# Quantum Computing


**Abbas Edalat**


**18 lectures + 9 tutorials**

Lecture Notes and Exercise Sheets
available from:

http://www.doc.ic.ac.uk/~ae


**Department of Computing**
**Imperial College London**
**180 Queen's Gate**
**London SW7 2BZ**


## Recommended Texts

- **Textbooks:**
  - Jozef Gruska, *Quantum Computing*, (McGraw-Hill, 1999).
  - Michael A. Nielsen and Issac L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, 2000). Course textbook.
  - N. David Mermin, *Quantum Computer Science*, (Cambridge University Press, 2007)
- **General Books:**
  - Colin P. Williams and Scott H. Clearwater, *Explorations in Quantum Computing*, (Springer-Verlag, 1998).
  - G. Berman, G. Doolen, R. Mainieri and V. Tsifrinovich, *An Introduction to Quantum Computers*, (World Scientific, 1998).
- **Survey Papers:** available on www, e.g. from http://xxx.lanl.gov/find/quant-ph http://www.qubit.org
  - Dorit Aharonov, *Quantum Computation*, December 15, 1998.
  - Arthur Ekert, Patrick Hayden and Hitoshi Inmori, *Basic Concepts in Quantum Computation*, 16 January 2000.
  - Eleanor Rieffel and Wolfgang Polak, *An Introduction to Quantum Computing for Non-Physicists*, 19 January 2000.



## Topics of the Course

- Introduction to Quantum Mechanics
- Quantum Bits and Complex Vector Spaces
- Quantum Evolution and Quantum Gates
- Quantum Registers, Universal Gates
- No-Cloning Theorem
- Quantum Entanglement and Teleportation
- Quantum Algorithms
- Quantum Search
- Quantum Fourier Transform
- Phase Estimation
- Quantum Counting
- Order Finding for Periodic Functions
- Quantum Factoring of Integers
- Physical Realization of Quantum Gates
- Quantum Error Correction



# 1 Introduction

A computer is a physical machine and any computation performed by such a machine is in essence a physical process. This is a simple factual statement but it has a profound consequence. It can be *logically* argued from this premise that:

- the laws of computation depend on the physical laws obeyed by the computer machine under consideration, and,
- there are no absolute laws of computation valid for all computational machines.

The prominent English logician/mathematician, Alan Turing, formulated the classical theory of computation in 1930's. He assumed that computation is performed by an idealized mechanical computer (with potentially infinite storage capacity) obeying the classical laws of physics. This model, now called the Turing model of computation, has proved to be adequate for describing the computational process performed by mechanical or the modern electronic computers.

However, with the current revolution in computer technology a new phenomenon is likely to emerge. In the past few decades, the miniaturization trend in computer manufacturing has followed the so-called Moore's law: Gordon Moore, one of the founders of Intel, observed in mid 1960's that the memory capacity of a typical chip doubles roughly every eighteen months while its physical size remains the same.

If this law, as it is expected, continues to hold in the coming years, the size of the memory component of a chip will, in less than 20 years, be reduced to the size of an atom. This means that in about 2020, a single atom will be sufficient to store a bit (0 or 1) of computer memory. At such atomic scales the laws of classical physics, which are the basis of Turing Machines and the classical theory of computation, collapse. Here, we will enter the realm of quantum physics. Every aspect of computing, including storing information, loading and running of programs and reading the output will be governed by laws of quantum physics which are quite different from those of classical physics.



## 1.1 Electrons: Particles or Waves?

We consider the experiment depicted in Figure 1. The gun in the bottom emits electrons which go through the two slits (slit 1 and slit 2) on the screen. A movable electron detector is placed on a wall behind the screen which counts the number of electrons arriving at the detector. The probability distribution of the electron position on the wall, when slit 2 is closed, is $P_1(x)$. When slit 1 is covered, the probability distribution is $P_2(x)$. So far the result of our experiment is as if we were shooting ordinary particles or bullets.

If indeed we were dealing with normal particles, then we would expect the distribution when both slits are open to be $\frac{1}{2}(P_1(x) + P_2(x))$. However, with electrons we get the distribution $P_{12}(x)$, which shows an interference between the two distributions $P_1(x)$ and $P_2(x)$. In fact, $P_{12}(x)$ oscillates between zero (destructive interference) and $P_1(x)+P_2(x)$ (constructive interference). This is the pattern one expects for waves (e.g. water waves or electromagnetic radiation), not particles.



## A quantum phenomenon

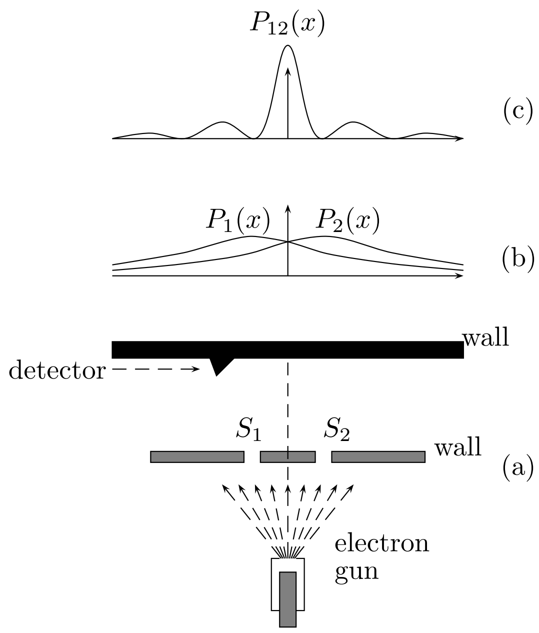



Figure 1: The two slit experiment



How can we understand our experimental results? Probabilities are always nonnegative and cannot cancel each other. This experiment proves therefore that ordinary probabilities are inadequate to describe the position of electrons. It is as if we also require *negative* probabilities. In quantum mechanics the distribution of electrons when slit 2 (respectively, slit 1) is closed is given by the *probability amplitude* $\psi_1(x)$ (respectively, $\psi_2(x)$), where $\psi_i(x)$ (for $i = 1, 2$) is a complex number [a] with $P_i(x) = |\psi_i(x)|^2$.

The distribution of electrons when both slits are open is given by $\psi_{12}(x) = c(\psi_1(x) + \psi_2(x))$, where $c$ is the normalization constant. From this it follows that

$$\begin{aligned} P_{12}(x) &\propto |\psi_1(x) + \psi_2(x)|^2 \\ &= |\psi_1(x)|^2 + |\psi_2(x)|^2 + 2\,\mathrm{Re}(\psi_1^*(x)\psi_2(x)) \\ &= P_1(x) + P_2(x) + 2\,\mathrm{Re}(\psi_1^*(x)\psi_2(x)). \end{aligned}$$

[a] Recall that a complex number is of the form $z = a + ib$ where $i = \sqrt{-1}$ and $a = \mathrm{Re}(z)$, the real part of $z$ and $b = \mathrm{Im}(z)$, the imaginary part of $z$, are real numbers. The *norm* of $z$ is $|z| = \sqrt{a^2 + b^2}$. The *conjugate* of $z$ is $z^* = a - ib$. The set of complex numbers is denoted by $\mathbb{C}$.

We see that the interference, and hence the oscillation, is due to the last term $2\,\mathrm{Re}(\psi_1^*(x)\psi_2(x))$. But there are yet more surprising facts.

Suppose, by placing a light source between the two slits, we try to observe through which slit any electron passes. Then, as expected, the probability distribution of the electron position, when slit 2 (respectively, slit 1), is closed, is $P_1(x)$ (respectively, $P_2(x)$). However, if both slits are open the distribution will be $\frac{1}{2}(P_1(x) + P_2(x))$ i.e. the electrons behave like particles again. Therefore, when we observe the electrons, the interference is completely lost. A quantum system behaves differently when it is observed from when it is not observed!

This is certainly against our intuition. In everyday life, we consider particles and waves as distinct entities. Furthermore, we always assume that the result of an experiment is independent of whether or not we observe how it takes place. Yet, the above experiment shows that electrons behave sometimes like particles and sometimes like waves. This is called *particle-wave duality*.



Furthermore, as we have seen, it does matter to the result of the experiment whether we observe the electrons in the process of experiment or not.

Nobody knows why electrons, or for that matter any quantum phenomenon, behaves like that. Quantum mechanics is a very accurate description of nature as it predicts quantum effects up to an astonishing precision of 14 decimal places. But we do not know why nature works like that and why quantum mechanics gives such a good description of nature. In other words, quantum mechanics tells us how things work but cannot tell us why they work that way.

The conceptual issues concerning quantum mechanics have been intensely debated by physicists as well as philosophers for over seventy years now and there is still no resolution of the controversies in sight. In this course, however, we will not be concerned with quantum mechanics or the philosophical issues surrounding it. We will rather accept the central principles of quantum physics to see how the quantum world and the quantum paradigm can be used for computation.



### First Principles of Quantum Mechanics

(i) The probability $P$ of a quantum event is given by the square $|\psi|^2$ of the absolute value of a complex number $\psi$ called the probability amplitude of the event.

(ii) When an event can take place in $n$ different ways, the probability amplitude is proportional to the sum or superposition of the probability amplitudes of each way considered separately: $\psi = c(\psi_1 + \psi_2 + \psi_3 + \cdots + \psi_n)$, where $c$ is the normalization constant.

(iii) If an experiment is carried out which can determine which of the alternative ways has actually taken place, the interference is lost, i.e. the probability will be proportional to the sum of the probabilities for each alternative way: $P = \frac{1}{n}(P_1 + P_2 + P_3 + \cdots + P_n)$.

A quantum state is thus specified by its probability amplitude. We will now find out how information can be represented by a quantum state.



## 2 Quantum Bits

Consider a two-state quantum system, such as

- the first two energy levels of the Hydrogen atom ($n = 0$ or the ground state and $n = 1$ or the first excited state) as in Figure 2,
- the spin (internal angular momentum) of an electron ($s = 1/2$ or $s = -1/2$), or,
- the vertical or horizontal polarization of light.

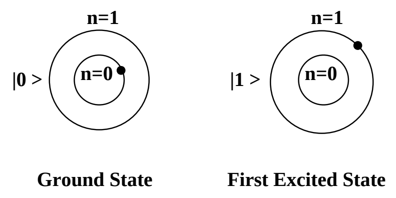



Figure 2: Hydrogen atom: a two-state system

It is convenient to denote the two basic states of the system by $|0\rangle$, denoting e.g. the ground state of the Hydrogen atom and $|1\rangle$, denoting e.g. the first excited state of the Hydrogen atom.

The notation $|.\rangle$ is called a *ket* and was invented by Dirac, the famous English physicist who was a pioneer in quantum theory. Later we will introduce the dual notion of a *bra* $\langle.|$ which can be used to denote expressions such as $\langle 0|0\rangle$ and $\langle 0|1\rangle$, called *brackets*.

Recall that in the two-slit experiment, the probability amplitude of the position of the electron was given by the sum of the probability amplitudes for the position of the electron when each of the slits was covered in turn.

In the same way, the probability amplitude $\psi$ of the Hydrogen atom (or the spin of the electron, etc.) is given by a superposition $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ of the two basic states $|0\rangle$ and $|1\rangle$. Here, $\alpha$ and $\beta$ are complex numbers with $|\alpha|^2 + |\beta|^2 = 1$ such that $|\alpha|^2$ and $|\beta|^2$ give the probabilities that the Hydrogen atom is in its ground and first excited state respectively. It is as if the Hydrogen atom has not yet decided to be in the ground or the first excited state and it is therefore in a superposition of these two states each with a given probability, i.e. it exists partially in both states!



**Definition 2.1** A *qubit* (quantum bit) is a quantum state of the form $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ where $\alpha$ and $\beta$ are complex numbers with $|\alpha|^2 + |\beta|^2 = 1$.

Qubits live in a two-dimensional complex vector space which we will now study.

## 2.1 Complex Vector Spaces

Recall that the two dimensional real vector space $\mathbb{R}^2$ is the collection of column vectors of the form

$$v = \begin{pmatrix} a \\ b \end{pmatrix}$$

where $a$ and $b$ are real numbers. The norm of $v$ is given by $\|v\| = \sqrt{a^2+b^2}$. The *transpose* of $v$ is the row vector $v^T = (a, b)$. The scalar product of two vectors

$$v_1 = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \qquad \text{and} \qquad v_2 = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}$$

is given by $v_1 \cdot v_2 \stackrel{\text{def}}{=} v_1^T v_2 = (a_1, b_1)\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = a_1a_2 + b_1b_2 = \|v_1\|\,\|v_2\|\cos\theta$, where $\theta$ is the angle between $v_1$ and $v_2$. If $v_1 \cdot v_2 = 0$ then the two vectors $v_1$ and $v_2$ are called orthogonal.



A map $L : \mathbb{R}^2 \to \mathbb{R}$ is *linear* if $L(a_1v_1 + a_2v_2) = a_1L(v_1) + a_2L(v_2)$ for all $a_i \in \mathbb{R}$ and $v_i \in \mathbb{R}^2$ $(i = 1, 2)$. For example, projection to the $x$-axis (i.e. $(x, y) \mapsto x$) is a linear map whereas a constant map (i.e. $v \mapsto c$) is not linear. For any $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in \mathbb{R}^2$, the *dual* of $u$ is the linear map $L_u : \mathbb{R}^2 \to \mathbb{R}$ given by $L_u(v) = u \cdot v$. We can identify $L_u$ with the row vector $u^T = (u_1, u_2)$. Then $L_u(v) = u^T v$ is obtained by the usual matrix multiplication.

If $u \in \mathbb{R}^2$ is a unit vector, then

$$L_u(v) = u \cdot v = \|v\| \cos\theta$$

gives the projection of $v$ on the unit vector $u$ as in Figure 3.

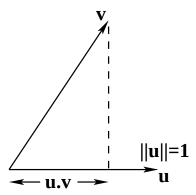



Figure 3: Projection of a vector on a unit vector



We say that the vectors

$$\{v_i \in \mathbb{R}^2 | i = 1, 2, \cdots k\}$$

are *linearly independent* if whenever

$$a_1v_1 + a_2v_2 + \cdots + a_kv_k = 0$$

for $a_i \in \mathbb{R}$ $(i = 1, \cdots, k)$ then $a_i = 0$ for all $i = 1, \cdots, k$. Otherwise they are called linearly dependent. Any three vectors in $\mathbb{R}^2$ are linearly dependent. A basis of $\mathbb{R}^2$ is any set of linearly independent vectors in $\mathbb{R}^2$ such that any vector in $\mathbb{R}^2$ can be expressed as a linear combination of vectors in the set. Any two linearly independent vectors $v_1$ and $v_2$ form a basis for $\mathbb{R}^2$. We say $v_1$ and $v_2$ form an *orthonormal basis* for $\mathbb{R}^2$ if $\|v_1\| = \|v_2\| = 1$ and $v_1 \cdot v_2 = 0$. The two vectors

$$r_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad r_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

form an orthonormal basis for $\mathbb{R}^2$, called the *standard* basis of $\mathbb{R}^2$.

**Exercise 2.2** Extend all the above definitions and properties to $\mathbb{R}^d$. $\square$

Similarly, the two dimensional *complex* vector space $\mathbb{C}^2$ is the collection of column vectors of the form

$$w = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

where $\alpha$ and $\beta$ are complex numbers.

The norm of $w$ is given by $\|w\| = \sqrt{|\alpha|^2 + |\beta|^2}$, where $|\alpha|$ is as before the norm of $\alpha$.

The linear structure of $\mathbb{C}^2$ is similar to that of $\mathbb{R}^2$. Vector addition and scalar multiplication by a complex number are given by:

$$\begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 \\ \beta_1 + \beta_2 \end{pmatrix} \qquad c\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} c\alpha \\ c\beta \end{pmatrix}.$$

The *adjoint* of $w$ is defined as the row vector $w^\dagger = (\alpha^*, \beta^*)$. In analogy with the scalar product of two real vectors, we define the scalar product or *inner product* of two complex vectors

$$w_1 = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} \qquad w_2 = \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}$$

as

$$(w_1, w_2) \stackrel{\text{def}}{=} w_1^\dagger w_2 = (\alpha_1^*, \beta_1^*) \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} = \alpha_1^* \alpha_2 + \beta_1^* \beta_2.$$



The vector space $\mathbb{C}^2$ with its inner product is called the two-dimensional *Hilbert Space.*

For any $u \in \mathbb{C}^2$, the dual of $u = \binom{u_1}{u_2}$ is the linear map $L_u : \mathbb{C}^2 \to \mathbb{C}$ defined by $L_u(w) = (u, w)$. We can identify $L_u$ with $u^\dagger$. Then $L_u(w) = (u, w) = u^\dagger w$. In the more abstract language, the dual of $|\psi\rangle$ is denoted by $\langle\psi|$, and the inner product of $|\psi\rangle$ and $|\phi\rangle$ by $\langle\psi|\phi\rangle$. The notation $\langle .|$ is called a *bra* and $\langle .|.\rangle$ a *bracket.* If $u \in \mathbb{C}^2$ has $\|u\| = 1$, then $L_u(w) = \langle u|w\rangle$ is the projection of $w$ in the direction of $u$.

We say that the vectors $\{w_i \in \mathbb{C}^2 | i = 1, 2, \cdots k\}$ are linearly independent if whenever

$$a_1 w_1 + a_2 w_2 + \cdots + a_k w_k = 0$$

for $a_i \in \mathbb{C}$ $(i = 1, \cdots, k)$ then $a_i = 0$ for all $i = 1, \cdots, k$. Otherwise they are called linearly dependent. Any three vectors in $\mathbb{C}^2$ are linearly dependent. A basis of $\mathbb{C}^2$ is a set of linearly independent vectors in $\mathbb{C}^2$ such that any vector in $\mathbb{C}^2$ can be expressed as a linear combination of the vectors in the set. Any two linearly independent vectors $w_1$ and $w_2$ form a basis for $\mathbb{C}^2$.



We say $w_1$ and $w_2$ form an orthonormal basis for $\mathbb{C}^2$ if $\|w_1\| = \|w_2\| = 1$ and $(w_1, w_2) = 0$.

The two vectors

$$s_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad s_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

form an orthonormal basis for $\mathbb{C}^2$.

**Exercise 2.3** Extend all the above definitions and properties to $\mathbb{C}^d$. $\square$

## 2.2 Qubits as Complex Unit Vectors

The qubit $\alpha|0\rangle + \beta|1\rangle$ can also be written in the vector notation,

$$\alpha|0\rangle + \beta|1\rangle = \alpha\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

where we have identified $|0\rangle$ and $|1\rangle$ respectively with the orthonormal basis vectors, $\binom{1}{0}$ and $\binom{0}{1}$.

In quantum computing, the basis consisting of $|0\rangle$ and $|1\rangle$ is called the *computational basis.* Since $|\alpha|^2 + |\beta|^2 = 1$, a qubit can therefore be identified with a normalized vector in $\mathbb{C}^2$.



**Exercise 2.4** Show that up to multiplication by a complex number with norm one, a qubit can be written in the form $\cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$ where $\theta$ and $\phi$ are real numbers. $\square$

Note a fundamental difference with the two slit experiment, where the position $x$ of the electron hitting the wall behind the screen is a continuous variable and hence the probability amplitudes $\phi_i$ are functions of a continuous variable.

In a two-state quantum system, however, there are only two basic states namely $|0\rangle$ and $|1\rangle$.

Of course, by Exercise 2.4 we know that a qubit has two continuous degrees of freedom provided by the real numbers $\theta$ and $\phi$. It may seem therefore that we can store an infinite amount of information, in fact the entire human knowledge, in a qubit $\cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$ by using the continuum range of $\theta$.

But the following quantum principle of measurement and observation shows that this is an illusion.

### 2.3 Basic Measurement Principle

Whenever a qubit $\alpha|0\rangle + \beta|1\rangle$ is measured in the computational basis $\{|0\rangle, |1\rangle\}$, the result of the observation is either the state $|0\rangle$ with probability $|\alpha|^2$ or the state $|1\rangle$ with probability $|\beta|^2$. Furthermore, the measurement process changes the state of the system, which collapses from the superposition $\alpha|0\rangle+\beta|1\rangle$ into the observed state (i.e. $|0\rangle$ or $|1\rangle$).

This means that as soon as we make a measurement, the quantum system will decide, with the appropriate probability, which state of the computational basis it wishes to turn to and it will immediately collapse into that basis state.

Hence from a single observation we can only obtain a single bit and all the information about $\alpha$ and $\beta$ are lost forever. Only if we have an infinite number of identical states $\cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$ and make an infinite number of observations, can we find the values of $\cos(\theta/2)$ and $\sin(\theta/2)$ whose squares are the probabilities of obtaining the basis vectors $|0\rangle$ and $|1\rangle$ respectively.



**Exercise 2.5** (i) Show that there is a one to one correspondence between qubits

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

and the points on the unit sphere in $\mathbb{R}^3$, called the *Bloch sphere*, with $\theta$ and $\phi$ as the spherical coordinates of a point of the sphere; see Figure 4.

(ii) Show that the two vectors $w_1$ and $w_2$ in $\mathbb{C}^4$ with $w_1^T = \frac{1}{2}(1,1,1,1)$ and $w_2^T = \frac{1}{2}(1,-1,1,-1)$ are orthogonal unit vectors. Find vectors $w_3$ and $w_4$ such that the collection $\{w_1, w_2, w_3, w_4\}$ forms an orthonormal basis for $\mathbb{C}^4$. □

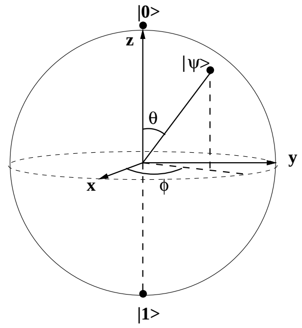



Figure 4: The Bloch sphere



### 2.4 Change of Basis

Any orthonormal basis of $\mathbb{C}^2$ can be regarded as a *computational basis.*

For example, the two qubits $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ also form a computational basis.

We have:

$$|0\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) \qquad |1\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle).$$

Hence, an arbitrary qubit

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

can be expressed in the basis $|+\rangle$, $|-\rangle$ as:

$$\begin{aligned}\alpha|0\rangle + \beta|1\rangle &= \tfrac{\alpha}{\sqrt{2}}(|+\rangle + |-\rangle) + \tfrac{\beta}{\sqrt{2}}(|+\rangle - |-\rangle) \\ &= \tfrac{\alpha+\beta}{\sqrt{2}}|+\rangle + \tfrac{\alpha-\beta}{\sqrt{2}}|-\rangle.\end{aligned}$$

It is in principle possible to make measurements with respect to other computational bases. A qubit which is measured with respect to the computational basis $|+\rangle$ and $|-\rangle$ will collapse to one of these vectors.



## 3 Quantum Gates

Classical circuits consist of wires which carry information (0 or 1) and gates which apply elementary operations on bits. The input to a circuit is a tuple of bits and the output is another such tuple. For example there is only one non-trivial classical gate with one input and one output bit, namely the NOT gate with $0 \mapsto 1$ and $1 \mapsto 0$.

A *quantum gate* is the quantum analogue of a classical gate. It takes a tuple of qubits and outputs another such tuple. Let's see how to define a quantum NOT gate on a single qubit with $|0\rangle \mapsto |1\rangle$ and $|1\rangle \mapsto |0\rangle$. We need to define the action of the gate on the general superposition $\alpha|0\rangle + \beta|1\rangle$. Clearly the easiest solution is to define this action as a linear map $\alpha|0\rangle+\beta|1\rangle \mapsto \beta|0\rangle+\alpha|1\rangle$.

It turns out that this is exactly how we should define the NOT gate. In fact, any quantum system, such as the input to a quantum gate, evolves according to a linear map which can be represented by matrices.

## 3.1 Representation of Linear Maps

A map $L : \mathbb{C}^2 \to \mathbb{C}^2$ is *linear* if $L(a_1 v_1 + a_2 v_2) = a_1 L(v_1) + a_2 L(v_2)$ for all $a_i \in \mathbb{C}$ and $v_i \in \mathbb{C}^2$ $(i = 1, 2)$.

**Exercise 3.1** (i) Show that composition of two linear maps is linear.

(ii) Show that if $L_1, L_2 : \mathbb{C}^2 \to \mathbb{C}^2$ are linear, so is $\alpha_1 L_1 + \alpha_2 L_2 : \mathbb{C}^2 \to \mathbb{C}^2$ defined by

$$(\alpha_1 L_1 + \alpha_2 L_2)(v) = \alpha_1 L_1(v) + \alpha_2 L_2(v).$$

(iii) Show that for $|\psi\rangle, |\phi\rangle \in \mathbb{C}^2$, the map $|\psi\rangle\langle\phi| : \mathbb{C}^2 \to \mathbb{C}^2$ defined, using the bracket notation, by $|\psi\rangle\langle\phi|(|x\rangle) = \langle\phi|x\rangle|\psi\rangle$ is linear. □

Suppose $L(|0\rangle) = a_{11}|0\rangle + a_{21}|1\rangle$ and $L(|1\rangle) = a_{12}|0\rangle + a_{22}|1\rangle$. Then we can write $L$ as

$$L = a_{11}|0\rangle\langle 0| + a_{21}|1\rangle\langle 0| + a_{12}|0\rangle\langle 1| + a_{22}|1\rangle\langle 1|.$$

The matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

is called the *matrix representation* of $L$ in the computational basis $|0\rangle$ and $|1\rangle$.



The action of $L$ on any qubit vector $\binom{\alpha}{\beta}$ is obtained by matrix multiplication:

$$L(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a_{11}\alpha + a_{12}\beta \\ a_{21}\alpha + a_{22}\beta \end{pmatrix}.$$

A change in the basis is also represented by a matrix as follows. Suppose $|\psi_1\rangle$ and $|\psi_2\rangle$ form a basis with $|0\rangle = b_{11}|\psi_1\rangle + b_{21}|\psi_2\rangle$ and $|1\rangle = b_{12}|\psi_1\rangle + b_{22}|\psi_2\rangle$. Then the matrix

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

gives the linear map for the change of coordinates. The vector $\binom{\alpha}{\beta}$, represented with respect to the basis $|0\rangle$ and $|1\rangle$, will have coordinates $B\binom{\alpha}{\beta}$ with respect to the basis $|\psi_1\rangle$ and $|\psi_2\rangle$.

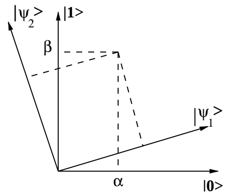



Figure 5: Change of basis



**Example 3.2** The matrix corresponding to the change of basis from $|0\rangle, |1\rangle$ to $|+\rangle, |-\rangle$ is given by

$$B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \ \square$$

What will the matrix representation of $L$ be in the new basis? If $\binom{\delta}{\lambda}$ gives the coordinates of a vector in the new basis, then $B^{-1}\binom{\delta}{\lambda}$ gives the coordinates of the vector in the old basis $|0\rangle$ and $|1\rangle$, where $B^{-1}$ is the matrix inverse of $B$. Hence the coordinates of the transformed vector in the old basis is $AB^{-1}\binom{\delta}{\lambda}$ and hence the coordinates in the new basis is $BAB^{-1}\binom{\delta}{\lambda}$. We conclude that the matrix representation of $L$ in the new basis is $BAB^{-1}$.

**Exercise 3.3** Show that the matrix representation of the NOT gate in the basis $|0\rangle, |1\rangle$ is given by $X = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. What is its matrix representation in the basis $|+\rangle, |-\rangle$? □

**Exercise 3.4** Extend the above notions and properties to $\mathbb{C}^d$. □



## 3.2 Unitary Matrices

We have seen that the evolution of quantum systems, such as the evolution of an input qubit through a quantum circuit to an output qubit, is governed by linear maps. Thus, we need to consider those linear maps which map qubits to qubits, i.e., unit vectors to unit vectors.

For any $n \times m$ matrix $A$ the *transpose* $A^T$ is defined by $(A^T)_{ij} = (A)_{ji}$, the *conjugate* $A^*$ of $A$ by $(A^*)_{ij} = (A_{ij})^*$ and the *adjoint* $A^\dagger$ of $A$ by $A^\dagger = (A^T)^*$. A matrix $A$ is called *unitary* if $A^\dagger = A^{-1}$. A linear map is *unitary* if it has a unitary matrix representation in some computational basis.

**Exercise 3.5** (i) Check that a matrix is unitary iff its columns (or rows) form an orthonormal basis.

(ii) Check that the matrix for a change in the computational basis is unitary.

(iii) Show that $(AB)^\dagger = B^\dagger A^\dagger$ and deduce that if a linear map has a unitary matrix representation, then its matrix representation in any computational basis is unitary. □

**Theorem 3.6** *A linear function maps a qubit to a qubit (i.e. it preserves normalized vectors) iff it is unitary.*

**Proof** Consider a linear map $L : \mathbb{C}^2 \to \mathbb{C}^2$ with a unitary matrix representation $M = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ in the computational basis $|0\rangle$ and $|1\rangle$. Then

$$L(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a_{11}\alpha + a_{12}\beta \\ a_{21}\alpha + a_{22}\beta \end{pmatrix}.$$

Hence, $\|L(\begin{pmatrix} \alpha \\ \beta \end{pmatrix})\|^2 = |\alpha|^2(|a_{11}|^2 + |a_{21}|^2) + |\beta|^2(|a_{12}|^2 + |a_{22}|^2) + (a_{11}^*a_{12} + a_{21}^*a_{22})\alpha^*\beta + (a_{11}a_{12}^* + a_{21}a_{22}^*)\alpha\beta^* = 1$, since the columns of $M$ form an orthonormal basis. Conversely, if $L$ preserves unit vectors, then by taking $\beta = 0$ (respectively, $\alpha = 0$) we obtain $|a_{11}|^2 + |a_{21}|^2 = 1$ (respectively, $|a_{12}|^2 + |a_{22}|^2 = 1$). From these, we also obtain $a_{11}^*a_{12} + a_{21}^*a_{22} = 0$ (why?). □

**Exercise 3.7** Show that

$$\langle u|Av\rangle = \langle A^\dagger u|v\rangle.$$

Deduce that a matrix $M$ is unitary iff it preserves all inner products, i.e. iff $\langle Mu|Mv\rangle = \langle u|v\rangle$ for all $u, v \in \mathbb{C}^2$. □



The NOT gate has the unitary matrix $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

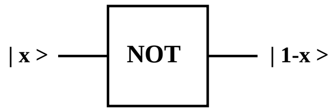



Figure 6: The NOT gate

The matrix $X$ and the two unitary matrices

$$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

are called the *Pauli* matrices. They give, respectively, the $x$, $y$ and $z$ components of the spin of the electron.

**Exercise 3.8** Show that any unitary matrix $U$ can be expressed as

$$U = \begin{pmatrix} e^{i(\alpha-\beta/2-\delta/2)}\cos\gamma/2 & -e^{i(\alpha-\beta/2+\delta/2)}\sin\gamma/2 \\ e^{i(\alpha+\beta/2-\delta/2)}\sin\gamma/2 & e^{i(\alpha+\beta/2+\delta/2)}\cos\gamma/2 \end{pmatrix}$$

where $\alpha$, $\beta$, $\delta$ and $\gamma$ are real numbers. □

It turns out that unitarity is the only requirement on linear maps for quantum evolution. Therefore, any unitary linear map defines a valid single qubit quantum circuit.



## 3.3 Quantum logic gates

**Definition 3.9** A *quantum logic gate* is a device that carries out a given unitary operation on its input qubits in a fixed period of time. A *quantum network* is a device consisting of a number of quantum logic gates whose computational steps are synchronized in time.

The most important of such gates are the following:

$$\begin{aligned} X &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} && \text{Pauli-}X\text{ gate} \\ Y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} && \text{Pauli-}Y\text{ gate} \\ Z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} && \text{Pauli-}Z\text{ gate} \\ H &= \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} && \text{Hadamard gate} \end{aligned}$$

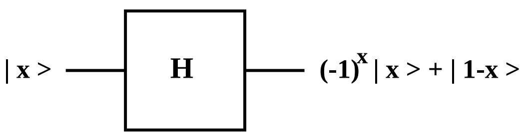



Figure 7: The Hadamard gate



$$\Phi = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} \qquad \text{the phase gate}$$

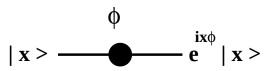
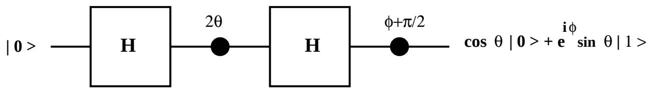



Figure 8: The Phase gate

**Exercise 3.10** Verify the output, up to a global phase, of the following quantum network. □




It follows from Exercise 3.10 that the Hadamard and phase gates are sufficient to construct any unitary operation on a single qubit.

Note that a fundamental property of quantum logic gates is that they are invertible as any unitary map has a unitary inverse which is also a valid quantum gate.

This is in contrast to classical logic gates, such as the AND gate, which may not be reversible.

# 4 Quantum Registers

Consider two qubits, each represented say by the ground state and the first excited state of a Hydrogen atom. Since we have a pair of two-state systems, classically there are four states, namely 00, 01, 10 and 11. Therefore the quantum system representing the two qubits is a superposition of the form:

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$$

with $\sum_{i\in\{0,1\}^2} |\alpha_i|^2 = 1$. This is called a *two-qubit quantum register*. Similar to the case of a single qubit, any $i \in \{0,1\}^2$ in this quantum register is measured with probability $|\alpha_i|^2$. We can however, make a measurement of a subset of the bits. For example if the first qubit is measured then one obtains 0 with probability $|\alpha_{00}|^2 + |\alpha_{01}|^2$ and the state then collapses to

$$\frac{\alpha_{00}|00\rangle + \alpha_{01}|01\rangle}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}}.$$

Note that the state is renormalized to have unit length.



A single qubit, as we have seen, lives in the two-dimensional complex (Hilbert) space. Where does a two-qubit quantum register lie? There are four basic two-qubit states

$$|00\rangle \qquad |01\rangle \qquad |10\rangle \qquad |11\rangle.$$

Hence, we can expect a general two-qubit quantum register be a normalized vector in $\mathbb{C}^4$, the four dimensional complex vector space. In fact, $\mathbb{C}^4$ can be constructed from $\mathbb{C}^2$ as follows.

## 4.1 Tensor Product

We will define a general construction, called the tensor product, on finite dimensional complex vector spaces. Note that, for any positive integer $m$, the $m$-dimensional complex vector space $\mathbb{C}^m$ has the standard basis

$$b_1^m, b_2^m, \cdots, b_m^m,$$

where $b_j^m$ is the column vector of dimension $m$ with all entries zero except for the $j$th entry which is one.



A vector $u \in \mathbb{C}^m$ can be written as $u = \sum_{j=1}^m u_j b_j^m$ where $u_j \in \mathbb{C}$ is the $j$th component of $u$:

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_j \\ \vdots \\ u_m \end{pmatrix}$$

Consider the complex vector spaces $\mathbb{C}^k$ and $\mathbb{C}^l$. We define the *tensor* product as a map

$$- \otimes - : \mathbb{C}^k \times \mathbb{C}^l \to \mathbb{C}^{kl}$$

with

$$v \otimes w = \begin{pmatrix} v_1 w \\ \vdots \\ v_j w \\ \vdots \\ v_k w \end{pmatrix}$$

where $v_j w$ $(1 \le j \le k)$ is the scalar multiplication of $v_j \in \mathbb{C}$ with the column vector $w \in \mathbb{C}^l$.



For $k = l = 2$, we have the tensor products $|i\rangle \otimes |j\rangle$ $(i, j = 0, 1)$ which we simply write, for convenience, as $|i\rangle|j\rangle$ or even as $|ij\rangle$, $(i, j = 0, 1)$ i.e. $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$.

We also note that the lexicographical order is the same as the numerical order when $ij$ is read as a binary number.

We can, for convenience, write the states in decimal notation $(00 \equiv 0, \quad 01 \equiv 1, \quad 10 \equiv 2, \quad 11 \equiv 3)$ to obtain:

$$|0\rangle \equiv |00\rangle \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad |1\rangle \equiv |01\rangle \equiv \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$|2\rangle \equiv |10\rangle \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad |3\rangle \equiv |11\rangle \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

We can therefore write the general two-qubit quantum register as

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle + \alpha_2|2\rangle + \alpha_3|3\rangle.$$

**Exercise 4.1** (i) Show the bilinearity property: $(\alpha v + \alpha' v') \otimes (\beta w + \beta' w') =$ $\alpha\beta v \otimes w + \alpha\beta' v \otimes w' + \alpha'\beta v' \otimes w + \alpha'\beta' v' \otimes w'$, where $\alpha, \alpha', \beta, \beta' \in \mathbb{C}$, $v, v' \in \mathbb{C}^k$, $w, w' \in \mathbb{C}^l$.

(ii) Show that $b_i^k \otimes b_j^l = b_{(i-1)l+j}^{kl}$.

(iii) Show that for $v, v' \in \mathbb{C}^k$ and $w, w' \in \mathbb{C}^l$ we have: $\langle v \otimes w | v' \otimes w' \rangle = \langle v | v' \rangle \langle w | w' \rangle$. □

In a similar way, we can define a quantum register of $n$-qubits which is an element of the $2^n$-dimensional complex vector (Hilbert) space $\mathbb{C}^{2^n}$.

A computational basis of $\mathbb{C}^{2^n}$ is given by the $2^n$ $n$-qubits

$$|i_1\rangle \otimes |i_2\rangle \otimes \cdots \otimes |i_n\rangle,$$

with $i_j \in \{0,1\}$ for $1 \le j \le n$.

We write this basis vector, for convenience, as $|i_1\rangle|i_2\rangle\cdots|i_n\rangle$ or simply as $|i_1 i_2 \cdots i_n\rangle$.

This basis vector can also be identified with the $2^n$ column vector in $\mathbb{C}^{2^n}$ all whose entries are zero except for the entry $i_1 i_2 \cdots i_n$ (read as an integer in the binary notation) which is 1.



## 4.2 Tensor Product of Matrices

The tensor product can be extended to matrices. If we have two linear operators with matrix representations

$$M : \mathbb{C}^k \to \mathbb{C}^k \qquad N : \mathbb{C}^l \to \mathbb{C}^l,$$

with respect to the standard bases of $\mathbb{C}^k$ and $\mathbb{C}^l$, we obtain a linear operator on $\mathbb{C}^{kl}$ with matrix representation

$$M \otimes N : \mathbb{C}^{kl} \to \mathbb{C}^{kl},$$

called the *tensor product* of $M$ and $N$, defined by:

$$M \otimes N = \begin{pmatrix} M_{11}N & M_{12}N & \cdots & M_{1k}N \\ M_{21}N & M_{22}N & \cdots & M_{2k}N \\ \vdots & \vdots & \vdots & \vdots \\ M_{k1}N & M_{k2}N & \cdots & M_{kk}N \end{pmatrix}$$

where, $M_{ij}$ is the $ij$ element of $M$ and each $M_{ij}N$ is the $l \times l$ matrix obtained by multiplying the matrix $N$ with the complex number $M_{ij}$. Thus, each block $M_{ij}N$ is an $l \times l$ matrix, so that $M \otimes N$ is a $kl \times kl$ matrix.



**Example 4.2** We have the following:

$$\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 3 \\ -1 & 2 & -3 & 6 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 2 \end{pmatrix} \square$$

**Exercise 4.3** (i) Show that:

$$(M \otimes N)(v \otimes w) = (Mv) \otimes (Nw).$$

(ii) Show the bilinearity property:

$$(\alpha M + \alpha' M') \otimes (\beta N + \beta' N') =$$

$$\alpha\beta M \otimes N + \alpha\beta' M \otimes N' + \alpha'\beta M' \otimes N + \alpha'\beta' M' \otimes N'.$$

(iii) Show that:

$$(M \otimes N)(M' \otimes N') = (MM') \otimes (NN').$$

(iv) Show that

$$\begin{aligned} (M \otimes N)^* &= M^* \otimes N^*, \\ (M \otimes N)^T &= M^T \otimes N^T \\ (M \otimes N)^\dagger &= M^\dagger \otimes N^\dagger. \end{aligned}$$

(v) Show that $M \otimes N$ is unitary (invertible) if $M$ and $N$ are unitary (invertible). □

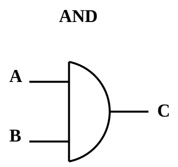


## 4.3 Two Qubit Gates

The Hadamard gate and the phase gate can be used, by Exercise 3.10, to transform the input state $|0\rangle|0\rangle$ into $|\psi_1\rangle|\psi_2\rangle$ where $|\psi_1\rangle$ and $|\psi_2\rangle$ are arbitrary qubits. However, these two-qubit quantum registers are special in that they are the tensor product of single qubits. They are called *separable* states. If a quantum register is not separable, then it is called *entangled*. For example, the state $\alpha|00\rangle + \beta|11\rangle$ (for $\alpha \ne 0 \ne \beta$) is entangled as it cannot be written as a tensor product.

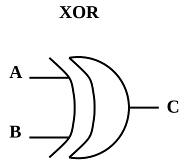
In order to produce entangled two-qubit quantum registers, we need to go beyond single qubit quantum gates. Multi-qubit circuits, like single qubit quantum gates, correspond to unitary operations on multi-qubit quantum registers.

Classically, the most important multi-bit gates are AND (Figure 9), XOR (exclusive-OR, Figure 10) and NAND (Figure 11). The NOT and the AND gates are a *universal* set of gates, which means that any Boolean function can be realized from some composition of these two gates.



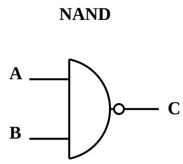

| A | B | C |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

Figure 9: The AND gate




| A | B | C |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

Figure 10: The XOR gate




| A | B | C |
|---|---|---|
| 0 | 0 | 1 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

Figure 11: The NAND gate



In fact, the NAND gate which is the AND gate followed by the NOT gate is itself universal.

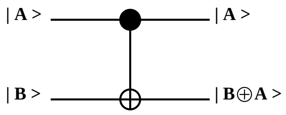
Note that XOR even with the NOT gate is not universal: any circuit involving these two gates only will either always preserve or always change the total parity of bits which restricts the class of functions representable by them.

The quantum analogue of the XOR gate is the CNOT or the *controlled*-NOT gate as in Figure 12. It has two input qubits: the first (top line in the figure) is called the *controlled* qubit, the second (bottom line) is the *target* qubit. If the controlled qubit is zero then the target qubit is intact. If the controlled qubit is one then the target qubit is flipped.

$$|00\rangle \mapsto |00\rangle,\ |01\rangle \mapsto |01\rangle,\ |10\rangle \mapsto |11\rangle,\ |11\rangle \mapsto |10\rangle.$$

This can be written as $|A, B\rangle \mapsto |A, B \oplus A\rangle$, where $B \oplus A$ is addition modulo 2, i.e. the action of XOR. Recall that for positive integers $x$ and $n$, the remainder of division of $x$ by $n$ is called $x$ modulo $n$ and is written $x \bmod n$.






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

Figure 12: The CNOT gate

**Exercise 4.4** Check that the matrix representing CNOT in the standard computational basis is the one given in Figure 12 and that it is unitary. □

Note that CNOT, like any unitary map, is an invertible transformation: given the output qubits $A$ and $A \oplus B$, we can obtain $B$ and hence the input qubits. The same is not true of the classical XOR or the NAND gates. Given $A \oplus B$ as the output of XOR, one cannot retrieve $A$ and $B$, which means that there is a loss of information in this irreversible classical gate.



**Exercise 4.5** Show that the circuit on the left in Figure 13, swaps the two qubits. It is denoted by the figure on the right. □

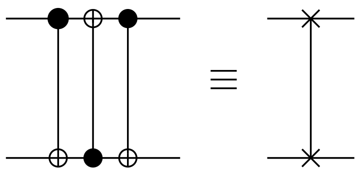


Figure 13: Swapping Gate

**Exercise 4.6** There is nothing special about the control bit in the CNOT gate to take value 1. The control gate in Figure 14 flips the target bit if the control bit is 0. This is represented pictorially by a hollow circle instead of the filled circle on the top line. What is its matrix representation? □

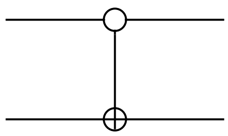


Figure 14: A variation of CNOT

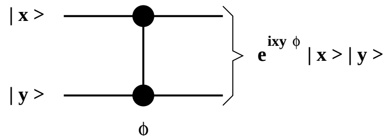



Figure 15: The controlled phase gate

Another useful two-qubit gate is the *controlled phase* gate $B(\phi)$, Figure 15, with matrix:

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

More generally, the *controlled-U transform* for any single qubit unitary transform $U$ maps $|0\rangle|y\rangle$ to $|0\rangle|y\rangle$ and $|1\rangle|y\rangle$ to $|1\rangle(U|y\rangle)$; see Figure 16.

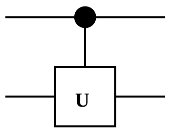



Figure 16: The controlled-$U$ gate



## 4.4 Classical Computations

A *quantum computer* is a family of quantum networks and a *quantum computation* is a unitary evolution of the network which sends any input qubits into some output qubits.

We will show that we can simulate classical gates with quantum circuits. We have already seen that unlike classical gates which may be irreversible (such as AND, XOR, NAND), quantum gates are always unitary and therefore reversible. We can make any irreversible classical gate into an equivalent but reversible gate by using the *Toffoli* gate, a reversible classical gate as in Figure 17.

| IN | | | OUT | | |
|---|---|---|---|---|---|
| a | b | c | a' | b' | c' |
| 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | 1 |
| 0 | 1 | 0 | 0 | 1 | 0 |
| 0 | 1 | 1 | 0 | 1 | 1 |
| 1 | 0 | 0 | 1 | 0 | 0 |
| 1 | 0 | 1 | 1 | 0 | 1 |
| 1 | 1 | 0 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 0 |

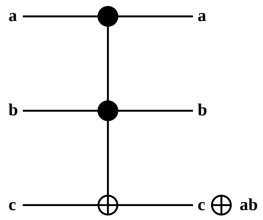



Figure 17: The Toffoli gate



The Toffoli gate can simulate a *reversible* NAND: by choosing $c = 1$, we get $a' = a$, $b' = b$ and $c' = 1 \oplus ab = \neg ab$. It can also simulate a FANOUT, Figure 18, which makes a copy of the input bit $a$.

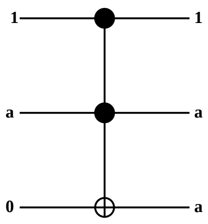



Figure 18: FANOUT using Toffoli

With NAND and FANOUT it is possible to construct any classical gate. The Toffoli gate can thus be used to construct, in a reversible way, any classical gate $f$. For this, as in NAND and FANOUT, we use some *ancilla* bits 0 or 1 as input to the Toffoli gate. We also have some *garbage* in the output which is not needed in the rest of computation. The ancilla and garbage bits are only important to make the computation reversible.



We can represent this by

$$(x, 0, 0) \mapsto (x, f(x), g(x)) \qquad (1)$$

where we have assumed that with the aid of the NOT gate, we can assume all ancilla bits are 0. Finally, we would like the garbage bits be in a standard state. This is done by adding a fourth register in an arbitrary state $y$ so that the effect of the reversible circuit so far is $(x, 0, 0, y) \mapsto (x, f(x), g(x), y)$. Next, we use the CNOT gate to induce $(x, f(x), g(x), y) \mapsto (x, f(x), g(x), y \oplus f(x))$. Finally we apply the inverse of the circuit 1 to the first three registers to obtain:

$$(x, f(x), g(x), y \oplus f(x)) \mapsto (x, 0, 0, y \oplus f(x)).$$

Deleting the ancilla bits, the overall evaluation is: $(x, y) \mapsto (x, y \oplus f(x))$, which we regard as the *standard reversible circuit* for evaluating $f$.

**Exercise 4.7** Check that the Toffoli gate transformation is unitary, and thus can be implemented as a quantum gate. □

Hence, using the quantum Toffoli gate we can simulate any classical circuit.

## 4.5 Randomized algorithms

In 1970's, Solavay and Strassen formulated a *randomized* algorithm to determine if an integer is prime or not. The algorithm, which uses a fair coin or a random number generator, succeeds to give the correct result only up to a certain probability. However, if it is repeated several times, one can determine with high probability if the number in question is prime or not. There is no efficient deterministic algorithm for this task.

Quantum circuits can also simulate randomized algorithms, i.e. algorithms which use a random bit in order to carry out a computation. In fact, no classical computer can generate a truly random number. But this is indeed possible in quantum computation.

A Hadamard gate with input $|0\rangle$ produces the output $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ which, when measured, gives one of the bits 0 and 1 each with probability $\frac{1}{2}$. Therefore the Hadamard gate can be used as a fair coin in order to simulate randomized algorithms.



## 4.6 A Universal Set of Matrices

The Hadamard gate, all phase gates and the CNOT gate form an *infinite universal set of gates*: more precisely any $n$-qubit unitary operation can be simulated exactly with $O(4^n n)$ such gates. Recall that $f(n) = O(g(n))$ means "$f(n) \leq kg(n)$ for some constant $k > 0$ for sufficiently large $n$".

Here we will show that the single qubit gates and the CNOT gate form an infinite universal set for quantum computation. First we prove that two-level unitary matrices, i.e. unitary matrices which act non-trivially on two or fewer components, are universal.

We first establish the result for $n = 3$; the general case is similar. Let

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

be a unitary matrix. We construct two-level unitary matrices $U_1$, $U_2$ and $U_3$ such that $U_3U_2U_1U = I$ where $I$ is the identity matrix.



It then follows that

$$U = U_1^\dagger U_2^\dagger U_3^\dagger,$$

which concludes the construction since the inverse of a two-level unitary matrix is another two-level unitary matrix.

The construction of $U_1$ is as follows.

If $b = 0$, let $U_1 = I$;

otherwise let

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

Then $U_1$ is a two-level matrix and

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

for some primed entries $a', b', c', d'$ etc., whose exact values are unimportant. The important point is the zero entry in the first column of $U_1U$.



Next, we construct $U_2$.

If $c' = 0$ (which implies $|a'| = 1$ by unitarity) put

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

otherwise, let

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

Then $U_2$ is a two-level unitary matrix and

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

for some double primed entries $a'', b'', c''$ etc., whose values are again unimportant. Since $U_2U_1U$ is unitary we must have $d'' = g'' = 0$ as the first row must have norm 1.

We finally put

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

This gives us $U_3U_2U_1U = I$ as required.

We now consider the general case.

**Theorem 4.8** *Any unitary matrix on $\mathbb{C}^d$ can be written as the composition of at most $d(d-1)/2$ two-level unitary matrices.*

**Proof** For the general case of $d$ dimensional unitary matrix $U$, we can find unitary two-level matrices $U_1, U_2, \cdots, U_{d-1}$ such that $U_{d-1}U_{d-2}\cdots U_1U$ has a one in the top left corner and zeros everywhere else in the first row and first column. Then we apply the same algorithm to the $(d-1)\times(d-1)$ matrix in the lower right corner of $U_{d-1}U_{d-2}\cdots U_1U$ etc. It follows that $U$ can be expressed as $U = V_1V_2\cdots V_k$ where the matrices $V_j$ are two-level unitary with $k \le (d-1)+(d-2)+\cdots+2+1 = d(d-1)/2$ □



Next we show that any two-level unitary matrix $U$, acting on $n$-qubits, can be implemented using single qubit gates and the $\mathsf{C}^{n-1}(\mathsf{X})$ gate. Suppose, first that the two non-trivial levels of $U$ are the last two rows and columns. This means that $U$ acts non-trivially only on the last qubit, when all other qubits are 1. Let $\hat{U}$ be the bottom right two-by-two matrix in $U$. Then $U$ can be simply implemented by a controlled-$\hat{U}$ gate conditioned on all but the last qubit to be equal to 1.

In the general case, assume that $U$ acts non-trivially on the computational basis $|s\rangle$ and $|t\rangle$ where $s = s_1s_2\cdots s_n$ and $t = t_1t_2\cdots t_n$. Again let $\hat{U}$ be the non-trivial submatrix of $U$ corresponding to the subspace generated by $|s\rangle$ and $|t\rangle$.

All we need to do is to find a change in the computational basis so that the non-trivial two levels correspond to a single qubit. Recall from Section 3.1 that the matrix representation in the new basis is $V = BUB^{-1}$ where $B$ is the matrix for changing the basis. From this we obtain $U = B^{-1}VB$. Our job is therefore to obtain $B$ and its inverse.



This is most easily done using *Gray Codes*. A Gray code connecting the two binary numbers $s = s_1s_2\cdots s_n$ and $t = t_1t_2\cdots t_n$ is a sequence of binary numbers such that each two consecutive numbers in the list differ in exactly one bit. For example 1011, 1001, 0001, 0000 is a Gray Code connecting 1011 and 0000.

**Algorithm 4.9** *Suppose we have the Gray code sequence $|s\rangle = |g_1\rangle \to |g_2\rangle \to |g_3\rangle \to \cdots \to |g_{m-1}\rangle \to |g_m\rangle = |t\rangle$. Assume $|g_1\rangle$ and $|g_2\rangle$ differ in their ith bit. We use a $\mathsf{C}^{n-1}(\mathsf{X})$ gate to flip the ith bit conditioned on the values of all other qubits to be the same as those of $|g_1\rangle$ and $|g_2\rangle$. This is repeated for swapping $|g_2\rangle$ and $|g_3\rangle$ and so on until we swap $|g_{m-2}\rangle$ and $|g_{m-1}\rangle$. Assume $|g_{m-1}\rangle$ and $|g_m\rangle$ differ in their jth bit. We use a controlled-$\hat{U}$ gate with the jth bit as the target and conditioned on the other qubits having the same values as those in $|g_{m-1}\rangle$ and $|g_m\rangle$. Finally, we use another sequence of $\mathsf{C}^{n-1}(\mathsf{X})$ gates to undo the swap operations in reverse order: $|g_{m-1}\rangle$ and $|g_{m-2}\rangle$, then $|g_{m-2}\rangle$ and $|g_{m-3}\rangle$ and so on until we swap $|g_2\rangle$ and $|g_1\rangle$.*



**Exercise 4.10** Show that the network in Figure 19 below implements the Toffoli gate using $H$, CNOT,

$$S = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \square$$

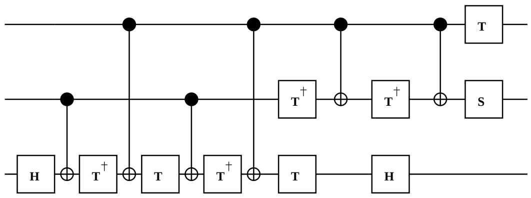



Figure 19: An implementation of the Toffoli gate

**Exercise 4.11** Show that if for a single qubit unitary operation $U$ we have $U = V^2$ where $V$ is another single qubit unitary operation, then the double controlled $U$ gate, called $\mathsf{C}^2(U)$ gate, can be implemented as in Figure 20. Show that for $V = (1-i)(I+iX)/2$ this implements the Toffoli gate. □

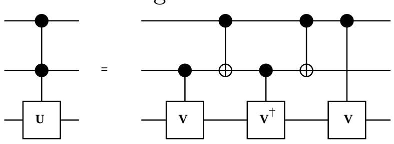



Figure 20: Implementation of $C^2(U)$

From Algorithm 4.9, we know how to implement an arbitrary unitary transformation on $\mathbb{C}^{2^n}$, i.e. a unitary operation on $n$ qubits, in terms of

(i) the controlled flip $\mathsf{C}^{n-1}(\mathsf{X})$, which flips the basis vectors in a given qubit conditioned on some specific values of all other qubits, and,

(ii) the controlled $\hat{U}$ gate $\mathsf{C}^{n-1}(\hat{\mathsf{U}})$, which applies the single qubit transformation $\hat{U}$ on a given qubit conditioned on some specific values of all other qubits.

We now show how to implement a controlled $U$ gate, $\mathsf{C}(\mathsf{U})$, for any single qubit unitary transformation $U$. For any matrix or linear map $A$ we formally write:

$$e^A = I + \frac{A^1}{1!} + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \cdots.$$

**Example 4.12** If $A^2 = I$, then we have:

$$\begin{aligned} e^{iAx} &= I + i\tfrac{x^1}{1!}A - \tfrac{x^2}{2!}I - i\tfrac{x^3}{3!}A + \cdots \\ &= \cos(x)I + i\sin(x)A. \end{aligned}$$

In particular, this holds for the Pauli matrices $X$, $Y$ and $Z$ since $X^2 = Y^2 = Z^2 = I$.



**Definition 4.13** The *rotation operators* around the $x$, $y$ and $z$ axes are respectively defined as:

$$R_x(\theta) = e^{-i\theta X/2} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}X,$$

$$R_y(\theta) = e^{-i\theta Y/2} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}Y,$$

$$R_z(\theta) = e^{-i\theta Z/2} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}Z.$$

**Exercise 4.14** Check that $XYX = -Y$ and $XZX = -Z$; then show that $XR_y(\theta)X = R_y(-\theta)$ and $XR_z(\theta)X = R_z(-\theta)$. □

**Proposition 4.15** *If $U$ is a single qubit unitary operation, then there exist $\alpha$, $\beta$, $\gamma$ and $\delta$ such that $U = e^{i\alpha}R_z(\beta)R_y(\gamma)R_z(\delta)$.*

**Proof** By Exercise 3.8, there exist $\alpha$, $\beta$, $\gamma$ and $\delta$ with:

$$U = \begin{pmatrix} e^{i(\alpha-\beta/2-\delta/2)}\cos\gamma/2 & -e^{i(\alpha-\beta/2+\delta/2)}\sin\gamma/2 \\ e^{i(\alpha+\beta/2-\delta/2)}\sin\gamma/2 & e^{i(\alpha+\beta/2+\delta/2)}\cos\gamma/2 \end{pmatrix}.$$

Then by simple matrix multiplication we get:

$$e^{i\alpha}R_z(\beta)R_y(\gamma)R_z(\delta) =$$

$$\begin{pmatrix} e^{i(\alpha-\beta/2-\delta/2)}\cos\gamma/2 & -e^{i(\alpha-\beta/2+\delta/2)}\sin\gamma/2 \\ e^{i(\alpha+\beta/2-\delta/2)}\sin\gamma/2 & e^{i(\alpha+\beta/2+\delta/2)}\cos\gamma/2 \end{pmatrix} = U \quad \square$$



For any single qubit unitary operation $U$ with $\alpha$, $\beta$, $\gamma$ and $\delta$ as in Proposition 4.15, put

$$A = R_z(\beta)R_y(\frac{\gamma}{2}), \qquad B = R_y(-\frac{\gamma}{2})R_z(-\frac{\delta+\beta}{2}),$$

$$C = R_z(\frac{\delta-\beta}{2}).$$

Then, it is easily checked that

$$ABC = I. \tag{2}$$

From Exercise 4.14, using $X^2 = I$ we get:

$$\begin{aligned} XBX &= XR_y(-\frac{\gamma}{2})XXR_z(-\frac{\delta+\beta}{2})X \\ &= R_y(\frac{\gamma}{2})R_z(\frac{\delta+\beta}{2}). \end{aligned}$$

Hence,

$$\begin{aligned} AXBXC &= R_z(\beta)R_y(\frac{\gamma}{2})R_y(\frac{\gamma}{2})R_z(\frac{\delta+\beta}{2})R_z(\frac{\delta-\beta}{2}) \\ &= R_z(\beta)R_y(\gamma)R_z(\delta). \end{aligned}$$

It follows that

$$U = e^{i\alpha}AXBXC. \tag{3}$$

We can now implement the $\mathsf{C}(\mathsf{U})$ gate using the single qubit unitary operations $A$, $B$, $C$, $X$ and the phase gate $\alpha$ as in the following exercise.



**Exercise 4.16** Assuming that the single qubit operation $U$ is given as in Equations 3 and 2 by $U = e^{i\alpha}AXBXC$ with $ABC = I$, show that the network in Figure 21 implements the controlled $U$ gate, $\mathsf{C}(\mathsf{U})$, using the single qubit gate $X$ the phase gate $\alpha$ and the single qubit gates $A$, $B$ and $C$. □

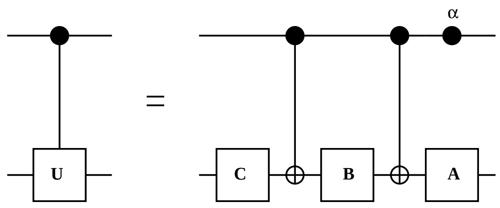



Figure 21: Implementation of $C(U)$

Therefore the $\mathsf{C}(\mathsf{U})$ gate as well as the Toffoli gate can be implemented using $O(1)$ single qubit gates and $\mathsf{CNOT}$ gates. In order to complete the proof that any quantum gate can be constructed using single qubit gates and the $\mathsf{CNOT}$ gate, we show how to implement for any $n \geq 2$ the controlled $U$ gate $\mathsf{C}^n(\mathsf{U})$ using the Toffoli gate and the $\mathsf{C}(\mathsf{U})$ gate. This is done as in the following exercise.

**Exercise 4.17** Show that the $\mathsf{C^n(U)}$ gate, where $U$ is a single qubit unitary operation, can be implemented using $n-1$ work qubits as in the network in Figure 22, depicted for $n=5$. Deduce that any such $\mathsf{C^n(U)}$ gate can be implemented by $O(n)$ single qubit and $\mathsf{CNOT}$ gates. □

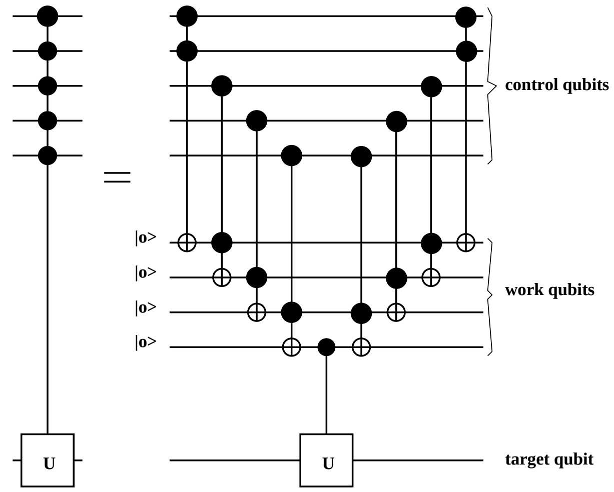



Figure 22: Implementation of $C^n(U)$



Returning to Algorithm 4.9, we need altogether at most $2(n-1)$ controlled operations to swap $|g_1\rangle$ and $|g_{n-1}\rangle$ and back again. By Exercise 4.17, each of these controlled operations needs $O(n)$ single qubit and $\mathsf{CNOT}$ gates to implement; each controlled-$\hat{U}$ gate also requires $O(n)$ such gates. Hence, implementing a two-level unitary operation requires $O(n^2)$ single qubit and $\mathsf{CNOT}$ gates.

**Example 4.18** Let

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

Then $s = 000$ and $t = 111$. With the sequence $000 \to 001 \to 011 \to 111$, we construct the network in Figure 23 which implements $U$. □



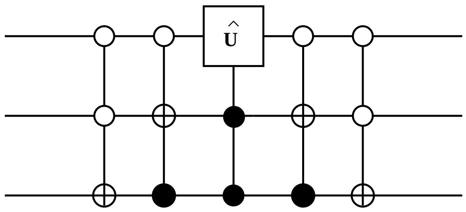



Figure 23: Network for a two-level unitary operation

Altogether, we have seen that any unitary operation on $n$-qubit quantum registers can be constructed by the composition of at most $2^n(2^n-1)/2$, i.e. $O(2^{2n})$, two-level unitary operations. Therefore implementing a unitary operation requires a circuit with $O(n^2 2^{2n})$ single qubit and $\mathsf{CNOT}$ gates. Clearly this is very inefficient. In practice we need a different technique.

In fact, it can be shown that the Hadamard gate, the phase gates $\pi/2$ and $\pi/4$, and the $\mathsf{CNOT}$ gate together form a *finite* universal set of gates: any unitary transformation on two or more qubits can be efficiently approximated as accurately as desired by a circuit with a finite number of these gates (see Nielsen & Chuang, page 194).



## 4.7 The No-Cloning Theorem

We can easily see that the $\mathsf{CNOT}$ gate induces: $|x0\rangle \mapsto |xx\rangle$ for $x=0$ or $x=1$, i.e. it produces a copy of such $x$. The question is if it can copy an arbitrary qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. The answer is no: we have $(\alpha|0\rangle + \beta|1\rangle)|0\rangle \mapsto \alpha|00\rangle + \beta|11\rangle$, and hence the output is an entangled state which cannot be the tensor product of $|\psi\rangle$ with itself. In fact we prove that, unlike the classical $\mathsf{FANOUT}$, no unitary matrix can copy an arbitrary qubit.

**Theorem 4.19** The No-Cloning Theorem. *There does not exist a unitary transformation $M$ such that $M|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle$ for all qubits $|\psi\rangle$.*

**Proof** Suppose there exists $M$ such that $M|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle$ for any qubit $|\psi\rangle$. Choose qubits $|\psi\rangle$ and $|\phi\rangle$ with $0 < \langle\psi|\phi\rangle < 1$, e.g. $|\psi\rangle = |0\rangle$ and $|\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Then we have $M|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle$ and $M|\phi\rangle|0\rangle = |\phi\rangle|\phi\rangle$. Since $M$, being unitary, preserves inner products, we get $\langle\psi|\phi\rangle\langle\psi|\phi\rangle = \langle\psi|\phi\rangle\langle 0|0\rangle = \langle\psi|\phi\rangle$, contradicting $0 < \langle\psi|\phi\rangle < 1$. Therefore $M$ cannot exist. □

# 5 Quantum Entanglement

As we have seen the CNOT gate can be used to create entangled states. The circuit in Figure 24 creates the four basic entangled states (depicted up to the normalization factor) which are called the *Bell* or *EPR* states. These are named after Bell, Einstein, Podolsky and Rosen who first discovered the amazing properties of these states. In fact, Einstein, Podolsky and Rosen tried to use these states in order to argue, unsuccessfully, that quantum mechanics is not a complete description of the physical world. Entangled states can be used to carry out a surprising act as follows.

| in | out |
|---|---|
| |00> | (|00> + |11>) |
| |01> | (|01> + |10>) |
| |10> | (|00> − |11>) |
| |11> | (|01> − |10>) |

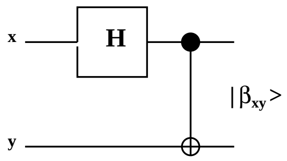



Figure 24: The Bell or EPR states



## 5.1 Quantum Teleportation

Alice and Bob meet, generate an EPR pair and each takes one qubit of the EPR state. Then they move far apart from each other. Several years later, Alice wants to send a qubit to Bob. She does not know the state of the qubit, cannot make copies of the qubit (by the No-Cloning Theorem), and can only send classical information to Bob. This certainly looks like an impossible task. Even if she knew the state of the qubit, she would have needed an infinite amount of time to describe its state to Bob since, as we have seen, a quantum state varies over a continuous space. Fortunately quantum entanglement makes this task possible.

Here is how it works. Assume the qubit Alice has to send to Bob is $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $\alpha$ and $\beta$ are unknown. Alice will combine $|\psi\rangle$ with her half of the EPR pair and then applies the CNOT gate and the Hadamard gate before measuring her two qubits. She then sends the result of the two bits to Bob who is, amazingly, able to retrieve $|\psi\rangle$ from this classical information.



More formally, Alice and Bob will make use of the three-qubit quantum circuit in Figure 25. Alice uses the top two lines, corresponding to the first two qubits, whereas Bob uses the bottom line, corresponding to the last qubit. The input is:

$$|\psi_0\rangle = |\psi\rangle|\beta_{00}\rangle = \frac{1}{\sqrt{2}}[\alpha|0\rangle(|00\rangle+|11\rangle)+\beta|1\rangle(|00\rangle+|11\rangle)]$$

where the state $|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ occupies the second qubit of Alice and the qubit of Bob.

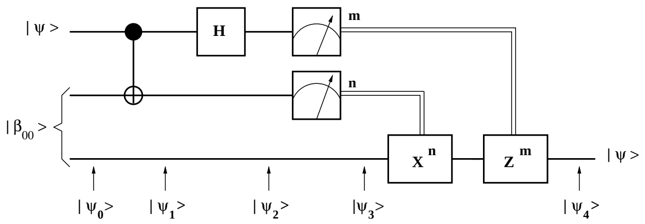



Figure 25: Quantum teleportation

Alice then sends her qubits through a CNOT gate which gives:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}[\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|10\rangle + |01\rangle)].$$



She then sends her first qubit through a Hadamard gate to obtain:

$$|\psi_2\rangle = \frac{1}{2}[\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle)+ \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)],$$

or

$$|\psi_2\rangle = \frac{1}{2}\left[|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\alpha|1\rangle + \beta|0\rangle) +|10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(\alpha|1\rangle - \beta|0\rangle)\right].$$

Thereafter, she measures her two qubits obtaining one of the following four pairs of bits:

00 01 10 11.

Each of these measurement results will collapse Bob's qubit as follows:

$$00 \mapsto |\psi_3(00)\rangle = \alpha|0\rangle + \beta|1\rangle$$

$$01 \mapsto |\psi_3(01)\rangle = \alpha|1\rangle + \beta|0\rangle$$

$$10 \mapsto |\psi_3(10)\rangle = \alpha|0\rangle - \beta|1\rangle$$

$$11 \mapsto |\psi_3(11)\rangle = \alpha|1\rangle - \beta|0\rangle$$

Alice communicates her two bits $mn$ with Bob over a classical channel.

Bob will then send his qubit through the circuit $X^n Z^m$ where

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

The final result can be easily checked to be $|\psi\rangle$.

**Exercise 5.1** Check that the final result is indeed the state $|\psi\rangle$. □

Note the following points:

- Quantum teleportation requires communication over a classical channel. Therefore there is no instantaneous transmission of information and, hence, the speed of light is not broken.
- Although Alice manages to send a copy of $|\psi\rangle$ to Bob the original version of $|\psi\rangle$ is destroyed. Therefore there is no violation of the No-Cloning theorem.
- Quantum teleportation was discovered theoretically in 1993 and has been confirmed by experiment since then.



# 6 Quantum Algorithms

We have seen that the state of a quantum system is a superposition of its computational basis elements and that the quantum state evolves, for example in a quantum circuit, according to a unitary operation which acts in parallel on all the basis elements in the superposition. The ability to compute in parallel is a fundamental feature of quantum circuits. In fact, we show here that a quantum circuit can in one step "compute" the values of a function on different inputs.

Let $f : \{0,1\} \to \{0,1\}$ be a Boolean function which outputs a bit for each input bit. We construct, using the method of Section 4.4, a two-qubit circuit which performs $U_f : |x, y\rangle \mapsto |x, y \oplus f(x)\rangle$.

**Exercise 6.1** Show that $U_f$ is unitary. □

We can now check that

$$U_f : \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)|0\rangle \mapsto \frac{|0, f(0)\rangle + |1, f(1)\rangle}{\sqrt{2}},$$

where $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ can be obtained by applying the Hadamard gate to $|0\rangle$.



We see that the $U_f$ gate "computes" both values of $f(0)$ and $f(1)$ in one step.

The above construction can be extended to Boolean functions with $n$ bits of input.

The $U_f$ gate now has $n+1$ input qubits, the first $n$ are input to $f$. It has also $n+1$ output qubits, the first $n$ qubits are identical with the first $n$ qubits of the input as in Figure 26.

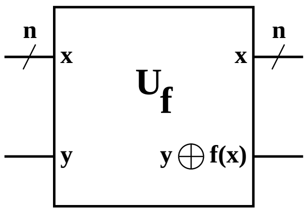



Figure 26: A Gate for Parallel Computation

In this and future figures, we use the convention that a stroked line with label $n$ indicates that the line stands for $n$ input or output wires.

**Exercise 6.2** Check that for any integer $n$ the operator $U_f$ is a unitary transformation. □



Instead of a single Hadamard operation, we use the $n$ tensor product

$$\underbrace{H \otimes H \otimes \cdots \otimes H}_{n \text{ times}}$$

of the Hadamard gates, which we denote by $H^{\otimes n}$. We also write the qubit

$$|\underbrace{00\cdots 0}_{n \text{ times}}\rangle$$

simply as $|0\rangle^{\otimes n}$.

**Exercise 6.3** Show that

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

We then have

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

Again we see that this gate is capable of "computing" all the $2^n$ values of $f$ in one step. Clearly for large $n$ this gives rise to a massive parallel computation. But does the circuit really compute all values of $f$ is one step?

The fact is that after any measurement of the output, the state vector collapses, we will only obtain one value $f(x)$ for a single $x \in \{0,1\}^n$ and all other values are lost for ever.

Therefore, in order to use the hidden information in this parallelism, we need to be able somehow to obtain more information from the superposition

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

than just one value of $f$. What is required is some appropriate *interference* between the values in the superposition in the output.

The great American physicist, Richard Feynman, was the first to observe that a Turing machine or a classical computer is not able to simulate a quantum system efficiently. He turned this observation around and raised the possibility that computers based on quantum mechanics can surpass electronic computers.

It was David Deutsch who in 1985 discovered the first quantum algorithm, generalized in 1992 to what is now called the Deutsch-Jozsa algorithm.



## 6.1 Deutsch's Algorithm

The first quantum algorithm, with a clear superiority over its classical counterpart, was discovered by Deutsch. Consider the simple $U_f$ gate of the previous section which uses the Hadamard gate on $|0\rangle$ to prepare its first qubit $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. This time a second Hadamard gate on input $|1\rangle$ is used to prepare its second qubit $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Furthermore, a third Hadamard gate is applied to the first qubit in the output as in Figure 27. We have: $|\psi_0\rangle = |01\rangle$ and

$$|\psi_1\rangle = (H \otimes H)|01\rangle = \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle).$$

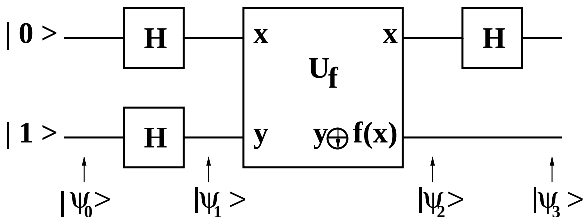



Figure 27: Deutsch's Circuit



Note that

$$U_f(|x\rangle(|0\rangle - |1\rangle)/\sqrt{2}) = (-1)^{f(x)}|x\rangle(|0\rangle - |1\rangle)/\sqrt{2}.$$

Hence,

$$|\psi_2\rangle = \begin{cases} \pm\frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) & \text{if } f(0) = f(1) \\ \pm\frac{1}{2}(|0\rangle - |1\rangle)(|0\rangle - |1\rangle) & \text{if } f(0) \neq f(1). \end{cases}$$

Finally, the action of the third Hadamard gate on the first qubit gives:

$$|\psi_3\rangle = \begin{cases} \pm|0\rangle(|0\rangle - |1\rangle)/\sqrt{2} & \text{if } f(0) = f(1) \\ \pm|1\rangle(|0\rangle - |1\rangle)/\sqrt{2} & \text{if } f(0) \neq f(1). \end{cases}$$

which can be written as:

$$|\psi_3\rangle = \pm|f(0) \oplus f(1)\rangle(|0\rangle - |1\rangle)/\sqrt{2}.$$

Therefore, with this quantum circuit we can, in a single step, evaluate $f(0) \oplus f(1)$ which a classical algorithm would need two steps to compute.

The above algorithm can be extended to Boolean functions on $n$ bits. Let $f : \{0,1\}^n \to \{0,1\}$ be either a constant function or a *balanced* function, i.e. $f(x) = 0$ for exactly $2^{n-1}$ values of $x$.



Classically it takes, in the worst case, $1 + 2^{n-1}$ queries to establish whether $f$ is constant or balanced.

Surprisingly, a quantum algorithm can decide this in only one single step. The circuit, named after Deutsch-Jozsa, is now as in Figure 28. The stroked line with label $n$ represents a set of $n$ qubits, which we call the query register.

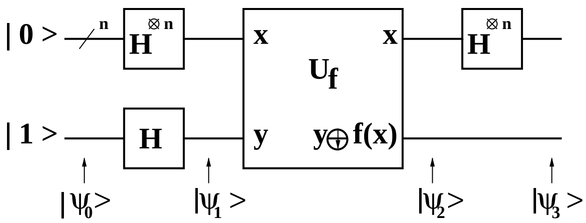



Figure 28: Deutsch-Jozsa Circuit

**Exercise 6.4** Show that

$$\begin{aligned} H|x\rangle &= \textstyle\sum_{y\in\{0,1\}}(-1)^{xy}|y\rangle/\sqrt{2} \\ H^{\otimes n}|x\rangle &= \textstyle\sum_{y\in\{0,1\}^n}(-1)^{x.y}|y\rangle/\sqrt{2^n} \end{aligned}$$

where $x.y$ is the bitwise inner product of $x$ and $y$ modulo 2. $\square$

We now have:

$|\psi_0\rangle = |0\rangle^{\otimes n}|1\rangle$
$|\psi_1\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} |x\rangle|(0\rangle - |1\rangle)/\sqrt{2}$
$|\psi_2\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} (-1)^{f(x)}|x\rangle(|0\rangle - |1\rangle)/\sqrt{2}$
$|\psi_3\rangle = \sum_{y\in\{0,1\}^n} \sum_{x\in\{0,1\}^n} \frac{(-1)^{x\cdot y+f(x)}|y\rangle}{2^n} \left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right).$

The amplitude of $|y\rangle = |0\rangle^{\otimes n}$ in $|\psi_3\rangle$ is $\sum_x (-1)^{f(x)}/2^n$. If $f$ is constant this will be 1 or $-1$ which means, by the normalization condition, that the amplitudes of all other states are zero. Any measurement of the query register produces the state $|0\rangle^{\otimes n}$ with probability 1. On the other hand if $f$ is balanced then the amplitude of $|y\rangle = |0\rangle^{\otimes n}$ will be zero, which means that the amplitude of at least one other $n$-qubit must be non-zero. In this case, any measurement of the query register produces at least one qubit 1. Hence, with a single measurement, one can decide whether $f$ is constant or balanced, an exponential reduction in the complexity of the classical algorithm. This is a massive improvement, but there are no known applications of the Deutsch-Jozsa algorithm.



# 7 Quantum Search

We will now turn to quantum algorithms which are both superior to their classical counterpart and also very useful in practice. The quantum search algorithm discovered by Grover in 1996 gives a fast solution to the following problem: In an unsorted list of $N$ items some are distinguished by satisfying a given condition; the task is to retrieve one of these distinguished states. More specifically, we assume, for convenience, that $N = 2^n$ and the items are the bit sequences of length $n$ with $M$ distinguished states or solutions. We further assume that there exists an oracle, a black box, which determines if a state is a solution or not. Formally, the oracle can be regarded as a Boolean function $f : \{0,1\}^n \to \{0,1\}$ such that $f(x) = 1$ if $x$ is a solution and $f(x) = 0$ otherwise.

On a classical computer, one needs on average $O(N/M)$ calls of the oracle $f$ to find a solution. Grover's algorithm determines a solution in $O(\sqrt{N/M})$ which is a quadratic improvement over classical computation.



Here is how it works. We prepare an $n$-qubit quantum register for the superposition of states in $\{0,1\}^n$. We also need a quantum implementation of the oracle $f$.

We use a single oracle qubit and a unitary map

$$O : |x\rangle|y\rangle \mapsto |x\rangle|y \oplus f(x)\rangle,$$

where $x \in \{0,1\}^n$ is a computational basis element in the quantum register and $y$ is the single oracle qubit.

If $y$ is prepared in state 0, then $O$ flips it if $x$ is a solution ($f(x) = 1$) and leaves it unchanged otherwise ($f(x) = 0$). However, as in the Deutsch-Jozsa algorithm, it is convenient to prepare the oracle qubit in the state $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ by applying the Hadamard gate to $|1\rangle$.

Then the oracle $O$ acts, for $x \in \{0,1\}^n$, as follows:

$$O : |x\rangle\frac{|0\rangle - |1\rangle}{\sqrt{2}} \quad \mapsto (-1)^{f(x)}|x\rangle\frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

Since the single oracle qubit stays in the state $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$ throughout the computation, we can simply remove it from now on for simplification.



Therefore, the action of $O$ on a general state of the quantum register is:

$$O : \sum_{x\in\{0,1\}^n} \alpha_x|x\rangle \quad \mapsto \quad \sum_{x\in\{0,1\}^n} (-1)^{f(x)}\alpha_x|x\rangle.$$

The quantum register is prepared in the state $|0\rangle^{\otimes n}$ which is then put, by applying the Hadamard transform $H^{\otimes n}$, in the superposition state

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle. \qquad (4)$$

We now apply the following sequence of operations called the *Grover* operator:

$$G = H^{\otimes n} P_0 H^{\otimes n} O$$

where the conditional phase shift $P_0$ is given by

$$P_0 : |x\rangle \mapsto \begin{cases} |x\rangle & x = 0 \\ -|x\rangle & x > 0, \end{cases}$$

for any computational basis state $|x\rangle$ (with $x$ in the range $0 \le x \le N-1$). By checking its action on basis states it is easy to check that $P_0 = 2|0\rangle\langle 0| - I$ where $I$ is the identity map.

Since the Hadamard operation is its own inverse, we get

$$H^{\otimes n} P_0 H^{\otimes n} = H^{\otimes n}(2|0\rangle\langle 0| - I)H^{\otimes n} = 2|\psi\rangle\langle\psi| - I.$$

Hence, the $G$ operator can be simply written as

$$G = (2|\psi\rangle\langle\psi| - I)O.$$

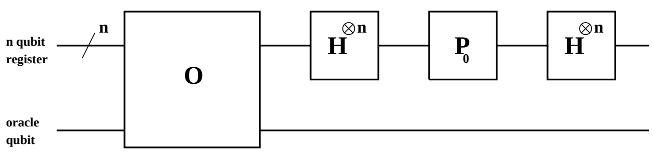



Figure 29: Circuit for Grover's Operation $G$

**Exercise 7.1** (i) Show that

$$(2|\psi\rangle\langle\psi| - I)(\sum_x \alpha_x |x\rangle) = \sum_x (-\alpha_x + 2\langle\alpha\rangle)|x\rangle,$$

where $\langle\alpha\rangle = \sum_x \alpha_x / N$.

(ii) Explain why the operation $2|\psi\rangle\langle\psi| - I$ is called *inversion about mean.* □



## 7.1 Geometric Interpretation of $G$

The Grover operation is a rotation in the real plane generated by the initial state $|\psi\rangle$ and the state obtained as the uniform superposition of the $M$ basis state solutions. To see this in detail, let $T = \{x \in \{0,1\}^n | f(x) = 1\}$ and $S = \{0,1\}^n \setminus T$. Put

$$|\sigma\rangle = \frac{1}{\sqrt{N-M}} \sum_{x \in S} |x\rangle, \qquad |\tau\rangle = \frac{1}{\sqrt{M}} \sum_{x \in T} |x\rangle.$$

Then it is easily checked that

$$|\psi\rangle = \sqrt{\frac{N-M}{N}}|\sigma\rangle + \sqrt{\frac{M}{N}}|\tau\rangle, \tag{5}$$

which shows that $|\psi\rangle$ is in the same two-dimensional real plane as $|\sigma\rangle$ and $|\tau\rangle$. Any state in this plane is in the form $a|\sigma\rangle + b|\tau\rangle$ with $a^2 + b^2 = 1$. We have $O(a|\sigma\rangle + b|\tau\rangle) = a|\sigma\rangle - b|\tau\rangle$, which shows that the action of $O$ on this plane is a reflection about $|\sigma\rangle$.

**Exercise 7.2** Show that the action of $2|\psi\rangle\langle\psi| - I$ in the $|\sigma\rangle$ and $|\tau\rangle$ plane is reflection about $|\psi\rangle$. □



Since the composition of two reflections is a rotation, it follows that $G$ rotates state vectors in the $|\sigma\rangle, |\tau\rangle$ plane by $\theta$ towards $|\tau\rangle$, where $\theta/2$ is the angle between $|\psi\rangle$ and $|\sigma\rangle$, i.e. $|\psi\rangle = \cos(\theta/2)|\sigma\rangle + \sin(\theta/2)|\tau\rangle$, with $\cos(\theta/2) = \sqrt{(N-M)/N}$, $\sin(\theta/2) = \sqrt{M/N}$. Assume, by replacing $N$ with $2N$ if necessary, that $M \leq N/2$ (Figure 30).

**Exercise 7.3** Show that if we change the computational basis so that $|\sigma\rangle$ and $|\tau\rangle$ are basis elements, then the matrix representation of $G$ will be;

$$G_{\sigma,\tau} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \ \square$$

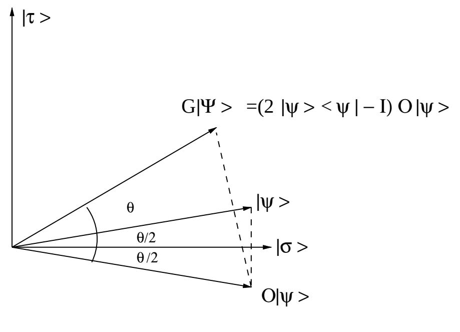



Figure 30: Action of Grover's Operation



Hence, after $m$ iterations we have:

$$G^m|\psi\rangle = \cos(\frac{2m+1}{2}\theta)|\sigma\rangle + \sin(\frac{2m+1}{2}\theta)|\tau\rangle.$$

It follows that when $\frac{2m+1}{2}\theta \approx \pi/2$, i.e. after

$$m = \lfloor \pi/2\theta - 1/2 \rceil \tag{6}$$

iterations, where $\lfloor d \rceil$ is the nearest integer to $d$, the state vector is within an angle $\theta/2 \leq \pi/4$ of $|\tau\rangle$. Measurement of the state vector now will give a solution with probability at least $\cos^2(\pi/4) = 1/2$.

In the case that $M \ll N$ (i.e. $M$ very small compared to $N$), we get $\theta \approx \sin\theta \approx 2\sqrt{M/N}$, so that $\theta/2 \approx \sqrt{M/N}$. Hence, measurement of the state vector in this case will produce a solution with probability at least $\cos^2(\theta/2) \approx 1 - M/N$.

In general, from Equation 6, we have $m \leq \lfloor \pi/2\theta \rfloor$. Since for any real $\theta$, we have

$$\frac{\theta}{2} \geq \sin\frac{\theta}{2} = \sqrt{\frac{M}{N}},$$

we obtain an upper-bound for the number of iterations of $G$ needed to find a solution:

$$m \leq \left\lfloor \tfrac{\pi}{4}\sqrt{\tfrac{N}{M}} \right\rfloor.$$

**Exercise 7.4** Let $n = 2$ so that $N = 2^n = 4$, and let $M = 1$. The oracle $f$ with $f(x) = 0$ for all $x \neq x_0$ and $f(x_0) = 1$ can be chosen from the four circuits in Figure 31, corresponding to $x_0 = 0, 1, 2, 3$ respectively from left to right.

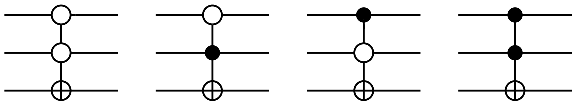


Figure 31: Four possible oracles

Show that the circuit in Figure 32 in effect implements the operation $G$. How many iterates of $G$ are needed to determine $x_0$? $\square$

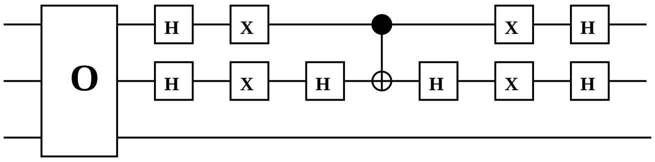



Figure 32: Circuit for $G$



# 8 Fourier Transform

The Fourier Transformation is one of the most fundamental tools in science and applied mathematics. It maps time dependent functions to their frequency domains in such a way that a periodic function of period $T > 0$ is mapped to a function whose frequency amplitude is non-vanishing only at frequencies which are integer multiples of $1/T$.

The *Quantum Fourier Transform* is the quantum analogue of the *Discrete Fourier Transform* (DFT).

Given a vector of complex numbers

$$(x_0, x_1, \cdots, x_{N-1})$$

as input, the DFT, denoted here by $F$, provides a vector

$$(y_0, y_1, \cdots, y_{N-1})$$

of complex numbers as output by the rule:

$$\begin{aligned} F: \quad (x_0, \cdots, x_{N-1}) &\quad\mapsto\quad (y_0, \cdots, y_{N-1}) \\ y_k &\quad=\quad \tfrac{1}{\sqrt{N}} \textstyle\sum_{j=0}^{N-1} x_j e^{2\pi i jk/N} \end{aligned}$$



In terms of the computational basis $|j\rangle$ ($j = 0, \cdots, N-1$), with $N = 2^n$, this can be written as

$$F : \sum_{j=0}^{N-1} x_j |j\rangle \mapsto \sum_{k=0}^{N-1} y_k |k\rangle \qquad (7)$$

which is the *quantum Fourier transform.*

**Exercise 8.1** What is the matrix for $F$ when $n = 1$ and $n = 2$? Show that $F$ is unitary for any $n$. $\square$

It follows from Equation 7, by writing the indices of states in binary form, that:

$$\begin{aligned}
&|j\rangle \\
&\mapsto \tfrac{1}{\sqrt{N}} \textstyle\sum_{k=0}^{N-1} e^{2\pi i jk/N} |k\rangle \\
&= \tfrac{1}{2^{n/2}} \textstyle\sum_{k_1=0}^{1} \cdots \sum_{k_n=0}^{1} e^{2\pi i j (\sum_{l=1}^{n} k_l 2^{-l})} |k_1 \cdots k_n\rangle \\
&= \tfrac{1}{2^{n/2}} \textstyle\sum_{k_1=0}^{1} \cdots \sum_{k_n=0}^{1} \bigotimes_{l=1}^{n} e^{2\pi i j k_l 2^{-l}} |k_l\rangle \\
&= \tfrac{1}{2^{n/2}} \textstyle\bigotimes_{l=1}^{n} \sum_{k_l=0}^{1} e^{2\pi i j k_l 2^{-l}} |k_l\rangle \\
&= \tfrac{1}{2^{n/2}} \textstyle\bigotimes_{l=1}^{n} (|0\rangle + e^{2\pi i j 2^{-l}} |1\rangle) \\
&= \tfrac{1}{2^{n/2}} (|0\rangle + e^{2\pi i 0.j_n} |1\rangle)(|0\rangle + e^{2\pi i 0.j_{n-1}j_n} |1\rangle) \cdots \\
&\cdots (|0\rangle + e^{2\pi i 0.j_2 \cdots j_{n-1}j_n} |1\rangle)(|0\rangle + e^{2\pi i 0.j_1 j_2 \cdots j_{n-1}j_n} |1\rangle)
\end{aligned}$$



Hence,

$$\begin{aligned} &F : |j_1 j_2 \cdots j_{n-1} j_n\rangle \mapsto \qquad (8) \\ &\frac{1}{2^{n/2}} (|0\rangle + e^{2\pi i 0.j_n} |1\rangle) \cdots (|0\rangle + e^{2\pi i 0.j_1 j_2 \cdots j_n} |1\rangle). \end{aligned}$$

This gives a simple way of implementing $F$ as in Figure 33 where

$$R_k = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{pmatrix}.$$

The $n$-qubit input on the left of the figure is $|j_1 j_2 \cdots j_{n-1} j_n\rangle$.

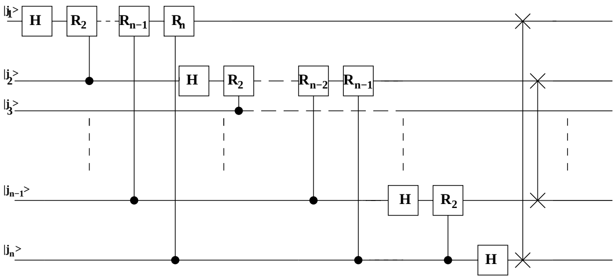


Figure 33: Circuit for Fourier Transform

We will now check that the circuit correctly implements $F$.

After the action of the Hadamard gate on the first qubit $|j_1\rangle$, we get:

$$\frac{1}{\sqrt{2^1}}(|0\rangle + e^{2\pi i 0.j_1}|1\rangle)|j_2 \cdots j_n\rangle$$

as $e^{2\pi i 0.j_1} = 1$ when $j_1 = 0$ and is $-1$ when $j_1 = 1$. Next applying the controlled-$R_2$ gate we obtain

$$\frac{1}{\sqrt{2^1}}(|0\rangle + e^{2\pi i 0.j_1 j_2}|1\rangle)|j_2 \cdots j_n\rangle.$$

Continuing in this way to apply the controlled-$R_3$ to $R_n$ gates, the final outcome is:

$$\frac{1}{\sqrt{2^1}}(|0\rangle + e^{2\pi i 0.j_1 j_2 \cdots j_n}|1\rangle)|j_2 \cdots j_n\rangle.$$

Next we consider the second qubit $|j_2\rangle$ which goes through the sequence of the Hadamard and the controlled-$R_2$ to $R_{n-1}$ gates. The result will be:

$$\frac{1}{\sqrt{2^2}}(|0\rangle+e^{2\pi i 0.j_1 j_2 \cdots j_n}|1\rangle)(|0\rangle+e^{2\pi i 0.j_2 \cdots j_n}|1\rangle)|j_3 \cdots j_n\rangle.$$

Repeating this scheme for qubits $|j_3\rangle \cdots |j_n\rangle$ gives:

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

Finally applying the swap operations, the order of all qubits is reversed and we end up with:

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

which is precisely the outcome of $F$. Since all the gates involved are unitary, it follows that $F$ is a unitary operation.

Note that there are $n(n+1)/2$ Hadamard and controlled-$R_k$ gates and $\lfloor n/2 \rfloor$ swapping gates in the circuit. Hence, this method provides an algorithm of complexity $O(n^2)$ for computing the Fourier transform.

**Exercise 8.2** (i) Work out the matrix for $F$ and the network which implements it for $n = 3$.

(ii) Show that the inverse of $F$ is given by:

$$F^\dagger : |j\rangle \mapsto \frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} e^{-2\pi i jk/2^n} |k\rangle. \qquad (9)$$

(iii) Work out the circuit for $F^\dagger$. □



# 9 Phase Estimation

The quantum Fourier transform gives rise to a technique for estimating the phase of an eigenvalue of a unitary matrix, which is the basis of a number of quantum algorithms. We first recall the notions of eigenvalue and eigenvector.

## 9.1 Eigenvalues and Eigenvectors

Let $A$ be an $m \times m$ matrix with complex entries. Then $\lambda \in \mathbb{C}$ is an *eigenvalue* for $A$ if $\det(A-\lambda I) = 0$, where det denotes determinant and $I$ is the identity $m \times m$ matrix. The expression $\det(A-\lambda I)$ is a polynomial of degree $m$ in $\lambda$ and (by the fundamental theorem of algebra) it has $m$, in general complex, roots which are the eigenvalues of $A$ (some of them may be multiple roots). Furthermore any non-zero vector $v \in \mathbb{C}^m$ with $Av = \lambda v$ is called an *eigenvector* of $A$ for the eigenvalue $\lambda$. The subspace generated by the eigenvectors of $A$ corresponding to the same eigenvalue $\lambda$ is called the *eigenspace* of $\lambda$.

If the eigenspace has dimension bigger than one then it is called *degenerate*.

**Exercise 9.1**
- Find the eigenvalues and eigenvectors of the Pauli matrices, $X$, $Y$, $Z$. Locate the eigenvectors on the Bloch sphere.
- Show that the eigenvalues of a unitary matrix have norm one.
- Show that a change of basis does not alter the eigenvalues of a matrix. How do the eigenvectors alter with a change in the basis? □

The matrix representation of a linear operator with respect to a given basis is *diagonal* if it is of the form $A = \sum_j \lambda_j |j\rangle\langle j|$ where the vectors $|j\rangle$ form an orthonormal basis with eigenvalues $\lambda_j$. A linear operator is *diagonalizable* if it has a diagonal matrix representation. A linear operator $L$ is *Hermitian* if $L^\dagger = L$.

**Theorem 9.2** *A linear operator $L$ is diagonalizable iff $L^\dagger L = LL^\dagger$. (Nielsen/Chuang, p. 72.)*

**Corollary 9.3** *Unitary and Hermitian operators are diagonalizable.*

## 9.2 Diagonalization of Linear Maps

Suppose we have a diagonalizable linear map $L : \mathbb{C}^n \to \mathbb{C}^n$ given by a matrix $A$ in some basis. By Corollary 9.3, $L$ can be, in particular, a unitary or a Hermitian map. How do we diagonalize the linear map $L$, equivalently the matrix $A$?

**Algorithm 9.4** *We construct an orthonormal basis with respect to which L is diagonal:*

*(i) Find all the eigenvalues* $\lambda_1, \lambda_2, \cdots, \lambda_k$ *(where* $1 \le k \le n$*) of* $A$*, i.e. the roots of* $\det(A-\lambda I) = 0$ *and determine the multiplicity* $n_j$ *of each, so that* $\sum_{j=1}^k n_j = n$

*(ii) For each eigenvalue* $\lambda_j$ *of* $A$*, find* $n_j$ *orthonormal vectors* $|v_{jm}\rangle$ $(1 \le m \le n_j)$ *which satisfy* $(A - \lambda_j I)|v_{jm}\rangle = 0$.

*(iii) The vectors* $\{|v_{jm}\rangle | 1 \le m \le n_j,\ 1 \le j \le k\}$ *will form an orthonormal basis with respect to which L is diagonal.*

**Exercise 9.5** Diagonalize the three Pauli matrices and the Hadamard matrix. □



Let $U$ be a unitary operator on, say, $l$ qubits with an eigenvalue $e^{2\pi i\phi}$ associated with an eigenvector $|u\rangle$. Assume that we do not know $U$, $\phi$ or $|u\rangle$, but we are given a single preparation of $|u\rangle$ and have devices, considered as black boxes, to perform controlled-$U$, controlled-$U^2$, controlled-$U^{2^2}$, and generally controlled-$U^{2^t}$ for a given positive integer $t$. Our goal is to estimate $\phi$. In the first stage, we construct the network of Figure 34.

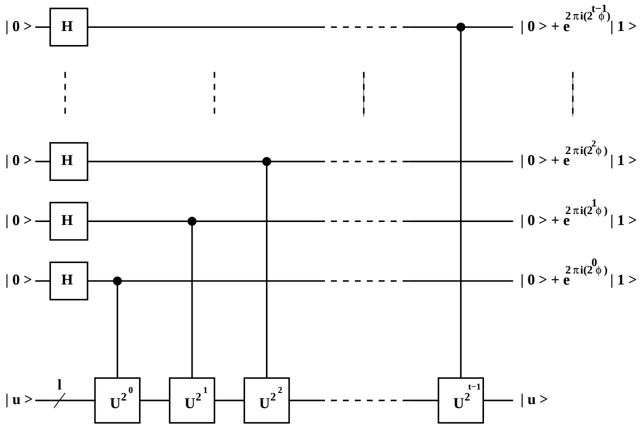


Figure 34: Circuit for Phase Estimation (stage 1)



The input to the second register is the eigenvector $|u\rangle$ which will not change under the action of the controlled-$U^{2^k}$ gates. The input to the first register is $|0\rangle^{\otimes t}$ which, after the $H^{\otimes t}$ gates becomes

$$[(|0\rangle + |1\rangle)/\sqrt{2}]^{\otimes t} = \frac{1}{\sqrt{2^t}} \sum_{k=0}^{2^t-1} |k\rangle. \qquad (10)$$

Now consider the $k$th qubit $(0 \le k \le t-1)$ counted from zero for the bottom qubit in the first register. After undergoing the controlled-$U^{2^k}$, it becomes $(|0\rangle + e^{2\pi i(2^k\phi)}|1\rangle)/\sqrt{2}$, as the phase factor $e^{2\pi i(2^k\phi)}$ is "kicked back" from the second register to this qubit. Thus, the output $|\psi\rangle$, say, of the first register (starting with the top qubit) is

$$|\psi\rangle = \bigotimes_{k=t-1}^{0} \frac{|0\rangle + e^{2\pi i(2^k\phi)}|1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2^t}} \sum_{y=0}^{2^t-1} e^{2\pi i\phi y}|y\rangle, \qquad (11)$$

where the last equality follows by induction on $t$.

In the next stage, we apply the inverse $F^\dagger$ of the quantum Fourier transform $F$ to the output of the first register and then measure the new output. Note that we have altogether $O(t^2)$ gates.



This gives a good estimate for $\phi$ as we will now show. First assume that $\phi$ can be expressed by exactly $t$ bits in binary representation: $\phi = 0.\phi_1\phi_2\cdots\phi_t$. Then the output of the first register after the first stage will be

$$|\psi\rangle = \frac{1}{\sqrt{2^t}}(|0\rangle + e^{2\pi i 0.\phi_t}|1\rangle)(|0\rangle + e^{2\pi i 0.\phi_{t-1}\phi_t}|1\rangle)\cdots$$
$$\cdots(|0\rangle + e^{2\pi i 0.\phi_2\cdots\phi_t}|1\rangle)(|0\rangle + e^{2\pi i 0.\phi_1\phi_2\cdots\phi_t}|1\rangle).$$

Comparing this result with the output of the quantum Fourier transform in Equation 8, we conclude immediately that, after the application of the inverse Fourier transform, we get

$$F^\dagger(|\psi\rangle) = |\phi_1\phi_2\cdots\phi_t\rangle,$$

i.e. measurement of the final output of the first register produces the value of $\phi$ exactly.

Now assume that $\phi \in [0,1]$ is any real number (rational or irrational). Let $\phi = a/2^t + \delta$ where $a = a_0a_1\cdots a_{t-1}$, with $a_k \in \{0,1\}$ for $0 \le k \le t-1$, such that $a/2^t$ is the best $t$-bit approximation to $\phi$, i.e. $0 \le |\delta| \le 1/2^{t+1}$. We have already dealt with $\delta = 0$, so we assume that $0 < |\delta| \le 1/2^{t+1}$.

Applying the inverse quantum Fourier transform, given by Equation 9, to Equation 11, we obtain:

$$|\hat{\phi}\rangle \stackrel{\text{def}}{=} F^{\dagger}|\psi\rangle = \frac{1}{2^t}\sum_{k=0}^{2^t-1}\sum_{y=0}^{2^t-1} e^{\frac{2\pi i}{2^t}(a-k)y} e^{2\pi i\delta y}|k\rangle. \tag{12}$$

The coefficient of $|k\rangle$ for $k = a$, which gives the probability of measuring $a$, is the geometric series:

$$\frac{1}{2^t}\sum_{y=0}^{2^t-1}(e^{2\pi i\delta})^y = \frac{1}{2^t}\left(\frac{1-(e^{2\pi i\delta})^{2^t}}{1-e^{2\pi i\delta}}\right).$$

From $|\delta| \le \frac{1}{2^{t+1}}$, we have $2^t|\delta| \le \frac{1}{2}$. Note that,

$$|1-e^{2\pi ix}| = 2|\sin \pi x| \tag{13}$$

and for $x \in [-1/2, 1/2]$,

$$2|x| \le |\sin \pi x| \le \pi|x|. \tag{14}$$

It follows that $|1-e^{2\pi i\delta 2^t}| = 2|\sin \pi\delta 2^t| \ge 4|\delta|2^t$. Furthermore, $|1-e^{2\pi i\delta}| = 2|\sin \pi\delta| \le 2\pi|\delta|$. Having a lower bound for the numerator and an upper bound for the denominator, we find that:

$$\left|\frac{1}{2^t}\left(\frac{1-(e^{2\pi i\delta})^{2^t}}{1-e^{2\pi i\delta}}\right)\right|^2 \ge \left(\frac{1}{2^t}\left(\frac{4|\delta|2^t}{2\pi|\delta|}\right)\right)^2 = \frac{4}{\pi^2}.$$



Hence, measurement of the output state will produce the best $t$-bit approximation $a$ to $\phi$ with probability at least $4/\pi^2 \approx 0.405$.

## 9.3 Bounding the Error

Now let $\phi = a/2^t + \delta$, where $a/2^t$, this time, is the best $t$-bit approximation less than or equal to $\phi$, i.e. $0 \le \delta < 1/2^t$.

By choosing $t$ large enough, we can actually bound the probability of obtaining a measurement $m$ which differs from $a$ by a tolerance constant $c$, i.e., $p(|m-a| > c)$.

Let $\alpha_k$ be the amplitude of $|(a+k)(\text{mod } 2^t)\rangle$ in $|\hat{\phi}\rangle$ of Equation 12. Then, we have:

$$\alpha_k = \frac{1}{2^t}\sum_{y=0}^{2^t-1}(e^{2\pi i(2^t\delta-k)/2^t})^y \tag{15}$$

$$= \frac{1}{2^t}\left(\frac{1-e^{2\pi i(2^t\delta-k)}}{1-e^{2\pi i(\delta-k/2^t)}}\right).$$



Since the set of integers

$$0, 1, \cdots, 2^t-1 \;(\text{mod } 2^t)$$

is precisely the set of integers

$$-2^{t-1}+1, -2^{t-1}+2, \cdots, -1, 0, 1, \cdots, 2^{t-1} \;(\text{mod } 2^t)$$

we have:

$$p(|m-a|>c) = \sum_{k=-2^{t-1}+1}^{-(c+1)}|\alpha_k|^2 + \sum_{k=c+1}^{2^{t-1}}|\alpha_k|^2. \tag{16}$$

As $|1-e^{ix}| \le 2$, for any real $x$, from Equation 15 we obtain:

$$|\alpha_k| \le \frac{2}{2^t|1-e^{2\pi i(\delta-k/2^t)}|}.$$

For the set of integers $k$ with:

$$-2^{t-1}+1 \le k \le 2^{t-1},$$

we have:

$$-1/2 \le \delta - k/2^t \le 1/2.$$

Therefore, from Equations 13 and 14, it follows that:

$$|\alpha_k| \le \frac{1}{2^{t+1}|\delta-k/2^t|}. \tag{17}$$



Using Equation 17 in Equation 16, and recalling that $0 \le 2^t\delta < 1$, we have for $c \ge 2$:

$$p(|m-a|>c)$$

$$\le \frac{1}{4}\left(\sum_{k=-2^{t-1}+1}^{-(c+1)}\frac{1}{(k-2^t\delta)^2} + \sum_{k=c+1}^{2^{t-1}}\frac{1}{(k-2^t\delta)^2}\right)$$

$$\le \frac{1}{4}\left(\sum_{k=-2^{t-1}+1}^{-(c+1)}\frac{1}{k^2} + \sum_{k=c+1}^{2^{t-1}}\frac{1}{(k-1)^2}\right)$$

$$\le \frac{1}{2}\sum_{k=c}^{2^{t-1}-1}\frac{1}{k^2} \le \frac{1}{2}\int_{c-1}^{2^{t-1}-1}\frac{dk}{k^2} \le \frac{1}{2(c-1)},$$

since the sum of reciprocals of squares of consecutive integers is less than the corresponding area under the graph of $x \mapsto x^{-2}$. We now aim to obtain a measurement $m$ close to $a$ up to an accuracy of $1/2^s$, for a given positive integer $s$, i.e. $|\frac{m}{2^t} - \frac{a}{2^t}| < \frac{1}{2^s}$ with probability at least $1-\epsilon$, for $0 < \epsilon < 1$. Thus, we take $c = 2^{t-s}-1$ and put $t = s+q$ where the positive integer $q$ is to be determined in terms of $s$ and $\epsilon$. Then:

$$p(|\frac{m}{2^t} - \frac{a}{2^t}| \ge \frac{1}{2^s}) = p(|m-a| > 2^q-1) \le \frac{1}{2(2^q-2)}.$$

Hence, it is sufficient to choose $q$ such that $\frac{1}{2(2^q-2)} \leq \epsilon$, i.e. $q = \lceil \log(2 + \frac{1}{2\epsilon}) \rceil$, which implies:

$$t = s + \left\lceil \log(2 + \frac{1}{2\epsilon}) \right\rceil . \tag{18}$$

**Exercise 9.6** More generally, suppose the states $|u\rangle$ for $u \in T$ are eigenstates of $U$ with eigenvalue $e^{2\pi i \phi_u}$. The phase estimation algorithm maps the normalized state

$$|0\rangle (\sum_{u \in T} d_u |u\rangle)$$

to the state

$$\sum_{u \in T} d_u |\hat{\phi}_u\rangle |u\rangle,$$

where the state $|\hat{\phi}_u\rangle$ gives a good estimate of $\phi_u$. Show that with $t$ chosen as in Equation 18, the probability of measuring $\phi_u$ accurate to $s$ bits in the output of the phase estimation algorithm is at least $|d_u|^2(1-\epsilon)$. $\square$

Phase estimation gives rise to a number of algorithms; we will study the following:

(i) quantum counting,

(ii) order finding, and,

(iii) prime factorization of large numbers.



# 10 Quantum Counting

We now return to Grover's search algorithm with the search space of size $N = 2^n$ and ask if can we compute the number $M$ of solutions if this number is not known. Clearly, it takes $N$ calls of the oracle to determine $M$ classically. There is surprisingly a much more efficient quantum algorithm for this.

Recall from Exercise 7.3 that the matrix representation of the operator $G$ in the plane generated by $|\sigma\rangle$ and $|\tau\rangle$ is

$$G_{\sigma,\tau} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

where

$$\sin\theta/2 = \sqrt{M/N}. \tag{19}$$

**Exercise 10.1** Show that the eigenvalues of $G_{\sigma,\tau}$ are $e^{i\theta}$ and $e^{-i\theta}$ corresponding to the eigenvectors $|a\rangle = i|\sigma\rangle + |\tau\rangle$ and $|b\rangle = |\sigma\rangle + i|\tau\rangle$. $\square$

We, therefore, apply the phase estimation algorithm to determine the eigenvalues of $G_{\sigma,\tau}$, from which we can calculate $M$. We construct the circuit in Figure 34, with $U = G_{\sigma,\tau}$.



The first register will contain

$$t = s + \lceil \log(2 + 1/(2\epsilon)) \rceil$$

qubits with input $|0\rangle^{\otimes t}$. The input to the second register is, as in Equation 4, the uniform superposition state

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

which by Equation 5 is a linear combination of $|\sigma\rangle$ and $|\tau\rangle$, hence a linear combination of $|a\rangle$ and $|b\rangle$.

By Exercise 9.6, measurement of the first register will provide, with a probability of at least $1 - \epsilon$, an approximation to $\theta$ or $-\theta$ correct to $s$ bits:

$$|\Delta\theta| \leq 2^{-s}. \tag{20}$$

From Equation 19 and our approximation for $\theta$ we get an approximation for $M$.

What is the bound on the error $\Delta M$ for $M$? Note that if we have the functional relation $g(y) = h(x)$ where $x$ and $y$ are dependent variables then:

$$\frac{\Delta y}{\Delta x} \approx \frac{dy}{dx} = h'(x)/g'(y).$$



We assume that $s$ is large and differentiate Equation 19 to obtain:

$$\frac{\Delta M}{\Delta\theta} \approx \sqrt{MN} \cos\theta/2,$$

from which it follows, by Equation 20, that

$$\Delta M \leq 2^{-s}\sqrt{MN}. \tag{21}$$

Consider an example. Let $s = \lfloor n/2 \rfloor + 1$ and $\epsilon = 1/6$; then $t = \lfloor n/2 \rfloor + 4$ and the algorithm requires $O(2^{n/2}) = O(\sqrt{N})$ iterations of $G$. Furthermore, for $M \geq 1$, Inequality 21 gives:

$$\Delta M < \sqrt{M} = O(\sqrt{M}). \tag{22}$$

We can also determine, using this algorithm, if a solution exists at all, i.e. whether $M = 0$ or $M \geq 1$. Consider the above example again.

- If $M = 0$, then, by Inequality 21, $\Delta M = 0$ so that the algorithm, with probability at least $5/6$, gives the correct answer $M = 0$.
- If $M \geq 1$, by Inequality 22, $\Delta M < M$ so that the algorithm proves that there exists a solution with probability at least $5/6$.

# 11 Order Finding

Suppose the positive integers $x$ and $N$ have no common factors and $x < N$. Then the *order* of $x$ modulo $N$ is defined to be the least positive integer $r$ such that $x^r = 1 (\text{mod } N)$. We put $L = \lceil \log N \rceil$ so that $N \leq 2^L$. There are no known classical polynomial time algorithms in $L$ to solve this problem. We will show below that there is an $O(L^4)$ quantum algorithm for this task.

Recall that for positive integers $x$ and $n$, the remainder of division of $x$ by $n$ is called $x$ modulo $n$ and is written $x \text{ mod } n$. More precisely, $x$ can be written uniquely as $x = kn + r$ where $k$ is nonnegative and $0 \leq r \leq n - 1$ is the remainder: $r = x \ (\text{mod } n)$. The great German mathematician Karl Friedrich Gauss introduced *Arithmetic modulo n* early in the nineteenth century. It is simply the operations of additions, subtraction, multiplication and division of integers when we take the result modulo $n$. It should be clear how one adds, subtracts and multiplies modulo $n$. But how do we divide modulo $n$?



What is the inverse of an integer modulo $n$? Note that in ordinary arithmetic of integers, the integers 1 and $-1$ are the only integers which have an inverse. In modular arithmetic, many more integers have an inverse as we will now see.

We denote by $\gcd(a, b)$ the greatest common divisor of integers $a$ and $b$. We aim to show that $x$ will have an inverse modulo $n$ iff $\gcd(x, n) = 1$.

**Proposition 11.1** *The greatest common divisor of two integers $a$ and $b$ is the least positive integer which can be written in the form $ar + bs$ where $r$ and $s$ are integers.*

**Proof** Assume $u = ar + bs$ is the least positive integer which can be written as such. Since $\gcd(a, b)$ divides both $a$ and $b$ it therefore divides $u$ and hence $\gcd(a, b) \leq u$. In order to show that $\gcd(a, b) \geq u$, we will prove that $u$ divides both $a$ and $b$. Suppose to the contrary that, for example, $u$ does not divide $a$. Then $a = mu + n$ where $1 \leq n \leq u - 1$. Therefore $n = a - m(ar + bs) = a(1 - mr) - bms$ which contradicts the assumption that $u$ is minimal. $\square$



**Corollary 11.2** *An integer $c$ divides both $a$ and $b$ iff $c$ divides $gcd(a, b)$.*

**Proof** The "if" part is easy. To show the "only if" part, assume $c$ divides both $a$ and $b$. Then $c$ divides any linear combination $ar + bs$, and in particular it divides $\gcd(a, b)$. $\square$

We say $x$ and $n$ are *co-prime* if $\gcd(x, n) = 1$.

**Corollary 11.3** *An integer $x$ has an inverse modulo $n > 1$ iff $x$ and $n$ are co-prime.*

**Proof** Suppose $x$ has an inverse modulo $n$ which we denote by $x^{-1}$. Then $xx^{-1} = 1 + nk$ for some integer $k$. Hence $1 = xx^{-1} - nk$ and therefore $\gcd(x, n) = 1$. On the other hand, if $\gcd(x, n) = 1$ then $xr + ns = 1$ for some integers $r$ and $s$. Hence $r$ is an inverse of $x$ modulo $n$. $\square$

**Exercise 11.4** (i) Find the inverse of 8 (mod 21).

(ii) Show that if $a$ and $b$ are both inverses of $x$ modulo $n$ then $a = b$ (mod $n$).

(iii) Show that the order $r$ of $x$ modulo $n$ satisfies $r \leq n$. $\square$



## 11.1 Quantum Order Finding

In order to find the order $r$ of $x$ modulo $N$ (with $\gcd(x, N) = 1$)), we apply the phase estimation algorithm to the operation $U$ defined by:

$$U|y\rangle = |xy (\text{mod } N)\rangle,$$

for $0 \leq y \leq N - 1$ and $U|y\rangle = |y\rangle$ for $N \leq y \leq 2^L - 1$. For $0 \leq s \leq r - 1$, let:

$$|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} \exp\left(\frac{-2\pi i s k}{r}\right) |x^k \ (\text{mod } N)\rangle.$$

Then, using $x^r = 1$ (mod $N$), we have

$$\begin{aligned} U|u_s\rangle &= \tfrac{1}{\sqrt{r}} \textstyle\sum_{k=0}^{r-1} \exp\left(\tfrac{-2\pi i s k}{r}\right) |x^{k+1} \ (\text{mod } N)\rangle \\ &= \exp\left(\tfrac{2\pi i s}{r}\right) |u_s\rangle, \end{aligned}$$

and hence $|u_s\rangle$ is an eigenvector of $U$ with eigenvalue $\exp(\frac{2\pi i s}{r})$ for each $s$ with $0 \leq s \leq r - 1$.

Our strategy will be to obtain, using the phase estimation algorithm, accurate approximation for the phase $s/r$ which will give us the value of $r$.

**Exercise 11.5** Show that $U$ is unitary. $\square$

However, we do not have a preparation for $|u_s\rangle$ as this requires the value of $r$ which is to be found. Instead, we note that

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

since the expression in the bracket is 1 if $k=0$ and is 0 otherwise. Therefore, we can use the input $|1\rangle$ in the second register which is a linear combination of eigenvectors of $U$. We take $n = 2L+1$ (for reasons which will become clear in Section 11.3) so that $t = 2L+1+\lceil\log(2+\frac{1}{2\epsilon})\rceil$. It follows from Exercise 9.6 that measurement of the first register will produce, for each $s$ with $0 \le s \le r-1$, an approximation $\phi_s$ to $s/r$ correct up to $2L+1$ bits with a probability of at least $(1-\epsilon)/r$. We still have the following two problems to resolve in the next two sections:

(i) How do we perform the sequence of controlled-$U^{2^k}$ gates?

(ii) Given the approximation $\phi_s$, how do we finally compute the order $r$?



## 11.2 Controlled-$U^{2^k}$ Sequence

We need to compute the sequence of controlled-$U^{2^k}$ operations, which had been regarded as black boxes in the phase estimation algorithm. By Equation 10, the state of the first register after the Hadamard gates is

$$\tfrac{1}{\sqrt{2^t}}\textstyle\sum_{k=0}^{2^t-1}|k\rangle.$$

Consider the state $|k\rangle = |k_t k_{t-1}\cdots k_1\rangle$ of the first register and the state $|u\rangle$ of the second register. Then, the sequence of the controlled-$U^{2^k}$ operations gives:

$$\begin{array}{rcl} |k\rangle|u\rangle & \mapsto & |k\rangle U^{k_t 2^{t-1}}U^{k_{t-1}2^{t-2}}\cdots U^{k_1 2^0}|u\rangle \\ & = & |k\rangle|x^{k_t 2^{t-1}}x^{k_{t-1}2^{t-2}}\cdots x^{k_1 2^0}u \pmod N\rangle \\ & = & |k\rangle|x^{k_t 2^{t-1}+k_{t-1}2^{t-2}\cdots+k_1 2^0}u \pmod N\rangle \\ & = & |k\rangle|x^k u \pmod N\rangle \end{array}$$

Therefore, the result of the controlled operations is equivalent to multiplying modulo $N$ the second register with $x$ to the power of the content of the first register. This is computed using the scheme of reversible computation in Section 4.4.



One first computes, in a reversible way, $x^k \pmod N$ in a third register and then multiplies, in a reversible way, the content of the second register with $x^k \pmod N$, deleting the content of the third register, the garbage. The reversible computation of $x^k \pmod N$ is done in two stages. First, we apply modular multiplication to compute $x^2 \pmod N$, then $x^4 \pmod N$, then $x^8 \pmod N$ and so on up to $x^{2^{t-1}} \pmod N$ where we use $t = 2L+1+\lceil\log(2+1/(2\epsilon))\rceil$.

This requires $t-1 = O(L)$ squaring operations each costing $O(L^2)$ if we use ordinary rules of multiplication. (There is a fast $O(L\log L\log\log L)$ algorithm for multiplication of large integers.) It takes therefore $O(L^3)$ basic operations to complete the first stage.

In the second stage, we use the identity

$$x^k u \pmod N = x^{k_t 2^{t-1}+k_{t-1}2^{t-2}\cdots+k_1 2^0}u \pmod N$$

to obtain $x^k u \pmod N$ by $t-1$ further multiplications, with a complexity of $O(L^3)$. Therefore, the operation $|k\rangle|u\rangle \mapsto |k\rangle|x^k u \pmod N\rangle$ can be done with $O(L^3)$ gates.



## 11.3 Continued fractions

Suppose we have made a measurement of the output of the first register and obtained the rational number $\phi_s$ which we know is an approximation to the phase $s/r$. We can obtain $r$ from $\phi_s$ by the technique of continued fractions.

Continued fractions enable us to approximate any real number with a sequence of rational numbers of the form

$$[a_0, a_1, a_2, \cdots, a_p] \stackrel{\text{def}}{=} a_0 + \cfrac{1}{a_1 + \frac{1}{a_2+\frac{1}{\cdots+\frac{1}{a_p}}}}, \quad (23)$$

where $a_j$ is a positive integer for $j \ge 1$. We will explain how to obtain the so-called *regular* continued fraction expansion of a real number $c$. We define two sequences $\langle a_j\rangle_{j\ge 0}$ and $\langle r_j\rangle_{j\ge 0}$ recursively. We put $a_0 = \lfloor c\rfloor$, $r_0 = c - a_0$ and define inductively a sequence of "split and invert" operations for $j \ge 1$:

$$a_j = \left\lfloor \frac{1}{r_{j-1}} \right\rfloor \qquad r_j = \frac{1}{r_{j-1}} - \left\lfloor \frac{1}{r_{j-1}} \right\rfloor.$$

For each $j \ge 0$ with $r_j > 0$, we have:

$$c = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_j + r_j}}}}$$

The rational number $[a_0, a_1, a_2, \cdots, a_j]$, is called the *jth convergent* of $c$. If $r_j = 0$ then the continued fraction terminates with $a_j$ and we obtain $c = [a_0, a_1, a_2, \cdots, a_j]$.

**Example 11.6** We find the continued fraction expansion of 47/13:

$$\begin{aligned}
\tfrac{47}{13} &= 3 + \tfrac{8}{13} = 3 + \tfrac{1}{\frac{13}{8}} \\
&= 3 + \tfrac{1}{1+\frac{5}{8}} = 3 + \tfrac{1}{1+\frac{1}{\frac{8}{5}}} \\
&= 3 + \tfrac{1}{1+\frac{1}{1+\frac{3}{5}}} = 3 + \tfrac{1}{1+\frac{1}{1+\frac{1}{\frac{5}{3}}}} \\
&= 3 + \tfrac{1}{1+\frac{1}{1+\frac{1}{1+\frac{2}{3}}}} = 3 + \tfrac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{\frac{3}{2}}}}} \\
&= 3 + \tfrac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{2}}}}}
\end{aligned}$$

It follows that $\frac{47}{13} = [3, 1, 1, 1, 1, 2]$. $\square$



The sequence of convergents of any real number converges to it. Here, we are only interested in the continued fraction expansion of rational numbers.

**Proposition 11.7** *The continued fraction expansion of a real number c terminates iff c is a rational number.*

**Proof** If the continued fraction terminates, then for some $j \ge 0$ we have $c = [a_0, a_1, a_2, \cdots, a_j]$ which is a rational number. On the other hand, assume, for a contradiction, that the continued fraction expansion of a rational number $c$ does not terminate. Thus, $r_j$ is a positive rational number for each $j \ge 0$. Let $r_j = u_j/v_j$ with $u_j > 0$ where $\gcd(u_j, v_j) = 1$ for all $j \ge 0$. Dividing $v_j$ by $u_j$, we have $v_j = m_j u_j + n_j$ where $0 \le n_j < u_j$. Then

$$\frac{u_{j+1}}{v_{j+1}} = r_{j+1} = \frac{1}{r_j} - \left\lfloor \frac{1}{r_j} \right\rfloor = \frac{m_j u_j + n_j}{u_j} - m_j = \frac{n_j}{u_j}.$$

Therefore, $u_{j+1} \le n_j < u_j$, i.e. $u_j$ is a strictly decreasing sequence of positive integers, which is a contradiction. $\square$

Since $[a_0, \cdots, a_{j-1}, a_j] = [a_0, \cdots, a_{j-1}, a_j - 1, 1]$, we can assume $j$ is even or odd as desired.



**Proposition 11.8** *We have* $[a_0, \cdots, a_j] = \frac{p_j}{q_j}$ *where* $p_0 = a_0$, $q_0 = 1$, $p_1 = 1 + a_0 a_1$, $q_1 = a_1$ *and for* $j \ge 2$:

$$p_j = a_j p_{j-1} + p_{j-2} \qquad q_j = a_j q_{j-1} + q_{j-2}.$$

**Proof** We use induction. The cases for $0 \le j \le 2$ are trivial to check. For $j \ge 3$, we have by definition:

$$[a_0, \cdots, a_j] = [a_0, \cdots, a_{j-1} + \frac{1}{a_j}]. \qquad (24)$$

Let $\hat{p}_l/\hat{q}_l$ be the sequence of convergents for the expansion on the right hand side of Equation 24, so that by the induction hypothesis,

$$[a_0, \cdots, a_{j-1} + \frac{1}{a_j}] = \frac{\hat{p}_{j-1}}{\hat{q}_{j-1}}. \qquad (25)$$

Since $\hat{p}_l = p_l$ and $\hat{q}_l = q_l$, for $0 \le l \le j-2$, we get:

$$\begin{aligned}
\tfrac{\hat{p}_{j-1}}{\hat{q}_{j-1}} &= \tfrac{(a_{j-1}+1/a_j)p_{j-2}+p_{j-3}}{(a_{j-1}+1/a_j)q_{j-2}+q_{j-3}} = \tfrac{p_{j-1}+p_{j-2}/a_j}{q_{j-1}+q_{j-2}/a_j} \\
&= \tfrac{a_j p_{j-1}+p_{j-2}}{a_j q_{j-1}+q_{j-2}} = \tfrac{p_j}{q_j}.
\end{aligned}$$

From this and Equations 24 and 25, we obtain: $p_j/q_j = [a_0, \cdots, a_j]$, as required. $\square$



**Corollary 11.9** *The continued fraction expansion of any positive rational number* $p/q$ *can be obtained in* $O(m^3)$ *operations if* $p$ *and* $q$ *are* $m$ *bit integers.*

**Proof** We can assume that $p/q \ge 1$ since if $0 < p/q < 1$ then $p/q = 0 + \frac{1}{q/p}$ and we can obtain the continued fraction expansion of $p/q$ from that of $q/p > 1$. Assume that the continued fraction expansion of $p/q$ terminates after $n$ split and invert operations, i.e. $p/q = p_n/q_n$. Since $a_j$ is positive for all $j$, it follows from the definition that $p_j$ and $q_j$ are nondecreasing sequences and therefore we have: $p_j = a_j p_{j-1} + p_{j-2} \ge 2p_{j-2}$ and similarly $q_j \ge 2q_{j-2}$. Hence, $p_j, q_j \ge 2^{\lfloor j/2 \rfloor}$ and therefore, $p \ge q \ge 2^{\lfloor n/2 \rfloor}$. It follows that $n = O(\log p) = O(m)$. Therefore, the continued fraction expansion of $p/q$ can be computed in $O(m)$ split and invert operations, each costing $O(m^2)$ for arithmetic gates. Hence, overall, it requires $O(m^3)$ basic operations. $\square$

**Exercise 11.10** Show by induction that

$$p_{j-1} q_j - q_{j-1} p_j = (-1)^j$$

and deduce that $\gcd(p_j, q_j) = 1$. $\square$

**Theorem 11.11** *Suppose $x$ and $p/q$ are rational numbers such that $|x - p/q| \leq 1/2q^2$. Then $p/q$ is a convergent of the continued fraction for $x$.*

**Proof** Let $p/q = [a_0, \cdots, a_n]$ with $n$ even and let $p_j, q_j$ be as in Proposition 11.8 so that $p/q = p_n/q_n$. By Exercise 11.10, $p_n$ and $q_n$ are co-prime; hence $q_n \leq q$. Let $\delta = 2q_n^2(x - p_n/q_n)$ so that $|\delta| \leq q_n^2/q^2 \leq 1$. The case $\delta = 0$ is trivial and we can assume, by replacing $\delta$ with $-\delta$ if necessary, that $0 < \delta \leq 1$. If we put

$$\lambda = 2\left(\frac{q_n p_{n-1} - p_n q_{n-1}}{\delta}\right) - \frac{q_{n-1}}{q_n},$$

then a simple calculation shows that

$$x = \frac{\lambda p_n + p_{n-1}}{\lambda q_n + q_{n-1}},$$

which implies $x = [a_0, \cdots, a_n, \lambda]$. By Exercise 11.10, $\lambda = 2/\delta - q_{n-1}/q_n$, since we have chosen $n$ to be even. Recalling that $q_j$ is increasing, we get: $\lambda = 2/\delta - q_{n-1}/q_n \geq 2 - 1 = 1$. Thus, $\lambda = [b_0, \cdots, b_l]$ with $b_0 \geq 1$. Hence, we have $x = [a_0, \cdots, a_n, b_0, \cdots, b_l]$ and $p/q$ is a convergent for the continued fraction expansion of $x$. □



We now return to the order finding algorithm. Recall that measurement of the first register will produce, for each $s$ with $0 \leq s \leq r-1$, a rational approximation $\phi_s$ to $s/r$ correct up to $2L+1$ bits with a probability of at least $(1-\epsilon)/r$ satisfying $|\phi_s - s/r| < 1/2^{2L+1} \leq 1/2r^2$ since $r \leq N \leq 2^L$. Hence, by Theorem 11.11, $s/r$ is a convergent of the rational number $\phi_s$, which can be computed in $O(L^3)$ basic operations.

**Proposition 11.12** *There is at most one fraction $p/q$ which satisfies both inequaities*

$$|\phi_s - p/q| < 1/2^{2L+1} \quad \& \quad q \leq 2^L. \tag{26}$$

**Proof** In fact, suppose we have a rational number $p'/q'$ satisfying $|\phi_s - p'/q'| < 1/2^{2L+1}$ with $q' \leq 2^L$ and $p/q \neq p'/q'$. Then

$$\begin{aligned} |p/q - p'/q'| &= |p/q - \phi_s + \phi_s - p'/q'| \\ &\leq |p/q - \phi_s| + |\phi_s - p'/q'| < 1/2^{2L}. \end{aligned}$$

But now $|pq' - qp'| \geq 1$ gives a contradiction:

$$|p/q - p'/q'| = \frac{|pq' - p'q|}{qq'} \geq \frac{|pq' - qp'|}{2^{2L}} \geq 1/2^{2L}. \ \square$$



Suppose now that $s$ and $r$ are co-prime. Then we try to find a convergent $p/q$ of $\phi_s$, with $q \leq 2^L$ and $|\phi_s - p/q| < 1/2^{2L+1}$. The denominator $q$ is a candidate for $r$.

The number, $\Phi(r)$, of positive numbers less than $r$ and co-prime with $r$ satisfies: $\Phi(r) \geq \frac{r}{2\ln\ln r}$ when $r$ is large. The function $\Phi$ is Euler's totient function.

Therefore, the probability that $\gcd(r,s) = 1$ is greater than $1/2\ln\ln r \geq 1/2\lceil\ln\ln N\rceil$.

Note that for large $n$ we have $(1-1/n)^n \approx e^{-1}$. Hence, if we repeat the algorithm $2\lceil\ln\ln N\rceil$ times, with probability $\geq 1-(1-1/2\lceil\ln\ln N\rceil)^{2\lceil\ln\ln N\rceil} \approx 1-1/e$, we will have an instance of $s/r$ with $s$ and $r$ co-prime.

If we do not succeed in finding a convergent $p/q$ of $\phi_s$, with $q \leq 2^L$ and $|\phi_s - p/q| < 1/2^{2L+1}$, then we start all over again. Note that the algorithm is probabilistic: The probablity of success after $2\lceil\ln\ln N\rceil$ rounds of trial is $(1-\epsilon)(1-1/e)$.



## 11.4 Summary

To find the order $r$ of the positive integer $x$ modulo $N$:

(i) Put $L = \lceil\log N\rceil$ and $t = 2L+1+\lceil\log(2+1/(2\epsilon))\rceil$, for small $\epsilon > 0$, in the phase estimation circuit:

(ii) Prepare the first register in the state $|0\rangle^{\otimes t}$ and the second register in the state $|1\rangle$.

(iii)
- Apply the Hadamard gates to the first register ($O(L)$ operations),
- then apply the sequence of controlled-$U$ operations, where $U|y\rangle = |xy \pmod N\rangle$, to both registers ($O(L^3)$ operations), and finally,
- apply the inverse Fourier transform to the first register ($O(L^2)$ operations).

(iv) Measure the first register to obtain $\phi_s$ an approximation to $s/r$.

(v) Apply the continued fraction algorithm and find $s/r$, the only fraction that satisfies the two inequalities 26, as a convergent of $\phi_s$ from which we can find $r$ if $\gcd(r,s) = 1$ ($O(L^3)$ operations).

(vi) The algorithm succeeds with probability $\geq (1-\epsilon)(1-1/e)$ if it is repeated $2\lceil\ln\ln N\rceil = O(\log L)$ times, i.e. for the cost of $O(L^3\log L)$ operations.

# 12 Quantum Factoring

Any positive integer is the product of powers of some prime numbers. The problem of finding the prime factors of a positive integer is thought to be intractable. Existing algorithms for factoring are exponential. For example, it took 1000 computers around the world two whole months in 1990 to factorize the ninth Fermat number $F_9 = 2^{2^9} + 1$, which has 155 decimal digits. It is estimated that it will take more than the age of the universe to factorize numbers with 200 digits.

The RSA encryption scheme which is the most important and widely used cryptographic technique is based on the assumption that factoring large numbers is an intractable problem.

Remarkably, the quantum order finding algorithm can be used to provide a polynomial algorithm for factoring. We need a few results from basic number theory. First, we will describe Euclid's algorithm for finding the gcd of two integers, which is based on the following result.



**Proposition 12.1** *Let $p$ and $q$ be integers and $r$ the remainder of division of $p$ by $q$. Assuming that $r \neq 0$ we have: $gcd(p, q) = gcd(q, r)$.*

**Proof** We show that each side divides the other side. There exists an integer $a$ such that $r = p - aq$. It follows that $\gcd(p, q)$, which divides $p$ and $q$, also divides $r$. Hence, by Corollary 11.2, $\gcd(p, q)$ divides $\gcd(q, r)$. Furthermore, $\gcd(q, r)$ divides $q$ and, since $p = aq + r$, it also divides $p$. Hence, by Corollary 11.2 again, $\gcd(q, r)$ divides $\gcd(p, q)$. □

Euclid's algorithm to find the gcd of integers $p$ and $q$ with $p > q$ is as follows. We divide $p$ by $q$ to find the remainder $r_1 < q$. By Proposition 12.1, $\gcd(p, q) = \gcd(q, r_1)$. We now divide $q$ by $r_1$ to find the remainder $r_2 < r_1$ with $\gcd(q, r_1) = \gcd(r_1, r_2)$. Inductively, we divide $r_{n-2}$ by $r_{n-1}$ to find the remainder $r_n < r_{n-1}$ with

$$\gcd(p, q) = \gcd(q, r_1) = \cdots = \gcd(r_n, r_{n+1}).$$

since $r_j$ is a strictly decreasing sequence of positive integers, there exists $n$ with $r_{n+1} = 0$, i.e. $r_{n-1} = ar_n$. Hence, the algorithm terminates with $\gcd(p, q) = \gcd(r_{n-1}, r_n) = r_n$.



What is the complexity of Euclid's algorithm? Suppose $p$ and $q$ are $L$ bit integers. Then $r_j$ will be an $L$ bit integer for all $j$. Each division in the algorithm costs $O(L^2)$. How many divisions are needed? We show that $r_{j+2} \le r_j/2$. This is clear if $r_{j+1} \le r_j/2$. Otherwise, if $r_{j+1} > r_j/2$, then $r_j = 1 \times r_{j+1} + r_{j+2}$ and hence $r_{j+2} = r_j - r_{j+1} \le r_j/2$. We conclude that at most $2\lceil \log p \rceil = O(L)$ divisions are required at the total cost of $O(L^3)$.

**Theorem 12.2** *Let $N$ be a composite number and $y \in \{1, \cdots, N\}$ a solution of $y^2 = 1 \pmod N$ with $y \neq 1 \pmod N$ and $y \neq N - 1 \pmod N$. Then, $gcd(y - 1, N)$ and $gcd(y + 1, N)$ are non-trivial factors of $N$ which can be computed in $O(\lceil \log N \rceil^3)$ operations.*

**Proof** Since $y^2 = 1 + Nk$, for some integer $k$, it follows that $N$ has a common factor with $y - 1$ or with $y + 1$. But $1 < y < N - 1$ implies $y - 1 < y + 1 < N$ so that $N$ itself cannot be the common factor. We therefore use Euclid's algorithm, in $O(\lceil \log N \rceil^3)$ operations, to find $\gcd(y - 1, N)$ and $\gcd(y + 1, N)$ which are non-trivial factors of $N$. □



We need one further result whose proof is given in Nielsen & Chuang, p. 634.

**Theorem 12.3** *Let $N = p_1^{n_1} \cdots p_m^{n_m}$ be the prime factorization of the composite odd integer $N$. Let $x$ be a randomly chosen integer in the range $1 \le x \le N$ subject to the condition that $x$ and $N$ are co-prime. If $r$ is the order of $x \pmod N$, then:*

$$p(r \text{ is even and } x^{r/2} \neq -1 \pmod N) \ge 1 - 1/2^m.$$

**Shor's quantum factoring algorithm.** Suppose the $L$ bit composite odd integer $N$ is given. To compute a non-trivial factor of $N$:

(i) Randomly choose $x$ with $1 \le x \le N - 1$.

(ii) If $\gcd(x, N) > 1$ then return $\gcd(x, N)$.

(iii) If $\gcd(x, N) = 1$, use the order finding algorithm to find the order $r$ of $x$ mod $N$.

(iv) If $r$ is odd, or if $r$ is even and $x^{r/2} = -1 \pmod N$, return to (i).

(v) Compute and return $\gcd(x^{r/2} - 1, N)$ and $\gcd(x^{r/2} + 1, N)$ using Euclid's algorithm.

The algorithm takes $O(L^3 \log L)$ basic operations.

# 13 Physical Realization

We will describe how the NOT gate, the simplest of all gates can be physically implemented as a two level quantum system. Quantum gates, some of which we have studied in detail, necessarily have an evolution in *discrete* time. But physical systems evolve in *continuous* time. Therefore, in order to see how to implement a quantum gate as a physical system, we need to know how quantum systems evolve in continuous time.

The time evolution of a closed quantum system is given by the *Schrödinger equation*, named after Erwin Schrödinger the prominent Austrian physicist who discovered it in 1920's:

$$i\hbar \frac{d|\psi\rangle}{dt} = \mathcal{H}|\psi\rangle \qquad (27)$$

where $\hbar$ is *Planck's constant*, a physical constant whose numerical value does not concern us, and $\mathcal{H}$ is the *Hamiltonian* of the system, a Hermitian operator (i.e. $\mathcal{H} = \mathcal{H}^\dagger$), which describes the energy of the system. The Hamiltonian gives all the information about the system.



Since $\mathcal{H}$ is diagonalizable by Corollary 9.3, it has with respect to some computational basis, a matrix representation of the form:

$$\mathcal{H} = \sum_i E_i |E_i\rangle\langle E_i|,$$

where $E_i$ is the eigenvalue, called the *energy level*, corresponding to the eigenvector $|E_i\rangle$, called the *energy eigenstate*.

**Theorem 13.1** *The eigenvalues of any Hermitian operator are real.*

**Proof** Let $L$ be any Hermitian operator with normalized eigenvector $v$ corresponding to eigenvalue $\lambda$. Then $Lv = \lambda v$ and hence $\langle v|Lv\rangle = \lambda\langle v|v\rangle = \lambda$. Since, for any vectors, $u$ and $w$, we have: $\langle u|Lw\rangle = \langle L^\dagger u|w\rangle$ and $\langle u|w\rangle = \langle w|u\rangle^*$, we deduce, using $L = L^\dagger$, that:

$$\lambda = \langle v|Lv\rangle = \langle L^\dagger v|v\rangle = \langle Lv|v\rangle = \langle v|Lv\rangle^* = \lambda^*. \ \square$$

The eigenvalues of a Hermitian matrix correspond to *observable* quantities, e.g. energy, spin etc. The lowest energy level is called the *ground energy level* and its corresponding eigenstate the *ground state*.



The evolution of an energy eigenstate is simple. From Schrödinger equation we get:

$$i\hbar \frac{d|E(t)\rangle}{dt} = \mathcal{H}|E(t)\rangle = E|E(t)\rangle,$$

or, $\frac{d|E(t)\rangle}{dt} = -\frac{iE}{\hbar}|E(t)\rangle$; We obtain:

$$|E(t)\rangle = e^{\frac{-iEt}{\hbar}}|E(0)\rangle; \qquad (28)$$

at $t$ the vector $|E(0)\rangle$ is simply multiplied by $e^{\frac{-iEt}{\hbar}}$. Notice that we can write $|E(t)\rangle$ as:

$$\begin{aligned}
&|E(t)\rangle \\
&= e^{\frac{-iEt}{\hbar}}|E(0)\rangle \\
&= (1 + \tfrac{(\frac{-iEt}{\hbar})^1}{1!} + \tfrac{(\frac{-iEt}{\hbar})^2}{2!} + \tfrac{(\frac{-iEt}{\hbar})^3}{3!} + \cdots)|E(0)\rangle \\
&= |E(0)\rangle + \tfrac{(\frac{-iEt}{\hbar})^1}{1!}|E(0)\rangle + \cdots + \tfrac{(\frac{-iEt}{\hbar})^n}{n!}|E(0)\rangle + \cdots \\
&= |E(0)\rangle + \tfrac{(\frac{-iEt}{\hbar})^1}{1!}|E(0)\rangle + \cdots + \tfrac{(\frac{-iEt}{\hbar})^n}{n!}|E(0)\rangle + \cdots \\
&= e^{\frac{-i\mathcal{H}t}{\hbar}}|E(0)\rangle.
\end{aligned}$$

Recall that for any operator $L$ and state $|\psi\rangle$, we formally define

$$e^L|\psi\rangle = |\psi\rangle + \frac{L^1}{1!}|\psi\rangle + \frac{L^2}{2!}|\psi\rangle + \cdots + \frac{L^n}{n!}|\psi\rangle + \cdots.$$



More generally, if $\mathcal{H}$ is any time independent Hamiltonian, we have

$$|E(t)\rangle = e^{\frac{-iE\mathcal{H}}{\hbar}}|E(0)\rangle = U(t)|E(0)\rangle. \qquad (29)$$

**Exercise 13.2** (i) Prove that $|E(t)\rangle = e^{\frac{-it\mathcal{H}}{\hbar}}|E(0)\rangle$ satisfies the Schrödinger equation for any time independent $\mathcal{H}$.

(ii) Show that the operator $U(t) = e^{\frac{-it\mathcal{H}}{\hbar}}$ is unitary.

(iii) For a unitary matrix $U$, define $\log U$, show that $-i\log U$ is Hermitian and deduce that $U = \exp(iK)$ for some Hermitian $K$. $\square$

Exercise 13.2 implies that there is a close relationship between Hermitian and unitary matrices:

- Any Hermitian matrix can be regarded as the Hamiltonian of a quantum system which gives rise via the Schrödinger equation to a unitary matrix describing the time evolution of the quantum system.
- Conversely, any unitary matrix describes the time evolution of a quantum system arising from some Hermitian matrix which is the Hamiltonian of the quantum system.

## 13.1 Measuring Physical Values

Suppose $A : \mathbb{C}^n \to \mathbb{C}^n$ is a Hermitian matrix which has eigenvalues $\lambda_j$ (where $1 \le j \le k$ and $k \le n$) with multiplicity $n_j$ . We know by Algorithm 9.4 that the collection $\{v_{jm} | 1 \le m \le n_j,\ 1 \le j \le k\}$ of eigenvectors of $A$ forms an orthonormal basis for $\mathbb{C}^n$. Recall that the eigenvectors $v_{jm}$ for $1 \le m \le n_j$ form an orthonormal basis for the eigenspace corresponding to eigenvalue $\lambda_j$. Let $P_j = \sum_{m=1}^{n_j} |v_{jm}\rangle\langle v_{jm}| : \mathbb{C}^n \to \mathbb{C}^n$.

**Definition 13.3** Let $S \subseteq \mathbb{C}^n$ be a $t$-dimensional subspace of $\mathbb{C}^n$ with orthonormal basis $|w_m\rangle$ ($1 \le m \le t$). Then the operator $P_S : \mathbb{C}^n \to \mathbb{C}^n$ defined by $P_S = \sum_{m=1}^{t} |w_m\rangle\langle w_m|$ is called *projection into* $S$. More abstractly, a Hermitian operator $P : \mathbb{C}^n \to \mathbb{C}^n$ is called a *projection* if $P \circ P = P$.

**Exercise 13.4** (i) Show that $P_S \circ P_S = P_S$ and that any projection operator $P$ is in fact projection into some subspace of $\mathbb{C}^n$.

(ii) Check that $P_j$ is the projection to the eigenspace corresponding to eigenvalue $\lambda_j$.

(iii) Show that $A = \sum_{j=1}^{k} \lambda_j P_j$. □



**Principle of measurement of an observable.** Given a state $|\psi\rangle$, a measurement of the observable (Hermitian operator) $A = \sum_{j=1}^{k} \lambda_j P_j$ has outcome $\lambda_j$ with probability

$$p(\lambda_j) = \langle\psi|P_j|\psi\rangle = \langle\psi|P_jP_j|\psi\rangle = \|P_j|\psi\rangle\|^2.$$

If the outcome is $\lambda_j$, then the state collapses into the state:

$$\frac{P_j|\psi\rangle}{\sqrt{p(\lambda_j)}}.$$

Let $|\psi\rangle = \sum_{jm} \alpha_{jm}|v_{jm}\rangle$ be an arbitrary state expressed in terms of the orthonormal basis.

**Exercise 13.5** Show that the probability of obtaining $\lambda_j$ is $p(\lambda_j) = \sum_{m=1}^{n_j} |\alpha_{jm}|^2$. □

We have:

$$\langle\psi|A|\psi\rangle = \sum_{j=1}^{k}\sum_{m=1}^{n_j} |\alpha_{jm}|^2\lambda_j = \sum_{j=1}^{k} p(\lambda_j)\lambda_j. \quad (30)$$

Since $p(\lambda_j)$ is the probability that $|\psi\rangle$ has value $\lambda_j$ for the observable $A$, Equation 30 implies that $\langle A\rangle \stackrel{\text{def}}{=} \langle\psi|A|\psi\rangle$ is the *average*, or *expected*, value of the operator $A$.



**Example 13.6** Suppose a single qubit has Hamiltonian

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

Then the energy levels are $-1$ (ground level) and 1 (excited level), corresponding respectively to the eigenstates $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$ (ground state) and $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ (excited state). Given a state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, the probability of measuring $-1$ is $|\langle -|\psi\rangle|^2 = |\alpha - \beta|^2/2$ whereas the probability of measuring 1 is $|\langle +|\psi\rangle|^2 = |\alpha + \beta|^2/2$. The average value of $\mathcal{H}$ is $\langle\psi|\mathcal{H}|\psi\rangle = \alpha\beta^* + \alpha^*\beta$. □

**Exercise 13.7** Find the energy levels and the energy eigenstates of the quantum systems with the following Hamiltonians:

(i) Pauli-$Y$: $Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$.

(ii) Hadamard: $H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

(ii) $\mathsf{CNOT}$:$\frac{1}{2}(|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 11| + |11\rangle\langle 10|)$.

In each case determine the probabilities of measuring the energy levels and the expected value of the Hamiltonian for an arbitrary state. □



## 13.2 Implementation of NOT

There are a number of two-level quantum systems with which one can implement the NOT gate.

We will take the proton spin in a uniform magnetic field $B$ in the $z-$direction.

The Schrödinger equation of the system is: $i\hbar\frac{d|\psi\rangle}{dt} = \mathcal{H}|\psi\rangle$, where the Hamiltonian is given by

$$\mathcal{H} = -\gamma\hbar B^z I^z = -\frac{1}{2}\gamma\hbar BZ = -\frac{1}{2}\hbar\omega_0 Z,$$

where $I^z = \frac{1}{2}Z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ describes the $z-$component of spin, $B^z = B$, $\gamma$ is the proton gyromagnetic ratio and $\omega_0 = \gamma B$ is the *eigenfrequency* of the system.

The two energy levels of the system are $-\hbar\omega_0/2$ (ground level) and $\hbar\omega_0/2$ (excited level), corresponding respectively to the eigenstates $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (ground state) and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (excited state).

We can write the solution of the Schrödinger equation at time $t$ as

$$|\psi(t)\rangle = c_0(t)|0\rangle + c_1(t)|1\rangle. \quad (31)$$

Substituting this in the Schrödinger equation, and using the notation $\dot{c} = \frac{dc}{dt}$, we get:

$$i\hbar(\dot{c_0}|0\rangle + \dot{c_1}|1\rangle) = -\frac{\hbar\omega_0}{2}(|0\rangle\langle 0| - |1\rangle\langle 1|)(c_0|0\rangle + c_1|1\rangle)$$

$$= -\frac{\hbar\omega_0}{2}(c_0|0\rangle - c_1|1\rangle).$$

Therefore, we get the two differential equations:

$$\dot{c_0} = \frac{i\omega_0}{2}c_0, \qquad \dot{c_1} = -\frac{i\omega_0}{2}c_1,$$

which have solutions:

$$c_0(t) = c_0(0)e^{i\omega_0 t/2}, \qquad c_1(t) = c_1(0)e^{-i\omega_0 t/2}.$$

**Exercise 13.8** Show that, at time $t$, the average values of the $x$, $y$ and $z$-component of spin, given respectively by the Hermitian matrices

$$I^x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad I^y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad I^z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

are as follows:

$$\begin{aligned} \langle\psi(t)|I^x|\psi(t)\rangle &= a\cos(\omega_0 t + \phi) \\ \langle\psi(t)|I^y|\psi(t)\rangle &= -a\sin(\omega_0 t + \phi) \\ \langle\psi(t)|I^z|\psi(t)\rangle &= \tfrac{1}{2}(|c_0(0)|^2 - |c_1(0)|^2), \end{aligned}$$

where $c_0(0)c_1^*(0) = ae^{i\phi}$ with real $a$ and $\phi$. $\square$



It follows from Exercise 13.8 that the average value of $I^z$ does not change with time. Also, the total length of spin is constant:

$$\langle I^x\rangle^2 + \langle I^y\rangle^2 + \langle I^z\rangle^2$$

$$= \frac{1}{4}(|c_0(0)|^4 + |c_1(0)|^4 + 2|c_0(0)|^2|c_1(0)|^2) = \frac{1}{4}.$$

The average value of the spin in the $xy$-plane rotates with frequency $\omega_0$ in the clockwise direction viewed from the top ($z > 0$).

We now apply a transverse circularly polarized magnetic field $(B^x, B^y)$ resonant with the precession of the vector $(\langle I^x\rangle, \langle I^y\rangle)$. Put $B^x = \hbar\cos\omega t$, $B^y = -\hbar\sin\omega t$. The new Hamiltonian is:

$$\begin{aligned} \mathcal{H} &= -\hbar\gamma(B^x I^x + B^y I^y + B^z I^z) \\ &= -\hbar\gamma(B^x I^x + B^y I^y) - \hbar\omega_0 I^z \\ &= -\frac{\hbar\gamma}{2}(B^+ I^- + B^- I^+) - \hbar\omega_0 I^z, \qquad (32) \end{aligned}$$

where

$$B^+ = B^x + iB^y = \hbar e^{-i\omega t}, \quad B^- = B^x - iB^y = \hbar e^{i\omega t},$$

$$I^+ = I^x + iI^y = |0\rangle\langle 1|, \qquad I^- = I^x - iI^y = |1\rangle\langle 0|.$$



Substituting the values of $B^{\pm}$ and $I^{\pm}$ into Equation 32, we obtain the Hamiltonian:

$$\mathcal{H} = \qquad (33)$$

$$-\frac{1}{2}\left[\omega_0(|0\rangle\langle 0| - |1\rangle\langle 1|) + \Omega\left(e^{i\omega t}|0\rangle\langle 1| + e^{-i\omega t}|1\rangle\langle 0|\right)\right],$$

where $\Omega = \gamma\hbar$. Using the Hamiltonian 33 in the Schrödinger equation, we obtain the following equations for the time dependent solution $|\psi(t)\rangle$:

$$\dot{c_0} = \frac{i}{2}(\omega_0 c_0 + \Omega e^{i\omega t}c_1),$$

$$\dot{c_1} = \frac{-i}{2}(\omega_0 c_1 - \Omega e^{-i\omega t}c_0).$$

These equations, which have time-periodic coefficients, reduce to equations with constant coefficients by going to a system of coordinates rotating with the resonant magnetic field:

$$c_0 = c_0' e^{i\omega t/2}, \qquad c_1 = c_1' e^{-i\omega t/2},$$

which give us:

$$\dot{c}_0' = \frac{i}{2}[\Omega c_1' - (\omega - \omega_0)c_0']$$

$$\dot{c}_1' = \frac{i}{2}[\Omega c_0' + (\omega - \omega_0)c_1'].$$



Using the resonant condition, $\omega = \omega_0$, these equations reduce to:

$$\dot{c}_0' = \frac{i\Omega}{2}c_1', \qquad \dot{c}_1' = \frac{i\Omega}{2}c_0'. \qquad (34)$$

In the new coordinates the precession about the $z$-axis is turned off; we only have the transverse constant magnetic field. Dropping the primes, differentiation of Equations 34 gives:

$$\ddot{c}_0 = -\frac{\Omega^2}{4}c_0, \qquad \ddot{c}_1 = -\frac{\Omega^2}{4}c_1,$$

from which we get the general solution:

$$c_0(t) = c_0(0)\cos\frac{\Omega t}{2} + ic_1(0)\sin\frac{\Omega t}{2},$$

$$c_1(t) = ic_0(0)\sin\frac{\Omega t}{2} + c_1(0)\cos\frac{\Omega t}{2}.$$

If initially the system is in the ground state, i.e. $c_0(0) = 1$ and $c_1(0) = 0$, then at time $t$:

$$c_0(t) = \cos\frac{\Omega t}{2}, \qquad c_1(t) = i\sin\frac{\Omega t}{2}.$$

At $t_1 = \pi/\Omega$, we have: $c_0(t_1) = 0$ and $c_1(t_1) = i$ so that $|c_0(t_1)|^2 = 0$ and $|c_1(t_1)|^2 = 1$. Hence, the system will be in the excited state at $t_1$.

Conversely, if the system is initially in the excited state: $c_0 = 0$ and $c_1 = 1$, then at time $t$ we have:

$$c_0(t) = i \sin \frac{\Omega t}{2}, \qquad c_1(t) = \cos \frac{\Omega t}{2}.$$

Hence at time $t_1 = \pi/\Omega$: $c_0(t_1) = i$ and $c_1(t_1) = 0$ so that $|c_0(t_1)|^2 = 1$ and $|c_1(t_1)|^2 = 0$. Therefore, the system will be in the ground state at $t_1$.

We conclude that a pulse of the resonant magnetic field for a period of $\pi/\Omega$, called a *$\pi$ pulse*, acts as a quantum NOT gate.

**Exercise 13.9** (i) Show that under the action of a resonant field the average values of the spin components at time $t$ for the state which is initially in the ground state are given by:

$$\langle I^x \rangle(t) = 0,\ \langle I^y \rangle(t) = \frac{1}{2} \sin \Omega t,\ \langle I^z \rangle(t) = \frac{1}{2} \cos \Omega t.$$

Hence, the average value of the $z$-component goes from $1/2$ at $t = 0$ to $-1/2$ at $t = \pi/\Omega$.

(ii) Show that one can construct, starting with the ground state, an equally weighted superposition of the ground and the excited states with a pulse of the resonant field applied for a suitable time interval. This therefore implements a fair coin.



# 14 Error Correction

We have seen how to implement the quantum NOT gate; other quantum gates can also be implemented in physical systems. In this way, we can build networks of quantum gates to implement quantum algorithms.

However, as we connect more and more quantum gates together we will soon face a serious practical problem. In general, there is no absolutely closed system: the surrounding environment inevitably interacts with any quantum network and destroys the interference pattern which is so crucial in quantum computing. This phenomenon is called *decoherence*, a fundamental issue in quantum computing.

The undesirable interaction of any coded message with its environment in a communication channel is called *noise*. It is also present in classical systems and is in general tackled by introducing redundancy to *encode* the message so that even after the effect of noise from the environment, the original information can be retrieved or *decoded*.



A basic encoding technique against noise is provided by simply multiplying the original information so that if some of the information is corrupted, there would be enough information to reconstruct the intended message.

**Example 14.1** Suppose we intend to send an information bit 0 or 1 through a classical communication channel which flips a bit with probability $p$ and leaves it intact with probability $1-p$. The probability of an error is therefore $p$. We replace the bit with three copies of itself i.e. $0 \mapsto 000$ and $1 \mapsto 111$ and then send the resulting three-bit. At the receiving end, some of the bits may have been flipped. We take a *majority vote* on the three bits to decide what was the original bit. This method succeeds if one or no bit was flipped and fails if two or all three bits were flipped. The probability of error is thus the probability that two or three bits are flipped: $p_e = 3p^2(1-p) + p^3 = 3p^2 - 2p^3$. We have $p_e < p$ if $3p^2 - 2p^3 < p$ which is equivalent to $p(1-p)(2p-1) < 0$, i.e. $p < 1/2$. Therefore, if $p < 1/2$, the three fold multiplication of the original bit will reduce the probability of error. □



## 14.1 Quantum error correction

In formulating error-correcting codes in quantum computation we face a number of problems absent in the classical sphere. These include: (i) By the No-Cloning theorem, we cannot duplicate a quantum state. (ii) Because measurement destroys quantum information, we cannot observe the output channel without losing information. It is however possible to effectively reduce the error by various encoding schemes.

We start with the simplest method which resembles the three fold multiplication of classical bits in Example 14.1. We encode the qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ as the three-qubit state $\alpha|000\rangle + \beta|111\rangle$ by the circuit in Figure 35. The error-correction method, applied after the effect of noise, uses the *general measurement principle*, explained below.

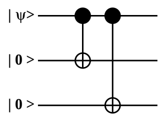



Figure 35: Circuit for qubit to three-qubit state

**The general measurement principle.** General quantum measurements are carried out by a collection of, say, $k$ linear *measurement operators* $M_j : \mathbb{C}^n \to \mathbb{C}^n$ $(1 \le j \le k)$ satisfying

$$\sum_{j=1}^{k} M_j^\dagger M_j = I,$$

where $I$ is the identity matrix. Each $j$ with $1 \le j \le k$ is a possible outcome of the measurement. If the state of the system is $|\psi\rangle$, then the probability of obtaining $j$ is

$$p(j) = \langle\psi|M_j^\dagger M_j|\psi\rangle,$$

and the state of the system after the measurement collapses into

$$\frac{M_j|\psi\rangle}{\sqrt{p(j)}}.$$

**Exercise 14.2** (i) Show that

$$\textstyle\sum_{j=1}^{k} p(j) = 1.$$

(ii) Show that both the basic measurement principle of Section 2.3 and the principle of measuring observables in Section 13.1 are special cases of the general measurement principle. □



We now return to the error-correction method. The initial state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is converted by the circuit of Figure 35 into $|\psi'\rangle = \alpha|000\rangle + \beta|111\rangle$ and then each of the three qubits is sent independently through the bit flip channel. We then apply a two-stage error correction method: We first make a measurement using the four measurement operators, the so-called *error syndromes*:

$$\begin{aligned}
M_0 &= |000\rangle\langle 000| + |111\rangle\langle 111| && \text{no error}\\
M_1 &= |100\rangle\langle 100| + |011\rangle\langle 011| && \text{bit flip on qubit 1}\\
M_2 &= |010\rangle\langle 010| + |101\rangle\langle 101| && \text{bit flip on qubit 2}\\
M_3 &= |001\rangle\langle 001| + |110\rangle\langle 110| && \text{bit flip on qubit 3,}
\end{aligned}$$

which are in fact projections and clearly satisfy $\sum_{j=0}^{3} M_j^\dagger M_j = I$. Hence, the general measurement principle is applicable. Depending on which error syndrome is measured we then recover the original state. Suppose, for example, that a bit flip occurred, on qubit two: $\alpha|000\rangle + \beta|111\rangle \mapsto |\psi''\rangle = \alpha|010\rangle + \beta|101\rangle$. Then $\langle\psi''|M_2^\dagger M_2|\psi''\rangle = 1$ and we will certainly measure $j = 2$. The state of system will not change: $M_2|\psi''\rangle = |\psi''\rangle$. Therefore, we finally flip qubit two to recover $|\psi'\rangle$ from $|\psi''\rangle$.