rangle + \sqrt{\frac{b}{N}} |B\rangle$$

If we rewrite $Z_0$ as $Z_0 = I - 2 |0^n\rangle \langle 0^n|$ then it follows that:

$$H^{\otimes n} Z_0 H^{\otimes n} = H^{\otimes n} \left(I - 2 \left|0^n\right\rangle \left\langle 0^n\right|\right) H^{\otimes n} = I - 2 |h\rangle \langle h|$$

Here we are using the fact that $H^\dagger = H$ and the fact that if $|\psi\rangle = U |\phi\rangle$ then $\langle\psi| = \langle\phi| U^\dagger$. Now we can examine the effects of $G$ on $|A\rangle$ and $|B\rangle$.

$$\begin{aligned}
G |A\rangle &= -H^{\otimes n} Z_0 H^{\otimes n} Z_f |A\rangle \\
&= (I - 2 |h\rangle \langle h|) \left(-Z_f\right) |A\rangle \\
&= (I - 2 |h\rangle \langle h|) |A\rangle \\
&= |A\rangle - 2 \langle h|A\rangle |h\rangle \\
&= |A\rangle - 2 \sqrt{\frac{a}{N}} \left( \sqrt{\frac{a}{N}} |A\rangle + \sqrt{\frac{b}{N}} |B\rangle \right) \\
&= \left(1 - \frac{2a}{N}\right) |A\rangle - \frac{2\sqrt{ab}}{N} |B\rangle
\end{aligned}$$

$$\begin{aligned}
G\,|B\rangle &= -H^{\otimes n} Z_0 H^{\otimes n} Z_f\,|B\rangle \\
&= -\left(I - 2\,|h\rangle\,\langle h|\right)\left(Z_f\right)|B\rangle \\
&= -\left(I - 2\,|h\rangle\,\langle h|\right)|B\rangle \\
&= -\,|B\rangle + 2\,\langle h|B\rangle\,|h\rangle \\
&= -\,|B\rangle + 2\sqrt{\frac{b}{N}}\left(\sqrt{\frac{a}{N}}\,|A\rangle + \sqrt{\frac{b}{N}}\,|B\rangle\right) \\
&= \frac{2\sqrt{ab}}{N}\,|A\rangle - \left(1 - \frac{2b}{N}\right)|B\rangle
\end{aligned}$$

This shows that the state of X after each application of $G$ will remain in the subspace spanned by $|A\rangle$ and $|B\rangle$.

If we let the first element of a vector written in the basis formed by $|A\rangle$ and $|B\rangle$ correspond to $|B\rangle$ and the second element correspond to $|A\rangle$ then the action of $G$ can be expressed as a matrix.

$$M = \begin{bmatrix} -\left(1 - \frac{2b}{N}\right) & -\frac{2\sqrt{ab}}{N} \\ \frac{2\sqrt{ab}}{N} & \left(1 - \frac{2a}{N}\right) \end{bmatrix} = \begin{bmatrix} \frac{b-a}{N} & -\frac{2\sqrt{ab}}{N} \\ \frac{2\sqrt{ab}}{N} & \frac{b-a}{N} \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{b}{N}} & -\sqrt{\frac{a}{N}} \\ \sqrt{\frac{a}{N}} & \sqrt{\frac{b}{N}} \end{bmatrix}^2$$

Comparing this with the matrix corresponding to a rotation transformation verifies that $M$ is a rotation.

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

If $\theta$ is the angle that satisfies

$$\sin\theta = \sqrt{\frac{a}{N}} \quad \text{and} \quad \cos\theta = \sqrt{\frac{b}{N}}$$

Then

$$R_{2\theta} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^2 = \begin{bmatrix} \sqrt{\frac{b}{N}} & -\sqrt{\frac{a}{N}} \\ \sqrt{\frac{a}{N}} & \sqrt{\frac{b}{N}} \end{bmatrix}^2 = M$$

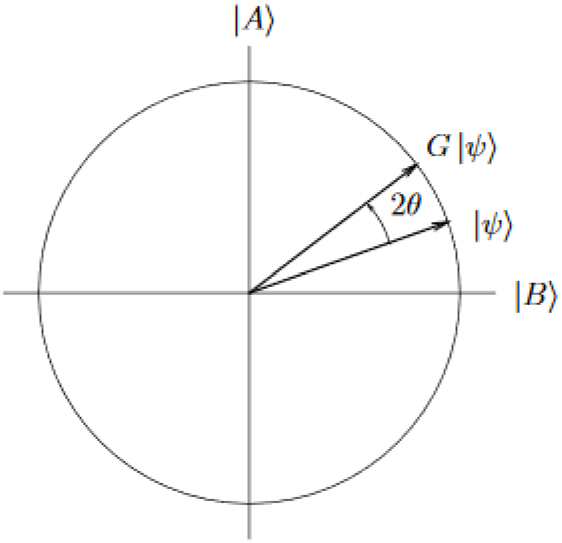
Thus $G$ causes a rotation by an angle $2\theta$ in the space spanned by $|A\rangle$ and $|B\rangle$ for

$$\theta = \sin^{-1}\sqrt{\frac{a}{N}}$$

The state of the register X can now be written as follows.

$$|h\rangle = \sqrt{\frac{b}{N}}\,|B\rangle + \sqrt{\frac{a}{N}}\,|A\rangle = \cos\theta\,|B\rangle + \sin\theta\,|A\rangle$$

The transformation corresponding to one application of $G$ is shown below.



After $k$ iterations of $G$ the state will be

$$\cos((2k+1)\theta)\,|B\rangle + \sin((2k+1)\theta)\,|A\rangle$$

The goal is to measure some element $x \in A$. Thus we would like the state of X to be as close to $|A\rangle$ as possible. If we want

$$\sin((2k+1)\theta) \approx 1$$

then

$$(2k+1)\theta \approx \frac{\pi}{2}$$

will suffice. Thus we choose

$$k \approx \frac{\pi}{4\theta} - \frac{1}{2}$$

Of course $k$ must be an integer. That is why we can only hope to approximate this quantity. Suppose $a = 1$. Then

$$\theta = \sin^{-1}\sqrt{\frac{1}{N}} \approx \frac{1}{\sqrt{N}}$$

Thus

$$k = \left\lfloor \frac{\pi}{4} \sqrt{N} \right\rfloor$$

With this choice of $k$ the probability of finding the single $x$ such that $f(x) = 1$ is:

$$\sin^2\left(\left(2\left\lfloor \pi\sqrt{N}/4 \right\rfloor + 1\right)\sin^{-1}\left(1/\sqrt{N}\right)\right)$$

The limit of this probability is 1 as $N$ goes to infinity and is always at least $1/2$. So even in the worst case, repeating the algorithm some small constant number of times and evaluating $f$ at the output each time will find the unique $x$ such that $f(x) = 1$ with very high probability.

In general case when we do not know that $a = 1$, the situation is more challenging. For instance if $a = 4$ but we still choose $k = \left\lfloor \pi\sqrt{N}/4 \right\rfloor$ then the success probability goes to 0 in the limit of large $N$. One strategy would be to simply choose a random value of $k$ in the range $\left\{1 ... \sqrt{N} + 1\right\}$. However the algorithm will fail to find an $x$ that satisfies $f(x) = 1$ with probability at most $3/4$. A somewhat better strategy would be to do the following:

1. Set $m = 1$

2. Choose $k \in \{1 ... m + 1\}$ uniformly and run Grover's algorithm for this choice of $k$. If the algorithm finds an $x$ such that $f(x) = 1$ then output $x$ and stop the procedure.

3. If $m > \sqrt{N}$ then fail. Else set $m = \lfloor (8/7)m \rfloor$ and return to step 2.

Note that 8/7 just happens to be the number that works for the analysis.

# 5 Quantum Cryptography

## 5.1 Quantum Key Distribution

Consider the scenario in which Alice and Bob wish to communicate while preventing an eavesdropper (Eve) from learning any information about their communication. Secure communication can be accomplished through **key distribution**. Key distribution is a method by which two parties that want to communicate privately share a key that it used to encrypt or scramble the message. Later the key can be used to recover or decrypt the message. A trivial way to encrypt a message is to generate a key $k$. We can add $k$ to each character in the message. When it reaches Bob, he subtracts $k$ to decrypt or recover the original message. Letters are encoded using binary digits.

$$
\begin{gathered}
A \to 0000 \\
B \to 0001 \\
C \to 0010 \\
D \to 0011
\end{gathered}
$$

Suppose that the message is $m$. The transmitted string is then $t = m + k$. Suppose that $k = 3$. We add this value to each character in the given message.

$$
\begin{gathered}
0000 \to 0011 \;\; (A \to D) \\
0001 \to 0100 \;\; (B \to E) \\
0010 \to 0101 \;\; (C \to F) \\
0011 \to 0110 \;\; (D \to G)
\end{gathered}
$$

Bob knows that he can decrypt the signal by $m = t - k$. It is very important to note that $k$ can only be used once. Even after the third there is a strong chance that Eve can learn something important about $k$. This method is called the **one time pad method**. The one time pad method is unbreakable but is very inefficient or impossible in many situations. For many pairs of parties including banks and online businesses, it is impossible to classically generate a secure private key over a public channel.

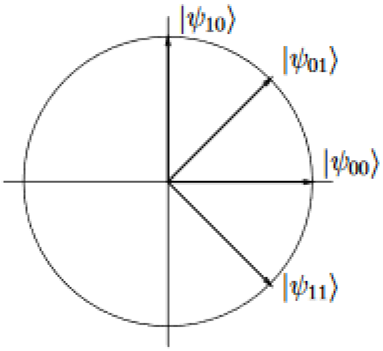
The aim of **quantum key distribution** or **QKD** is to allow Alice and Bob to generate a secure private key that can be used for the one time pad method without having to meet privately. While there are many different schemes for doing this, we will only be discussing the **BB84 protocol**. It is important to note that we are assuming in these situations that Alice and Bob are communicating over a classical channel that is authenticated and cannot be tampered with but is readable by Eve. We call this the **public channel**. They also send qubits over a quantum channel. Eve has full access to this channel meaning that she can intercept these messages and perform measurements on these messages and transmit the modified qubits to whichever party was to receive them in the first place.

## 5.2 The BB84 Protocol

Let us first define the following four states:

$$|\psi_{00}\rangle = |0\rangle$$
$$|\psi_{10}\rangle = |1\rangle$$
$$|\psi_{01}\rangle = |+\rangle$$
$$|\psi_{11}\rangle = |-\rangle$$

These four states are shown in the image below.



Alice begins by randomly generating two strings of bits $x \in \{0,1\}^m$ and $y \in \{0,1\}^m$. Alice prepares $m$ qubits in the state $\left|\psi_{xy}\right\rangle = \left|\psi_{x_1y_1}\right\rangle\left|\psi_{x_2y_2}\right\rangle\cdots\left|\psi_{x_my_m}\right\rangle$ and sends these $m$ qubits over a quantum channel to Bob.

Bob randomly chooses $y' \in \{0,1\}^m$ and measures each qubit received from Alice as follows:

- If $y'_i = 0$ then Bob measures qubit $i$ with respect to the standard basis.
- If $y'_i = 1$ then Bob measures qubit $i$ with respect to the diagonal basis.

Let $x' \in \{0,1\}^m$ be the string corresponding to the results of Bob's measurements. The important thing to note is that if $y_i = y'_i$ and there was no noise or eavesdropping then it is certain that $x_i = x'_i$.

Finally Alice and Bob publicly compare $y$ and $y'$. They discard all bits $x_i$ and $x'_i$ for which $y_i \neq y'_i$. Of course the difficulty arises when Eve tries to extract information about $x$. The general principle working for Alice and Bob is the **uncertainty principle**. Eve cannot learn something about $\left|\psi_{xy}\right\rangle$ without disturbing it. She does not know $y$ and thus she cannot hope to learn the bits of $x$ without making some mistakes and causing some disturbance to $|\psi\rangle$.

What we have described now is the first state of the protocol. In the second stage of the protocol, Alice and Bob need to estimate how much Eve might know about $x$ and $x'$. They do this by sacrificing some of the remaining bits of $x$ and $x'$ ($1/2$ of them selected randomly for example). By comparing these bits publicly, they can estimate the error rate with high accuracy and if it is too large then they abort (Eve potentially has too much information for them to succeed). The maximum error rate that can be tolerated is about 11%. If they have an acceptable error rate then Alice and Bob will have two strings $x$ and $x'$ that agree in a high percentage of positions with high probability.

**Example**: Suppose $m = 8$. Alice randomly chooses

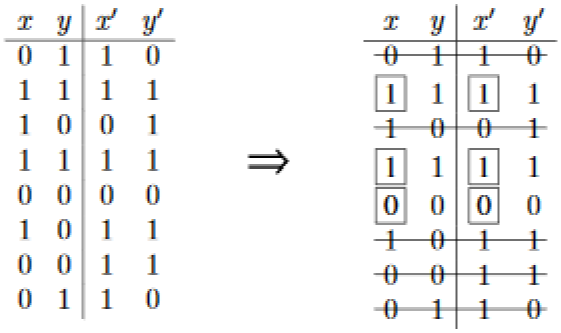
$$x = 01110100$$

$$y = 11010001$$

She sends $\left|\psi_{xy}\right\rangle$ to Bob. Suppose there is no tampering with the message so that Bob receives precisely $\left|\psi_{xy}\right\rangle$. He randomly chooses

$$y' = 01110110$$

and measures. The first qubit is in the state $|\psi_{01}\rangle = |+\rangle$ and because $y'_1 = 0$ Bob measures in the standard basis. He will get a uniform random bit; say $x'_1 = 1$. The second qubit is in the state $|\psi_{11}\rangle = |-\rangle$ and because $y'_2 = 1$ Bob measures in the diagonal basis. He gets $x'_2 = 1$ with certainty this time. Continuing this way we might get the following table:



| $x$ | $y$ | $x'$ | $y'$ |
|---|---|---|---|
| 0 | 1 | 1 | 0 |
| 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 1 |
| 1 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 |
| 0 | 0 | 1 | 1 |
| 0 | 1 | 1 | 0 |

$\Rightarrow$

| $x$ | $y$ | $x'$ | $y'$ |
|---|---|---|---|
| ~~0~~ | ~~1~~ | ~~1~~ | ~~0~~ |
| $\boxed{1}$ | 1 | $\boxed{1}$ | 1 |
| ~~1~~ | ~~0~~ | ~~0~~ | ~~1~~ |
| $\boxed{1}$ | 1 | $\boxed{1}$ | 1 |
| $\boxed{0}$ | 0 | $\boxed{0}$ | 0 |
| ~~1~~ | ~~0~~ | ~~1~~ | ~~1~~ |
| ~~0~~ | ~~0~~ | ~~1~~ | ~~1~~ |
| ~~0~~ | ~~1~~ | ~~1~~ | ~~0~~ |

Alice and Bob compare publicly compare $y$ and $y'$. They agree at positions 2 4 and 5. Thus they keep the bits of $x$ and $x'$ at these positions and discard the rest. The three remaining bits of $x$ and $x'$ agree.

# 6 Quantum Noise and Error Correction

## 6.1 Classical Repetition Codes

We will start with a very simple classical error correcting code known as the **3 bit repetition code**. Consider a channel from Alice to Bob that carries 1 bit at a time. There is noise in this channel so that the bit sent by Alice is not always received correctly by Bob. We suppose that the noise is parameterized by some real number $p \in [0, 1]$ that represents the probability of an error. If Alice sends the bit $b$ then Bob receives $b$ with probability $1 - p$ and $\neg b$ with probability $p$.

$$b \mapsto \begin{cases} b & \text{with probability } 1 - p \\ \neg b & \text{with probability } p \end{cases}$$

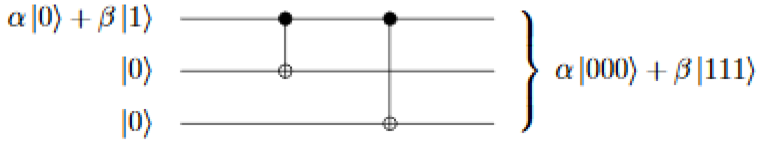
Suppose that Alice needs to send a very important bit to Bob and the two of them cannot accept that Bob will receive the wrong bit with probability $p$. What can they do? Well they can use an error correcting code to decrease the probability of error. Possibly the simplest example o an error correcting code is the 3 bit repetition code. The encoding is to simply send the same bit 3 times and the decoding is to simply take the most frequently appearing bit. Overall the probability of error shrinks from $p$ to $3p^2 - 2p^3$. Note that this type of channel is called a **binary symmetric channel** and this code protects against **one bit flip**.

## 6.2 Bit Flips

Can we encode quantum information in a similar way to the previous example? Suppose Alice has a qubit $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$ that she wishes to send to Bob through a noisy channel. A sensible analogue of the repetition code is to perform this encoding:

$$\alpha |0\rangle + \beta |1\rangle \mapsto \alpha |000\rangle + \beta |111\rangle$$

This encoding could be performed using the following simple circuit:


$\alpha |0\rangle + \beta |1\rangle$
$|0\rangle$
$|0\rangle$
$\alpha |000\rangle + \beta |111\rangle$

This encoding corrects against bit flip errors. Consider a quantum channel that is again parameterized by the number $p \in [0, 1]$. With probability $1 - p$ the channel acts as the identity. With the remaining probability $p$ the channel applies a bit flip error. This means that the following unitary operation is performed:

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

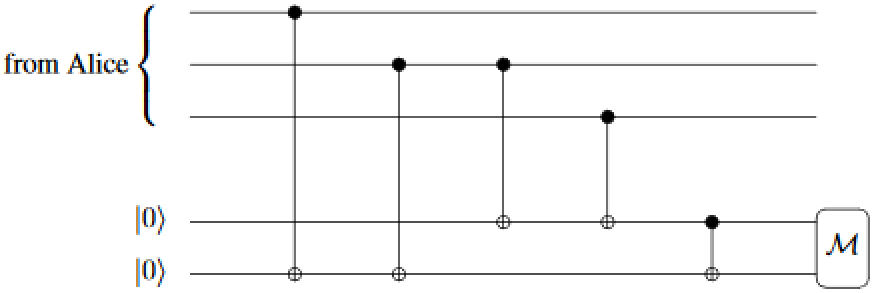
This channel can therefore be described by the admissible operation

$$\Phi(\rho) = (1 - p)\rho + p\sigma_x\rho\sigma_x^\dagger = (1 - p)\rho + p\sigma_x\rho\sigma_x$$

Suppose that Alice encodes $\alpha\,|0\rangle + \beta\,|1\rangle$ as $\alpha\,|000\rangle + \beta\,|111\rangle$ and sends the three qubits through the channel one at a time. Imagine that the second qubit experiences a bit flip error but the first and third are unaffected. The state of the three qubits becomes

$$\alpha\,|010\rangle + \beta\,|101\rangle$$

Bob could decode this using the following circuit:



Consider what this circuit does in the present case:

$$\begin{aligned}(\alpha\,|010\rangle + \beta\,|101\rangle)\,|00\rangle &= \alpha\,|010\rangle\,|00\rangle + \beta\,|101\rangle\,|00\rangle \\ &\mapsto \alpha\,|010\rangle\,|10\rangle + \beta\,|101\rangle\,|10\rangle \\ &= (\alpha\,|010\rangle + \beta\,|101\rangle)\,|10\rangle\end{aligned}$$

The measurement would give outcome 10 with certainty. The output string is called the **syndrome** and in this case it tells us that the bit flip error occurred on the second qubit (or 10 is binary). Thus Bob corrects the error by applying $\sigma_z$ to the second qubit:

$$\alpha\,|010\rangle + \beta\,|101\rangle \mapsto \alpha\,|000\rangle + \beta\,|111\rangle$$

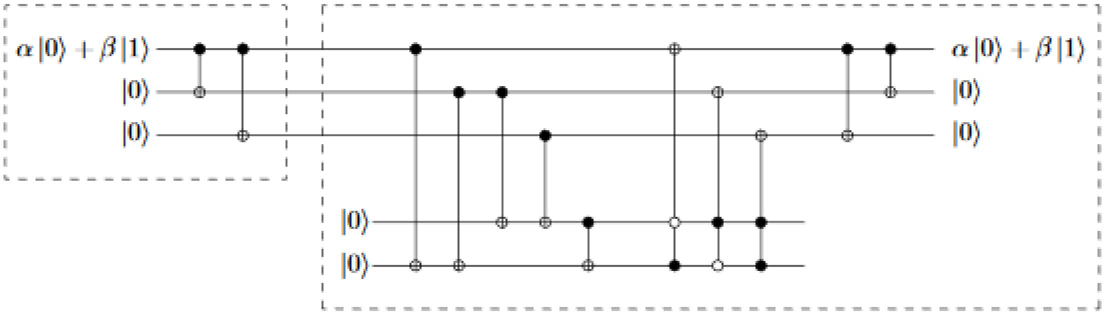
Now Bob can apply the inverse of the encoding procedure to recover $\alpha\,|0\rangle + \beta\,|1\rangle$:

$$\alpha\,|000\rangle + \beta\,|111\rangle \mapsto (\alpha\,|0\rangle + \beta\,|1\rangle)\,|00\rangle$$

The same procedure works in the case that the first or third qubit experiences a bit flip as well. In general if the syndrome is 00 then Bob will do nothing. Otherwise if the syndrome is $j \in \{1\ 2\ 3\}$ in binary then Bob will apply a $\sigma_x$ operation to qubit number $j$. Bob will then apply the reverse of the encoding procedure.

| Classical state | Syndrome |
|---|---|
| $\|000\rangle$ | 00 |
| $\|001\rangle$ | 11 |
| $\|010\rangle$ | 10 |
| $\|011\rangle$ | 01 |
| $\|100\rangle$ | 01 |
| $\|101\rangle$ | 10 |
| $\|110\rangle$ | 11 |
| $\|111\rangle$ | 00 |

The entire code can be represented by the following diagram.



$\alpha|0\rangle + \beta|1\rangle$
$|0\rangle$
$|0\rangle$
$|0\rangle$
$|0\rangle$
$\alpha|0\rangle + \beta|1\rangle$
$|0\rangle$
$|0\rangle$

Just as with the classical case this code corrects up to one bit flip error.

## 6.3 Phase Flips

Another example of an error that could occur is the **phase flip error**. This occurs when $|a\rangle \mapsto (-1)^a |a\rangle$ for $a \in \{0, 1\}$. This error is represented by the $\sigma_z$ matrix.

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

Suppose that we are interested in protecting against phase errors. We could change the encoding slightly in order to do this:

$$\alpha|0\rangle + \beta|1\rangle \mapsto \alpha|+\rangle|+\rangle|+\rangle + \beta|-\rangle|-\rangle|-\rangle$$

This is easily done by encoding $\alpha|0\rangle + \beta|1\rangle$ as $\alpha|000\rangle + \beta|111\rangle$ just as before and then applying a Hadamard transform on each qubit. The effect of a phase flip on the diagonal basis is similar to the effect of a bit flip on the standard basis.

$$\sigma_z |+\rangle = |-\rangle$$

$$\sigma_z |-\rangle = |+\rangle$$

Therefore Bob can easily correct against phase flips on a single qubit by first applying Hadamard transforms to all three qubits and then correcting as before. For instance if a phase flip happens on the first qubit then encoding

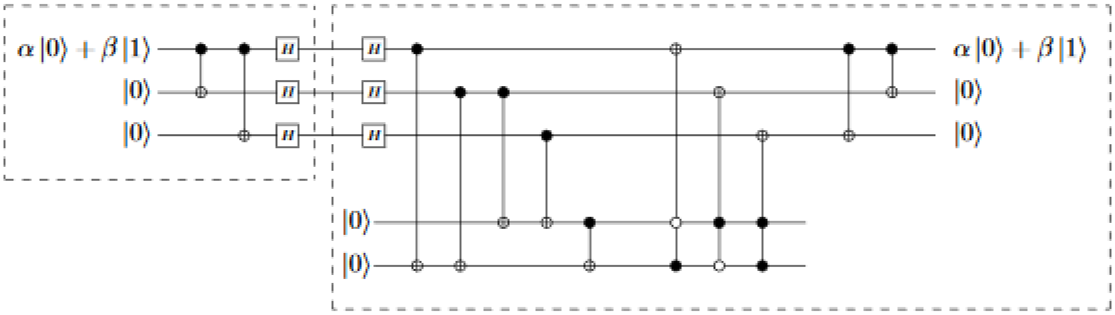
$$\alpha |+\rangle |+\rangle |+\rangle + \beta |-\rangle |-\rangle |-\rangle$$

$$\alpha |-\rangle |+\rangle |+\rangle + \beta |+\rangle |-\rangle |-\rangle$$

Applying Hadamard transforms to all three qubits would then give:

$$\alpha |100\rangle + \beta |011\rangle$$

The same procedure can then be applied to obtain $\alpha |0\rangle + \beta |1\rangle$. This process can be described by the following quantum circuit diagram.

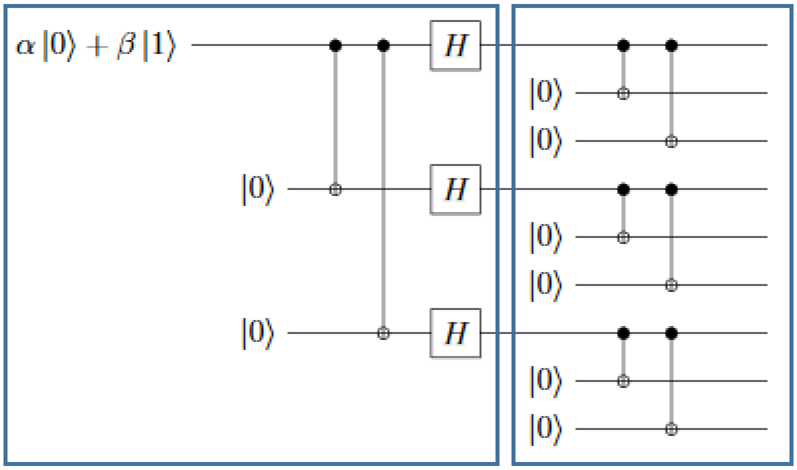


## 6.4 Shor's 9 Qubit Code

Is there anyway to protect against both bit flips and phase flips simultaneously? We can do this by concatenating our two previous codes. We first apply the encoding that protects against phase flips and then encode each of three resulting qubits using the code that protects against bit flips. Mathematically we will encode $\alpha |0\rangle + \beta |1\rangle$ as

$$\alpha \frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}}$$
$$+ \beta \frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}$$

This is known as **Shor's 9 qubit code**. A circuit for the encoding procedure is shown below. Notice that the first block is equivalent to the encoding procedure for phase flips and the second block is equivalent to the encoding procedure for bit flips.



**Example**: Suppose that the error $\sigma_x\sigma_z$ occurs on qubit number four. The encoded state becomes

$$\alpha\frac{(|000\rangle + |111\rangle)(|100\rangle - |011\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}}$$
$$+\beta\frac{(|000\rangle - |111\rangle)(|100\rangle + |011\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}$$

The bit flip code correcting procedure is applied to each block. This result in three 2 bit syndromes which in this case are 00 01 and 00. The first and third blocks have no bit flip errors while the second has a bit flip error in position 1. The encode state after the bit flip correction is

$$\alpha\frac{(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}}$$
$$+\beta\frac{(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}$$

Now the phase flip code correcting procedure is applied and gives a single 2 bit syndrome. In this case it is 10. This implies that a phase flip occurred somewhere in block 2. Applying a $\sigma_z$ gate to any qubit in this block results in

$$\alpha\frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}}$$
$$+\beta\frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}$$

which is the original encoding.