AN INTRODUCTION TO QUANTUM COMPUTING

NOSON S. YANOFSKY

ABSTRACT. Quantum Computing is a new and exciting field at the intersection of mathematics, computer science and physics. It concerns a utilization of quantum mechanics to improve the efficiency of computation. Here we present a gentle introduction to some of the ideas in quantum computing. The paper begins by motivating the central ideas of quantum mechanics and quantum computation with simple toy models. From there we move on to a formal presentation of the small fraction of (finite dimensional) quantum mechanics that we will need for basic quantum computation. Central notions of quantum architecture (qubits and quantum gates) are described. The paper ends with a presentation of one of the simplest quantum algorithms: Deutsch's algorithm. Our presentation demands neither advanced mathematics nor advanced physics.

CONTENTS

1. INTUITION

Quantum Computing is a fascinating new field at the intersection of computer science, mathematics and physics. This field studies how to harness some of the strange aspects of quantum physics for use in computer science. Many of the texts to this field require knowledge of a large corpus of advanced mathematics or physics. We try to remedy this situation by presenting the basic ideas of quantum computing

understandable to anyone who has had a course in pre-calculus or discrete structures. (A good course in linear algebra would help, but, the reader is reminded of many definitions in the footnotes.)

The reason why we are able to ignore the higher mathematics and physics is that we do not aim to teach the reader all of quantum mechanics and all of quantum computing. Rather, we lower our aim to simply present that part necessary to offer a taste of what quantum computing is all about. What makes this possible is that we only need finite dimensional quantum mechanics, i.e., the vector spaces that represent the states of the system will only be of finite dimension. Such vector spaces consist of finite vectors with complex entries. These vectors will change by being multiplied by operators or matrices. These matrices will be finite and have complex entries. We do not do any more quantum computing than what is needed for our final goal: Deutsch's algorithm. We stress that the reader does not need more than the ability to do matrix multiplication in order to understand this paper.

To motivate our use of vectors to describe states and matrices as ways of describing dynamics, we show that it is understandable if one looks at a basic toy models. Our models deal with childrens' marbles moving along the edges of a graph. Every computer scientist and logician who has taken a class in discrete structures knows how to represent a (weighted) graph as an adjacency matrix. We shall take this basic idea and generalize it in several straightforward ways. While doing this, we shall present many concepts that are at the very core of quantum mechanics.

We begin with graphs that are without weights and progress to graphs that are weighted with real numbers, and finally to graphs that are weighted with complex numbers. With this in hand, we present a graph theoretic version of the double-slit experiment. This is *the* most important experiment in quantum mechanics. We conclude with a discussion of ways of combining systems to yield larger systems.
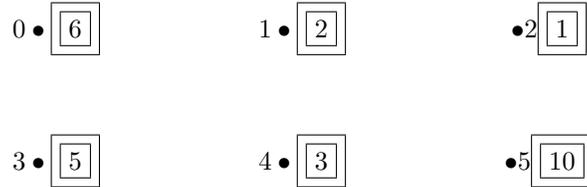
Throughout this chapter, we shall present an idea in a toy model, then generalize it to an abstract point and lastly discuss the connection with quantum mechanics before moving on to the next idea.

This paper is based on a forthcoming text *Quantum Computing for Computer Scientists* coauthored with Mirco Mannucci. The text was accepted for publication by Cambridge University Press and should see the light of day in the beginning of 2008. In the text we take the reader through the same material and go much further. The reader who appreciates this paper, will definitely gain from the text.
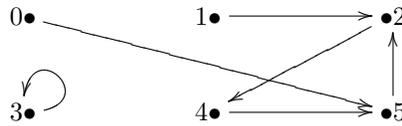
1.1. **Classical Deterministic Systems.** We begin with a simple system described by a graph along with some children's marbles. Imagine the marbles as being on the vertices of the graph. The state of a system is described by how many marbles are on each vertex. For example, say that there are six vertices in the graph and a total of 27 marbles.

We might place six marbles on vertex 0, two marbles on vertex 1 and the rest as described by this picture.

$$0 \bullet \boxed{6} \qquad 1 \bullet \boxed{2} \qquad \bullet 2\ \boxed{1}$$

$$3 \bullet \boxed{5} \qquad 4 \bullet \boxed{3} \qquad \bullet 5\ \boxed{10}$$

We shall denote this state as $X = [6, 2, 1, 5, 3, 10]^T$. The states of such a system will simply be a collum vector of size 6.

We are not only interested in states of the system, but also in the way that the states change — or the "dynamics" of the system. This can be represented by a graph with directed edges. The dynamics might be described by the following directed graph:

The idea is that if there exists an arrow from vertex i to vertex j, then in one time click, all the marbles on vertex i will move to vertex j. We place the following restriction on the types of graphs we shall be concerned with: graphs with exactly one outgoing edge from each vertex. This will correspond to the notion of a classical determinstic system. At each time click the marbles will have exactly one place to go.

This graph is equivalent to the matrix, M (for "marbles"):

$$
M \quad = \quad
\begin{array}{c}
\\
0 \\
1 \\
2 \\
3 \\
4 \\
5
\end{array}
\begin{array}{c}
\begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\
\left[
\begin{array}{cccccc}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0
\end{array}
\right]
\end{array}
$$

where $M[i, j] = 1$ if and only if there is an arrow from vertex j to vertex i.[1] Our restricted class of graphs are related to the restricted class of boolean matrices that have exactly one 1 in each column.

---

[1] Although most texts might have $M[i, j] = 1$ if and only if there is an arrow from vertex i to vertex j, we shall need it to be the other way for reasons which will become apparent later. The difference is trivial.

Let us say we multiply M by a state of the system $X = [6, 2, 1, 5, 3, 10]^T$. Then we have

$$
MX \;=\;
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
6 \\ 2 \\ 1 \\ 5 \\ 3 \\ 10
\end{bmatrix}
\;=\;
\begin{bmatrix}
0 \\ 0 \\ 12 \\ 5 \\ 1 \\ 9
\end{bmatrix}
\;=\; Y
$$

What does this correspond to? If X describes the state of the system at time t, then Y is the state at time $t + 1$, i.e., after one time click. We can see this clearly by looking at the formula for matrix multiplication:

$$
Y[i] = (M \star X)[i] \quad = \quad \sum_{k=0}^{5} M[i, k] X[k].
$$

In plain English, this states that the number of marbles that will reach vertex i after one time step is the sum of all the marbles that are on vertices with edges connecting to vertex i. Notice that the top two entries of Y are zero. This corresponds to the fact that there are no arrows going to vertex 0 or 1.

In general if $X = [x_0, x_1, \ldots, x_{n-1}]^T$ is a column vector corresponding to having x_i marbles on vertex i, M is a n by n Boolean matrix, and if $MX = Y$ where $Y = [y_0, y_1, \ldots, y_{n-1}]^T$, then there are y_j marbles on vertex j in one time click. M is thus a way of describing how the state of the marbles can change from time t to time $t + 1$.

As we shall soon see, (finite dimensional) quantum mechanics works in the same way. States of a system are represented by column vectors and the way in which the system changes in one time click is represented by matrices. Multiplying a matrix with a column vector yields a new state of the system. Quantum mechanics explores the way states of similar systems evolve over time.

Returning to our marbles, let's multiply M by itself. $MM = M^2$. However, since our entries are Boolean, we shall multiply the matrices as Boolean, i.e., $1 + 1 = 1 \vee 1 = 1$.

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

Looking at the formula for Boolean matrix multiplication

$$M^2[i,j] \quad = \quad \bigvee_{k=0}^{n-1} M[i,k] \wedge M[k,j]$$

we can see that this formula really shows us how to go from vertex j to vertex i in *two* time clicks.

And so we have that

$$M^2[i,j] \quad = \quad 1 \text{ if and only if there is a path of length 2 from vertex } j \text{ to vertex } i.$$

For an arbitrary k we have

$$M^k[i,j] \quad = \quad 1 \text{ if and only if there is a path of length } k \text{ from vertex } j \text{ to vertex } i.$$

In general, multiplying an n by n matrix by itself several times will correspond to the whether there is a path after several time clicks. Consider $X = [x_0, x_1, \ldots, x_{n-1}]^T$ to be the state where one has x_0 marbles on vertex 0, x_1 marbles on vertex 1, $\ldots$, x_{n-1} marbles on vertex $n-1$. Then, after k steps, the state of the marbles is Y where $Y = [y_0, y_1, \ldots, y_{n-1}]^T = M^k X$. In other words, y_j is the number of marbles on vertex j after k steps.

In quantum mechanics, if there are two or more matrices or operators that manipulate states, then the action of one followed by another is described by their matrix product. We shall take different states of systems and multiply the states by various matrices (of the appropriate type) to obtain other states. These other states will again be multiplied by other matrices until we attain the desired state.
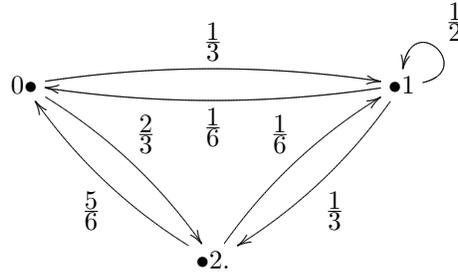
1.2. **Classical Probabilistic Systems.** In quantum mechanics, neither the state of a system nor the dynamics of a system are deterministic. There is an indeterminacy in our knowledge of a state. Furthermore, the states change with probabilistic laws as opposed to deterministic laws. That means that states do not change in set ways. Rather, the laws are given by stating that states will change from one state to another state with a certain likelihood.

In order to capture these probabilistic scenarios, let us generalize what we did in the last subsection. Instead of dealing with a bunch of marbles moving around, we shall deal with a single marble. The state of the system will tell us the probabilities of the single marble being on each vertex. For a three-vertex graph, a typical state might look like this $X = [\frac{1}{5}, \frac{3}{10}, \frac{1}{2}]^T$. This will correspond to the fact that there is a one-fifth[2] chance that the marble is on vertex 0; a three-tenth chance that the marble is on vertex 1; and a half chance that the marble is on vertex 2. Since the marble must be somewhere on the graph, the sum of the probabilities is 1.

We must generalize the dynamics as well. Rather than exactly one arrow leaving each vertex, we will have several arrows leaving each vertex with non-negative real numbers between 0 and 1 as weights. These weights will describe the probability of the single marble going from one vertex to another in one time click. We shall restrict our attention to weighted graphs that satisfy the following two conditions: a) the sum of all the weights leaving a vertex is 1 and b) the sum of all the weights entering a vertex is 1. This will correspond to the fact that a marble must go someplace (there might be loops) and a marble must come from someplace. An example of such a graph is

---

[2] Although the theory works with any $r \in [0,1]$, we shall deal only with fractions.

The matrix for this graph is

$$
M \;=\; \begin{bmatrix} 0 & \frac{1}{6} & \frac{5}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix}
$$

The adjacency matrices for our graphs will have non-negative real entries where the sums of the rows and the sums of the columns are all 1. Such matrices are called "doubly stochastic matrices."

Let us see how the states interact with the dynamics. Suppose we have a state $X = \left[\frac{1}{6}, \frac{1}{6}, \frac{2}{3}\right]^T$. We will calculate how a state changes: $MX = Y$

$$
\begin{bmatrix} 0 & \frac{1}{6} & \frac{5}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{6} \\ \frac{1}{6} \\ \frac{2}{3} \end{bmatrix} \;=\; \begin{bmatrix} \frac{21}{36} \\ \frac{9}{36} \\ \frac{6}{36} \end{bmatrix}
$$

Notice that the sum of the entries of Y is 1. If we have X expressing the probability of the position of a marble, and M expressing the probability of the way the marble moves around, then $MX = Y = \left[\frac{21}{36}, \frac{9}{36}, \frac{6}{36}\right]^T$ is to be interpreted as expressing the probability of the marble's location after moving. In other words, if X is the probability of the marble at time t, then MX is the probability of the marble at time $t+1$.

If M is an n by n doubly stochastic matrix and X is an n by 1 column vector whose entries sum to 1, then $M^k X = Y$ can be interpreted as expressing the probability of the position of a marble after k time clicks. That is, if $X = [x_0, x_1, \ldots, x_{n-1}]^T$ means that there is an x_i chance that a marble is on vertex i, then $M^k X = Y = [y_0, y_1, \ldots, y_{n-1}]^T$ means that after k time clicks, there is a y_j chance that the marble is on vertex j.

We are not constrained to multiply M by itself. We may also multiply M by another doubly stochastic matrix. Let M and N be two n by n doubly stochastic matrices. $M \star N$ will then describe a probability transition of going from time t to $t+1$ to $t+2$.

In quantum computers, quantum systems are generally in a probabilistic state. Manipulating the system will correspond to multiplying the state by matrices. Each time click will correspond to one matrix multiplication. At the end of the computation, the resulting vector will describe the state of the system.

Before moving on to the next section, we shall examine an interesting example. This shall be known as the "probabilistic double slit experiment." Consider the following picture of a shooting gun.
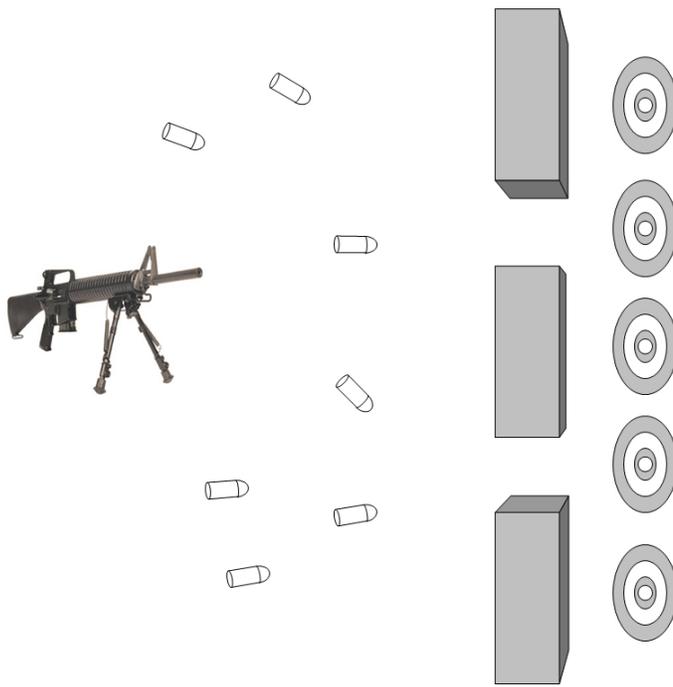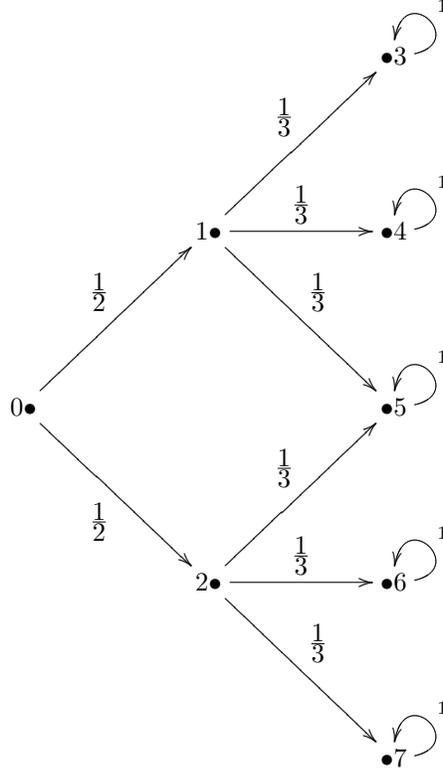

FIGURE 1. Double slit experiment with bullets.

There are two slits in the wall. The shooter is a good enough shot to always get the bullets through one of the two slits. There is a 50-50 chance of which slit the bullet will go through. Once a bullet is through a slit, there are three targets to the right of each slit that the bullet can hit with equal probability. The middle target can get hit in one of two ways: from the top slit going down, or from the bottom slit going up. It is assumed that it takes the bullet one time click to go from the gun to the wall and one time click to go from the wall to the targets. The picture

correspond to the following weighted graph.

Notice that the vertex marked 5 can receive bullets from either of the two slits.

Corresponding to this graph is the matrix B (for "bullets")

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

In words, B describes the way a bullet will move after one time click.[3]

———————

[3] The matrix B is not a doubly stochastic matrix. The sum of the weights entering vertex 0 is not 1. The sum of weights leaving vertices 3, 4, 5, 6, and 7 are more than 1. This fact should not bother you. We are interested in demonstrating the way probabilities behave with respect to these matrices.

Let us calculate the probabilities for the bullet's position after two time clicks.

$$B \star B \quad = \quad B^2 \quad = \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{3} & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{3} & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 1 & 0 & 0 \\ \frac{1}{6} & 0 & \frac{1}{3} & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{6} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

So B^2 indicates the probabilities of the bullet's position after two time clicks.

If we are sure that we start with the bullet in position 0, i.e., $X = [1, 0, 0, 0, 0, 0, 0, 0]^T$, then, after two time clicks, the state of the bullets will be

$$B^2 X = [0, 0, 0, \frac{1}{6}, \frac{1}{6}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}]^T$$

The key idea is to notice that $B^2[5, 0] = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$ because the bullets start from position 0, then there are two possible ways of the bullet getting to position 5. The possibilities sum to $\frac{1}{3}$. This is what we would expect. We shall revisit this example in the next subsection where strange things start happening!

1.3. **Quantum Systems.** We are now ready to leave the world of classical probabilities and enter the world of the quantum. One of the central facts about quantum mechanics is that complex numbers play a major role in the calculations. Probabilities of states and transitions are not given as a real numbers p between 0 and 1. Rather, they are given is a complex numbers c such that $|c|^2$ is a real number[4] between 0 and 1.

What is the difference how the probabilities are given? What does it matter if a probability is given as a real number between 0 and 1, or as a complex number whose modulus squared is a real number between 0 and 1? The difference is — and this is the core of quantum theory — that real number probabilities can be added to obtain larger real numbers. In contrast, complex numbers can cancel each other and lower their probability. In detail, if p_1 and p_2 are two real numbers between 0 and 1, then $(p_1 + p_2) \geq p_1$ and $(p_1 + p_2) \geq p_2$. Now let's look at the complex case. Let c_1 and c_2 be two complex numbers with their squares of modulus $|c_1|^2$ and $|c_2|^2$. $|c_1 + c_2|^2$ need not be bigger then $|c_1|^2$ and it also does not need to be bigger than $|c_2|^2$.

For example[5] , if $c_1 = 5 + 3i$ and $c_2 = -3 - 2i$, then $|c_1|^2 = 34$ and $|c_2|^2 = 13$ but $|c_1 + c_2|^2 = |2 + i|^2 = 5$. 5 is less than 34 and 5 is less than 13.

---

[4]We remind the reader that if $c = a + bi$ is a complex number, then its modulus is $|c| = \sqrt{a^2 + b^2}$ and $|c|^2 = a^2 + b^2$.
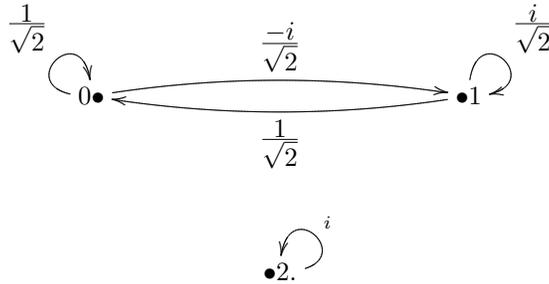
[5]The important point here is that the modulus squared is positive. For simplicity of calculations, we have chosen easy complex numbers.

This possibility of canceling out complex numbers corresponds to something called "interference" in quantum mechanics. One complex number might interfere with another. It is one of the most important ideas in quantum theory.

Let us generalize our states and graphs from the previous subsection. Rather than insisting that the sum of the entries in the column vector is 1, we insist that the sum of the modulus squared of the entries is 1. This makes sense since we are considering the probability as the modulus squared.

For dynamics, rather than talking about graphs with real number weights, we shall talk about graphs with complex number weights. Instead of insisting that the adjacency matrix of such a graph be a doubly stochastic matrix, we ask instead that the adjacency matrix be unitary.[6]

For example, consider the graph

The corresponding unitary adjacency matrix is

$$U \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{-i}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & i \end{bmatrix}.$$

Unitary matrices are related to doubly stochastic matrices as follows. The modulus squared of the all the complex entries in U forms a doubly stochastic matrix. The i,jth element in U is denoted $U[i,j]$, and its modulus squared is denoted $|U[i,j]|^2$. By abuse of notation, we shall denote the entire matrix of modulus squares as $|U[i,j]|^2$:

$$|U[i,j]|^2 \quad = \quad \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$
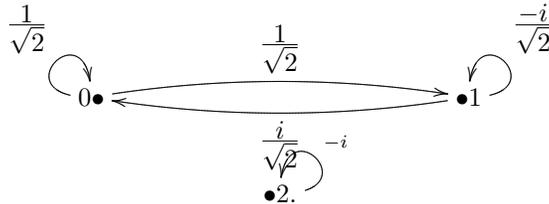
It is easy to see that this is a doubly stochastic matrix.

---

[6]Let us just remember: a matrix U is unitary if $U \star U^\dagger = I = U^\dagger \star U$. The adjoint of U, denoted as $U^\dagger$, is defined as $U^\dagger = (\overline{U})^T = \overline{(U^T)}$ or $U^\dagger[j,k] = \overline{U[k,j]}$.
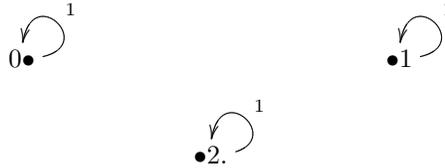
From this graph-theoretic point of view, it is easy to see what unitary means: the conjugate transpose of the U matrix is

$$
U^\dagger \;\; = \;\;
\begin{bmatrix}
\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\
\frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \\
0 & 0 & -i
\end{bmatrix}.
$$

This matrix corresponds to the graph

If U is the matrix that takes a state from time t to time $t+1$, then $U^\dagger$ is the matrix that takes a state from time t to time $t-1$. If we multiply U and $U^\dagger$, we get the identity matrix I_3 which corresponds to the graph

This means that if we perform some operation and then "undo" the operation, we will find ourselves in the same state as we began with probability 1. It is important to note that unitary does not only mean invertible. It means invertible in a very easy way, i.e., the inverse is the dagger of the matrix. This "invertibility" is again an important issue in quantum mechanics. Most of the dynamics will be invertible (except measurements).

In order to see the interference phenomenon and the related "superposition" phenomenon, we will revisit the double-slit experiment from the last subsection. Rather than looking at bullets, which are relatively large objects and hence adhere to the laws of classical physics, we shall look at microscopic objects such as photons that follow the laws of quantum physics. Rather than having a gun, we shall have a laser shoot photons through two slits as in figure 2.

Again we shall make the assumption that a photon will pass through one of the two slits. Each slit has a 50% chance of the photon going through it. To the left of each slit there are three measuring devices. It is assumed that it takes one time click to go from the laser to the slits and one time click to go from the slits to the targets. We are not interested in how large the slits are or how far the measuring positions are from the slits. Physicists are very adapt at calculating many different aspects of this experiment. We are only interested in the set-up.
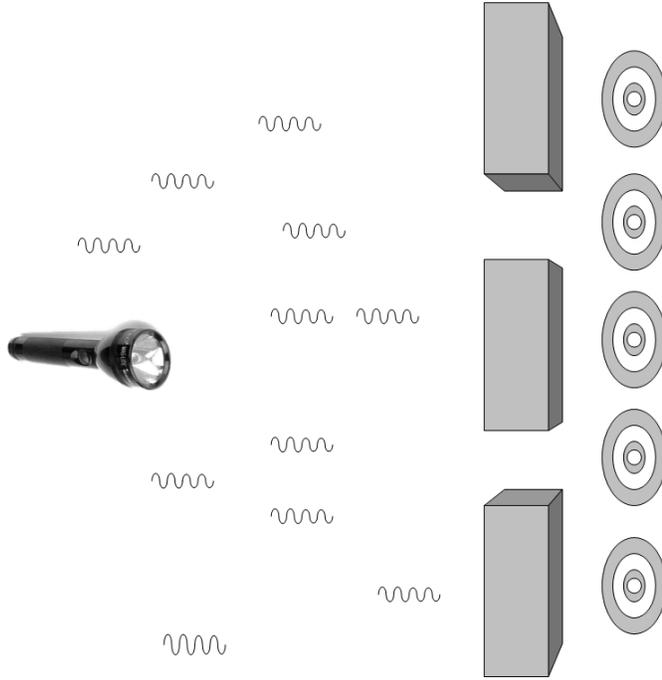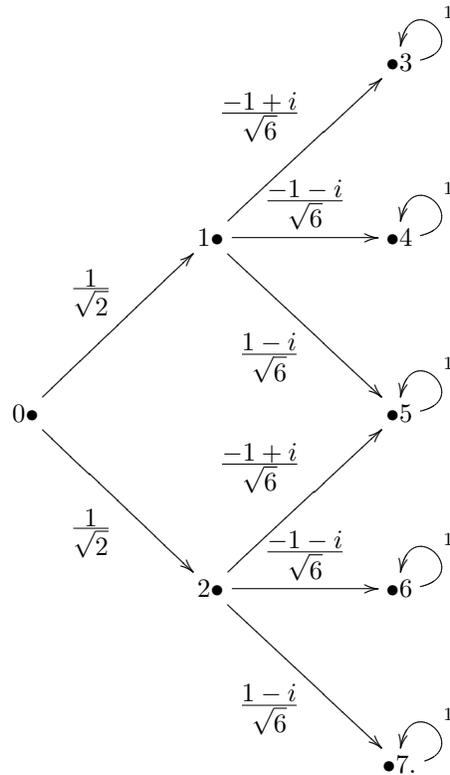
FIGURE 2. Double slit experiment with photons.

Consider the following weighted graph that describes the set-up of the experiment:

The modulus squared of $\frac{1}{\sqrt{2}}$ is $\frac{1}{2}$, which corresponds to the fact that there is a 50-50 chance of the photon going through either slit. $\left|\frac{\pm 1 \pm i}{\sqrt{6}}\right|^2 = \frac{1}{3}$ which corresponds to the fact that whichever slit the photon goes through, there is a $\frac{1}{3}$ of a chance of its hitting any of the three measuring positions to the right of that slit. [7]

The adjacency matrix for this graph is P (for "photons")[8]

$$
P \quad = \quad
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{-1+i}{\sqrt{6}} & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & \frac{-1-i}{\sqrt{6}} & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & \frac{1-i}{\sqrt{6}} & \frac{-1+i}{\sqrt{6}} & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & \frac{-1-i}{\sqrt{6}} & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & \frac{1-i}{\sqrt{6}} & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
.
$$

The modulus squared of the P matrix is exactly the same as the bullets matrix i.e., $|P[i,j]|^2 \quad = \quad B$. Let us see what happens if we calculate the transitions

---

[7]The actual complex number weights are not our interest here. If we wanted to calculate the actual numbers, we would have to measure the width of the slits, the distance between the slits, the distance from the slits to the measuring devices etc. However, our goal here is to clearly demonstrate the interference phenomenon. And so we chose the above complex numbers simply because the modulus squared are exactly the same as the bullets case.

[8]This matrix is not a unitary matrix. Looking carefully at row 0, one can immediately see that P is not unitary. In our graph, there is nothing entering vertex 0. The reason why this matrix fails to be unitary is because we have not put in all the arrows in our graph. There are many more possible ways the photon can travel in a real-life physical situation. In particular, the photon might go from the right to the left. The diagram and matrix would become too complicated if we put in all the transitions. We are simply trying to demonstrate the interference phenomenon and we can accomplish that even with a matrix that is not quite unitary.

matrix after *two* time clicks.

$$P^2 \;\; = \;\; \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \dfrac{-1+i}{\sqrt{12}} & \dfrac{-1+i}{\sqrt{6}} & 0 & 1 & 0 & 0 & 0 & 0 \\ \dfrac{-1-i}{\sqrt{12}} & \dfrac{-1-i}{\sqrt{6}} & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \dfrac{-1+i}{\sqrt{6}} & \dfrac{1-i}{\sqrt{6}} & 0 & 0 & 1 & 0 & 0 \\ \dfrac{-1-i}{\sqrt{12}} & 0 & \dfrac{-1-i}{\sqrt{6}} & 0 & 0 & 0 & 1 & 0 \\ \dfrac{-1+i}{\sqrt{12}} & 0 & \dfrac{1-i}{\sqrt{6}} & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

How do we interpret this in terms of probability? Let us look at the modulus squared of each of the entries.

$$|P^2[i,j]|^2 \;\; = \;\; \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{3} & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{3} & 0 & 0 & 1 & 0 & 0 & 0 \\ \mathbf{0} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 1 & 0 & 0 \\ \frac{1}{6} & 0 & \frac{1}{3} & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{6} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

This matrix is almost exactly the same as B^2 but with one glaring difference. $B^2[5,0] = \frac{1}{3}$ because of the two ways of starting at position 0 and ending at position 5. We added the nonnegative probabilities $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$. However with a photon that follows the laws of quantum mechanics, the complex numbers are added as opposed to their probabilities.

$$\frac{1}{\sqrt{2}}\left(\frac{-1+i}{\sqrt{6}}\right) + \frac{1}{\sqrt{2}}\left(\frac{1-i}{\sqrt{6}}\right) \;\; = \;\; \frac{-1+i}{\sqrt{12}} + \frac{1-i}{\sqrt{12}} \;\; = \;\; \frac{0}{\sqrt{12}} \;\; = \;\; 0.$$

Thus giving us $|P^2[5,0]|^2 = 0$. In other words, although there are two ways of a photon going from vertex 0 to vertex 5, there will be no photon at vertex 5.

How is one to understand this phenomenon? Physicists have a simple explanation for interference: waves. A familiar observation such as a pebble thrown into a pool of water will easily convince us that waves interfere, sometimes reinforcing each other, sometimes cancelling each other. In our experiment, photons are going through both slits at one time and they are canceling each other out at the middle measuring device. Thus, the double-slit experiment points to the wave-like nature of light.

The experiment can be done with only one photon shot out from vertex 0. This ensures that there will not be another wave for it to cancel out with. And yet, when only one photon goes through a slit, there is still an interference phenomenon. What is going on here?

The naive probabilistic interpretation of the position of the photon following the bullet metaphor of the previous subsection is thus not entirely adequate. Let the state of the system be given by $X = [c_0, c_1, \ldots, c_{n-1}]^T \in \mathbb{C}^n$. It is incorrect to say that the probability of the photon being in position k is $|c_k|^2$. Rather, a system in state X means that the particle is in *all* positions simultaneously. It is only after we measure the photon that the chances of it being found in position k is $|c_k|^2$. The photon (or its associated wave) goes through the top slit *and* the bottom slit simultaneously. And when the photon exits both slits, it can cancel *itself* out. A photon is not in *a* singe position, rather it is in *many* positions or a "superposition".

This might cause some justifiable disbelief. After all, we do not see things in a superposition of states. Our everyday experience tells us that things are in one position or (exclusive or!) another position. How can this be? The reason why we see particles in exactly one position is because we have performed a measurement. When we measure something at the quantum level, the quantum object that we have measured is no longer in a superposition of states, rather it collapses to a single classical state. So we have to redefine what a state of a quantum system means: A system in state X means that *after measuring* the photon it will be in position k with probability $|c_k|^2$.

What are we to make of these strange ideas? Are we really to believe them? Richard Feynman in discussing the double-slit experiment [**?**] (Vol. III, Page 1-1) waxes lyrical:

> We choose to examine a phenomenon which is impossible, *absolutely* impossible, to explain in any classical way, and which has in it the heart of quantum mechanics. In reality, it contains the *only* mystery. We can not make the mystery go away by "explaining" how it works. We will just tell you how it works.

It is exactly this superposition of states that is the real power behind quantum computing. Classical computers are in one state at every moment. Quantum computers can be put in a superposition of states. Imagine putting a computer in many different classical states at one time and then processing with *all* the states. This is the ultimate in parallel processing! Such a computer can only be conceived in the quantum world.

1.4. **Combining Systems.** Quantum mechanics can also deal with systems that have more than one part. In this subsection we learn how to combine several systems into one. We shall talk about combining classical probabilistic systems. However, whatever is stated about classical probabilistic systems is also true for quantum systems.

Consider two different marbles. Imagine that a red marble follows the probabilities of the graph whose corresponding adjacency matrix is

$$M \quad = \quad \begin{bmatrix} 0 & \frac{1}{6} & \frac{5}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix}.$$

Consider also a blue marble that follows the transitions given by the graph

$$G_N \quad = \quad$$

i.e., the matrix

$$N \quad = \quad \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

What does a state for a *two* marble system look like? Since the red marble can be on one of three vertices and the blue marble can be on one of two vertices, there are $3 \times 2 = 6$ possible states of the combined system. This is the tensor product of a 3 by 1 vector with a 2 by 1 vector. A typical state might look like:

$$X \quad = \quad \begin{matrix} 0a \\ 0b \\ 1a \\ 1b \\ 2a \\ 2b \end{matrix} \begin{bmatrix} \frac{1}{18} \\ 0 \\ \frac{2}{18} \\ \frac{1}{3} \\ 0 \\ \frac{1}{2} \end{bmatrix},$$

which would correspond to the fact that there is a

$\frac{1}{18}$ chance of the red marble being on vertex 1 and the blue marble being on vertex a,

0 chance of the red marble being on vertex 1 and the blue marble being on vertex b,

$\vdots$

$\frac{1}{2}$ chance of the red marble being on vertex 3 and the blue marble being on vertex b.

Now we may ask, how does a system with these *two* marbles change? What is its dynamics? Imagine that the red marble is on vertex 1 and the blue marble is on vertex a. We may write this as "$1a$." What is the probability of the state going from state $1a$ to state $2b$? Obviously, the red marble must go from vertex 1 to vertex 2 and (multiply) the blue marble must go from vertex a to vertex b. The probability is $\frac{1}{3} \times \frac{2}{3} = \frac{2}{9}$. In general, for a system to go from state ij to a

state $i'j'$ we must multiply the probability of going from state i to state i' with the probability of going from state j to state j'.

$$ij \xrightarrow{\quad M[i',i] \times N[j',j] \quad} i'j' \ .$$

For the changes of all the states, we have to do this for all the entries. We are really giving the tensor product of two matrices[9]:

$$M \otimes N \ = \ \begin{array}{c} \\ 0 \\ \\ 1 \\ \\ 2 \end{array} \left[\begin{array}{ccc} 0 \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} & \frac{1}{6} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} & \frac{5}{6} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} \\[2em] \frac{1}{3} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} & \frac{1}{2} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} & \frac{1}{6} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} \\[2em] \frac{2}{3} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} & \frac{1}{3} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} & 0 \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} \end{array} \right].$$

$$ = \begin{array}{c} \\ 0a \\ 0b \\ 1a \\ 1b \\ 2a \\ 2b \end{array} \begin{array}{c} \begin{array}{cccccc} 0a & 0b & 1a & 1b & 2a & 2b \end{array} \\ \left[\begin{array}{cccccc} 0 & 0 & \frac{1}{18} & \frac{2}{18} & \frac{5}{18} & \frac{10}{18} \\ 0 & 0 & \frac{2}{18} & \frac{1}{18} & \frac{10}{18} & \frac{5}{18} \\ \frac{1}{9} & \frac{2}{9} & \frac{1}{6} & \frac{2}{6} & \frac{1}{18} & \frac{2}{18} \\ \frac{2}{9} & \frac{1}{9} & \frac{2}{6} & \frac{1}{6} & \frac{2}{18} & \frac{1}{18} \\ \frac{2}{9} & \frac{4}{9} & \frac{1}{9} & \frac{2}{9} & 0 & 0 \\ \frac{4}{9} & \frac{2}{9} & \frac{2}{9} & \frac{1}{9} & 0 & 0 \end{array} \right] \end{array}$$

The graph that corresponds to this matrix, $G_M \times G_N$ is called the Cartesian product of two weighted graphs. In quantum theory, the states of two systems are combined using the tensor product of two vectors and the dynamics of two systems are combined by using the tensor product of two matrices. The tensor product of the matrices will then act on the tensor product of the vectors.

In general, the Cartesian product of an n-vertex graph with an n'-vertex graph is an $(n \times n')$-vertex graph. If we have an n-vertex graph G and we are interested in m different marbles moving in this system, we would need to look at the graph

$$G^m \ = \ \underbrace{G \times G \times \cdots \times G}_{m \text{ times}}$$

---

[9]Formally, the tensor product of matrices is a function

$$\otimes : \mathbb{C}^{m \times m} \times \mathbb{C}^{n \times n} \longrightarrow \mathbb{C}^{m \times m} \otimes \mathbb{C}^{n \times n} = \mathbb{C}^{mn \times mn}$$

and it is defined as

$$(A \otimes B)[j, k] = A[j/n, k/m] \times B[j \text{ MOD } n, k \text{ MOD } m].$$

which will have n^m vertices. If M_G is the associated adjacency matrix, then we will be interested in

$$M_G^{\otimes m} \quad = \quad \underbrace{M_G \otimes M_G \otimes \cdots \otimes M_G}_{m \text{ times}}$$

which will be a n^m by n^m matrix.

One might think of a bit as a two vertex graph with a marble on the 0 vertex or a marble on the 1 vertex. If one were then interested in m bits, one would need a 2^m vertex graph or equivalently a 2^m by 2^m matrix. So there is exponential growth of the resources needed for the amount of bits discussed.

This exponential growth for a system is actually one of the main reasons why Richard Feynman started thinking about quantum computing. He realized that because of this exponential growth, it would be hard for a classical computer to simulate such a system. He asked whether a quantum computer with its ability to do massive parallel processing, might be able to simulate such a system.

2. Basic Quantum Theory

Armed with the intuition, we tackle the formal statement of quantum theory. A disclaimer is in order: We are only presenting a small part of finite dimensional quantum physics. There is no way that we can give more than a minute fraction of this magnificent subject in these few pages. It is a sincere hope that this will inspire the reader to go on and study more. For a mathematician, the best book to start reading about quantum mechanics is, of course, Dirac's classic text [?]. However, there are many other primers available, e.g., [?, ?, ?] and [?]. The more advanced mathematician might want to look at [?].

2.1. **States.** An n dimensional quantum system is a system that can be observed in one of n possible states. Examples of such systems are

- a particle can be in one of n positions;
- a system might have one of n energy levels;
- a photon might have one of n polarization directions.

For clarity, lets talk of the first example. Lets say we have a particle that can be in one of n positions.

The states of such a system shall be represented by collum vectors of n complex numbers. We shall denote these vectors with the "ket" $|\ \ \rangle$ [10] notation:

$$|\varphi\rangle = [c_0, c_1, \ldots, c_j, \ldots, c_{n-1}]^T.$$

How is one to interpret these kets? Let us look at simple cases. The state

$$|\psi\rangle = [0, 1, \ldots, 0, \ldots, 0]^T$$

is to be thought of as saying that our particle will be found in position 1. The state

$$|\psi'\rangle = [0, 0, \ldots, 1, \ldots, 0]^T$$

is to be interpreted that the particle is in position j. These two states are examples of what are called "pure states".

How is one to interpret an arbitrary

$$|\varphi\rangle = [c_0, c_1, \ldots, c_j, \ldots, c_{n-1}]^T?$$

---

[10]"Ket" is the second half of "bracket". However we shall not use the "bra" part in our exposition.

Let S be the sum of the squares of modulus of the c_j, i.e.,

$$S = |c_0|^2 + |c_1|^2 + \ldots + |c_{n-1}|^2.$$

This is the length of the vector $|\varphi\rangle$. Then $|\varphi\rangle$ is to be interpreted that if one was to measure the state described by $|\varphi\rangle$ we would find the particle in position 0 with probability $|c_0|^2/S$, in position 1 with probability $|c_1|^2/S$, in position 2 with probability $|c_2|^2/S$, $\ldots$, in position $n-1$ with probability $|c_{n-1}|^2/S$ Such states are called "superpositions". They say that the particle is in more than one "position" at a time. It is important to stress that $|\varphi\rangle$ means that the particle is in *all* positions simultaneously. It does not mean that the particle is in some single position and the c_j are giving us probabilities of which position.

These superpositions can be added: if

$$|\varphi\rangle = [c_0, c_1, \ldots, c_j, \ldots, c_{n-1}]^T$$

and

$$|\varphi'\rangle = [c_0', c_1', \ldots, c_j', \ldots, c_{n-1}']^T,$$

then

$$|\varphi\rangle + |\varphi'\rangle = [c_0 + c_0', c_1 + c_1', \ldots, c_j + c_j', \ldots, c_{n-1} + c_{n-1}']^T.$$

Also, if there is a complex number $c \in \mathbb{C}$, we can multiply a ket by this c:

$$c|\varphi\rangle = [c \times c_0, c \times c_1, \ldots, c \times c_j, \ldots, c \times c_{n-1}]^T$$

These operation satisfy all the properties of being a complex vector space. So the states of an n dimensional quantum system are represented by the complex vector space $\mathbb{C}^n$.

Let us add a superposition to itself.

$$|\varphi\rangle + |\varphi\rangle = 2|\varphi\rangle = [c_0 + c_0, c_1 + c_1, \ldots, c_j + c_j, \ldots, c_{n-1} + c_{n-1}]^T$$
$$= [2c_0, 2c_1, \ldots, 2c_j, \ldots, 2c_{n-1}]^T.$$

The sum of the modulus squared is

$$S' = |2c_0|^2 + |2c_1|^2 + \ldots + |2c_{n-1}|^2 = 2^2|c_0|^2 + 2^2|c_1|^2 + \ldots + 2^2|c_{n-1}|^2 = 2^2(|c_0|^2 + |c_1|^2 + \ldots + |c_{n-1}|^2).$$

The chances of a particle being found in position j is

$$\frac{|2c_j|^2}{S'} = \frac{2^2|c_j|^2}{2^2(|c_0|^2 + |c_1|^2 + \ldots + |c_{n-1}|^2)} = \frac{|c_j|^2}{|c_0|^2 + |c_1|^2 + \ldots + |c_{n-1}|^2}.$$

In other words, the ket $2|\varphi\rangle$ describes the same physical system as $|\varphi\rangle$. This makes sense, after all when we add two of the same superpositions, we do not expect any interference. We expect that they should reinforce each other. A similar analysis shows that for an arbitrary $c \in \mathbb{C}$ we have that the ket $|\varphi\rangle$ and the ket $c|\varphi\rangle$ describe the same physical state. Geometrically, we can say that the vector $|\varphi\rangle$ and the extension $c|\varphi\rangle$ describe the same physical state. So the only thing that is important is the direction of $|\varphi\rangle$ not the length of $|\varphi\rangle$. We might as well work with a "normalized" $|\varphi\rangle$, i.e.,

$$\frac{|\varphi\rangle}{||\varphi\rangle|}$$

which has length 1. In fact, in section 1, we only worked with normalized vectors.

Given an n dimensional quantum system represented by $\mathbb{C}^n$ and an m dimensional quantum system represented by $\mathbb{C}^m$, we can combine these two systems to

form one system. This one system is to represented by the tensor product of the two vector spaces:

$$\mathbb{C}^n \otimes \mathbb{C}^m \cong \mathbb{C}^{n \times m}.$$

If $|\varphi\rangle$ is in the first system and $|\varphi'\rangle$ is in the second system, then we represent the combined system as

$$|\varphi\rangle \otimes |\varphi'\rangle = |\varphi, \varphi'\rangle = |\varphi\varphi'\rangle.$$

It is important to realize that, in general, there are more elements in the tensor product of the two systems than in the union of each of the two systems. States in $\mathbb{C}^n \otimes \mathbb{C}^m$ that cannot be represented simply as an element in $\mathbb{C}^n$ and an element in $\mathbb{C}^m$ are said to be "entangled".

2.2. Dynamics.

Quantum systems are not static. The states of the system are constantly changing. Changes, or "operators", on an n dimensional quantum system are represented by n by n unitary matrices. Given a state $|\varphi\rangle$ that represents a system at time t, then the system will be in state $U|\varphi\rangle$ at time $t + 1$.

What does unitary really mean? If $U|\varphi\rangle = |\varphi'\rangle$ then we can easily form $U^\dagger$ and multiply both sides of the equation by $U^\dagger$ to get $U^\dagger U|\varphi\rangle = U^\dagger|\varphi'\rangle$ or $|\varphi\rangle = U^\dagger|\varphi'\rangle$. In other words, because U is unitary, there is a related matrix that can "undo" the action that U does. $U^\dagger$ takes the result of U's action and gets the original vector back. In the quantum world, most actions actions are "undoable" or "reversible" in such a manner.

If U operates on $\mathbb{C}^n$ and U' operates on $\mathbb{C}^m$, then $U \otimes U'$ will operate on $\mathbb{C}^n \otimes \mathbb{C}^m$ in the following way:

$$(U \otimes U')(|\varphi\rangle \otimes |\varphi'\rangle) = U|\varphi\rangle \otimes U'|\varphi'\rangle.$$

2.3. Observables.

There are other types of operations that one can do to a n dimensional quantum system: one can observe, or "measure", the system. When we measure a system, it is no longer in a superposition. The superposition is said to "collapse" to a pure state.

$$|\varphi\rangle = [c_0, c_1, \ldots, c_j, \ldots, c_{n-1}]^T \rightsquigarrow |\varphi'\rangle = [0, 0, \ldots, 1, \ldots, 0].^T$$

The question is which of the n pure states will the state collapse to? The answer is that it is random. Let S be the sum of all the squares of the modulus, i.e.,

$$S = |c_0|^2 + |c_1|^2 + \cdots + |c_j|^2 + \cdots + |c_{n-1}|^2.$$

There is a $|c_0|^2/S$ of a chance of the superposition collapsing to the 0th pure state. There is a $|c_1|^2/S$ of a chance of the superposition collapsing to the 1th pure state. etc. There is no known mathematical way to decide which pure state the system will, in-fact, collapse to.

An observable, or "measurement" on an n dimensional system is represented by an n by n hermitian matrix. We remind the reader that a n by n matrix A is hermitian, or "self-adjoint" if $A^\dagger = A$. In other words, $A[j, k] = \overline{A[k, j]}$. Equivalently A is hermitian if and only if $A^T = \overline{A}$. For example, the matrix

$$\begin{bmatrix} 5 & 4 + 5i & 6 - 16i \\ 4 - 5i & 13 & 7 \\ 6 + 16i & 7 & -2.1 \end{bmatrix}$$

is hermitian.

For a matrix A in $\mathbb{C}^{n \times n}$, if there is a number c in $\mathbb{C}$ and a vector $|\psi\rangle$ in $\mathbb{C}^n$ such that

$$A|\psi\rangle = c|\psi\rangle$$

then c is called an "eigenvalue" of A and $|\psi\rangle$ is called an "eigenvector" of A associated to c. The eigenvalues of a hermitian matrix are all real numbers. Furthermore, distinct eigenvectors which have distinct eigenvalues of any hermitian matrix are orthogonal. The hermitian matrices that represent observables for an n dimensional system have the further property that there are n distinct eigenvalues and n distinct eigenvectors. That means that the set of eigenvectors form a basis for the entire complex vector space that represents the quantum system we are interested in. Hence if we have an observable A and $|\varphi\rangle$ an eigenvalue of A then $A|\varphi\rangle = c|\varphi\rangle$ for some $c \in \mathbb{C}$. $c|\varphi\rangle$ represents the same state as $|\varphi\rangle$ as we said before. So if the system is in an eigenvector of the basis, then the system will not change.

3. Architecture

In this section we are going to show how the ideas of quantum mechanics are going to influence our construction of quantum computers. In this paper we do not dwell on actual hardware implementations. Rather we shall look at the quantum generalizations of bits and logical gates. In section 3.1 we go from bits to qubits. We also discuss the notation that is needed for this. In section 3.2 we show how to look at classical computing as matrix manipulations. From the view afforded by this perspective, we easily generalize the notion of logical gate to quantum gate. There are many quantum gates, but we shall only look at a few that will be needed in the next section.

3.1. **Bits and Qubits.** What is a bit? A bit is an atom of information that represents one of two disjoint situations. An example of a bit is electricity going through a circuit or electricity not going through a circuit; a switch turned on or a switch turned off; a way of saying true or false. All these examples are saying the same thing: a bit is a way of describing a system whose set of states is of size two.

A bit can be represented by two 2 by 1 matrices. We shall represent 0—or better the state $|0\rangle$ as

$$|0\rangle \quad = \quad \begin{matrix} 0 \\ 1 \end{matrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

We shall represent a 1, or state $|1\rangle$ as:

$$|1\rangle \quad = \quad \begin{matrix} 0 \\ 1 \end{matrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Since these are two different representations, we have an honest-to-goodness bit.

A bit is either in state $|0\rangle$ or in state $|1\rangle$. That was sufficient for the classical world. But that is not sufficient for the quantum world. In that world we have situations where we are in one state *and* in the other state simultaneously. In the quantum world we have systems where a switch is in a superposition of states on *and* off. So we define a "quantum bit" or a "qubit" as a way of describing a

quantum system of dimension two. We shall represent any such qubit as a two by one matrix with complex numbers:

$$\begin{matrix} 0 \\ 1 \end{matrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$$

where $|c_0|^2 + |c_1|^2 = 1$. Notice that a classical bit is a special type of qubit. $|c_0|^2$ is to be interpreted as the probability that after measuring the qubit, it will be found in state $|0\rangle$. $|c_1|^2$ is to be interpreted as the probability that after measuring the qubit it will be found in state $|1\rangle$. Whenever we measure a qubit, it automatically becomes a bit. So we shall never "see" a general qubit. Nevertheless, they do exist and they are the core of our tale. The power of quantum computers is due to the fact that a system can be in many states at the same time.

Following the normalization procedure that we learned in the last section, any element of $\mathbb{C}^2$ can be converted into a qubit. For example, the vector

$$V = \begin{bmatrix} 5 + 3i \\ 6i \end{bmatrix}$$

has norm

$$|V| = \sqrt{\langle V, V \rangle} = \sqrt{[5 - 3i, -6i] \begin{bmatrix} 5 + 3i \\ 6i \end{bmatrix}} = \sqrt{34 + 36} = \sqrt{70}.$$

So V describes the same physical state as the qubit

$$\frac{V}{\sqrt{70}} = \begin{bmatrix} \dfrac{5 + 3i}{\sqrt{70}} \\ \dfrac{6i}{\sqrt{70}} \end{bmatrix}.$$

After measuring the qubit $\dfrac{V}{\sqrt{70}}$, the probability of being in state $|0\rangle$ is $\frac{34}{70}$, and the probability of being in state $|1\rangle$ is $\frac{36}{70}$.

It is easy to see that the bits $|0\rangle$ and $|1\rangle$ are the canonical basis of $\mathbb{C}^2$. So any qubit can be written as

$$\begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = c_0 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = c_0 |0\rangle + c_1 |1\rangle.$$

Let us look at several ways of writing different qubits. $\dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ can be written as

$$\begin{bmatrix} \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}} |0\rangle + \frac{1}{\sqrt{2}} |1\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}.$$

Similarly $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ can be written as

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

It is important to realize that

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} = \frac{|1\rangle + |0\rangle}{\sqrt{2}}.$$

These are both ways of writing $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$. In contrast,

$$\frac{|0\rangle - |1\rangle}{\sqrt{2}} \neq \frac{|1\rangle - |0\rangle}{\sqrt{2}}$$

The first state is $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$ and the second one is $\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$. However the two states are related:

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

How are qubits to be implemented? While this is not our focus, some examples of qubit implementations are given:

- An electron might be in one of two different orbits around a nucleus of an atom. (Ground state and excited state.)
- A photon might be in one of two different polarized states.
- A subatomic particle might be in spinning in one of two different directions.

There will be enough quantum indeterminacy and quantum superposition effects within all these systems to represent a qubit.

Computers with only one bit of storage are not very interesting. Similarly, we will need quantum devices with more then one qubit. Consider a byte, or eight bits. A typical byte might look like

$$01101011.$$

We might also write it as

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Previously, we learned that in order to combine systems one should use the tensor product. We can describe the above byte as

$$|0\rangle \otimes |1\rangle \otimes |1\rangle \otimes |0\rangle \otimes |1\rangle \otimes |0\rangle \otimes |1\rangle \otimes |1\rangle.$$

As a qubit, this is an element of

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2.$$

This vector space can be written as $(\mathbb{C}^2)^{\otimes 8}$. This is a complex vector space of dimension $2^8 = 256$. Since there is only one complex vector space (up to isomorphism) of this dimension, this space is isomorphic to $\mathbb{C}^{256}$.

Our byte can be described as in yet another way: As a $2^8 = 256$ row vector

$$
\begin{array}{l}
00000000 \\
00000001 \\
\vdots \\
01101010 \\
01101011 \\
01101100 \\
\vdots \\
11111110 \\
11111111
\end{array}
\begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
1 \\
0 \\
\vdots \\
0 \\
0
\end{bmatrix}.
$$

This is fine for the classical world. However, for the quantum world, a general qubit can be written as

$$
\begin{array}{l}
00000000 \\
00000001 \\
\vdots \\
01101010 \\
01101011 \\
01101100 \\
\vdots \\
11111110 \\
11111111
\end{array}
\begin{bmatrix}
c_0 \\
c_1 \\
\vdots \\
c_{106} \\
c_{107} \\
c_{108} \\
\vdots \\
c_{254} \\
c_{255}
\end{bmatrix}
$$

where $\sum_{i=0}^{255} |c_i|^2 = 1$.

In the classical world, you need to write the state of each of the eight bits. This amounts to writing eight bits. In the quantum world, a state of eight qubits is given by writing 256 complex numbers. This exponential growth was one of the reasons why researchers started thinking about quantum computing. If you wanted to emulate a quantum computer with a 64 qubit register, you would need to store $2^{64} = 18,446,744,073,709,551,616$ complex numbers. This is beyond our current ability.

Let us practice writing two qubits in ket notation. The qubits

$$\begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

can be written as

$$|0\rangle \otimes |1\rangle.$$

Since the tensor product is understood, we might also write these qubits as $|0,1\rangle$ or $|01\rangle$. The qubit corresponding to

$$\frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

can be written as

$$\frac{1}{\sqrt{3}}|00\rangle - \frac{1}{\sqrt{3}}|10\rangle + \frac{1}{\sqrt{3}}|11\rangle = \frac{|00\rangle - |10\rangle + |11\rangle}{\sqrt{3}}.$$

The tensor product of two states is not commutative.

$$|0\rangle \otimes |1\rangle = |0,1\rangle = |01\rangle \neq |10\rangle = |1,0\rangle = |1\rangle \otimes |0\rangle.$$

The first ket says that the first qubit is in state 0 and the second qubit is in state 1. The second ket says that first qubit is in state 1 and the second state is in state 0.

3.2. **Classical Gates.** Classical logical gates are ways of manipulating bits. Bits go into logical gates and bits come out. We represent n input bits as a 2^n by 1 matrix and m output bits as a 2^m by 1 matrix. How should we represent our logical gates? A 2^m by 2^n matrix takes a 2^n by 1 matrix and outputs a 2^m by 1 matrix. In symbols:

$$(2^m \times 2^n)(2^n \times 1) = (2^m \times 1).$$

So bits will be represented by column vectors and logic gates will be represented by matrices.

Let us do a simple example. Consider the NOT gate. NOT takes as input one bit, or a 2 by 1 matrix, and outputs one bit, or a 2 by 1 matrix. NOT of $|0\rangle$ equals $|1\rangle$ and NOT of $|1\rangle$ equals $|0\rangle$. The matrix

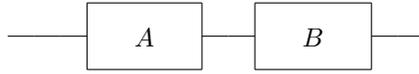$$\text{NOT} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

This matrix satisfies

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

which is exactly what we want.

What about the other gates? The other gates will be given by the following matrices:
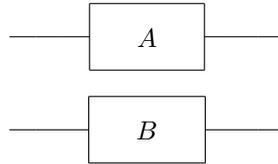
	AND				NAND				OR				NOR		
1	1	1	0	0	0	0	1	1	0	0	0	0	1	1	1
0	0	0	1	1	1	1	0	0	1	1	1	1	0	0	0

When we perform a computation, often we have to carry out one operation followed by another.

We shall call this performing "sequential" operations. If matrix A corresponds to performing an operation and matrix B corresponds to performing another operation, then the multiplication matrix $B \star A$ corresponds to performing the operation sequentially.

Besides sequential operations, there are "parallel" operations:

Here we are doing A to some bits and B to other bits. This will be represented by

$A \otimes B$, the tensor product of two matrices.

Combination of sequential and parallel operations gates/matrices will be circuits. A starting point is the realization that the circuit

can be realized as

$$(B \star A) \otimes (B' \star A') = (B \otimes B') \star (A \otimes A').$$

We will of course have some really complicated matrices, but they will all be decomposable into the sequential and parallel composition of simple gates.
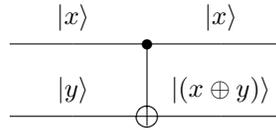
3.3. **Quantum Gates.** A quantum gate is simply any unitary matrix that manipulates qubits. There are some simple gates that are quantum gates.

The "Hadamard matrix"

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \end{bmatrix}.$$

The Hadamard matrix is also its own inverse. As it turns out, the Hadamard matrix is one of the most important matrices in quantum computing.
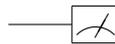
Consider the following "controlled-not" gate.

This gate has two inputs and has two outputs. The top input is the control bit. It controls what the output will be. If $|x\rangle = |0\rangle$, then the output of $|y\rangle$ will be the same as the input. If $|x\rangle = |1\rangle$ then the output of $|y\rangle$ will be the opposite. If we write the top qubit first and then the bottom qubit, then the controlled-not gate takes $|x, y\rangle$ to $|x, x \oplus y\rangle$ where $\oplus$ is the binary exclusive or operation.

The matrix that corresponds to this reversible gate is

$$\begin{array}{c c} & \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{array}.$$
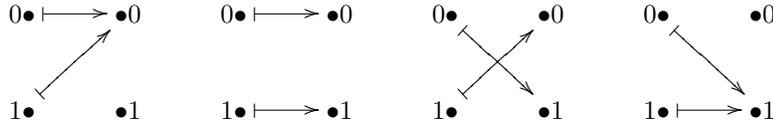
One last piece of notation: we shall denote an observation (measurement) by the following "meter":

There are many other quantum gates, but we shall not need them for our work in the next section.

4. Deutsch's Algorithm

The simplest quantum algorithm is Deutsch's algorithm which is a nice short algorithm that solves a slightly contrived problem. This algorithm is concerned with functions from the set $\{0, 1\}$ to the set $\{0, 1\}$. There are four such functions which we might visualize as
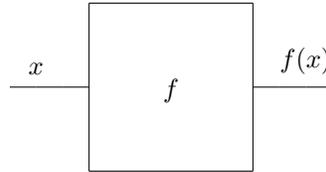

Call a function $f : \{0, 1\} \longrightarrow \{0, 1\}$, "balanced" if $f(0) \neq f(1)$). In contrast, call a function "constant" if $f(0) = f(1)$. Of the four functions, two are balanced and two are constant.

Deutsch's algorithm solves the following problem: Given a function $f : \{0, 1\} \longrightarrow \{0, 1\}$ as a black-box, where one can evaluate an input, but cannot "look inside" and "see" how the function is defined, tell if the function is balanced or constant.
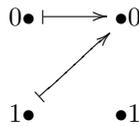
With a classical computer, one would have to first evaluate f on an input, then evaluate f on the second input and then compare the outputs. The point is that with a classical computer, f must be evaluated twice. Can we do better?

A quantum computer can be in two states at one time. We shall use this superposition of states to evaluate both inputs at one time.

In classical computing, evaluating a given function f would correspond to performing the following operation

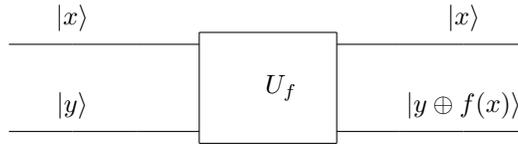Such a function could be thought of as a matrix. The function
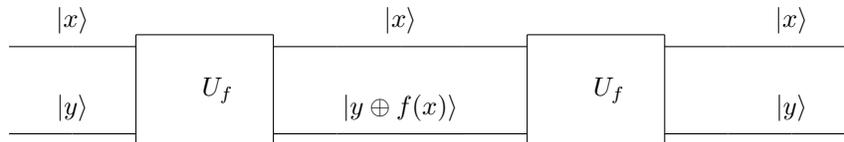

is equivalent to the matrix

$$
\begin{array}{c c}
 & \begin{array}{c c} 0 & 1 \end{array} \\
\begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.
\end{array}
$$

Multiplying either state $|0\rangle$ or state $|1\rangle$ on the right of this matrix would result in state $|0\rangle$. The column name is to be thought of as the input and the row name is to be thought of as the output.

However, this will not be good for a quantum system. For a quantum system we need a little something extra. A quantum system must be unitary (reversible). Given the output, we must be able to find the input. If f is the name of the function, then the following black-box U_f system will be the quantum gate that we shall employ to evaluate input:
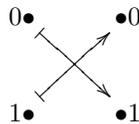
$$|x\rangle \qquad\qquad\qquad |x\rangle$$
$$\boxed{U_f}$$
$$|y\rangle \qquad\qquad\qquad |y \oplus f(x)\rangle$$

The top input, $|x\rangle$, will be the qubit value that one wishes to evaluate and the bottom input, $|y\rangle$, controls the output. The top output will be the same as the input qubit $|x\rangle$ and the bottom output will be the qubit $|y \oplus f(x)\rangle$ where $\oplus$ is XOR, the exclusive or operation. We are going to write from left to right the top qubit first and then the bottom. So we say that this function takes the state $|x, y\rangle$ to the state $|x, y \oplus f(x)\rangle$. If $y = 0$ this simplifies: $|x, 0\rangle$ to $|x, 0 \oplus f(x)\rangle = |x, f(x)\rangle$. This gate is reversible by simply looking at the following system

$$|x\rangle \qquad\qquad |x\rangle \qquad\qquad |x\rangle$$
$$\boxed{U_f} \qquad\qquad \boxed{U_f}$$
$$|y\rangle \qquad\qquad |y \oplus f(x)\rangle \qquad\qquad |y\rangle$$

State $|x, y\rangle$ goes to $|x, y \oplus f(x)\rangle$ which further goes to

$$|x, (y \oplus f(x)) \oplus f(x)\rangle = |x, y \oplus (f(x) \oplus f(x))\rangle = |x, y \oplus 0\rangle = |x, y\rangle.$$

In quantum systems, evaluating f is equivalent to multiplying a state of the input by a unitary matrix U_f. For the function

$$
\begin{array}{ccc}
0\bullet & & \bullet 0 \\
& \times & \\
1\bullet & & \bullet 1
\end{array}
$$

the corresponding unitary matrix is

$$
\begin{array}{c c}
 & \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\
\begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array} &
\left[\begin{array}{cccc}
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{array} \right]
\end{array}.
$$

Remember that the top column name correspond to the input $|x, y\rangle$ and the left-hand row name corresponds to the outputs $|x, y \oplus f(x)\rangle$. A 1 in the xy column and the $x'y'$ row means for input $|x, y\rangle$ the output will be $|x', y'\rangle$.

So we are given such a matrix that expresses a function but we cannot "look inside" the matrix to "see" how it is defined. We are asked to determine if the function is balanced or constant.
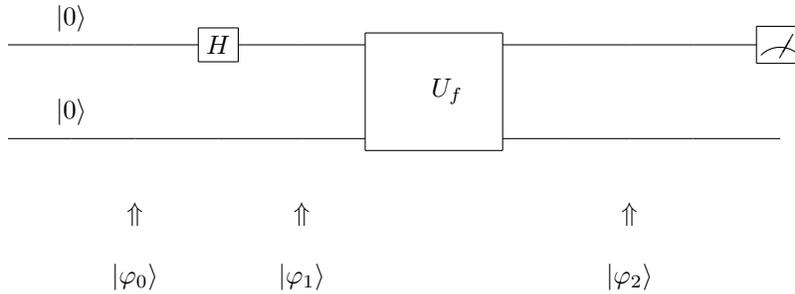
4.1. **First Attempt.** Let us take a first stab at a quantum algorithm to solve this problem. Rather than evaluating f twice, let us try our trick of superposition of states. Instead of having the top input to be either in state $|0\rangle$ or in state $|1\rangle$, we shall put the top input in state

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

which is "half-way" $|0\rangle$ and "half-way" $|1\rangle$. The Hadamard matrix can place a qubit in such a state.

$$
H|0\rangle = \left[\begin{array}{cc} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{array} \right] \left[\begin{array}{c} 1 \\ 0 \end{array} \right] = \left[\begin{array}{c} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{array} \right] = \frac{|0\rangle + |1\rangle}{\sqrt{2}}.
$$

The obvious (and not necessarily correct) state to put the bottom input is as state $|0\rangle$. And so we have:

In terms of matrices this corresponds to

$$
U_f(H \otimes I)(|0\rangle \otimes |0\rangle).
$$

We shall carefully examine the states of the system at every point. The system starts in

$$|\varphi_0\rangle \quad = \quad (|0\rangle \otimes |0\rangle) \quad = \quad |0,0\rangle.$$

We then apply the Hadamard matrix only to the top input — leaving the bottom input alone — to get

$$|\varphi_1\rangle \quad = \quad \left[\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right]|0\rangle \quad = \quad \frac{|0,0\rangle + |1,0\rangle}{\sqrt{2}}.$$

After multiplying with U_f we have

$$|\varphi_2\rangle \quad = \quad \frac{|0, f(0)\rangle + |1, f(1)\rangle}{\sqrt{2}}$$

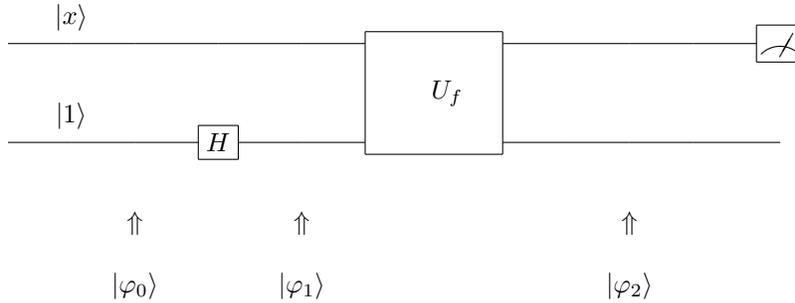For the function $0 \mapsto 1$ and $1 \mapsto 0$ the state $|\varphi_2\rangle$ would be

$$|\varphi_2\rangle = \begin{matrix} & \begin{matrix} 00 & 01 & 10 & 11 \end{matrix} \\ \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix} \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix}\begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} = \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix}\begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} = \frac{|0,1\rangle + |1,0\rangle}{\sqrt{2}}$$

If we measure the top qubit, there will be a 50-50 chance of finding it in state $|0\rangle$ and a 50-50 chance of finding it in state $|1\rangle$. Similarly there is no real information to be gotten by measuring the bottom qubit. So the obvious algorithm does not work. We need another trick.

4.2. **Second Attempt.** Let us take another stab at solving our problem. Rather then leaving the bottom qubit in state $|0\rangle$, let us put it in the superposition state:

$$\frac{|0\rangle - |1\rangle}{\sqrt{2}} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Notice the minus sign. Even though there is a negation, this state is also "half-way" in state $|0\rangle$ and "half-way" in state $|1\rangle$. The change of phase will help us get our desired results. We can get to this superposition of states by multiplying state $|1\rangle$ with the Hadamard matrix. We shall leave the top qubit as an ambiguous $|x\rangle$.

In terms of matrices, this becomes:

$$U_f(I \otimes H)|x, 1\rangle.$$

Let us look carefully at how the states of the qubits change.

$$|\varphi_0\rangle \quad = \quad |x, 1\rangle.$$

After the Hadamard matrix, we have

$$|\varphi_1\rangle \quad = \quad |x\rangle \left[\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right] \quad = \quad \frac{|x, 0\rangle - |x, 1\rangle}{\sqrt{2}}.$$

Applying U_f we get

$$|\varphi_2\rangle \quad = \quad |x\rangle \left[\frac{|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle}{\sqrt{2}} \right] \quad = \quad |x\rangle \left[\frac{|f(x)\rangle - |\overline{f(x)}\rangle}{\sqrt{2}} \right]$$

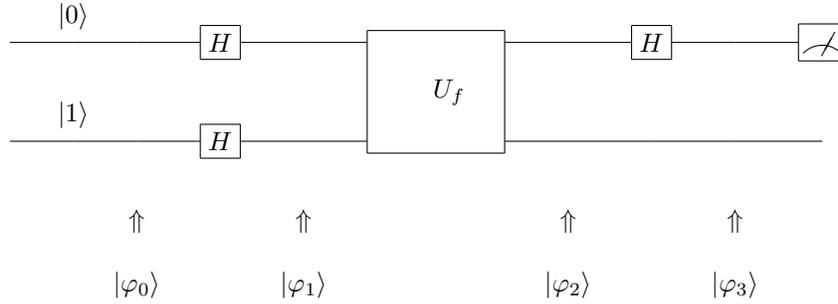where $\overline{f(x)}$ means the opposite of $f(x)$. And so we have

$$|\varphi_2\rangle = \begin{cases} |x\rangle \left[\dfrac{|0\rangle - |1\rangle}{\sqrt{2}} \right] & \text{if } f(x) = 0 \\[3ex] |x\rangle \left[\dfrac{|1\rangle - |0\rangle}{\sqrt{2}} \right] & \text{if } f(x) = 1 \end{cases}.$$

Remembering that $a - b = (-1)(b - a)$ we might write this as

$$|\varphi_2\rangle = (-1)^{f(x)} |x\rangle \left[\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right].$$

What would happen if we evaluate either the top or the bottom state? Again, this does not really help us. We do not gain any information. The top qubit will be in state $|x\rangle$ and the bottom qubit will — with equal probability — be either in state $|0\rangle$ or in state $|1\rangle$. We need something more.

4.3. **Deutsch's Algorithm.** Now let us combine both of these attempts to actually give Deutsch's algorithm. Deutsch's algorithm works by putting *both* the top and bottom qubits into a superposition. We will also put the results of the top qubit through a Hadamard matrix.

In terms of matrices this becomes:

$$(I \otimes H)U_f(H \otimes H)|0,1\rangle$$

At each point of the algorithm the states are as follows.

$$|\varphi_0\rangle = |0,1\rangle.$$

$$|\varphi_1\rangle = \left[\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right]\left[\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right] = \frac{+|0,0\rangle - |0,1\rangle + |1,0\rangle - |1,1\rangle}{2} = \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix}\begin{bmatrix} +\frac{1}{2} \\ -\frac{1}{2} \\ +\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}.$$

We saw from our last attempt at solving this problem, that when we put the bottom qubit into a superposition and then multiply by U_f we will be in a superposition

$$(-1)^{f(x)}|x\rangle\left[\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right].$$

Now with $|x\rangle$ in a superposition, we have

$$|\varphi_2\rangle = \left[\frac{(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle}{\sqrt{2}}\right]\left[\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right].$$

Let us carefully look at

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

If f is constant this becomes either

$$+1(|0\rangle + |1\rangle) \text{ or } -1(|0\rangle + |1\rangle)$$

(depending on being constantly 0 or constantly 1.) If f is balanced it becomes either

$$+1(|0\rangle - |1\rangle) \text{ or } -1(|0\rangle - |1\rangle)$$

(depending on which way it is balanced.) Summing up, we have that

$$|\varphi_2\rangle = \begin{cases} (\pm 1) \left[\dfrac{|0\rangle + |1\rangle}{\sqrt{2}} \right] \left[\dfrac{|0\rangle - |1\rangle}{\sqrt{2}} \right] & \text{if } f \text{ is constant} \\[4ex] (\pm 1) \left[\dfrac{|0\rangle - |1\rangle}{\sqrt{2}} \right] \left[\dfrac{|0\rangle - |1\rangle}{\sqrt{2}} \right] & \text{if } f \text{ is balanced} \end{cases}$$

Remembering that the Hadamard matrix is its own inverse and takes $\dfrac{|0\rangle + |1\rangle}{\sqrt{2}}$ to $|0\rangle$ and takes $\dfrac{|0\rangle - |1\rangle}{\sqrt{2}}$ to $|1\rangle$, we apply the Hadamard matrix to the top qubit to get

$$|\varphi_3\rangle = \begin{cases} (\pm 1)|0\rangle \left[\dfrac{|0\rangle - |1\rangle}{\sqrt{2}} \right] & \text{if } f \text{ is constant} \\[4ex] (\pm 1)|1\rangle \left[\dfrac{|0\rangle - |1\rangle}{\sqrt{2}} \right] & \text{if } f \text{ is balanced} \end{cases}$$

Now we simply measure the top qubit. If it is in state $|0\rangle$, then we know that f is a constant function, otherwise it is a balanced function. And we did this all with only one evaluation as opposed to the two evaluations that a classical algorithm demands.

Notice that although the ± 1 tells us even more information, namely which of the two balanced functions or which of the two constant functions, we are not privy to this information. Upon measuring, if the function is balanced, we will measure $|1\rangle$ regardless if the state was $(-1)|1\rangle$ or $(+1)|1\rangle$.

In conclusion, we have solved a problem that a classical computer would require two function evaluations. Deutsch's quantum algorithm solved the same problem with one function evaluation. Other quantum algorithms are built-up with similar ideas to the ones presented here.

Department of Computer and Information Science,
Brooklyn College, CUNY,
Brooklyn, N.Y. 11210.

and

Computer Science Department,
The Graduate Center, CUNY,
New York, N.Y. 10016.

e-mail: noson@sci.brooklyn.cuny.edu