# Chapter 1

# Introduction

Nature at the sub-atomic scale behaves totally differently from anything that our experience with the physical world prepares us for. **Quantum mechanics** is the name of the branch of physics that governs the world of elementary particles such as electrons and photons, and it is paradoxical, unintuitive, and radically strange. Below is a sampling of a few such odd features:

- Complete knowledge of a system's state is forbidden - A measurement reveals only a small amount of information about the quantum state of the system.
- The act of measuring a particle fundamentally disturbs its state.
- Quantum entities do not have trajectories. All that we can say for an elementary particle is that it started at A and was measured later at B. We cannot say anything about the trajectory or path it took from A to B.
- Quantum Mechanics is inherently probabilistic. If we prepare two elementary particles in identical states and measure them, the results may be different for each particle.
- Quantum entities behave in some ways like particles and in others like waves. But they really behave in their own unique way, neither particles nor waves.

These features are truly strange, and difficult to accept. To quote the great physicist Niels Bohr, "Anyone who is not shocked by quantum theory has not understood it." We start by describing a simple experiment that highlights many differences between quantum mechanics and our classical intuition. It

is the famous double slit experiment. The intuition gained in the process will help us as we define qubits, and more generally in the study of quantum computing.

## 1.1 The Double Slit Experiment

What is the nature of light? You may have learned light is electromagnetic *waves* propagating through space. Also, you may have learned that light is made of a rain of individual *particles* called photons. But these two notions seem contradictory, how can it be both?

The debate over the nature of light goes deep into the history of science. The eminent physicist Isaac Newton believed that light was a rain of particles, called corpuscles. At the beginning of the nineteenth century, Thomas Young demonstrated with his famous double-slit interference experiment that light propagates as waves. With Maxwell's formulation of electromagnetism at the end of the nineteenth century, it was generally accepted that light is propagated as electromagnetic waves, and the debate seemed to be over. However, in 1905, Einstein was able to explain the photoelectric effect, by using the idea of light quanta, or particles which we now call photons.

Similar confusion reigned over the nature of electrons, which behaved like particles, but then it was discovered in electron diffraction experiments, performed in 1927, that they exhibit wave behavior. So do electrons behave like particles or waves? And what about photons? This great challenge was resolved with the discovery of the equations of quantum mechanics. But the theory is not intuitive, and its description of matter is very different from our common experience.

To understand what seems to be a paradox, we look to Young's double-slit experiment. Here's the set up: a source of light is shone at a screen with two very thin, identical slits cut into it. Some light passes through the two slits and lands upon a subsequent screen. Take a look at Figure 1.1 for a diagram of the experiment setup.

First, think about what would happen to a stream of bullets going through this double slit experiment. The source, which we think of as a machine gun, is unsteady and sprays the bullets in the general direction of the two slits. Some bullets pass through one slit, some pass through the other slit, and others don't make it through the slits. The bullets that do go through the slits then land on the observing screen behind them. Now suppose we closed slit 2. Then the bullets can only go through slit 1 and land in a small spread behind slit 1. If we graphed the number of times a bullet that went through slit 1 landed at

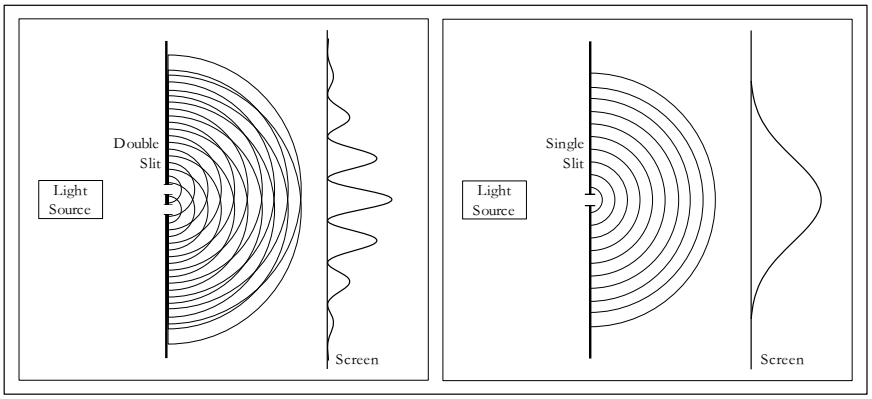



Figure 1.1: Double- and single-slit diffraction. Notice that in the double-slit experiment the two paths interfere with one another. This experiment gives evidence that light propagates as a wave.

the position $y$ on the observation screen, we would see a normal distribution centered directly behind slit 1. That is, most land directly behind the slit, but some stray off a little due to the small amount randomness inherent in the gun, and because of they ricochet off the edges of the slit. If we now close slit 1 and open slit 2, we would see a normal distribution centered directly behind slit 2.

Now let's repeat the experiment with both slits open. If we graph the number of times a bullet that went through *either* slit landed at the position $y$, we should see the sum of the graph we made for slit 1 and a the graph for slit 2.

Another way we can think of the graphs we made is as graphs of the probability that a bullet will land at a particular spot $y$ on the screen. Let $P_1(y)$ denote the probability that the bullet lands at point $y$ when only slit 1 is open, and similarly for $P_2(y)$. And let $P_{12}(y)$ denote the probability that the bullet lands at point $y$ when both slits are open. Then $P_{12}(y) = P_1(y) + P_2(y)$.

Next, we consider the situation for waves, for example water waves. A water wave doesn't go through *either* slit 1 or slit 2, it goes through both. You should imagine the crest of 1 water wave as it approaches the slits. As it hits the slits, the wave is blocked at all places but the two slits, and waves on the other side are generated at each slit as depicted in Figure 1.1.

When the new waves generated at each slit run into each other, interference occurs. We can see this by plotting the intensity (that is, the amount of energy

carried by the waves) at each point $y$ along the viewing screen. What we see is the familiar interference pattern seen in Figure 1.1. The dark patches of the interference pattern occur where the wave from the first slit arrives perfectly out of sync with wave from the second slit, while the bright points are where the two arrive in sync. For example, the bright spot right in the middle is bright because each wave travels the exact same distance from their respective slit to the screen, so they arrive in sync. The first dark spots are where the wave from one slit traveled exactly half of a wavelength longer than the other wave, thus they arrive at opposite points in their cycle and cancel. Here, it is not the intensities coming from each slit that add, but height of the wave. This differs from the case of bullets: $I_{12}(y) \neq I_1(y) + I_2(y)$, but $h_{12}(y) = h_1(y) + h_2(y)$, and $I_{12}(y) = h(y)^2$, where $h(y)$ is the height of the wave and $I(y)$ is the intensity, or energy, of the wave.

Before we can say what light does, we need one more crucial piece of information. What happens when we turn down the intensity in both of these examples?

In the case of bullets, turning down the intensity means turning down the rate at which the bullets are fired. When we turn down the intensity, each time a bullet hits the screen it transfers the same amount of energy, but the frequency at which bullets hit the screen becomes less.

With water waves, turning down the intensity means making the wave amplitudes smaller. Each time a wave hits the screen it transfers *less* energy, but the frequency of the waves hitting the screen is unchanged.

Now, what happens when we do this experiment with light. As Young observed in 1802, light makes an interference pattern on the screen. From this observation he concluded that the nature of light is wavelike, and reasonably so! However, Young was unable at the time to turn down the intensity of light enough to see the problem with the wave explanation.

Picture now that the observation screen is made of thousands of tiny little photo-detectors that can detect the energy they absorb. For high intensities the photo-detectors individually are picking up a lot of energy, and when we plot the intensity against the position $y$ along the screen we see the same interference pattern described earlier. Now, turn the intensity of the light very very very low. At first, the intensity scales down lower and lower everywhere, just like with a wave. But as soon as we get low enough, the energy that the photo-detectors report reaches a minimum energy, and all of the detectors are reporting the same energy, call it $E_0$, just at different rates. This energy corresponds to the energy carried by an individual photon, and at this stage we see what is called the quantization of light.

Photo-detectors that are in the bright spots of the interference pattern

report the energy $E_0$ very frequently, while darker areas report the energy $E_0$ at lower rates. Totally dark points still report nothing. This behavior is the behavior of *bullets*, not waves! We now see that photons behave unlike either bullets or waves, but like something entirely different.

Turn down the intensity so low that only one photo-detector reports something each second. In other words, the source only sends one photon at a time. Each time a detector receives a photon, we record where on the array it landed and plot it on a graph. The distribution we draw will reflect the *probability* that a single photon will land at a particular point.

Logically we think that the photon will either go through one slit or the other. Then, like the bullets, the probability that the photon lands at a point should be $y$ is $P_{12}(y) = P_1(y) + P_2(y)$ and the distribution we expect to see is the two peaked distribution of the bullets. *But this not what we see at all.*

What we actually see is the same interference pattern from before! But how can this be? For there to be an interference pattern, light coming from one slit must interfere with light from the other slit; but there is only one photon going through at a time! The modern explanation is that the photon actually goes through *both* slits at the same time, and interferes with *itself.* The mathematics is analogous to that in the case of water waves. We say that the probability $P(y)$ that a photon is detected at $y$ is proportional to the square of some quantity $a(y)$, which we call a probability amplitude. Now probability amplitudes for different alternatives add up. So $a_{12}(y) = a_1(y) + a_2(y)$. But $P_{12}(y) = |a_{12}(y)|^2 \neq |a_1(y)|^2 + |a_2(y)|^2 = P_1(y) + P_2(y)$.

Logically, we can ask which slit the photon went through, and try to measure it. Thus, we might construct a double slit experiment where we put a photodetector at each slit, so that each time a photon comes through the experiment we see which slit it went through and where it hits on the screen. But when such an experiment is preformed, the interference pattern gets completely washed out! The very fact that we know which slit the photon goes through makes the interference pattern go away. This is the first example we see of how measuring a quantum system alters the system.

Here the photon looks both like a particle, a discreet package, and a wave that can have interference. It seems that the photon acts like both a wave and a particle, but at the same time it doesn't exactly behave like either. This is what is commonly known as the wave-particle duality, usually thought of as a paradox. The resolution is that the quantum mechanical behavior of matter is unique, something entirely new.

What may be more mind blowing still is that if we conduct the exact same experiment with *electrons* instead of light, we get the exact same results! Although it is common to imagine electrons as tiny little charged spheres,

they are actually quantum entities, neither wave nor particle but understood by their wavefunction.

The truth is that there is no paradox, just an absence of intuition for quantum entities. Why should they be intuitive? Things on our scale do not behave like wavefunctions, and unless we conduct wild experiments like this we do not see the effects of quantum mechanics. The following sections describe in more detail some of the basic truths of quantum mechanics, so that we can build an intuition for a new behavior of matter.

## 1.2 The Axioms of Quantum Mechanics

"I think I can safely say that nobody understands quantum mechanics."

-Richard Feynman

Paradoxically, the fundamental principles of quantum mechanics can be stated very concisely and simply. The challenge lies in understanding and applying these principles, which is the goal of the rest of the book. Here are two basic

- The superposition principle explains how a particle can be superimposed between two states at the same time.
- The measurement principle tells us how measuring a particle changes its state, and how much information we can access from a particle.
- The unitary evolution axoim governs how the state of the quantum system evolves in time.

In keeping with our philosophy, we will introduce the basic axioms gradually, starting with simple finite systems, and simplified basis state measurements, and building our way up to the more general formulations. This should allow the reader a chance to develop some intuition about these topics.

## 1.3 The Superposition Principle

Consider a system with $k$ distinguishable (classical) states. For example, the electron in a hydrogen atom is only allowed to be in one of a discrete set of energy levels, starting with the ground state, the first excited state, the second excited state, and so on. If we assume a suitable upper bound on the total

energy, then the electron is restricted to being in one of $k$ different energy levels — the ground state or one of $k-1$ excited states. As a classical system, we might use the state of this system to store a number between 0 and $k-1$. The superposition principle says that if a quantum system can be in one of two states then it can also be placed in a linear superposition of these states with complex coefficients.

Let us introduce some notation. We denote the ground state of our $k$-state system by $|0\rangle$, and the succesive excited states by $|1\rangle, \ldots, |k-1\rangle$. These are the $k$ possible classical states of the electron. The superposition principle tells us that, in general, the quantum state of the electron is $\alpha_0 |0\rangle + \alpha_1 |1\rangle + \cdots + \alpha_{k-1} |k-1\rangle$, where $\alpha_0, \alpha_1, \ldots, \alpha_{k-1}$ are complex numbers normalized so that $\sum_j |\alpha_j|^2 = 1$. $\alpha_j$ is called the *amplitude of the state* $|j\rangle$. For instance, if $k = 3$, the state of the electron could be

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

or

$$|\psi\rangle = \frac{1}{\sqrt{2}} |0\rangle - \frac{1}{2} |1\rangle + \frac{i}{2} |2\rangle$$

or

$$|\psi\rangle = \frac{1+i}{3} |0\rangle - \frac{1-i}{3} |1\rangle + \frac{1+2i}{3} |2\rangle .$$

The superposition principle is one of the most mysterious aspects about quantum physics — it flies in the face of our intuitions about the physical world. One way to think about a superposition is that the electron does not make up its mind about whether it is in the ground state or each of the $k-1$ excited states, and the amplitude $\alpha_0$ is a measure of its inclination towards the ground state. Of course we cannot think of $\alpha_0$ as the probability that an electron is in the ground state — remember that $\alpha_0$ can be negative or imaginary. The measurement priniciple, which we will see shortly, will make this interpretation of $\alpha_0$ more precise.

## 1.4 The Geometry of Hilbert Space

We saw above that the quantum state of the $k$-state system is described by a sequence of $k$ complex numbers $\alpha_0, \ldots, \alpha_{k-1} \in \mathbb{C}$, normalized so that $\sum_j |\alpha_j|^2 = 1$. So it is natural to write the state of the system as a $k$ dimen-

sional vector:

$$|\psi\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{k-1} \end{pmatrix}$$

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

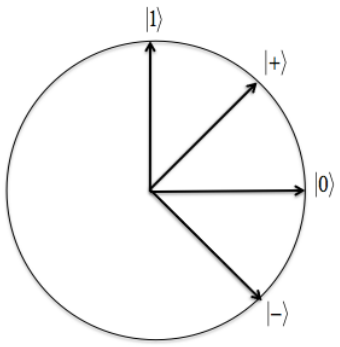



Figure 1.2: Representation of qubit states as vectors in a Hilbert space.

But hold on! Earlier we wrote the quantum state in a very different (and simpler) way as: $\alpha_0 |0\rangle + \alpha_1 |1\rangle + \cdots + \alpha_{k-1} |k-1\rangle$. Actually this notation, called Dirac's ket notation, is just another way of writing a vector. Thus

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |k-1\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

So we have an underlying geometry to the possible states of a quantum system: the $k$ distinguishable (classical) states $|0\rangle, \ldots, |k-1\rangle$ are represented by mutually orthogonal unit vectors in a $k$-dimensional complex vector space. i.e. they form an orthonormal basis for that space (called the standard basis). Moreover, given any two states, $\alpha_0 |0\rangle + \alpha_1 |1\rangle + \cdots + \alpha_{k-1} |k-1\rangle$, and $\beta |0\rangle + \beta |1\rangle + \cdots + \beta |k-1\rangle$, we can compute the inner product of these two vectors, which is $\sum_{j=0}^{k-1} \alpha_j^* \beta_j$. The absolute value of the inner product is the cosine of the angle between these two vectors in Hilbert space. You should verify that

the inner product of any two basis vectors in the standard basis is 0, showing that they are orthogonal.

The advantage of the ket notation is that the it labels the basis vectors explicitly. This is very convenient because the notation expresses both that the state of the quantum system is a vector, while at the same time explicitly writing out the physical quantity of interest (energy level, position, spin, polarization, etc).

## 1.5 Bra-ket Notation

In this section we detail the notation that we will use to describe a quantum state, $|\psi\rangle$. This notation is due to Dirac and, while it takes some time to get used to, is incredibly convenient.

### Inner Products

We saw earlier that all of our quantum states live inside a Hilbert space. A Hilbert space is a special kind of vector space that, in addition to all the usual rules with vector spaces, is also endowed with an inner product. And an inner product is a way of taking two states (vectors in the Hilbert space) and getting a number out. For instance, define

$$|\psi\rangle = \sum_k a_k \, |k\rangle \, ,$$

where the kets $|k\rangle$ form a basis, so are orthogonal. If we instead write this state as a column vector,

$$|\psi\rangle = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{N-1} \end{pmatrix}$$

Then the inner product of $|\psi\rangle$ with itself is

$$\langle \psi, \psi \rangle = \begin{pmatrix} a_0^* & a_1^* & \cdots & a_{N_1}^* \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{N-1} \end{pmatrix} = \sum_{k=0}^{N-1} a_k^* a_k = \sum_{k=0}^{N-1} |a_k|^2$$

The complex conjugation step is important so that when we take the inner product of a vector with itself we get a real number which we can associate

with a length. Dirac noticed that there could be an easier way to write this by defining an object, called a "bra," that is the conjugate-transpose of a ket,

$$\langle\psi| = |\psi\rangle^\dagger = \sum_k a_k^* \langle k| \,.$$

This object acts on a ket to give a number, as long as we remember the rule,

$$\langle j| \, |k\rangle \equiv \langle j|k\rangle = \delta_{jk}$$

Now we can write the inner product of $|\psi\rangle$ with itself as

$$\begin{aligned}
\langle\psi|\psi\rangle &= \left(\sum_j a_j^* \langle j|\right)\left(\sum_k a_k |k\rangle\right) \\
&= \sum_{j,k} a_j^* a_k \langle j|k\rangle \\
&= \sum_{j,k} a_j^* a_k \delta_{jk} \\
&= \sum_k |a_k|^2
\end{aligned}$$

Now we can use the same tools to write the inner product of any two states, $|\psi\rangle$ and $|\phi\rangle$, where

$$|\phi\rangle = \sum_k b_k |k\rangle \,.$$

Their inner product is,

$$\langle\psi|\phi\rangle = \sum_{j,k} a_j^* b_k \langle j|k\rangle = \sum_k a_k^* b_k$$

Notice that there is no reason for the inner product of two states to be real (unless they are the same state), and that

$$\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^* \in \mathbb{C}$$

In this way, a bra vector may be considered as a "functional." We feed it a ket, and it spits out a complex number.

**The Dual Space**

We mentioned above that a bra vector is a *functional* on the Hilbert space. In fact, the set of all bra vectors forms what is known as the *dual space*. This space is the set of *all* linear functionals that can act on the Hilbert space.

## 1.6 The Measurement Principle

This linear superposition $|\psi\rangle = \sum_{j=0}^{k-1} \alpha_j |j\rangle$ is part of the private world of the electron. Access to the information describing this state is severely limited — in particular, we cannot actually measure the complex amplitudes $\alpha_j$. This is not just a practical limitation; it is enshrined in the measurement postulate of quantum physics.

A measurement on this $k$ state system yields one of at most $k$ possible outcomes: i.e. an integer between 0 and $k-1$. Measuring $|\psi\rangle$ in the standard basis yields $j$ with probability $|\alpha_j|^2$.

One important aspect of the measurement process is that it alters the state of the quantum system: the effect of the measurement is that the new state is exactly the outcome of the measurement. I.e., if the outcome of the measurement is $j$, then following the measurement, the qubit is in state $|j\rangle$. This implies that you cannot collect any additional information about the amplitudes $\alpha_j$ by repeating the measurement.

Intuitively, a measurement provides the only way of reaching into the Hilbert space to probe the quantum state vector. In general this is done by selecting an orthonormal basis $|e_0\rangle, \ldots, |e_{k-1}\rangle$. The outcome of the measurement is $|e_j\rangle$ with probability equal to the square of the length of the projection of the state vector $\psi$ on $|e_j\rangle$. A consequence of performing the measurement is that the new state vector is $|e_j\rangle$. Thus measurement may be regarded as a probabilistic rule for projecting the state vector onto one of the vectors of the orthonormal measurement basis.

Some of you might be puzzled about how a measurement is carried out physically? We will get to that soon when we give more explicit examples of quantum systems.

## 1.7 Qubits

Qubits (pronounced "cue-bit") or quantum bits are basic building blocks that encompass all fundamental quantum phenomena. They provide a mathemat-

ically simple framework in which to introduce the basic concepts of quantum physics. Qubits are 2-state quantum systems. For example, if we set $k = 2$, the electron in the Hydrogen atom can be in the ground state or the first excited state, or any superposition of the two. We shall see more examples of qubits soon.

The state of a qubit can be written as a unit (column) vector $\binom{\alpha}{\beta} \in \mathbb{C}^2$. In Dirac notation, this may be written as:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \qquad \text{with} \quad \alpha, \beta \in \mathbb{C} \quad \text{and} \quad |\alpha|^2 + |\beta|^2 = 1.$$

This linear superposition $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ is part of the private world of the electron. For us to know the electron's state, we must make a measurement. Making a measurement gives us a single classical bit of information — 0 or 1. The simplest measurement is in the standard basis, and measuring $|\psi\rangle$ in this $\{|0\rangle, |1\rangle\}$ basis yields 0 with probability $|\alpha|^2$, and 1 with probability $|\beta|^2$.

One important aspect of the measurement process is that it alters the state of the qubit: the effect of the measurement is that the new state is exactly the outcome of the measurement. *I.e.*, if the outcome of the measurement of $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ yields 0, then following the measurement, the qubit is in state $|0\rangle$. This implies that you cannot collect any additional information about $\alpha$, $\beta$ by repeating the measurement.

More generally, we may choose any orthogonal basis $\{|v\rangle, |w\rangle\}$ and measure the qubit in that basis. To do this, we rewrite our state in that basis: $|\psi\rangle = \alpha'\,|v\rangle + \beta'\,|w\rangle$. The outcome is $v$ with probability $|\alpha'|^2$, and $|w\rangle$ with probability $|\beta'|^2$. If the outcome of the measurement on $|\psi\rangle$ yields $|v\rangle$, then as before, the the qubit is then in state $|v\rangle$.

## Examples of Qubits

### Atomic Orbitals

The electrons within an atom exist in quantized energy levels. Qualitatively these electronic orbits (or "orbitals" as we like to call them) can be thought of as resonating standing waves, in close analogy to the vibrating waves one observes on a tightly held piece of string. Two such individual levels can be isolated to configure the basis states for a qubit.

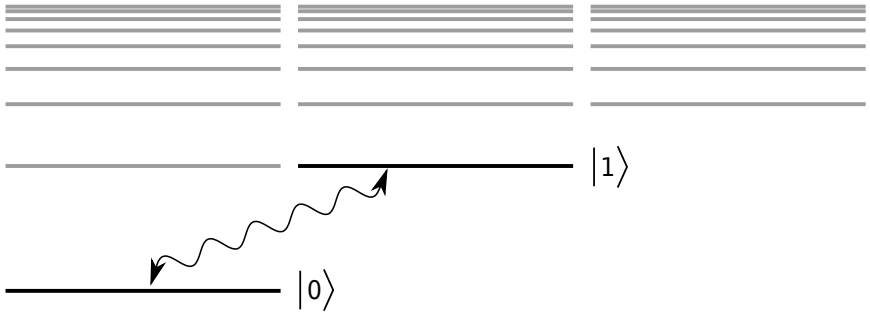



Figure 1.3: Energy level diagram of an atom. Ground state and first excited state correspond to qubit levels, $|0\rangle$ and $|1\rangle$, respectively.

### Photon Polarization

Classically, a photon may be described as a traveling electromagnetic wave. This description can be fleshed out using Maxwell's equations, but for our purposes we will focus simply on the fact that an electromagnetic wave has a *polarization* which describes the orientation of the electric field oscillations (see Fig. 1.4). So, for a given direction of photon motion, the photon's polarization axis might lie anywhere in a 2-d plane perpendicular to that motion. It is thus natural to pick an orthonormal 2-d basis (such as $\vec{x}$ and $\vec{y}$, or "vertical" and "horizontal") to describe the polarization state (i.e. polarization direction) of a photon. In a quantum mechanical description, this 2-d nature of the photon polarization is represented by a qubit, where the amplitude of the overall polarization state in each basis vector is just the projection of the polarization in that direction.

The polarization of a photon can be measured by using a polaroid film or a calcite crystal. A suitably oriented polaroid sheet transmits x-polarized photons and absorbs y-polarized photons. Thus a photon that is in a superposition $|\phi\rangle = \alpha\,|\mathsf{x}\rangle + \beta\,|\mathsf{y}\rangle$ is transmitted with probability $|\alpha|^2$. If the photon now encounters another polariod sheet with the same orientation, then it is transmitted with probability 1. On the other hand, if the second polaroid sheet has its axes crossed at right angles to the first one, then if the photon is transmitted by the first polaroid, then it is definitely absorbed by the second sheet. This pair of polarized sheets at right angles thus blocks all the light. A somewhat counter-intuitive result is now obtained by interposing a third polariod sheet at a 45 degree angle between the first two. Now a photon that is transmitted by the first sheet makes it through the next two with probability

1/4.

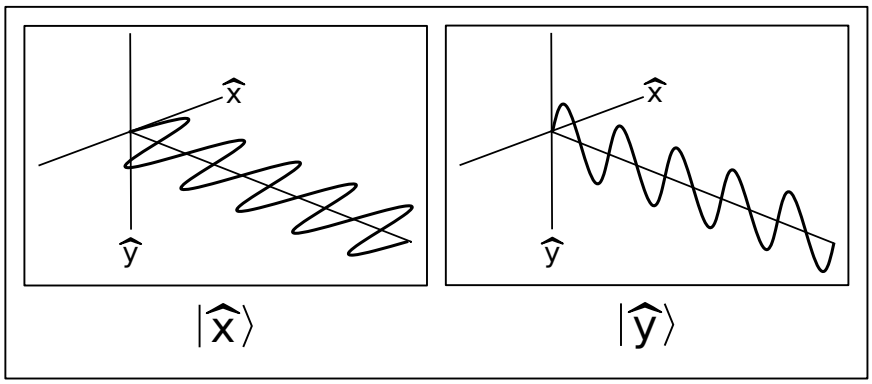



Figure 1.4: Using the polarization state of light as the qubit. Horizontal polarization corresponds to qubit state, $|\hat{x}\rangle$, while vertical polarization corresponds to qubit state, $|\hat{y}\rangle$.

To see this first observe that any photon transmitted through the first filter is in the state, $|0\rangle$. The probability this photon is transmitted through the second filter is $1/2$ since it is exactly the probability that a qubit in the state $|0\rangle$ ends up in the state $|+\rangle$ when measured in the $|+\rangle, |-\rangle$ basis. We can repeat this reasoning for the third filter, except now we have a qubit in state $|+\rangle$ being measured in the $|0\rangle, |1\rangle$-basis — the chance that the outcome is $|0\rangle$ is once again $1/2$.

**Spins**

Like photon polarization, the spin of a (spin-1/2) particle is a two-state system, and can be described by a qubit. Very roughly speaking, the spin is a quantum description of the magnetic moment of an electron which behaves like a spinning charge. The two allowed states can roughly be thought of as clockwise rotations ("spin-up") and counter clockwise rotations ("spin-down"). We will say much more about the spin of an elementary particle later in the course.

**Measurement Example I: Phase Estimation**

Now that we have discussed qubits in some detail, we can are prepared to look more closesly at the measurement principle. Consider the quantum state,

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{e^{i\theta}}{\sqrt{2}}|1\rangle .$$

If we were to measure this qubit in the standard basis, the outcome would be 0 with probability 1/2 and 1 with probability 1/2. This measurement tells us only about the norms of the state amplitudes. Is there any measurement that yields information about the phase, $\theta$?

To see if we can gather any phase information, let us consider a measurement in a basis other than the standard basis, namely

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

What does $|\phi\rangle$ look like in this new basis? This can be expressed by first writing,

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

Now we are equipped to rewrite $|\psi\rangle$ in the $\{|+\rangle, |-\rangle\}$-basis,

$$\begin{aligned} |\psi\rangle &= \frac{1}{\sqrt{2}}|0\rangle + \frac{e^{i\theta}}{\sqrt{2}}|1\rangle) \\ &= \frac{1}{2}(|+\rangle + |-\rangle) + \frac{e^{i\theta}}{2}(|+\rangle - |-\rangle) \\ &= \frac{1+e^{i\theta}}{2}|+\rangle + \frac{1-e^{i\theta}}{2}|-\rangle \quad . \end{aligned}$$

Recalling the Euler relation, $e^{i\theta} = \cos\theta + i\sin\theta$, we see that the probability of measuring $|+\rangle$ is $\frac{1}{4}((1+\cos\theta)^2 + \sin^2\theta) = \cos^2(\theta/2)$. A similar calculation reveals that the probability of measuring $|-\rangle$ is $\sin^2(\theta/2)$. Measuring in the $(|+\rangle, |-\rangle)$-basis therefore reveals some information about the phase $\theta$.

Later we shall show how to analyze the measurement of a qubit in a general basis.

**Measurement example II: General Qubit Bases**

What is the result of measuring a general qubit state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$, in a general orthonormal basis $|v\rangle\,, \left|v^{\perp}\right\rangle$, where $|v\rangle = a|0\rangle + b|1\rangle$ and $|v^{\perp}\rangle = b^*|0\rangle - a^*|1\rangle$? You should also check that $|v\rangle$ and $\left|v^{\perp}\right\rangle$ are orthogonal by showing that $\left\langle v^{\perp}|v\right\rangle = 0$.

To answer this question, let us make use of our recently acquired bra-ket notation. We first show that the states $|v\rangle$ and $\left|v^{\perp}\right\rangle$ are orthogonal, that is, that their inner product is zero:

$$\begin{aligned}
\left\langle v^{\perp}|v\right\rangle &= \left(b^*\,|0\rangle - a^*\,|1\rangle\right)^{\dagger}\left(a\,|0\rangle + b\,|1\rangle\right) \\
&= \left(b\,\langle 0| - a\,\langle 1|\right)^{\dagger}\left(a\,|0\rangle + b\,|1\rangle\right) \\
&= ba\,\langle 0|0\rangle - a^2\,\langle 1|0\rangle + b^2\,\langle 0|1\rangle - ab\,\langle 1|1\rangle \\
&= ba - 0 + 0 - ab \\
&= 0
\end{aligned}$$

Here we have used the fact that $\langle i|j\rangle = \delta_{ij}$.

Now, the probability of measuring the state $|\psi\rangle$ and getting $|v\rangle$ as a result is,

$$\begin{aligned}
P_{\psi}(v) &= |\langle v|\psi\rangle|^2 \\
&= \left|\left(a^*\,\langle 0| + b^*\,\langle 1|\right)\left(\alpha\,|0\rangle + \beta\,|1\rangle\right)\right|^2 \\
&= |a^*\alpha + b^*\beta|^2
\end{aligned}$$

Similarly,

$$\begin{aligned}
P_{\psi}(v^{\perp}) &= \left|\left\langle v^{\perp}|\psi\right\rangle\right|^2 \\
&= \left|\left(b\,\langle 0| - a\,\langle 1|\right)\left(\alpha\,|0\rangle + \beta\,|1\rangle\right)\right|^2 \\
&= |b\alpha - a\beta|^2
\end{aligned}$$

# Chapter 2

# Entanglement

What are the allowable quantum states of systems of several particles? The answer to this is enshrined in the addendum to the first postulate of quantum mechanics: the superposition principle. In this chapter we will consider a special case, systems of two qubits. In keeping with our philosophy, we will first approach this subject naively, without the formalism of the formal postulate. This will facilitate an intuitive understanding of the phenomenon of quantum entanglement — a phenomenon which is responsible for much of the "quantum weirdness" that makes quantum mechanics so counter-intuitive and fascinating.

## 2.1 Two qubits

Now let us examine a system of two qubits. Consider the two electrons in two hydrogen atoms, each regarded as a 2-state quantum system:

Since each electron can be in either of the ground or excited state, classically the two electrons are in one of four states – 00, 01, 10, or 11 – and represent 2 bits of classical information. By the superposition principle, the quantum state of the two electrons can be any linear combination of these four classical states:

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

where $\alpha_{ij} \leq \mathbb{C}$, $\sum_{ij} |\alpha_{ij}|^2 = 1$. Of course, this is just Dirac notation for the unit vector in $\mathbb{C}^4$:

$$\begin{pmatrix} \alpha_{00} \\ \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{pmatrix}$$

### Measurement

As in the case of a single qubit, even though the state of two qubits is specified by four complex numbers, most of this information is not accessible by measurement. In fact, a measurement of a two qubit system can only reveal two bits of information. The probability that the outcome of the measurement is the two bit string $x \in \{0,1\}^2$ is $|\alpha_x|^2$. Moreover, following the measurement the state of the two qubits is $|x\rangle$. i.e. if the first bit of $x$ is $j$ and the second bit $k$, then following the measurement, the state of the first qubit is $|j\rangle$ and the state of the second is $|k\rangle$.

An interesting question comes up here: what if we measure just the first qubit? What is the probability that the outcome is 0? This is simple. It is exactly the same as it would have been if we had measured both qubits: $\Pr\{\text{1st bit } = 0\} = \Pr\{00\} + \Pr\{01\} = |\alpha_{00}|^2 + |\alpha_{01}|^2$. Ok, but how does this partial measurement disturb the state of the system?

The answer is obtained by an elegant generalization of our previous rule for obtaining the new state after a measurement. The new superposition is obtained by crossing out all those terms of $|\psi\rangle$ that are inconsistent with the outcome of the measurement (i.e. those whose first bit is 1). Of course, the sum of the squared amplitudes is no longer 1, so we must renormalize to obtain a unit vector:

$$|\phi_{\text{new}}\rangle = \frac{\alpha_{00}|00\rangle + \alpha_{01}|01\rangle}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}}$$

### Entanglement

Suppose the first qubit is in the state $3/5\,|0\rangle + 4/5\,|1\rangle$ and the second qubit is in the state $1/\sqrt{2}\,|0\rangle - 1/\sqrt{2}\,|1\rangle$, then the joint state of the two qubits is $(3/5\,|0\rangle + 4/5\,|1\rangle)(1/\sqrt{2}\,|0\rangle - 1/\sqrt{2}\,|1\rangle) = 3/5\sqrt{2}\,|00\rangle - 3/5\sqrt{2}\,|01\rangle + 4/5\sqrt{2}\,|10\rangle - 4/5\sqrt{2}\,|11\rangle$.

More generally, if the state of the first qubit is $\alpha_0|0\rangle + \alpha_1|1\rangle$ and the state of the second qubit is $\beta_0|0\rangle + \beta_1|1\rangle$, then the joint state of the two qubits is $\alpha_0\beta_0|00\rangle + \alpha_0\beta_1|01\rangle + \alpha_1\beta_0|10\rangle + \alpha_1\beta_1|11\rangle$.

Can every state of two qubits be decomposed in this way? Our classical intuition would suggest that the answer is obviously affirmative. After all each of the two qubits must be in some state $\alpha|0\rangle + \beta|1\rangle$, and so the state of the two qubits must be the product. In fact, there are states such as $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ which cannot be decomposed in this way as a state of the first qubit and that of the second qubit. Can you see why? Such a state is called an entangled state. When the two qubits are entangled, we cannot determine the state of each qubit separately. The state of the qubits

has as much to do with the relationship of the two qubits as it does with their individual states.

If the first (resp. second) qubit of $|\Phi^+\rangle$ is measured then the outcome is 0 with probability 1/2 and 1 with probability 1/2. However if the outcome is 0, then a measurement of the second qubit results in 0 with certainty. This is true no matter how large the spatial separation between the two particles.

The state $|\Phi^+\rangle$, which is one of the Bell basis states, has a property which is even more strange and wonderful. The particular correlation between the measurement outcomes on the two qubits holds true no matter which rotated basis a rotated basis $|v\rangle, |v^\perp\rangle$ the two qubits are measured in, where $|0\rangle = \alpha|v\rangle + \beta|v^\perp\rangle$ and $|1\rangle = -\beta|v\rangle + \alpha|v^\perp\rangle$. This can bee seen as,

$$\begin{aligned}
|\Phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\
&= \frac{1}{\sqrt{2}}\left(\left(\alpha|v\rangle + \beta|v^\perp\rangle\right) \otimes \left(\alpha|v\rangle + \beta|v^\perp\rangle\right)\right) \\
&\quad - \frac{1}{\sqrt{2}}\left(\left(-\beta|v\rangle + \alpha|v^\perp\rangle\right) \otimes \left(-\beta|v\rangle + \alpha|v^\perp\rangle\right)\right) \\
&= \frac{1}{\sqrt{2}}\left(\left(\alpha^2+\beta^2\right)|vv\rangle + \left(\alpha^2+\beta^2\right)|v^\perp v^\perp\rangle\right) \\
&= \frac{1}{\sqrt{2}}\left(|vv\rangle + |v^\perp v^\perp\rangle\right)
\end{aligned}$$

### EPR Paradox:

Everyone has heard Einstein's famous quote "God does not play dice with the Universe". The quote is a summary of the following passage from Einstein's 1926 letter to Max Born: "Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not really bring us any closer to the secret of the Old One. I, at any rate, am convinced that He does not throw dice." Even to the end of his life, Einstein held on to the view that quantum physics is an incomplete theory and that some day we would learn a more complete and satisfactory theory that describes nature.

In what sense did Einstein consider quantum mechanics to be incomplete? Think about flipping a coin. For all common purposes, the outcome of a coin toss is random — heads half the time, and tails the other half. And this lines up exactly with our observations, but we know that randomness isn't the whole story. A more complete theory would say that if we knew *all* of the

initial conditions of the coin *exactly* (position, momentum), then we could use Newton's laws of classical physics to figure out exactly how the coin would land, and therefore the outcome of the coin flip. Another way to say this is that the coin flip amplifies our lack of knowledge about the state of the system, and makes the outcome seem completely random. In the same way, Einstein believed that the randomness of quantum measurements reflected our lack of knowledge about additional degrees of freedom, or "hidden variables," of the quantum system.

Einstein sharpened this line of reasoning in a paper he wrote with Podolsky and Rosen in 1935, where they introduced the famous Bell states. The EPR argument works like this. For Bell state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, when you measure first qubit (in the bit basis), the second qubit is determined (in the bit basis). What is even more remarkable is that if you measure the first qubit in the sign basis, the second qubit is determined in the sign basis. You should verify that in the sign basis $(|+\rangle, |-\rangle)$, the state $|\Phi^+\rangle$ can be written as $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|++\rangle + |--\rangle)$. It follows that knowledge of the sign of one qubit completely determines the other.

Now lets suppose the qubits are very far apart, say one light-second. If we measure qubit 1 in the standard basis, then measure qubit 2 a half second later in the same basis, the two measurements must agree. Then qubit 2 must have been in a definite state for a half second before it was measured: from the instant we measured qubit 1, we knew qubit 2. But the qubits couldn't have communicated any information in that time.

What if we had measured the first qubit in the $|+\rangle, |-\rangle$ basis instead? Then similarly for half a second before we measure qubit 2, it was in a definite state in the $|+\rangle, |-\rangle$ basis. But qubit 2 could not possibly know which basis qubit 1 was measured in until a full second *after* we measure qubit 1! This is because we assumed that light takes a one second to travel from qubit 1 to qubit 2. This appears to contradict the uncertainty principle for $|\pm\rangle$ and $|0\rangle, ket1$ states says that there is no definite $|\pm\rangle$ state that is also a definite $|0\rangle, |1\rangle$ state.

Einstein, Podolsky, and Rosen concluded that since qubit 2 cannot have any information about which basis qubit 1 was measured in, its state in both bit and sign bases is simultaneously determined, something that quantum mechanics does not allow. EPR therefore suggested that quantum mechanics is an incomplete theory, and there is a more complete theory where "God does not throw dice." Until his death in 1955, Einstein tried to formulate a more complete "local hidden variable theory" that would describe the predictions of quantum mechanics, but without resorting to probabilistic outcomes.

But in 1964, almost three decades after the EPR paper, John Bell showed that properties of Bell (EPR) states were not merely fodder for a philosophical discussion, but had verifiable consequences: local hidden variables[1] are not the answer. He described an experiment to be performed on two qubits entangled in a Bell state such that a local hidden variable theory would disagree with quantum mechanics about the outcome. The Bell experiment has been performed to increasing accuracy, originally by Aspect, and the results have always been consistent with the predictions of quantum mechanics and inconsistent with local hidden variable theories.

## 2.2 Bell's Thought Experiment

Bell considered the following experiment: let us assume that two particles are produced in the Bell state $|\Phi^+\rangle$ in a laboratory, and the fly in opposite directions to two distant laboratories. Upon arrival, each of the two qubits is subject to one of two measurements. The decision about which of the two experiments is to be performed at each lab is made randomly at the last moment, so that speed of light considerations rule out information about the choice at one lab being transmitted to the other. The measurements are cleverly chosen to distinguish between the predictions of quantum mechanics and any local hidden variable theory. Concretely, the experiment measures the correlation between the outcomes of the two experiments. The choice of measurements is such that any classical hidden variable theory predicts that the correlation between the two outcomes can be at most 0.75, whereas quantum mechanics predicts that the correlation is $\cos^2 \pi/8 \approx 0.85$. Thus the experiment allows us to distinguish between the predictions of quantum mechanics and any local hidden variable theory! We now describe the experiment in more detail.

The two experimenters A and B (for Alice and Bob) each receives one qubit of a Bell state $|\Phi^+\rangle$, and measures it in one of two bases depending upon the value of a random bit $r_A$ and $r_B$ respectively. Denote by $a$ and $b$ respectively the outcomes of the measurements. We are interested in the highest achievable correlation between the two quantities $r_A \times r_B$ and $a + b(mod 2)$. We will see below that there is a particular choice of bases for the quantum measurements made by A and B such that $P[r_A \times r_B = a+b(mod 2)] = \cos^2 \pi/8 \approx .85$. Before we do so, let us see why no classical hidden variable theory allows a correlation of over 0.75. i.e. $P[r_A \times r_B = a + b(mod 2)] \leq 0.75$.

[1] We will describe what we mean by a local hidden variable theory below after we start describing the actual experiment

We can no longer postpone a discussion about what a local hidden variable theory is. Let us do so in the context of the Bell experiment. In a local hidden variable theory, when the Bell state was created, the two particles might share an arbitrary amount of classical information, $x$. This information could help them coordinate their responses to any measurements they are subjected to in the future. By design, the Bell experiment selects the random bits $r_A$ an $r_B$ only after the two particles are too far apart to exchange any further information before they are measured. Thus we are in the setting, where A and B share some arbitrary classical information $x$, and are given as input independent, random bits $x_A$ an $x_B$ as input, and must output bits $a$ and $b$ respectively to maximize their chance of achieving $r_A \times r_B = a + b(mod 2)$. It can be shown that the shared information $x$ is of no use in increasing this correlation, and indeed, the best they can do is to always output $a = b = 0$. This gives $P[r_A \times r_B = a + b(mod 2)] \leq .75$.

Let us now describe the quantum measurements that achieve greater correlation. They are remarkably simple to describe:

- if $r_A = 0$, then Alice measures in the standard $|0\rangle / |1\rangle$ basis.
- if $r_A = 1$, then Alice measures in the $\pi/4$ basis (i.e. standard basis rotated by $\pi/4$).
- if $r_B = 0$, then Bob measures in the $\pi/8$ basis.
- if $r_B = 1$, then Bob measures in the $-\pi/8$ basis.

The analysis of the success probability of this experiment is also beautifully simple. We will show that in each of the four cases $r_A = r_B = 0$, etc, the success probability $P[r_A \times r_B = a + b(mod 2)] = \cos^2 \pi/8$.

We first note that if Alice and Bob measure in bases that make an angle $\theta$ with each other, then the chance that their measurement outcomes are the same (bit) is exactly $\cos^2 \theta$. This follows from the rotational invariance of $|\Phi^+\rangle$ and the following observation: if the first qubit is measured in the standard basis, then the outcome is outcome is an unbiased bit. Moreover the state of the second qubit is exactly equal to the outcome of the measurement — $|0\rangle$ if the measurement outcome is 0, say. But now if the second qubit is measured in a basis rotated by $\theta$, then the probability that the outcome is also 0 is exactly $\cos^2 \theta$.

Now observe that in three of the four cases, where $x_A \cdot x_B = 0$, Alice and Bob measure in bases that make an angle of $\pi/8$ with each other. By our observation above, $P[a + b \equiv 0 \mod 2] = P[a = b] = \cos^2 \pi/8$.

In the last case $x_A \cdot x_B = 1$, and they measure in bases that make an angle of $3\pi/8$ with each other. So, $P[a + b \equiv 0 \mod 2] = P[a \neq b] = \cos^2 3\pi/8 = \sin^2(\pi/2 - 3\pi/8) = sin^2\pi/8$. Therefore, $P[a + b \equiv 1 \mod 2] = 1 - \sin^2 \pi/8 = \cos^2 \pi/8$. So in each of the four cases, the chance that Alice and Bob succeed is $\cos^2 \pi/8 \approx .85$

# Chapter 3

# Quantum Gates, Circuits & Teleportation

### Unitary Operators

The third postulate of quantum physics states that the evolution of a quantum system is necessarily unitary. Geometrically, a unitary transformation is a rigid body rotation of the Hilbert space, thus resulting in a transformation of the state vector that doesn't change its length.

Let us consider what this means for the evolution of a qubit. A unitary transformation on the Hilbert space $\mathbb{C}^2$ is specified by mapping the basis states $|0\rangle$ and $|1\rangle$ to orthonormal states $|v_0\rangle = a\,|0\rangle + b\,|1\rangle$ and $|v_1\rangle = c\,|0\rangle + d\,|1\rangle$. It is specified by the linear transformation on $\mathbb{C}^2$:

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

If we denote by $U^\dagger$ the conjugate transpose of this matrix:

$$U^\dagger = \begin{pmatrix} a^* & b^* \\ c^* & d^* \end{pmatrix}$$

then it is easily verified that $UU^\dagger = U^\dagger U = I$. Indeed, we can turn this around and say that a linear transformation $U$ is unitary if and only if it satisfies this condition, that

$$UU^\dagger = U^\dagger U = I.$$

Let us now consider some examples of unitary transformations on single qubits or equivalently single qubit quantum gates:

- Hadamard Gate. Can be viewed as a reflection around $\pi/8$ in the real plane. In the complex plane it is actually a $\pi$-rotation about the $\pi/8$ axis.

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

  The Hadamard Gate is one of the most important gates. Note that $H^\dagger = H$ – since $H$ is real and symmetric – and $H^2 = I$.

- Rotation Gate. This rotates the plane by $\theta$.

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

- NOT Gate. This flips a bit from 0 to 1 and vice versa.

$$NOT = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

- Phase Flip.

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

  The phase flip is a NOT gate acting in the $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ basis. Indeed, $Z|+\rangle = |-\rangle$ and $Z|-\rangle = |+\rangle$.

How do we physically effect such a (unitary) transformation on a quantum system? To explain this we must first introduce the notion of the Hamiltonian acting on a system; you will have to wait for three to four lectures before we get to those concepts.

### Two Qubit Gates

Recall that the third axiom of quantum physics states that the evolution of a quantum system is necessarily unitary. Intuitively, a unitary transformation is a rigid body rotation of the Hilbert space. In particular it does not change the length of the state vector.

Let us consider what this means for the evolution of a two qubit system. A unitary transformation on the Hilbert space $\mathbb{C}^4$ is specified by a $4x4$ matrix $U$ that satisfies the condition $UU^\dagger = U^\dagger U = I$. The four columns of $U$ specify the four orthonormal vectors $|v_{00}\rangle$, $|v_{01}\rangle$, $|v_{10}\rangle$ and $|v_{11}\rangle$ that the basis states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$ are mapped to by $U$.

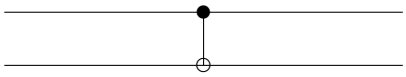
A very basic two qubit gate is the controlled-not gate or the CNOT:

**Controlled Not (CNOT)**

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

The first bit of a CNOT gate is called the "control bit," and the second the "target bit." This is because (in the standard basis) the control bit does not change, while the target bit flips if and only if the control bit is 1.

The CNOT gate is usually drawn as follows, with the control bit on top and the target bit on the bottom:



Though the CNOT gate looks very simple, any unitary transformation on two qubits can be closely approximated by a sequence of CNOT gates and single qubit gates. This brings us to an important point. What happens to the quantum state of two qubits when we apply a single qubit gate to one of them, say the first? Let's do an example. Suppose we apply a Hadamard gate to the superposition: $|\psi\rangle = 1/2\,|00\rangle - i/\sqrt{2}\,|01\rangle + 1/\sqrt{2}\,|11\rangle$. Then this maps the first qubit as follows:

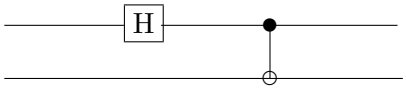
$$\begin{aligned} |0\rangle &\to 1/\sqrt{2}\,|0\rangle + 1/\sqrt{2}\,|1\rangle \\ |1\rangle &\to 1/\sqrt{2}\,|0\rangle - 1/\sqrt{2}\,|1\rangle\,. \end{aligned}$$

So

$$\begin{aligned} |\psi\rangle &\to 1/2\sqrt{2}\,|00\rangle + 1/2\sqrt{2}\,|01\rangle - i/2\,|00\rangle + i/2\,|01\rangle + 1/2\,|10\rangle - 1/2\,|11\rangle \\ &= (1/2\sqrt{2} - i/2)\,|00\rangle + (1/2\sqrt{2} + i/2)\,|01\rangle + 1/2\,|10\rangle - 1/2\,|11\rangle\,. \end{aligned}$$

**Bell states**

We can generate the Bell states $|\Phi^+\rangle = \frac{1}{\sqrt{2}}\,(|00\rangle + |11\rangle)$ with the following simple quantum circuit consisting of a Hadamard and CNOT gate:



The first qubit is passed through a Hadamard gate and then both qubits are entangled by a CNOT gate.

If the input to the system is $|0\rangle \otimes |0\rangle$, then the Hadamard gate changes the state to

$$\tfrac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle = \tfrac{1}{\sqrt{2}}|00\rangle + \tfrac{1}{\sqrt{2}}|10\rangle \ ,$$

and after the CNOT gate the state becomes $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, the Bell state $|\Phi^+\rangle$.

Notice that the action of the CNOT gate is not so much copying, as our classical intuition would suggest, but rather to entangle.

The state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is one of four Bell basis states:

$$\begin{aligned} \left|\Phi^{\pm}\right\rangle &= \tfrac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \\ \left|\Psi^{\pm}\right\rangle &= \tfrac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \ . \end{aligned}$$

These maximally entangled states on two qubits form an orthonormal basis for $\mathbb{C}^4$. Exercise: give a simple quantum circuit for generating each of these states, and prove that the Bell basis states form an orthonormal basis for $\mathbb{C}^4$.

### Tensor Products

So far we have avoided a discussion of the addendum to the superposition axiom, which tells us the allowable states of a composite quantum system consisting of two subsystems. The basic question for our example of a two qubit system is this: how do the 2-dimensional Hilbert spaces corresponding to each of the two qubits relate to the 4-dimensional Hilbert space corresponding to the composite system? i.e. how do we glue two 2-dimensional Hilbert spaces to get a 4-dimensional Hilbert space? This is done by taking a tensor product of the two spaces.

Let us describe this operation of taking tensor products in a slightly more general setting. Suppose we have two quantum systems - a $k$-state system with associated $k$-dimensional Hilbert space $V$ with orthonormal basis $|0\rangle, \ldots, |k-1\rangle$ and a $l$-state system with associated $l$-dimensional Hilbert space $W$ with orthonormal basis $|0\rangle, \ldots, |l-1\rangle$. What is resulting Hilbert space obtained by gluing these two Hilbert spaces together? We can answer this question as follows: there are $kl$ distinguishable states of the composite system — one for each choice of basis state $|i\rangle$ of the first system and basis state $|j\rangle$ of the second system. We denote the resulting of dimension $kl$ Hilbert

space by $V \otimes W$ (pronounced "$V$ tensor $W$"). The orthonormal basis for this new Hilbert space is given by:

$$\{|i\rangle \otimes |j\rangle : 0 \leq i \leq k-1, 0 \leq j \leq l-1\},$$

So a typical element of $V \otimes W$ will be of the form $\sum_{ij} \alpha_{ij}(|i\rangle \otimes |j\rangle)$.

In our example of a two qubit system, the Hilbert space is $\mathbb{C}^2 \otimes \mathbb{C}^2$, which is isomorphic to the four dimensional Hilbert space $\mathbb{C}^4$. Here we are identifying $|0\rangle \otimes |0\rangle$ with $|00\rangle$.

### Tensor product of operators

Suppose we have two quantum systems: a $k$-state system with associated Hilbert space $V$ and a $l$-state system with associated Hilbert space $W$. Suppose we apply a unitary transformation $A$ to the first system and $B$ to the second system. What is the resulting transformation on the combined system $V \otimes W$? To figure this out, let us first see how the combined transformation acts on basis states of $V \otimes W$. Consider a basis state $|i\rangle \otimes \{ketj$ where $0 \leq i \leq k-1$ and $0 \leq j \leq l-1$. Since $A$ is only acting on $V$ and $B$ only on $W$, this state is transformed to $A|i\rangle \otimes B|j\rangle$.

Suppose $|v\rangle$ and $|w\rangle$ are unentangled states on $\mathbb{C}^m$ and $\mathbb{C}^n$, respectively. The state of the combined system is $|v\rangle \otimes |w\rangle$ on $\mathbb{C}^{mn}$. If the unitary operator $A$ is applied to the first subsystem, and $B$ to the second subsystem, the combined state becomes $A|v\rangle \otimes B|w\rangle$.

In general, the two subsystems will be entangled with each other, so the combined state is not a tensor-product state. We can still apply $A$ to the first subsystem and $B$ to the second subsystem. This gives the operator $A \otimes B$ on the combined system, defined on entangled states by linearly extending its action on unentangled states.

(For example, $(A \otimes B)(|0\rangle \otimes |0\rangle) = A|0\rangle \otimes B|0\rangle$. $(A \otimes B)(|1\rangle \otimes |1\rangle) = A|1\rangle \otimes B|1\rangle$. Therefore, we define $(A \otimes B)(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle)$ to be $\frac{1}{\sqrt{2}}(A \otimes B)|00\rangle + \frac{1}{\sqrt{2}}(A \otimes B)|11\rangle = \frac{1}{\sqrt{2}}(A|0\rangle \otimes B|0\rangle + A|1\rangle \otimes B|1\rangle)$.)

Let $|e_1\rangle, \ldots, |e_m\rangle$ be a basis for the first subsystem, and write $A = \sum_{i,j=1}^{m} a_{ij}|e_i\rangle\langle e_j|$ (the $i$,$j$th element of $A$ is $a_{ij}$). Let $|f_1\rangle, \ldots, |f_n\rangle$ be a basis for the second subsystem, and write $B = \sum_{k,l=1}^{n} b_{kl}|f_k\rangle\langle f_l|$. Then a basis for the combined system is $|e_i\rangle \otimes |f_j\rangle$, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. The operator $A \otimes B$ is

$$\begin{aligned} A \otimes B = & \left( \sum_{ij} a_{ij} |e_i\rangle\langle e_j| \right) \otimes \left( \sum_{kl} b_{kl} |f_k\rangle\langle f_l| \right) \\ = & \sum_{ijkl} a_{ij} b_{kl} |e_i\rangle\langle e_j| \otimes |f_k\rangle\langle f_l| \\ = & \sum_{ijkl} a_{ij} b_{kl} (|e_i\rangle \otimes |f_k\rangle)(\langle e_j| \otimes \langle f_l|) \ . \end{aligned}$$

Therefore the $(i,k),(j,l)$th element of $A \otimes B$ is $a_{ij}b_{kl}$. If we order the basis $|e_i\rangle \otimes |f_j\rangle$ lexicographically, then the matrix for $A \otimes B$ is

$$\begin{pmatrix} a_{11}B & a_{12}B & \cdots \\ a_{21}B & a_{22}B & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \ ;$$

in the $i,j$th sub-block, we multiply $a_{ij}$ by the matrix for $B$.

## 3.1 No Cloning Theorem and Quantum Teleportation

The axioms of quantum mechanics are deceptively simple. Our view is that to begin to understand and appreciate them you have to be exposed to some of their most counterintuitive consequences. Paradoxically, this will help you build a better intuition for quantum mechanics.

In this chapter we will study three very simple but counterintuitive consequences of the laws of quantum mechanics. The theme of all three vignettes is the copying or transmission of quantum information.

### No Cloning Theorem

Given a quantum bit in an unknown state $|\phi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$, is it possible to make a copy of this quantum state? i.e. create the state $|\phi\rangle \otimes |\phi\rangle = (\alpha_0 |0\rangle + \alpha_1 |1\rangle) \otimes (\alpha_0 |0\rangle + \alpha_1 |1\rangle)$? The axioms of quantum mechanics forbid this very basic operation, and the proof of the no cloning theorem helps gain insight into this.

To be more precise, we are asking whether it is possible to start with two qubits in state $|\phi\rangle \otimes |0\rangle$ and transform them to the state $|\phi\rangle \otimes |\phi\rangle$? By the third axiom of quantum mechanics, for this to be possible there must be a

unitary transformation $U$ such that $U|\phi\rangle \otimes |0\rangle = |\phi\rangle \otimes |\phi\rangle$. We will show that no unitary transformation can achieve this simultaneously for two orthogonal states $|\phi\rangle$ and $|\psi\rangle$.

Recall that a unitary transformation is a rotation of the Hilbert space, and therefore necessarily preserves angles. Let us make this more precise. Consider two quantum states (say on a single qubit): $|\phi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ and $|\psi\rangle = \beta_0|0\rangle + \beta_1|1\rangle$. The cosine of the angle between them is given by (the absolute value of) their inner product: $\alpha_0^*\beta_0 + \alpha_1^*\beta_1$.

Now consider the quantum states (on two qubits) $|\phi\rangle \otimes |\phi\rangle = (\alpha_0|0\rangle + \alpha_1|1\rangle)(\alpha_0|0\rangle + \alpha_1|1\rangle)$ and $|\psi\rangle \otimes |\phi\rangle = (\beta_0|0\rangle + \beta_1|1\rangle)(\beta_0|0\rangle + \beta_1|1\rangle)$. Their inner product is: $(\alpha_0^*\beta_0 + \alpha_1^*\beta_1)^2$. i.e. $\langle\phi|\,\psi\rangle^2 = \langle\phi\phi|\,\psi\psi\rangle$.

We are now ready to state and prove the no cloning theorem:

Assume we have a unitary operator $U$ and two quantum states $|\phi\rangle$ and $|\psi\rangle$:

$$\begin{aligned} |\phi\rangle \otimes |0\rangle &\xrightarrow{U} |\phi\rangle \otimes |\phi\rangle \\ |\psi\rangle \otimes |0\rangle &\xrightarrow{U} |\psi\rangle \otimes |\psi\rangle \quad . \end{aligned}$$

Then $\langle\phi|\psi\rangle$ is 0 or 1.

$\langle\phi|\psi\rangle = (\langle\phi|\otimes\langle 0|)(|\psi\rangle\otimes|0\rangle) = (\langle\phi|\otimes\langle\phi|)(|\psi\rangle\otimes|\psi\rangle) = \langle\phi|\psi\rangle^2$. In the second equality we used the fact that $U$, being unitary, preserves inner products.

### Superdense Coding

Suppose Alice and Bob are connected by a *quantum* communications channel. By this we mean, for example, that they can communicate qubits over an optical fibre using polarized photons. Is this much more powerful than a classical communication channel, over which only classical bits may be transmitted? The answer seems obvious, since a classical bit is a special case of a quantum bit. And a qubit appears to encode an infinite number of bits of information, since to specify its state we must specify two complex numbers. However, the truth is a little more subtle, since the axioms of quantum mechanics also severely restrict how we may access information about the quantum state by a measurement.

So the question we wish to ask is "how many classical bits can Alice transmit to Bob in a message consisting of a single qubit?" We will show that if Alice and Bob share entanglement in the form of a Bell state, then Alice can transmit two classical bits by transmitting just one qubit over the quantum channel.

The overall idea is this: say Alice and Bob share $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Alice can transform this shared state to any of the four Bell basis states $|\Phi^+\rangle$, $|\Phi^-\rangle$, $|\Psi^+\rangle$, $|\Psi^-\rangle$ by applying a suitable quantum gate just to her qubit. Now if she transmits her qubit to Bob, he holds both qubits of of a Bell basis state and can perform a measurement in the Bell basis to distinguish which of the four states he holds.

Let's now see the details of Alice's protocol: if Alice wishes to transmit the two bit message $b_1 b_2$, she applies a bit flip $X$ to her qubit if $_1 = 1$ and a phase flip $Z$ to her qubit if $b_2 = 1$. You should verify that in the four cases 00, 01, 10, 11 this results in the two qubits being in the state $|\Phi^+\rangle$, $|\Phi^-\rangle$, $|\Psi^+\rangle$, $|\Psi^-\rangle$ respectively.

After receiving Alice's qubit, Bob measures the two qubits in the Bell basis by running the circuit we saw in chapter 2 backwards (i.e., applying $(H \otimes I) \circ CNOT$), then measuring in the standard basis.

Note that Alice really did use two qubits total to transmit the two classical bits. After all, Alice and Bob somehow had to start with a shared Bell state. However, the first qubit – Bob's half of the Bell state – could have been sent well before Alice had decided what message she wished to send to Bob.

One can show that it is not possible to do any better. No more than two classical bits can be transmitted by sending just one qubit. To see why you will have to understand our next example.

### Quantum Teleportation

After months of effort, Alice has managed to synthesize a special qubit, which she strongly suspects has some wonderful physical properties. Unfortunately, she doesn't explicitly know the state vector $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$. And she does not have the equipment in her lab to carry out a crucial next phase of her experiment. Luckily Bob's lab has the right equipment, though it is at the other end of town. Is there a way for Alice to safely transport her qubit to Bob's lab?

If Alice and Bob share a Bell state, then there is a remarkable method for Alice to transmit her qubit to Bob. The method requires her to make a certain measurement on her two qubits: the qubit she wishes to transmit and her share of the Bell state. She then calls up Bob on the phone and tells him the outcome of her measurement — just two classical bits. Depending upon which of four outcomes Alice announces to him on the phone, Bob performs one of four operations on his qubit, and voila, his qubit is in the state $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$!

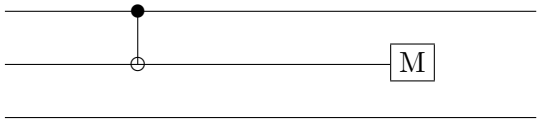
But hold on a moment, doesn't this violate the no cloning theorem?! No, because Alice's qubit was destroyed by measurement before Bob created his copy. Let us build our way to the teleportation protocol in a couple of simple stages:

Let us start with the following scenario. Alice and Bob share two qubits in the state $a\,|00\rangle + b\,|11\rangle$. Alice and Bob don't know the amplitudes $a$ and $b$. How can Bob end up with the state $a\,|0\rangle + b\,|1\rangle$? An easy way to achieve this is to perform a CNOT gate on the two qubits with Bob's qubit as the control, and Alice's qubit as the target. But this requires an exchange of quantum information. What if Alice and Bob can only exchange classical information?

Here is a way. Alice performs a Hadamard on her qubit. The state of the two qubits is now $a/\sqrt{2}(|0\rangle + |1\rangle)\,|0\rangle + b/\sqrt{2}(|0\rangle - |1\rangle)\,|1\rangle = 1/\sqrt{2}\,|0\rangle\,(a\,|0\rangle + b\,|1\rangle) + 1/\sqrt{2}\,|1\rangle\,(a\,|1\rangle - b\,|1\rangle)$. Now if Alice measures her qubit in the standard basis, if the measurement outcome is 0, then Bob's qubit is the desired $a\,|0\rangle + b\,|1\rangle$. If the measurement outcome is 1, then Bob's qubit is $a\,|0\rangle - b\,|1\rangle$. But in this case if Bob were to apply a phase flip gate (Z) to his qubit, it would end up in the desired state $a\,|0\rangle + b\,|1\rangle$.

Back to teleportation. Alice has a qubit in state $a\,|0\rangle + b\,|1\rangle$, and Alice and Bob share a Bell state. Is there any way for them to convert their joint state to $a\,|00\rangle + b\,|11\rangle$, without exchanging any quantum information? If they succeed, then by our previous discussion Alice can teleport her qubit to Bob.

Consider what happens if Alice applies a CNOT gate with her qubit $a\,|0\rangle + b\,|1\rangle$ as the control qubit, and her share of the Bell state as the target qubit.




$$|\phi\rangle \otimes |\psi\rangle = \sum_{i=0,1} a_i |i\rangle \otimes \sum_{j=0,1} \frac{1}{\sqrt{2}} |j, j\rangle.$$

After passing through the CNOT gate this becomes

$$\sum_{i,j} a_i \big|i, i \oplus j, j\big\rangle.$$

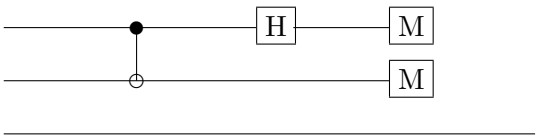
Now $A$ measures the middle qubit. Suppose it is measured as $l$; then $l = i \oplus j \implies j = i \oplus l$. The state is now

$$\sum_{i} a_i \big|i, i \oplus l\big\rangle.$$

Next, $A$ transmits $l$ to $B$. If $l = 0$, $B$ takes no action, while if $l = 1$, then $B$ performs a bit flip, $X$, on his qubit, resulting in the desired state

$$\sum_i a_i \big| i, i \big\rangle.$$

We already saw how to teleport once we achieved this state. Putting it all together, the following quantum circuit describes the resulting teleportation protocol. The topmost qubit is the unknown qubit that Alice wishes to teleport, the second and third qubits are initially in a Bell state:




The measurement of the middle qubit after performing the CNOT gate sets up the state $\sum_j a_j \big| j, j \big\rangle$ on the first and third qubit. Moreover, $A$'s application of the Hadamard gate on the first qubit induces the transformation

$$\sum_j a_j \big| j, j \big\rangle \longrightarrow \sum_{ij} a_j (-1)^{ij} \big| i, j \big\rangle.$$

Finally $A$ measures $i$ and sends the measurement to $B$. The state is now:

$$\sum_j a_j (-1)^{ij} |j\rangle.$$

If $i = 0$ then we are done; if $i = 1$ then $B$ applies a phase flip. In either case the state is now $a_0|0\rangle + a_1|1\rangle$.

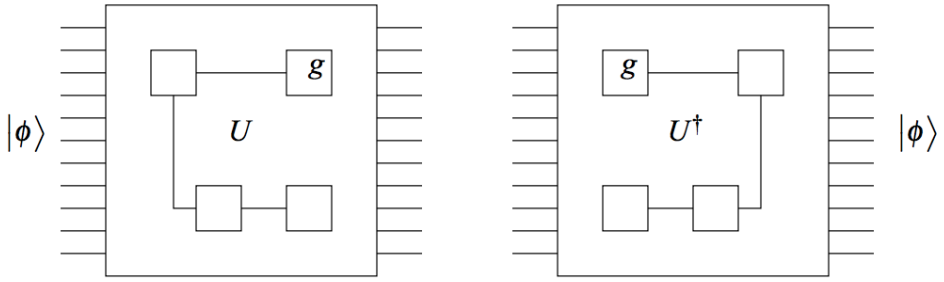
So $A$ has transported the quantum state to $B$ simply by sending two classical bits.

# Chapter 4

# Fourier Sampling & Simon's Algorithm

## 4.1 Reversible Computation

A quantum circuit acting on $n$ qubits is described by an $2^n \times 2^n$ unitary operator $U$. Since $U$ is unitary, $UU^\dagger = U^\dagger U = I$. This implies that each quantum circuit has an inverse circuit which is the mirror image of the original circuit and which carries out the inverse operator $U^\dagger$.




The circuits for $U$ and $U^\dagger$ are the same size and have mirror image gates. Examples:

$$H = H^\dagger$$

$$\text{CNOT} = \text{CNOT}^\dagger$$

$$R_\theta = R_{-\theta}^\dagger$$

## 4.2 Simulating Classical Circuits

Let us first consider whether given any classical circuit there is an equivalent quantum circuit. More concretely, suppose there is a classical circuit that computes a function $f(x) \in \{0,1\}^m$ on input $x \in \{0,1\}^n$, is there a quantum circuit that does the same? Obviously such a quantum circuit must map computational basis states to computational basis states (i.e. it must map each state of the form $|x\rangle$ to the state $|f(x)\rangle$). A unitary transformation taking basis states to basis states must be a permutation. (Indeed, if $U|x\rangle = |u\rangle$ and $U|y\rangle = |u\rangle$, then $|x\rangle = U^{-1}|u\rangle = |y\rangle$.) Therefore we need the input, or domain, to be the exact same number of bits as the range: $m = n$. What is more, the function $f(x)$ must be a permutation on the $n$-bit strings. Since this must hold after every application of a quantum gate, it follows that if a quantum circuit computes a classical function, then it must be *reversible*: it must have an inverse.

How can a classical circuit $C$ which takes an $n$ bit input $x$ and computes $f(x)$ be made into a reversible quantum circuit that computes the same function? The circuit must never lose any information, so how could it compute a function mapping $n$ bits to $m < n$ bits (e.g. a boolean function, where $m = 1$)?

The solution to this problem is to have the circuit take the $n$ input qubits in the state $|x\rangle$ and send them to $|x\rangle$, while in the process taking some $m$ qubits in the $|0\rangle$ state to $|f(x)\rangle$. Then the inverse map is simple: the $n$-bit string $|x\rangle$ goes back to $\mathbf{x}$, and $|f(x)\rangle$ goes to an $m$ bit string of 0's: $|0\rangle$.

However, this is not always perfectly easy, some times to make the circuit work it needs scratch qubits in the input. A scratch qubit is a qubit that starts out in the $|0\rangle$ state, and ends in the $|0\rangle$ state. Its purpose is to be used in computations inside of the circuit. Of course, since the quantum circuit does not alter these qubits, the inverse circuit also leaves them alone. While these bits are a necessary ingredient to a reversible quantum circuit, they are not the main character and are often let out of circuit diagrams. Take a look at Figure 4.2 for the full picture.

How is this done? It is a fact that any classical AND and OR gates can be simulated with a C-SWAP gate and some scratch $|0\rangle$ qubits (Figure 4.2). For example, if we want to make $a \wedge b$, then we input $a$ as the control bit and $b$ as one of the swap bits, with $c = 0$ as the other swap bit. In the end, we measure the third register where $c$ went in, and this will be $a \wedge b$ Now look what happens: if $a = b = 1$, then $b$ and $c$ swap, so that the third register reads $b = 1$: true! If $a$ is one but $b$ is 0, then $b$ and $c$ swap, but the third register is $b = 0$: false. And clearly, if $a = 0$ then so to will read the third register.

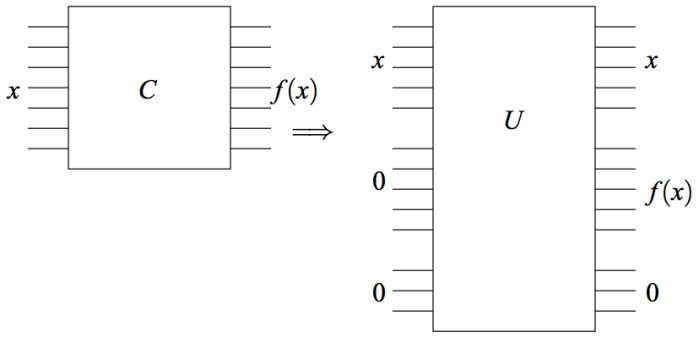
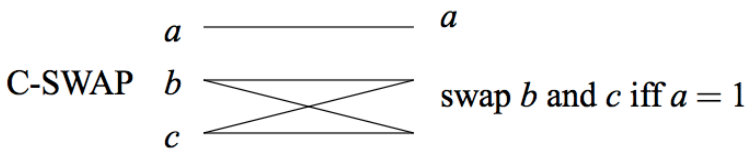



Figure 4.1: Note that the input and output have the same number of qubits in the reversible quantum circuit.



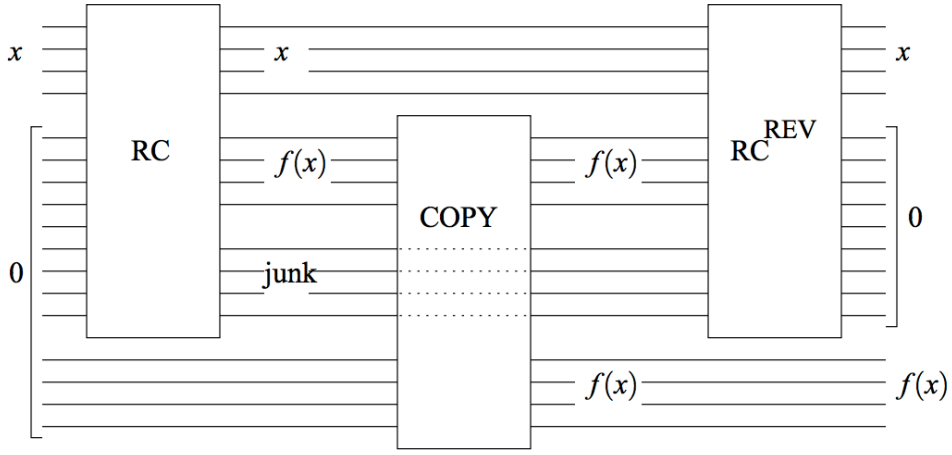

If we construct the corresponding reversible circuit (lets call it RC), we have a small problem. The CSWAP gates end up converting input scratch bits to garbage. Why is this a problem? We have our output $|f\rangle(x)$, don't we. This seems like it should be good enough. But in fact it is not. All the junk that gets made in the in-between steps can be entangled with the output qubits. Thus if it gets measured, it will screw alter our function. Furthermore, there is a principle which states that any unmeasured, thrown away qubits are just as good as measured. This is called the principle of deferred measurement, and it means that junk qubits are no good.

So how do we restore the scratch bits to 0 on output? We use the fact that RC is a reversible circuit. We use the CSWAP gates, for example, to produce the output $f(x)$. We can then copy the output onto some scratch qubits, which we will keep as our output. We then use the reverse of our circuit RC on the input, $f(x)$, and the junk to turn it all back to 0's and $x$. But because we copied $f(x)$ in the middle, we keep it at the end.

The sequence of steps for the overall circuit is

$$(x, 0^k, 0^m, 0^k, 1) \xrightarrow{C'} (x, y, \text{garbage}_x, 0^k, 1) \xrightarrow{\text{copy } y} (x, y, \text{garbage}_x, y, 1) \xrightarrow{(C')^{-1}} (x, 0^k, 0^m, y, 1) \ .$$

Overall, this gives us a clean reversible circuit $\hat{C}$ corresponding to $C$.




## 4.3 Fourier Sampling

Consider a quantum circuit acting on $n$ qubits, which applies a Hadamard gate to each qubit. i.e. the circuit applies the unitary transformation $H^{\otimes n}$, or $H$ tensored with itself $n$ times.

Another way to define this unitary transformation $H_{2^n}$ is as the $2^n \times 2^n$ matrix in which the $(x, y)$ entry is $2^{-n/2}(-1)^{x \cdot y}$.

Applying the Hadamard transform (or the Fourier transform over $Z_2^n$) to the state of all zeros gives an equal superposition over all $2^n$ states

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

In general, applying the Hadamard transform to the computational basis state $|u\rangle$ yields:

$$\mathcal{H}_{2^n} |u\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} (-1)^{u \cdot x} |x\rangle$$

We define the Fourier sampling problem as follows: Input an $n$ qubit state $|\phi\rangle = \sum_{x \in \{0,1\}^n} \alpha_x |x\rangle$. Compute $H^{\otimes n} |\phi\rangle$ and measure the resulting state $\sum_y \hat{\alpha}_y |y\rangle$ to output $y$ with probability $|\hat{\alpha}_y|^2$.

Fourier sampling is probably the most fundamental primitive we use in quantum algorithms (where in place of the Hadamard transform, we will use a more general form of type of Fourier transform). Fourier sampling is easy on a quantum computer, but appears to be difficult to carry out on a classical

computer. In what follows, we will explore some of the power of Fourier sampling.

## 4.4 Phase State

We will now see how to set up an interesting state for fourier sampling. Given a classical circuit for computing a boolean function $f : \{0,1\}^n \to \{0,1\}$, this procedure due to Deutsch and Jozsa, shows how to transform it into a quantum circuit that produces the quantum state $|\phi\rangle = 1/2^{n/2} \sum_x (-1)^{f(x)} |x\rangle$.

The quantum algorithm to carry out this task uses two quantum registers, the first consisting of $n$ qubits, and the second consisting of a single qubit.

- Start with the registers in the state $|0^n\rangle |0\rangle$
- Compute the Fourier transform on the first register to get $\sum_{x\in\{0,1\}^n} |x\rangle \otimes |0\rangle$.
- Compute $f$ to get $\sum_x |x\rangle |f(x)\rangle$.
- Apply a conditional phase based on $f(x)$ to get $\sum_x (-1)^{f(x)} |x\rangle |f(x)\rangle$.
- Uncompute $f$ to get $\sum_x (-1)^{f(x)} |x\rangle \otimes |0\rangle$.

## 4.5 Extracting $n$ bits with 2 evaluations of Boolean Function

Suppose we are given a black box (or an obfuscated classical circuit) that computes the function function $f_s : \{0,1\}^n \to \{1,-1\}$, where $f(x) = s \cdot x$. $s \cdot x$ denotes the dot product $s_1 x_1 + \cdots + s_n x_n \bmod 2$. The challenge is to use this black box to efficiently determine $s$.

It is easy to see how to perform this task with $n$ queries to the black box: simply input in turn the $n$ inputs $x$ of Hamming weight 1. The outputs of the black box are the bits of $s$. Since each query reveals only one bit of information, it is clear that $n$ queries are necessary.

Remarkably there is a quantum algorithm (the base case of the Bernstein-Vazirani algorithm) that requires only two (quantum) queries to the black box:

- Use the black box to set up the phase state $|\phi\rangle = 1/2^{n/2} \sum_x (-1)^{f(x)} |x\rangle$.

- Apply the Fourier transform $H^{\otimes n}$ and measure. The outcome of the measurement is $s$.

To see that the outcome of the measurement is $s$, recall that $H^{\otimes n}|s\rangle = 1/2^{n/2}\sum_x(-1)^{s\cdot x}|x\rangle = |\phi\rangle$. Since $H^{\otimes n}$ is its own inverse, it follows that $H^{\otimes n}|\phi\rangle = |s\rangle$.

More generally, the transformation $H^{\otimes n}$ maps the standard basis $|s\rangle$ to the fourier basis $|\phi_s\rangle = 1/2^{n/2}\sum_x(-1)^{s\cdot x}|x\rangle$ and vice-versa.

We have shown that a quantum algorithm can be more efficient than any probabilistic algorithm in terms of the number of queries. One way to use this difference in the number of queries in order to demonstrate a gap between quantum and probabilistic algorithms is to make the queries very expensive. Then the quantum algorithm would be $n/2$ times faster than any probabilistic algorithm for the given task. But this does not help us in our goal, which is to show that quantum computers violate the extended Church-Turing thesis. The idea behind proving a superpolynomial gap (which we will outline below) is to make each query itself be the answer to a Fourier sampling problem. Now each query itself is much easier for the quantum algorithm than for any probabilistic algorithm. Carrying this out recursively for $\log n$ levels leads to the superpolynomial speedup for quantum algorithms.

## 4.6 Recursive Fourier Sampling

Our goal is to give a superpolynomial separation between quantum computation and classical probabilistic computation. The idea is to define a recursive version of the fourier sampling problem, where each query to the function (on an input of length n) is itself the answer to a recursive fourier sampling problem (on an input of length n/2). Intuitively a classical algorithm would need to solve $n$ subproblems to solve a problem on an input of length $n$ (since it must make $n$ queries). Thus its running time satisfies the recurrence $T(n) \geq nT(n/2) + O(n)$ which has solution $T(n) = \Omega(n^{\log n})$. The quantum algorithm needs to make only two queries and thus its running time satisfies the recurrence $T(n) = 2T(n/2) + O(n)$, which solves to $T(n) = O(n \log n)$.

Here is how it works for two levels: we are given a black box computing a function $f : \{0,1\}^{3n/2} \to \{0,1\}$, with the promise that for every n bit string $x$, the function $f_x : \{0,1\}^{n/2} \to \{0,1\}$ defined by $f_x(y) = f(xy)$ ($xy$ is denotes the concatenation of $x$ and $y$) satisfies $f_x(y) = s_x \cdot y$ for some $s_x \in \{0,1\}^{n/2}$. We are also given a black box $g : \{0,1\}^{n/2} \to \{0,1\}$ which satisfies the condition that if we construct a boolean function $h$ on $n$ bits as $h(x) = g(s_x)$, then $h(x) = s \cdot x$ for some $n$-bit string $s$. The challenge is to figure out $s$.

The proof that no classical probabilistic algorithm can reconstruct $s$ is somewhat technical, and establishes that for a random $g$ satisfying the promise, any algorithm (deterministic or probabilistic) that makes $n^{o(\log n)}$ queries to $g$ must give the wrong answer on at least $1/2 - o(1)$ fraction of $g$'s. This lemma continues to hold even if the actual queries are chosen by a helpful(but untrusted) genie who knows the answer.

For those with a background in computational complexity theory – this establishes that relative to an oracle $BQP \not\subseteq MA$. $MA$ is the probabilistic generalization of $NP$. It is conjectured that recursive fourier sampling does not lie in the polynomial hierarchy. In particular, it is an open question to show that, relative to an oracle, recusive fourier sampling does not lie $AM$ or in $BPP^{NP}$.

## 4.7 Simon's Algorithm

The Problem

Suppose we are given a black box for computing a 2-to-1 function $f : \mathbf{Z}_2^n \to \mathbf{Z}_2^n$ (from $n$-bit strings to $n$-bit strings), with the promise that there is a non-zero string $s \in \mathbf{Z}_2^n \setminus \{0\}$ such that

$$\text{for all } x \neq y,\ f(x) = f(y) \text{ if and only if } x \oplus y = s.$$

Here $\oplus$ is the bitwise direct sum modulo 2. For example, $\begin{array}{r} 1101 \\ \oplus 0111 \\ \hline 1010 \end{array}$ or $\begin{array}{r} 01 \\ \oplus 01 \\ \hline 00 \end{array}$.

A quick example of such a function with $n = 3$ is

$$f(x) = \begin{cases} 001 & \text{if } x = 000 \text{ or } 011 \\ 010 & \text{if } x = 001 \text{ or } 010 \\ 100 & \text{if } x = 111 \text{ or } 100 \\ 111 & \text{if } x = 110 \text{ or } 101 \end{cases}$$

where $s = 011$.

The *problem* of Simon's algorithm is to determine $s$.

Classically

A simple way to solve this problem classically would be to randomly input values to the black box until we find two inputs that produce the same output, and compute their direct sum. But there are $2^{n-1}$ possible outputs, so despite the help from the birthday paradox, we expect it to take $\sqrt{2^{n-1}} = 2^{(n-1)/2}$ attempts to find $s$: still exponential time.

Furthermore, it can be shown that no classical computer can find $s$ faster than exponential time. We will use the power of quantum computing to find a faster way.

Quantum

To utilize the power of quantum computing, we will access the function in superposition. So suppose instead of a black box we are given the circuit $C_f$ for computing $|f\rangle$, from which we can construct the unitary transformation $U_f$:

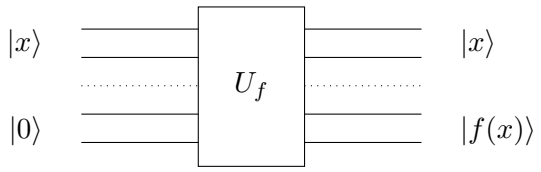



Figure 4.2: Black Box Circuit

The point here is that the input can be a superposition over all $n$-bit strings $\sum_{x=0}^{N-1} \alpha_x |x\rangle$ $(N = 2^n)$, yielding the output $\sum_{x=0}^{N-1} \alpha_x |x\rangle |f(x)\rangle$. This can be thought of as querying $f$ in superposition.

Simon's Algorithm consists of 3 main steps.

**Step 1**: Prepare the random superposition $\frac{1}{\sqrt{2}}(|x_0\rangle + |x_0 \oplus s\rangle)$

**Step 2**: Use Fourier sampling to produce a $y$ such that $y \cdot s = 0$

**Step 3**: Repeat until there are enough such $y$'s that we can classically solve for $s$.

Now lets see the details on how to do each step.

**Step 1**: Prepare the random superposition $\frac{1}{\sqrt{2}}(|x_0\rangle + |x_0 \oplus s\rangle)$

First query the function with a uniform superposition of the n-bit strings. To prepare this uniform superposition, start with the state $|0\rangle$ then apply the Hadamard transform. With $N = 2^n$, this is written:

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

Next we will use the unitary transformation $U_f$ to query $f$ in uniform superposition.

$$\sqrt{\frac{1}{N}} \sum_{x=0}^{N-1} |x\rangle |0\rangle \xrightarrow{U_f} \sqrt{\frac{1}{N}} \sum_{x=0}^{N-1} |x\rangle |f(x)\rangle$$

Now what happens if we measure the second register containing $|f(x)\rangle$? It must collapse into $|f(x_0)\rangle$ for some $x_0 \in \mathbf{Z}_2^n$. But this reveals information

about the first register, and it will also collapse into the pre-images of $f(x_0)$: $x_0$ and $x_0 \oplus s$.

$$\sqrt{\frac{1}{N}} \sum_{x=0}^{N-1} |x\rangle |f(x)\rangle \stackrel{\text{measure } f}{\longrightarrow} \frac{1}{\sqrt{2}}(|x_0\rangle + |x_0 \oplus s\rangle) |f(x_0)\rangle$$

The first register is now the state $\frac{1}{\sqrt{2}}(|x_0\rangle + |x_0 \oplus s\rangle)$ where $x_0$ is a random n-bit string. The challenge is to read off $s$ from this superposition. We cannot simply measure the state because the superposition will be destroyed, and the result we get will have no information about $s$.

**Step 2**: Use Fourier sampling to find a $y$ such that $y \cdot s = 0$.

We now show that $H^{\otimes 2}(\frac{1}{\sqrt{2}} |x_0\rangle + \frac{1}{\sqrt{2}} |x_0 \oplus s\rangle$ is a uniform superposition over all states $|y\rangle$ such that $y \cdot s = 0$. This means that Fourier sampling $(\frac{1}{\sqrt{2}} |x_0\rangle + \frac{1}{\sqrt{2}} |x_0 \oplus s\rangle$ results in a uniform superposition of $y$ such that $y \cdot s = 0$. Recall that

$$H^{\otimes n} |x\rangle = \sqrt{\frac{1}{N}} \sum_y \alpha_y |y\rangle \text{ where } \alpha_y = (-1)^{x \cdot y}$$

So $H^{\otimes 2} \frac{1}{\sqrt{2}}(|x_0\rangle + |x_0 \oplus s\rangle) = \frac{1}{2} \sum_y \alpha_y |y\rangle$ where

$$\begin{aligned} \alpha_y &= \frac{1}{\sqrt{2}}(-1)^{x_0 \cdot y} + (-1)^{(x_0 \oplus s) \cdot y} \\ &= \frac{1}{\sqrt{2}}(-1)^{x_0 \cdot y}(1 + (-1)^{s \cdot y}) \end{aligned}$$

Now it is easy to see that if $s \cdot y = 0$, $\alpha_y = \pm \frac{1}{\sqrt{2}}$, but if $s \cdot y = 1$, $\alpha_y = 0$.

Therefore when we measure the first register, we will measure a $y$ such that $y \cdot s = 0$.

**Step 3**: Repeat until there are enough such $y$'s that we can classically solve for $s$.

There are exactly $n$ linearly independent values of $y$ such that $y \cdot s = y_1 s_1 + y_2 s_2 + \cdots y_n s_n = 0$, and one of these is the trivial solution $y = 0$. Therefore, there are $n-1$ non-trivial, linearly independent solutions to $y \cdot s = 0$. But if $y_1$ and $y_2$ are linearly independent solutions, $(y_1 + y_2) \cdot s = y_1 \cdot s + y_2 \cdot s = 0$ so linear combinations of solutions are also solutions. This gives us a total of $2^{n-1}$ $y$'s such that $y \cdot s = 0$. To solve for $s$, we need to find exactly $n - 1$ non-trivial, linearly independent $y$ such that $y \cdot s = 0$.

For example, if $s = 010$, then $y_0 = 000$, $y_1 = 001$, and $y_2 = 100$ are linearly independent solutions to $y \cdot s = 0$. But the linear combination $y_1 + y_2 = 101$ is

also a solution. We need only find two of $\{y_1, y_2, y_1+y_2\}$ in order to classically solve for s.

How long should we expect this to take? The probability that we fail on the first run is the probability that we find $y = 0$, which is one value out of $2^{n-1}$. So $P_1 = 1/2^{n-1}$, where $P_1$ denotes the probability of failing on the first run. Lets call the first nontrivial solution $y_1$.

We fail when looking for $y_2$ if we find 0 or $y_1$, so $P_2 = 2/2^{n-1} = 1/2^{n-2}$. When looking for $y_3$, we fail if we find any of $\{0, y_1, y_2, y_1 + y_2\}$, so $P_3 = 4/2^{n-1} = 1/2^{n-3}$. Carrying on in this way, the probability of failing to find $y_i$ is $P_i = 1/2^{n-i}$.

The chance that we fail *up to and including* $y_i$ can be approximated by $P < 1/2^{n-1} + 1/2^{n-2} + \cdots + 1/2^{n-i}$. If we push this approximation all the way to $i = n - 1$, we see that we fail with probability less than 1 (compute the geometric sum). That's not a strong enough approximation, so instead notice that our probability of failure up to and including $i = n - 2$ is less than 1/2. Then our probability of success up to the $n - 2$st run is greater than 1/2. We find the final linearly independent term on the last run with probability 1/2 (if you don't believe this, notice that half of the solutions are linear combinations that include $y_{n-1}$). Finally our total probability of success is $P(success) > 1/2 * 1/2 = 1/4$. Therefore, we expect our process to take $O(n)$ steps (our limit says 4 by $n$ runs of the algorithm should be enough for success).

Simon's algorithm is summed up by the following circuit.

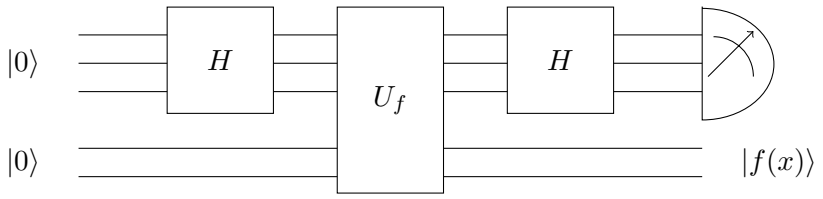



Figure 4.3: Circuit for Simon's Algorithm

In summary, the above circuit for Simon's algorithm corresponds to the

following sequence of transformations.

$$\begin{aligned}
|0\rangle\,|0\rangle &\xrightarrow{H^{\otimes n}} \frac{1}{\sqrt{2^n}}\sum_x |x\rangle\,|0\rangle \\
&\xrightarrow{U_f} \frac{1}{\sqrt{2^n}}\sum_x |x\rangle\,|f(x)\rangle \\
&\xrightarrow{\text{measure}} \frac{1}{\sqrt{2}}(|x_0\rangle + |x_0 \oplus s\rangle) \otimes |f(x_0)\rangle \\
&\xrightarrow{H^{\otimes n}} \frac{1}{\sqrt{2^n}}\sum_y \alpha_y\,|y\rangle\,|f(x_0)\rangle
\end{aligned}$$

for some numbers $\alpha_y$.

As above, for each $y$, if $s \cdot y = 1$, then $\alpha_y = 0$, whereas if $s \cdot y = 0$, then $\alpha_y = (-1)^{x_0 \cdot y}\sqrt{2}$.

When we observe the first register, we get a uniformly random $y$ such that $s \cdot y = s_1 y_1 + \cdots + s_n y_n = 0$. We repeat to collect more and more equations, and recover $s$ from $n-1$ linearly independent equations.

Example

Let $n = 2$ and $f(x) = \begin{cases} 00 \text{ if } x = 00 \text{ or } 10 \\ 01 \text{ if } x = 01 \text{ or } 11 \end{cases}$ so that $s = 10$.

First, apply the Hadamard transform to prepare $|x\rangle$, then $U_f$ to prepare $|f(x)\rangle$

$$|0\rangle\,|0\rangle \xrightarrow{H^{\otimes 2}} \frac{1}{2}\sum_{x=0}^{3} |x\rangle\,|0\rangle \xrightarrow{U_f} \frac{1}{2}\sum_{x=0}^{3} |x\rangle\,|f(x)\rangle$$

Then measure the second register to finalize the first step of the process. For the purpose of argument, lets suppose we measure $f(x) = 01$, so that $x_0 = 01$ and $x_0 \oplus s = 11$:

$$\frac{1}{2}\sum_{x=0}^{3} |x\rangle\,|f(x)\rangle \xrightarrow{\text{measure}} \frac{1}{\sqrt{2}}(|01\rangle + |11\rangle) \otimes |01\rangle$$

We then impose the Hadamard transform to achieve:

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

We expect it will take 2 runs through the above process to measure $y = 01$. Then the only nonzero solution for $s$ is $s = 10$.

# Chapter 5

# Quantum Fourier Transform

Quantum Fourier Transform is a *quantum* implementation of the discrete Fourier transform. You might be familiar with the discrete Fourier Transform or Fourier Analysis from the context of signal processing, linear algebra, or one of its many other applications. In short, Fourier Analysis is a tool to describe the internal frequencies of a function.

Here we will present a quantum algorithm for computing the discrete Fourier transform which is exponentially faster than the famous Fast Fourier Transform of classical computers. However, this algorithm provides an excellent example of the tension between exponentially faster quantum algorithms and the problems of measurement. While we can carry out the QFT algorithm to transform the $n$ qubit state vector $|\alpha\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle + \cdots + \alpha_n |n\rangle$ to its Fourier transform $|\beta\rangle = \beta_0 |0\rangle + \beta_1 |1\rangle + \cdots + \beta_n |n\rangle$, a measurement on $|\beta\rangle$ will only return one of its $n$ components, and we are not able to recover all the information of the Fourier transform. For this reason, we describe this algorithm as quantum Fourier *sampling*.

The Quantum Fourier Transform is a generalization of the Hadamard transform. It is very similar, with the exception that QFT introduces phase. The specific kinds of phases introduced are what we call primitive roots of unity, $\omega$. Before defining the Fourier Transform, we will take a quick look at these primitive roots of unity.

Recall that in the complex plane, there exist $n$ solutions to the equation $z^n = 1$. For example if $n = 2$, $z$ could be 1 or -1. If $n = 4$, $z$ could be $1, i, -1$, or $-i$. You can easily check that these roots can be written as powers of $\omega = e^{2\pi i/n}$. This number $\omega$ is called the primitive $n$th root of unity. In Figure 1, $\omega$ is drawn along with the other complex roots of unity for n=5.

In this figure, we see that $\omega$ lies on the unit circle so $\mid \omega \mid = 1$, and the line from the origin to $\omega$ makes the angle $\phi = 2\pi/M$ with the real line. If we square $\omega$, we double the angle. Furthermore, if we raise $\omega$ to the $j$th power, $\omega^j$ has phase angle $\phi = 2j\pi/M$ and is still an $M$th root of unity.

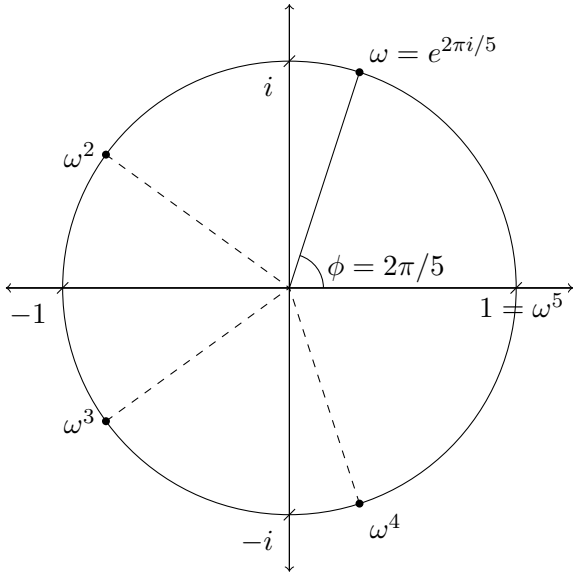



Figure 5.1: The 5 complex 5th roots of 1.

Now we can move in to the Fourier Transform itself. The discrete Fourier transform is defined by

$$QFT_M = \frac{1}{\sqrt{M}} \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \cdots & \omega^{M-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \cdots & \omega^{2M-2} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \cdots & \omega^{3M-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{M-1} & \omega^{2M-2} & \omega^{3M-3} & \cdots & \omega^{(M-1)(M-1)} \end{pmatrix}$$

Another way of writing this is to say that the $jk$th entry of $QFT_M$ is $\omega^{jk}$. The transform takes the vector $\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{n-1} \end{pmatrix}$ to its Fourier transform $\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{n-1} \end{pmatrix}$ as specified by the above matrix.

## Examples

### Ex. 1

Lets take a look at $QFT_2$. Because $M = 2$, $\omega = e^{\pi i} = -1$ Therefore we have

$$QFT_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & \omega \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

As you can see, $QFT_2$ is simply equal to $H^{\otimes 2}$.

How about $QFT_4$? The primitive 4th root of unity is $i$, so that

$$QFT_4 = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}$$

**Ex. 2**

Find the quantum Fourier transform for $M = 4$ of the functions $|f\rangle = \frac{1}{2}(|0\rangle+|1\rangle+|2\rangle+|3\rangle) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$;

$|g\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, and $|h\rangle = |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

The corresponding Fourier transforms are given by:

1. $|f\rangle$:
$$\left|\hat{f}\right\rangle = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$
2. $|g\rangle$:
$$|\hat{g}\rangle = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$
3. $|h\rangle$:
$$\left|\hat{h}\right\rangle = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix}$$

Lets do a bit of analysis of these examples. In example 1, you might notice that the columns of $QFT_4$ are orthogonal. For example, the inner product of the first column with the second column is $\frac{1}{2}[(1*1)+(1*-i)+(1*-1)+(1*i)] = \frac{1}{2}(1-i-1+i) = 0$. You should also notice that, by design, the columns of $QFT_4$ have magnitude 1. Thus $QFT_4$ is unitary.

In example 2 you should notice is that the vectors like $|f\rangle$ that had a lot of zeros (large spread) had Fourier transforms with few zeros (narrow spread), and vice-versa.

Finally, in example 2 notice how the only difference between the Fourier transforms of $|g\rangle$ and $|h\rangle$ is a difference of relative phase shifts.

We would like to be able to make some statements to solidify these ideas about Fourier transforms, so lets prove them.

## Properties of QFT

Studying the properties of a mathematical object gives us insight into the way it works. These properties will not only be important to our use of the Fourier transform later on, but they also provide a foundation of how to understand the discrete Fourier transform.

1. $QFT_M$ is unitary.

   It is well known that an operator is unitary if its columns are orthonormal. Denote the $i$th and $j$th columns of $QFT_M$ as $F_i$ and $F_j$. Then $F_i = \frac{1}{\sqrt{M}} \begin{pmatrix} 1 \\ \omega^{i*1} \\ \vdots \\ \omega^{i*(M-1)} \end{pmatrix}$ and $F_j = \frac{1}{\sqrt{M}} \begin{pmatrix} 1 \\ \omega^{i*j} \\ \vdots \\ \omega^{j*(M-1)} \end{pmatrix}$.

   Thus

   $$\langle F_i | F_j \rangle = \frac{1}{M} \sum_{n=0}^{M-1} \omega^{ni} \overline{\omega^{nj}} = \frac{1}{M} \sum_{n=0}^{M-1} (\omega^{i-j})^n$$

   From here it is easy to see that if $i = j$, $\langle F_i | F_j \rangle = 1$.

   For the case $i \neq j$, we will notice that $\frac{1}{M} \sum_{n=0}^{M-1} (\omega^{i-j})^n$ is a geometric series, and expand the sum. Thus

   $$\frac{1}{M} \sum_{n=0}^{M-1} (\omega^{i-j})^n = \frac{1}{M} \frac{\omega^{M(i-j)} - 1}{\omega^{i-j} - 1} = \frac{1}{M} \frac{1 - 1}{\omega^{i-j} - 1} = 0$$

   where $\omega^{M(i-j)} = 1$ because $\omega$ is an $M$th root of unity.

   Because the Fourier transform is a unitary operator, we can implement it in a quantum circuit. Thus if $N = 2^n$, we can apply the Fourier transform $QFT_N$ to a n-qubit system.

2. *Linear Shift*

   This, property noted in the above examples, states that linear shifts of state-vectors cause relative phase shifts of their Fourier transform. This is expressed mathematically by saying if $|f(x)\rangle$, $x \in \mathbf{Z}_M$, has Fourier transform $\left|\hat{f}(x)\right\rangle$, then $|f(x+j)\rangle$ has Fourier transform $\left|\hat{f}(x)\right\rangle \phi_j$, where $\phi_j$ is a phase shift and $\phi_j = e^{\frac{2\pi}{M} xj}$. Furthermore, because $QFT_M$ is unitary and $QFT_M QFT_M^\dagger = \mathbb{I}$, the converse is true. A linear phase shift on $|f\rangle$ produces a linear shift in $\left|\hat{f}\right\rangle$.

So if $QFT_N \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{N-1} \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{N-1} \end{pmatrix}$, then $QFT_N \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_0 \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \omega\beta_1 \\ \vdots \\ \omega^{N-1}\beta_{N-1} \end{pmatrix}$ and $QFT_N \begin{pmatrix} \alpha_0 \\ \omega\alpha_1 \\ \vdots \\ \omega^{N-1}\alpha_{N-1} \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_0 \end{pmatrix}$.

If you have never seen this property before, it should be shocking. We will not offer a proof of this in general here, but below is an example with $N = 4$.

**Example**:

Let $|\Theta\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$ and $|\Phi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_0 \end{pmatrix}$. Then

$$\left|\hat{\Theta}\right\rangle = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 \\ \alpha_0 + i\alpha_1 - \alpha_2 - i\alpha_3 \\ \alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 \\ \alpha_0 - i\alpha_1 - \alpha_2 + i\alpha_3 \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

$$\left|\hat{\Phi}\right\rangle = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 \\ -i\alpha_0 + \alpha_1 + i\alpha_2 - \alpha_3 \\ -\alpha_0 + \alpha_1 - \alpha_2 + \alpha_3 \\ i\alpha_0 + \alpha_1 - i\alpha_2 - \alpha_3 \end{pmatrix} = \begin{pmatrix} \beta_0 \\ -i\beta_1 \\ -\beta_2 \\ i\beta_3 \end{pmatrix}$$

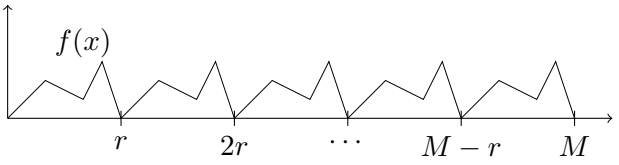
The important point here is that the only difference between $\left|\hat{\Theta}\right\rangle$ and $\left|\hat{\Phi}\right\rangle$ is a relative phase shift. But does this matter?

If we are going to measure a state then the phases don't matter at all, because if the phase is $\phi$, then $\langle\phi|\phi\rangle = 1$. Therefore the phase of a given state does not effect the probability of measuring that state. However, there is a way we can gather information about the pahses.

We won't be able to tell by measuruing the difference between $\frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ and $\frac{1}{2}\begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix}$ by making a measurment. However, if we apply QFT, we see that $QFT_4\frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ and $QFT_4\frac{1}{2}\begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$. Thus, **measuring the Fourier Transform of the states will reveal the relative phases**.

3. *Period/Wavelength Relationship*

Suppose $f$ is periodic with period $r$, for example




Then $\hat{f}$ (the Fourier transform of $f$) is supported only on multiples of $M/r$. Thus, $\hat{f}(x) = 0$ unless $x = kM/r$.

If $r$ is the period of $f$, we can think of $M/r$ as the wavelength of $f$. If you already have intuition for Fourier transform this should come as no surprise. In general, the wider the range of a function, the sharper the range in the Fourier domain; and vise versa. In example, the fourier transform of a delta function is an even spread, while the transform of an even spread is a delta function.

We will prove this in a special case, where $f(j) = \begin{cases} \sqrt{\frac{r}{M}} & \text{if } j = 0 \pmod r \\ 0 & \text{otherwise.} \end{cases}$ While this is a very special case, it is actually the only case that we will need to develop Shor's algorithm. Furthermore this property *can* be proved in general.

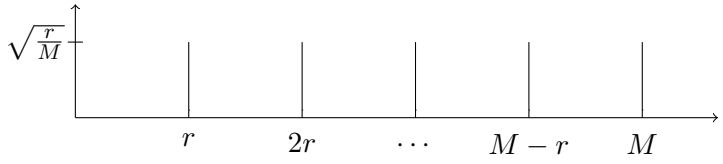



Figure 5.2: $f(j)$ in special case proof.

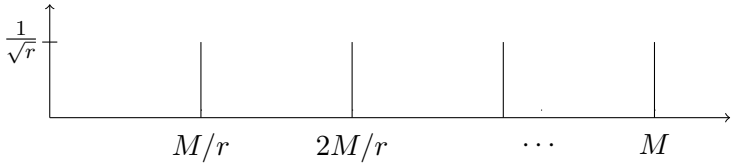



Figure 5.3: $\hat{f}(j)$

Because this function is relatively simple, we can prove the desired relationship by brute force. Suppose $|\alpha\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix}$ has Fourier transform $|\beta\rangle = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_N \end{pmatrix}$, then the $j$th component of its Fourier transform is given by

$$\beta_j = \frac{1}{\sqrt{M}} \sum_{i=0}^{N-1} \alpha_i \omega^{ij} \tag{5.1}$$

This expression comes from matrix multiplication, you should take a look at the definition of $F_M$ above to verify this if it looks unfamiliar. In our special case, $f(j) = \begin{cases} \sqrt{\frac{r}{M}} & \text{if } j = 0 \pmod r \\ 0 & \text{otherwise.} \end{cases}$. Using (**??**) we can calculate its Fourier transform

$$\hat{f}(j) = \frac{\sqrt{r}}{M} \sum_{i=0}^{\frac{M}{r}-1} \omega^{rij} \tag{5.2}$$

We have seen sums like this before. It is just a geometric series, and it is not too difficult to compute.

$$\sum_{i=0}^{\frac{M}{r}-1} \omega^{rij} = \frac{\omega^{Mj} - 1}{\omega^{rj} - 1}$$

But $\omega^{Mj} = 1$ (recall $\omega^M = 1$), so the numerator is always equal to 0. The only time the denominator can equal zero is when $rj = kM$ for some integer $k$. In this case, the numerator and the denominator are equivalently 0, so we must compute the limit using l'Hospitals rule.

$$\begin{aligned} \lim_{j \to k\frac{M}{r}} \frac{\omega^{Mj} - 1}{\omega^{rj} - 1} &= \lim_{j \to k\frac{M}{r}} \frac{M}{r} \frac{\omega^{Mj-1}}{\omega^{rj-1}} \\ &= \lim_{j \to k\frac{M}{r}} \frac{M}{r} \frac{\omega^{Mj}}{\omega^{rj}} \frac{\omega^{-1}}{\omega^{-1}} \\ &= \frac{M}{r} \end{aligned}$$

When we plug this result back into (**??**), the outcome is the desired result.

$$\hat{f}(j) = \begin{cases} \frac{1}{\sqrt{r}} & \text{if } j = 0 \pmod{\frac{M}{r}} \\ 0 & \text{otherwise.} \end{cases}$$

Notice that the normalization factor makes good sense.

To get a look at how we could prove this property in general, imagine periodic functions (in $r$) of this type as a basis for any periodic function. Allow the possibility of relative phase shifts, and you can prove this property in general.

### Classical Fast Fourier Transform

The FFT was a major breakthrough for classical computers. Because the Fourier transform is an $M * M$ matrix, straightforward multiplication by $F_M$ would take $O(M^2)$ steps to carry out, because multiplication of $f$ on each row takes $M$ multiplications. The FFT reduced this to $O(M \log M)$ steps. The FFT is incredibly important in signal processing that essentially all of your electronics rely on it. Without the FFT, modern electronics would have far fewer capabilities and would be much slower than they are today.

The FFT requires only that $M = 2^m$ for some integer $m$, but this is a relatively easy requirement because the computer can simply choose their domain.

The *fast* Fourier transform uses the symmetry of the Fourier transform to reduce the computation time. Simply put, we rewrite the Fourier transform of size $M$ as two Fourier transforms of size $M/2$ - the odd and the even terms. We then repeat this over and over again to exponentially reduce the time. To see how this works in detail, we turn to the matrix of the Fourier transform. While we go through this, it might be helpful to have $QFT_8$ in front of you to take a look at. Note that the exponents have been written modulo 8, since $\omega^8 = 1$. Take a look at some of the symmetries and patterns.

$$\begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & \omega & \omega^2 & \omega^3 & \omega^4 & \omega^5 & \omega^6 & \omega^7 \\
1 & \omega^2 & \omega^4 & \omega^6 & 1 & \omega^2 & \omega^4 & \omega^6 \\
1 & \omega^3 & \omega^6 & \omega & \omega^4 & \omega^7 & \omega^2 & \omega^5 \\
1 & \omega^4 & 1 & \omega^4 & 1 & \omega^4 & 1 & \omega^4 \\
1 & \omega^5 & \omega^2 & \omega^7 & \omega^4 & \omega & \omega^6 & \omega^3 \\
1 & \omega^6 & \omega^4 & \omega^2 & 1 & \omega^6 & \omega^4 & \omega^2 \\
1 & \omega^7 & \omega^6 & \omega^5 & \omega^4 & \omega^3 & \omega^2 & \omega^1
\end{pmatrix}$$

Notice how row $j$ is very similar to row $j + 4$. Also, notice how column $j$ is very similar to column $j + 4$. Motivated by this, we are going to split the Fourier transform up into its even and odd columns.

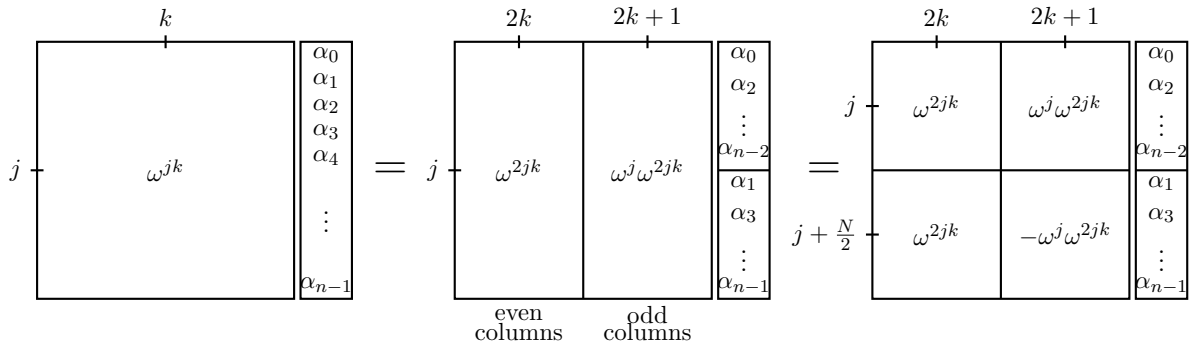



Figure 5.4: Breaking the Fourier transform up by symmetries.

In the first frame, we have represented the whole Fourier transform matrix by describing the $j$th row and $k$th column: $\omega^{jk}$. In the next frame, we separate the odd and even columns, and similarly separate the vector that is to be transformed. You should convince yourself that the first equality really is an equality either by carrying out the matrix multiplication, or by noticing that all we did was re-organize the basis vectors. In the third frame, we add a little symmetry by noticing that $\omega^{j+N/2} = -\omega^j$ (since $\omega^{N/2} = -1$).

Notice that both the odd side and even side contain the term $\omega^{2jk}$. But if $\omega$ is the primitive $N$th root of unity, then $\omega^2$ is the primitive $N/2$nd root of unity. Therefore, the matrices whose $j,k$th entry is $\omega^{2jk}$ are really just $QFT_{N/2}$! Now we can write $QFT_N$ as in Figure 5.

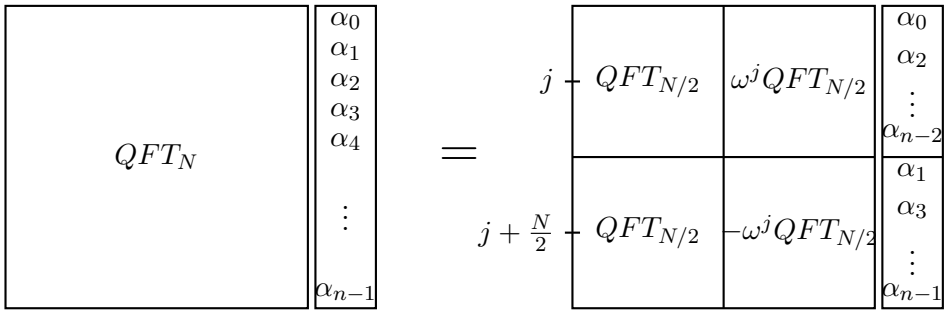



Figure 5.5: Simplified Fourier transform

Now suppose we are calculating the Fourier transform of the function $f(x)$. We can write the above manipulations as an equation that computes the $j$th term $\hat{f}(j)$.

$$\hat{f}(j) = \left(F_{M/2}\overrightarrow{f_{\text{even}}}\right)(j) + \omega^j \left(F_{M/2}\overrightarrow{f_{\text{odd}}}\right)(j)$$

This turns our calculation of $QFT_N$ into two applications of $QFT_{N/2}$. We can turn this into four applications of $QFT_{N/4}$, and so forth. As long as $N = 2^n$ for some $n$, we can break down our calculation of $QFT_N$ into $N$ calculations of $QFT_1 = 1$. This greatly simplifies our calculation.

### Quantum Fourier Transform w/ quantum gates

The strength of the the FFT is that we are able to use the symmetry of the discrete Fourier transform to our advantage. The circuit application of QFT uses the same principle, but because of the power of superposition QFT is even faster.

The QFT is motivated by the FFT so we will follow the same steps, but because this is a quantum algorithm the implementation of the steps will be different. That is, we first take the Fourier transform of the odd and even parts, then multiply the odd terms by the phase $\omega^j$.

The first step is to separate the odd and even terms. But in a quantum algorithm, the odd and even terms are already together in superposition: the odd terms are those whose least significant bit is 1, and the even with 0. Therefore, we can apply $QFT_{M/2}$ to both the odd and even terms

together. We do this by applying $QFT_{M/2}$ to the m-1 most significant bits, then recombining the odd and even appropriately by applying the Hadamard to the least significant bit.

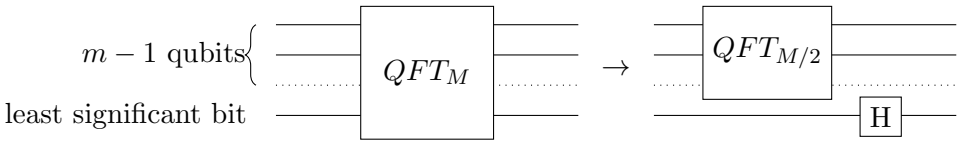



Figure 5.6: $QFT_{M/2}$ and a Hadamard gate correspond to $FFT_M/2$ on the odd and even terms

Now to carry out the phase multiplication, we need to multiply each odd term $j$ by the phase $\omega^j$. But remember, an odd number in binary ends with a 1 while an even ends with a 0. Thus we can use the *controlled phase shift*, where the least significant bit is the control, to multiply only the odd terms by the phase without doing anything to the even terms. Recall that the controlled phase shift is similar to the CNOT gate in that it only applies a phase to the target if the control bit is one.

The phase associated with each controlled phase shift should be equal to $\omega^j$ where $j$ is associated to the $k$th bit by $j = 2^k$.

Thus, apply the controlled phase shift to each of the first $m-1$ qubits, with the least significant bit as the control. With the controlled phase shift and the Hadamard transform, $QFT_M$ has been reduced to $QFT_M/2$.

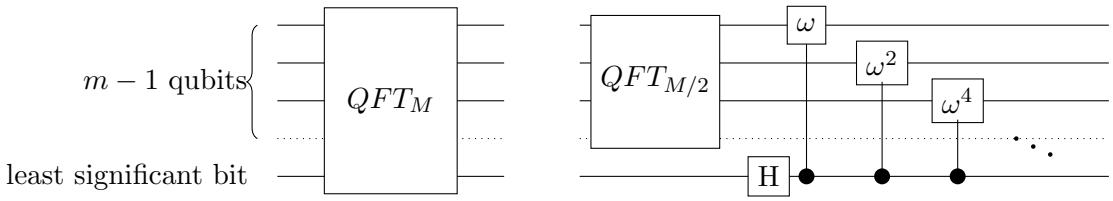



Figure 5.7: $QFT_M$ is reduced to $QFT_{M/2}$ and M additional gates

### Example

Lets construct $QFT_3$. Following the algorithm, we will trun $QFT_3$ into $QFT_2$ and a few quantum gates. Then continuing on this way we turn $QFT_2$ into $QFT_1$ (which is just a Hadamard gate) and another few gates. Controlled phase gates will be represented by $R_\phi$.

then run through another iteration to get rid of $QFT_2$

You should now be able to visualize the circuit for $QFT$ on more qubits easily. Furthermore, you can see that the number of gates necessary to carry out $QFT_M$ it takes exactly $\sum_{i=1}^{\log M} i = \log M(\log M + 1)/2 = O(\log^2 M)$.

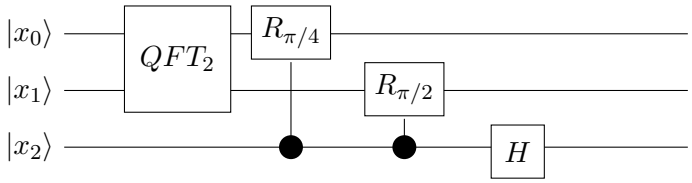



Figure 5.8: First Iteration

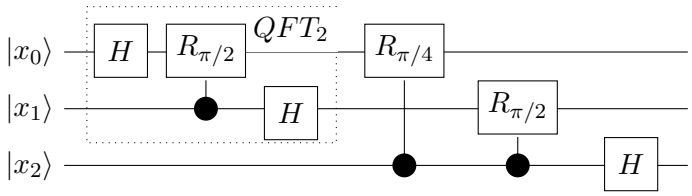



Figure 5.9: Second Iteration. Recall that $H = QFT_1$

# Chapter 6

# Period Finding and Factoring

Let's say we are given a black box for computing a periodic function, i.e. a function where

$$f(x) = f(y) \text{ if and only if } x \equiv y \pmod{r}$$

The goal of a period finding algorithm is to find $r$.

The algorithm for period finding is very similar to Simon's algorithm, in fact we can think of it as a generalization of Simon's algorithm. The steps we follow are very similar.

*Classically*, we could solve this problem by querying our function with subsequent inputs until the function repeats. This takes $O(r) = O(2^n)$ queries to the function. There are other ways to solve this problem, but it can be shown that all classical algorithms solve this problem in exponential time.

With a *Quantum* computer, we can access the function in *superposition* to query the function with $N = 2^n$ inputs for each $n$ qubits at the same time. The key ingredients to our approach will be the period/wavelength and linear shift properties of the Fourier transform. We first access the function in superposition to create a periodic superposition that may not start from 0 (it includes alinear shift), and then take its Fourier transform to get rid of the linear shift.

This approach is very similar to the approach of Simon's algorithm, and is the historical motivation for the period finding algorithm. Lets examine the details.

**Step 1**: Prepare the periodic superposition $\sqrt{\frac{r}{N}} \sum_{j=0}^{N/r-1} |x_0 + jr\rangle$

**Step 2**: Fourier *sample* to measure some $y = \frac{kN}{r}$ for $k \in \{0, 1, \cdots, r-1\}$.

**Step 3**: Repeat until there are enough such $y$'s so that we can compute their greatest common divisor and solve for $r$.

**Step 1**. It is best to start a quantum algorithm with the easily prepared state $|0\rangle$, but we want to access our function in superposition: we need the state $\frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle |0\rangle$. To prepare this state, we

just implement $QFT_N$ on the first $n$[1] qubits:

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

Next, as in Simon's algorithm, we access our function in superposition. Let $U_f$ be the unitary transformation that carries out our function, and implement it:

$$\frac{1}{\sqrt{N}}\sum_{x=0}^{N-1}|x\rangle|0\rangle \xrightarrow{U_f} \frac{1}{\sqrt{N}}\sum_{x=0}^{N-1}|x\rangle|f(x)\rangle$$

To get a periodic superposition out of this, we measure $|f\rangle$. Then $|f\rangle$ must collapse into some value $f(x_0)$. Furthermore, because measuring $|f\rangle$ reveals information about $|x\rangle$, the state $|x\rangle$ will also collapse into the pre-image of $f(x_0)$. But because $f$ is periodic, the pre-image of $f(x_0)$ is $\{x_0,\ x_0+r,\ x_0+2r,\ \cdots,\ x_0+(\frac{N}{r}-1)r\}$.

$$\frac{1}{\sqrt{N}}\sum_{x=0}^{N-1}|x\rangle\otimes|f(x)\rangle \xrightarrow{\text{measure}|f\rangle} \sqrt{\frac{r}{N}}\sum_{i=0}^{N/r-1}|x_0+ir\rangle|f(x_0)\rangle$$

Now our first register is in a periodic superposition, where the period is the same as the period of the function! But we can't just measure, because each time we run the algorithm, we might measure a different value of $|f\rangle$, thus obtaining a periodic superposition that is linearly shifted by some other $x_0$

**Step 2**: We can't just measure our superposition right away, because that would destroy the superposition. And because of the random linear shift $x_0$, a measurement wouldn't reveal any useful information. Instead, we will rely on the properties of the Fourier transform to retrieve the information we want. Remember that if $f$ is periodic with period $r$, then $\hat{f}$ is periodic with period $N/r$. Furthermore, remember that we only see the effect of the linear shift $x_0$ in the phase of $\hat{f}$. Therefore if we take the Fourier transform of the first register, we will be left only with states that are multiples of $N/r$.

$$\sqrt{\frac{r}{N}}\sum_{i=0}^{N/r-1}|ir+x_0\rangle \xrightarrow{QFT_N} \frac{1}{\sqrt{r}}\sum_{i=0}^{r-1}\left|i\frac{N}{r}\right\rangle\phi_i$$

where $\phi_i$ is the (unimportant) phase associated with each term due to the linear shift $x_0$.

Now we can measure and retrieve $k\frac{N}{r}$ for some integer $k$!

**Step 3** Now we repeat the algorithm to retrieve several distinct multiples of $N/r$. Once we have enough values, we can compute their GCD to retrieve $N/r$. $N$ is a given in the problem, so it is easy to compute $r$. Computing $GCD$ is easy thanks to Euclid's algorithm.

[1] Don't let the different n's confuse you. If there are $n$ qubits, then we need $N = 2^n$ complex numbers to describe the system.

How long should we expect this to take? Let us compute the chance of finding the correct period after $t$ samples. Suppose after finding $t$ distinct multiples of $N/r$, we have not found the desired period $N/r$, but instead an integer multiple of it, say $\lambda N/r$. This means that each of the $t$ samples must be a multiple of $\lambda N/r$. There are exactly $N/(\lambda N/r) = r/\lambda$ such multiples of $\lambda N/r$. And since there are $r$ multiples in total, the probability of measuring a multiple of $\lambda N/r$ is $1/\lambda$. Therefore,

$$\Pr[\text{gcd is a multiple of } N/r] = \left(\frac{1}{\lambda}\right)^t \leq \left(\frac{1}{2}\right)^t,$$

and we err with probability

$$\Pr[\text{gcd} > N/r \text{ after } t \text{ samples}] \leq N\left(\frac{1}{2}\right)^t.$$

Therefore we must repeat the period finiding circuit $O(\log N)$ times to be confident in our solution.

The above algorithm can be summed up by Figure 1:

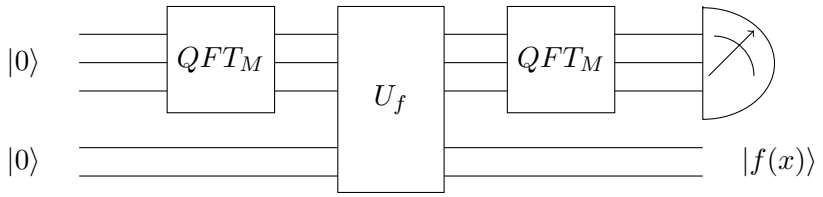



Figure 6.1: Circuit for period finding

### Example

In this example we find the period of the function $f(x) = x \pmod 2$. It is easy to see that the period of this function is $r = 2$

We will use a 3-qubit system so that $N = 8$. It is a good rule of thumb to choose $N \gg r$. The first step is to apply the quantum Fourier transform:

$$|0\rangle|0\rangle \xrightarrow{QFT_8} \frac{1}{\sqrt{8}}\sum_{x=0}^{7}|x\rangle|0\rangle$$

Next we apply our function.

$$\frac{1}{\sqrt{8}}\sum_{x=0}^{7}|x\rangle|0\rangle \xrightarrow{U_f} \frac{1}{\sqrt{8}}\sum_{x=0}^{7}|x\rangle|x \bmod 2\rangle$$

The next step is to measure $|f\rangle$. Then $|f\rangle$ must collapse into either $|0\rangle$ or $|1\rangle$. For the purpose of demonstration, lets say our measurement returns $|f(x)\rangle = |1\rangle$. Then x must be odd.

$$\frac{1}{\sqrt{8}}\sum_{x=0}^{7}|x\rangle \otimes |f(x)\rangle \xrightarrow{measure|f\rangle} \frac{1}{2}(|1\rangle + |3\rangle + |5\rangle + |7\rangle) \otimes |1\rangle$$

Now we need to extract the period of the first register without the obnoxious linear shift. So once again we apply the Fourier transform.

$$\frac{1}{2}(|1\rangle + |3\rangle + |5\rangle + |7\rangle) \xrightarrow{QFT_8} \frac{1}{\sqrt{2}}(|0\rangle - |4\rangle)$$

**Note:** If instead of measuring $|f\rangle = |1\rangle$ we had measured $|f\rangle = |0\rangle$, there would be a different linear shift. But the properties of Fourier transform dictate that this only effects the *phase* of the Fourier transform. In other words, that last step would have looked like $\frac{1}{2}(|0\rangle + |2\rangle + |4\rangle + |6\rangle) \xrightarrow{QFT_8} \frac{1}{\sqrt{2}}(|0\rangle + |4\rangle)$. This agrees with what we know about the principal of deferred measurement.

Finally, if we take a few measurements we will be sure to measure both $|0\rangle$ and $|4\rangle$. Therefore $N/r = 4$, and since $N = 8$, it is clear that $r = 2$.

## Summary

Now that we understand how the algorithm works, we can write it without some of the fluff.

$$|0\rangle\,|0\rangle \xrightarrow{QFT_M} \frac{1}{\sqrt{M}} \sum_{x \in \mathbf{Z}_M} |x\rangle\,|0\rangle \tag{6.1}$$

$$\xrightarrow{f} \frac{1}{\sqrt{M}} \sum_{x \in \mathbf{Z}_M} |x\rangle\,|f(x)\rangle \tag{6.2}$$

$$\xrightarrow{\text{measure 2nd register}} \sqrt{\frac{r}{M}} \sum_{k=0}^{\frac{M}{r}-1} |x_0 + kr\rangle\,|f(x_0)\rangle \tag{6.3}$$

$$\xrightarrow{QFT_M} \sqrt{\frac{r}{M}} \frac{1}{\sqrt{M}} \sum_{y \in \mathbf{Z}_M} \alpha_y\,|y\rangle \tag{6.4}$$

where $\alpha_y = \sum_{k=0}^{M/r-1} \omega^{(x_0+kn)y} = \omega^{x_0 y} \sum_k \omega^{kry}$.

There are two cases for $y$:

1. Case 1: $y$ is a multiple of $\frac{M}{r}$.

   In this case, then $\omega^{kry} = e^{2\pi i ry/M} = e^{n2\pi i} = 1$. So $\alpha_y = \frac{\sqrt{r}}{M}\frac{M}{r} = \frac{1}{\sqrt{r}}$. This should be thought of as *constructive interference* due to the final $QFT_M$.

   Note that there are $r$ multiples of $M/r$. Because $\sum_1^r (\frac{1}{\sqrt{r}})^2 = 1$, we know that $\alpha_y = 0$ for any $y$ that is *not* a multiple of $\frac{M}{r}$ by normality.

2. Case 2: $y$ is not a multiple of $\frac{M}{r}$.

   We already know that $\alpha_y$ must be 0 from the previous case. Furthermore, note that $\omega^{ry}, \omega^{2ry}, \ldots$ are evenly spaced vectors in the complex plain of unit length around the origin. Summing over these vectors we see that $\alpha_y$ is 0. This can be viewed as *destructive interference* due to the final $QFT_M$.

The interference that occurs in the final step is one reason quantum computers are so well equipped for period finding. We call it interference because it is additions in the *phase* that cause the cancellations.

## 6.1 Shor's Quantum Factoring Algorithm

One of the most celebrated algorithms for quantum computers is Shor's Algorithm for factoring. The time it takes for a classical computer to factor some number with n digits grows exponentially with n, meaning that numbers with many digits take a very long time for a classical computer to factor. RSA cryptography and other cryptography algorithms take advantage of this difficulty, and as a result a large amount of information is protected by large semi prime numbers (products of two primes).

The time it takes a quantum computer to factor an n-digit number grows as a polynomial in n.

The reason we focused so much attention on period finding is because the problem of factoring can be reduced to the problem of period finding thanks to modular arithmetic. This isn't obvious, but with a little setup we can understand why.

### Setup

In modular arithmetic, we call a number $x$ a non-trivial square root of 1 modulo $N$ if $x^2 \equiv 1 \pmod{N}$ and $x \neq \pm 1$. For example, 2 is a non-trivial square root of unity modulo 3 because $2^2 = 4 \equiv 1 \pmod{3}$. It turns out that if we can find such an $x$, we can factor $N$. Later we will see that we can use period finding to find $x$. This idea is summed up in the following lemmas.

Factoring is equivalent to finding a nontrivial squareroot of $1 \mod N$. Let $x \neq \pm 1 \mod N$ and $x^2 = 1 \mod N$. Then $x^2 - 1 = 0 \pmod{N}$ so that $x^2 - 1$ is a multiple of $N$. Factoring, we see that $N \mid (x+1)(x-1)$, but because $x \neq \pm 1 \pmod{N}$, $N \nmid (x \pm 1)$.

Therefore, $\gcd(N, x+1)$ and $\gcd(N, x-1)$ are factors of N, and greatest common divisor is easy to compute with Euclid's algorithm.

**Example:** Suppose we want to factor the number 15. It is easy to see that $4^2 = 16 \equiv 1 \mod 15$, but $4 \neq \pm 1 \mod 15$. So 4 is a non-trivial square root of unity modulo 15. Then gcd(15,5) and gcd(15,3) are factors of 15. Sure enough we see that $5 \cdot 3 = 15$.

Now, all we need to do is find this nontrivial squareroot of unity, and we can factor whatever number we need. As promised, we can do this with period finding, specifically by computing the order of a random integer.

The **order** of some integer $x$ modulo $N$ is the smallest integer $r$ such that $x^r = 1 \mod N$. For example, the order of 2 modulo 3 is 2 since $2^2 \equiv 1$, the order of 3 modulo 5 is 4 since $3^2 = 9 \equiv 4$; $3^3 = 25 \equiv 2$; and $3^4 = 81 \equiv 1 \pmod{5}$. Another way to say this is that the order of $x$ is just the period of the function $f(i) = x^i \mod N$.

Suppose $N = p \cdot q$, and $x \in \mathbf{Z}_N$, $x \neq p, q$. Then with probability $\geq 1/2$, the order $s$ of $x$ is even, and $x^{s/2}$ is a nontrivial square root of 1 mod N.

The proof of this statement requires results from number theory (Fermat's little Theorem, Chinese remainder Theorem) that are outside the scope of this course, so we will state it without proof. However, it should be intuitive: if you imagine the order of a number to vary randomly from one number to the next, you expect the order of a number to be even with probability about half.

**Example:** Find the order of 2 (mod 63), and use it to factor 63.

1. $2 = 2$
2. $2^2 = 4$
3. $2^3 = 8$
4. $2^4 = 16$
5. $2^5 = 32$
6. $2^6 = 64 \equiv 1 \pmod{63}$

so that the order of 2 is 6. Note that a quantum computer wouldn't have to compute each of these powers, it would simply use the period finding algorithm described earlier. Now we compute $2^3 = 8 \neq \pm 1$, so that $\gcd(63, 8+1) = 9$ and $\gcd(63, 8-1) = 7$ are factors of 63.

### The Algorithm

When finding order using the period finding algorithm, it is important to use enough qubits. A sensible rule is that you need to use $m$ qubits so that $2^m \gg N^2$, where $N$ is the number we are trying to factor, because the order of a random number might be as large as $N$.

We now have all the necessary tools to carry out Shor's algorithm. Start by picking a random number, then use the period finding algorithm to compute its order. If the order is even, we can use it to find a nontrivial square root of unity. If the order is odd or $x^{s/2} = -1$, throw it out and start with a new number.

Because we know that the order of $x$ will be even and $x^{s/2}$ will be a nontrivial square root with probability at least 1/2, we can be confident that we will be able to factor $N$ in just a few runs of the algorithm. Because the time it takes to find the period grows as a polynomial in the number of bits, and the number of bits grows like $2 \log N$(by the above requirement), we expect the time it takes to factor $N$ to grow as a polynomial in $\log N$.

Here is the circuit for Shor's Algorithm. It relies heavily on period finding, and so the circuit looks a lot like the circuit for period finding. The key difference is that we are finding the period of $f(i) = x^i$, and the number of bits we need to input is very large.

### Example

Here's an example that's a little more fun. Lets factor 119. Suppose we pick the number 16 to start with.

First, we compute it's order.

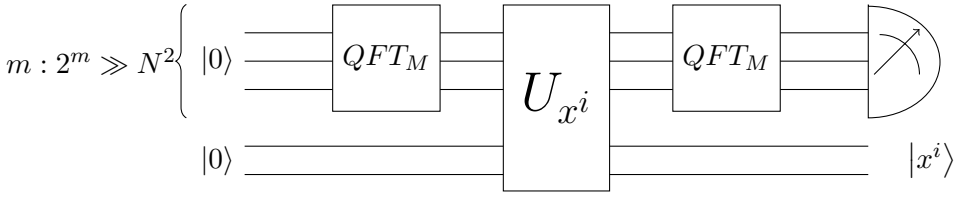



Figure 6.2: Circuit for factoring

1. $16 = 16$
2. $16 \cdot 16 = 256 \equiv 18$
3. $18 \cdot 16 = 288 \equiv 50$
4. $50 \cdot 16 = 800 \equiv 86$
5. $86 \cdot 16 = 1376 \equiv 67$
6. $67 \cdot 16 = 1072 = 119 \cdot 7 + 1 \equiv 1$

so that the order of 16 mod 119 is 6. Now, we compute $16^3 \equiv 50$. Gcd(49,119) = 7, so 7 is a factor of 119, and gcd(51, 119) = 17 which is another factor of 119.

# Chapter 7

# Grover's Search Algorithm

Searching an item in an unsorted table or array of size $N$ costs a classical computer $O(N)$ running time. If $N$ is large, this is like searching for a needle in a haystack. Can a quantum computer search for a needle in a haystack more efficiently than its classical counterpart? Grover, in 1995, affirmatively answered this question by proposing a search algorithm that consults the table only $O(\sqrt{N})$ times. In contrast to algorithms like quantum factoring which provide exponential speedups, the search algorithm only provides a quadratic improvement. However, the algorithm is quite important because it has broad applications, and because the same technique can be used to speedup algorithms for NP-complete problems.

One might wonder whether there are even faster quantum algorithms for search. However, it turns out that the quadratic speedup is optimal. This was proved in 1994, even before Grover's algorithm. Any quantum algorithm for search must consult the table at least some constant times $\sqrt{N}$ times. There are two ways to think about Grover's algorithm, and we will describe both ways below.

## 7.1 Quantum Search

### Idea of the Algorithm

The Grover search algorithm strives to solve this exact problem: We are given a boolean function $f : \{1, \ldots, N\} \to \{0, 1\}$, and are promised that for exactly one $a \in \{1, \ldots, N\}$, $f(a) = 1$. Of course, $a$ is the item we are searching for.

The basic idea of the Grover search algorithm is best described geometrically. Because our black box function has only two outcomes, we can identify two important states: $|a\rangle$, the state we are looking for; and everything else, call it $|e\rangle = \sum_{x \neq a} \frac{1}{\sqrt{N-1}} |x\rangle$. These two vectors span a two dimensional subspace, which contains the uniform superposition $|\psi_0\rangle = \sum_x \frac{1}{\sqrt{N}} |x\rangle$. Furthermore, $|a\rangle$ and $|e\rangle$ are orthogonal. We can represent this two dimensional subspace geometrically.

Because $|\psi_0\rangle$ is $N - 1$ parts $|e\rangle$ and only one part $|a\rangle$, it lies very close to $|e\rangle$. Grover's algorithm works by starting with the state $|\psi_0\rangle$ and successively increasing the angle between it and $|e\rangle$, to eventually get closer and closer to $|a\rangle$. It does this by a sequence of reflections: first by reflecting about $|e\rangle$, and then by reflecting about $|\psi_0\rangle$. The net effect of these two reflections, as we will

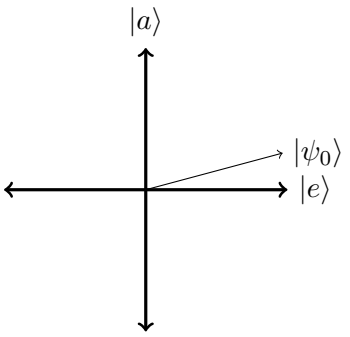



Figure 7.1: Two dimensional space spanned by $|a\rangle$ and $|e\rangle$, and the uniform superposition $|\psi_0\rangle$.

see, is to increase the angle between the state and $|e\rangle$. Repeating this pair of reflections moves the state farther and farther from $|e\rangle$, and therefore closer and closer to $|a\rangle$. Once it is close enough, measuring the state results in outcome $a$ with good probability.

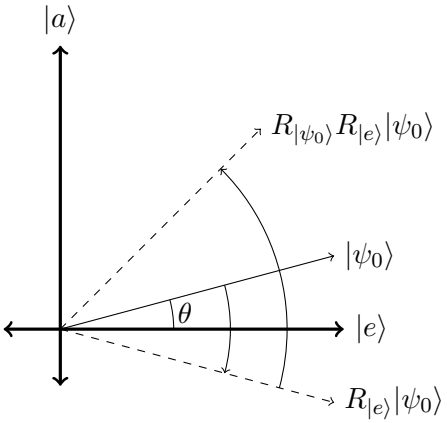



Figure 7.2: First reflect about $|e\rangle$, then reflect about $|\psi_0\rangle$.

Now the question is just how exactly to carry out these two reflections.

### The quantum oracle

From our boolean function $f : \{1, \ldots, N\} \to \{0, 1\}$, we can construct a quantum circuit $U_f$ to carry out this computation. Since we know $f$ can be computed classically in polynomial time, we can also compute it in superposition:

$$\sum_x \alpha_x |x\rangle|0\rangle \to \sum_x \alpha_x |x\rangle|f(x)\rangle$$

There is also a tricky way to put our result into a form that equally contains all of the information relevant to our problem. We can put the answer register in the superposition $|-\rangle$, so that when we implement $f$ the information is stored in the phase or sign of the result:

$$\sum_x \alpha_x |x\rangle|-\rangle \to \sum_x (-1)^{f(x)} \alpha_x |x\rangle|-\rangle$$

In more detail:

$$\begin{aligned} U_f : \sum_x \alpha_x |x\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) &\mapsto \sum_x \alpha_x \left(\frac{|x\rangle|f(x)\rangle - |x\rangle|\overline{f(x)}\rangle}{\sqrt{2}}\right) \\ &= \sum_x \alpha_x |x\rangle \left(\frac{|f(x)\rangle - |\overline{f(x)}\rangle}{\sqrt{2}}\right) \\ &= \sum_x \alpha_x |x\rangle (-1)^{f(x)} \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \end{aligned}$$

$U_f$ has the property that when $x = a$, the phase of the state will be multiplied $-1$. We will see that this implementation of the circuit is equivalent to a reflection over the vector $|e\rangle$.

### Grover's algorithm

Grover's algorithm finds $a$ in $O(\sqrt{N})$ steps. As before, consider the two dimensional subspace spanned by the two states: $|a\rangle$ and $|e\rangle$, where $|e\rangle$ is as above. Let $\theta$ be the angle between $|e\rangle$ and $|\psi_0\rangle = \sum_x \frac{1}{\sqrt{N}}|x\rangle$. See Figure 7.1 for an illustration of these vectors.

Since $|a\rangle$ is the target, we want to increase $\theta$: that is, rotate our input. One way to rotate a vector is to make two reflections. In particular, we can rotate a vector $|v\rangle$ by $2\theta$ by reflecting about $|e\rangle$ and then reflecting about $|\psi_0\rangle$. This transformation is also illustrated in Figure 7.2.

You can see that each time we implement these two reflections, we rotate by $2\theta$. This means that we need $\frac{\pi}{2}/2\theta = \pi/\theta$ iterations for the algorithm to complete. But what is $\theta$?

$$\langle\psi_0|a\rangle = \cos(\pi/2 - \theta) = \sin(\theta), \qquad \langle\psi_0|a\rangle = \sum_x \frac{1}{\sqrt{N}}\langle x|a\rangle = \frac{1}{\sqrt{N}}$$

so that

$$\sin(\theta) = \frac{1}{\sqrt{N}}$$

Since $1/\sqrt{N}$ is very small, $\sin\theta \approx \theta$, and $\theta \approx \frac{1}{\sqrt{N}}$. Thus, we need $O(\sqrt{N})$ iterations for the algorithm to complete. In the end, we get very close to $|a\rangle$, and then with high probability, a measurement of the state yields $a$.

This gets us everything except for the exact mechanism for each reflection. Because the reflection over $|\psi_0\rangle$ is not dependent on knowledge of $|a\rangle$, we should be able to construct it purely from our

regular unitary gates. The reflection over $|e\rangle$, however, requires knowledge of $|a\rangle$, so we will need to use our oracle $U_f$ to construct this reflection.

1. As it turns out, reflection about $|e\rangle$ is easy. All we need to do is flip the phase of the component in the direction of $|a\rangle$. We already saw how to achieve this using the second implementation of $f$ that we showed earlier.

2. For the reflection about $|\psi_0\rangle$, we use the Diffusion operator $D$ (assume $N = 2^n$), which works as follows. First, apply $H_{2^n}$, which maps $|\psi_0\rangle \mapsto |00\ldots0\rangle$. Then reflect around $|00\ldots0\rangle$ (this is accomplished by the circuit $U_g$, where $g$ is a function such that $g(00\ldots0) = 0$ and $g(x) = 1$ for $x \neq 00\ldots0$. Finally, apply $H_{2^n}$ to return to the original basis. (Note that this is simply a reflection around the zero vector in the Hadamard basis. You might understand this operation better after the second description of the algorithm below.)

### Another approach

Let's look at the search algorithm differently, with all superpositions.

**Claim .1** *The Diffusion operator $D$ has two properties:*

1. *It is unitary and can be efficiently realized.*

2. *It can be seen as an "inversion about the mean."*

Proof:

1. For $N = 2^n$, $D$ can be decomposed and rewritten as:

$$
\begin{aligned}
D &= H_N \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix} H_N \\
&= H_N \left( \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} - I \right) H_N \\
&= H_N \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} H_N - I \\
&= \begin{pmatrix} 2/N & 2/N & \cdots & 2/N \\ 2/N & 2/N & \cdots & 2/N \\ \vdots & \vdots & \ddots & \vdots \\ 2/N & 2/N & \cdots & 2/N \end{pmatrix} 1I \\
&= \begin{pmatrix} 2/N-1 & 2/N & \cdots & 2/N \\ 2/N & 2/N-1 & \cdots & 2/N \\ \vdots & \vdots & \ddots & \vdots \\ 2/N & 2/N & \cdots & 2/N-1 \end{pmatrix}
\end{aligned}
$$

Observe that $D$ is expressed as the product of three unitary matrices (two Hadamard matrices separated by a conditional phase shift matrix). Therefore, $D$ is also unitary. Regarding the implementation, both the Hadamard and the conditional phase shift transforms can be efficiently realized within $O(n)$ gates.

2. Consider $D$ operating on a vector $|\alpha\rangle$ to generate another vector $|\beta\rangle$:

$$
D \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_N \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_N \end{pmatrix}
$$

If we let $\mu = 1/N \sum_j \alpha_j$ be the mean amplitude, then the expression $2\mu - \alpha_i$ describes a reflection of $\alpha_i$ about the mean. This might be easier to see by writing it as $\mu + (\mu - \alpha_i)$. Thus, the amplitude of $\beta_i = -\frac{2}{N} \sum_j \alpha_j + \alpha_i = -2\mu + \alpha_i$ can be considered an "inversion about the mean" with respect to $\alpha_i$.

■

The quantum search algorithm iteratively improves the probability of measuring a solution. Here's how:

1. Start state is $|\psi_0\rangle = \sum_x \frac{1}{\sqrt{N}}|x\rangle$
2. Invert the phase of $|a\rangle$ using $f$
3. Then invert about the mean using $D$
4. Repeat steps 2 and 3 $O(\sqrt{N})$ times, so in each iteration $\alpha_a$ increases by $\frac{2}{\sqrt{N}}$

This process is illustrated in Figure 7.3.

Notice that at any point in the algorithm, the state can be described by two numbers, the amplitude $\alpha_a$ of $a$, and the amplitude $\alpha'$ of any $x \neq a$. Initially $\alpha' = \alpha_x = 1/\sqrt{N}$. As we run the algorithm $\alpha_a$ increases and $\alpha'$ decreases. Suppose we just want to find $a$ with probability $\frac{1}{2}$. Then we only need to run the algorithm until $\alpha_a \approx 1/\sqrt{2}$. At this point, $\alpha' \approx \frac{1}{\sqrt{2N}}$. This means that $\alpha' \geq \frac{1}{\sqrt{2N}}$ during the entire execution of the algorithm. Since in each iteration of the algorithm, $\alpha_a$ increases by at least $2\alpha'$ it follows that the increase is at least $\frac{2}{\sqrt{2N}} = \sqrt{\frac{2}{N}}$. Since our target is $\alpha_a = \frac{1}{\sqrt{2}}$, the number of iterations $\leq \frac{1}{\sqrt{2}}/\sqrt{\frac{2}{N}} = \sqrt{N}$.

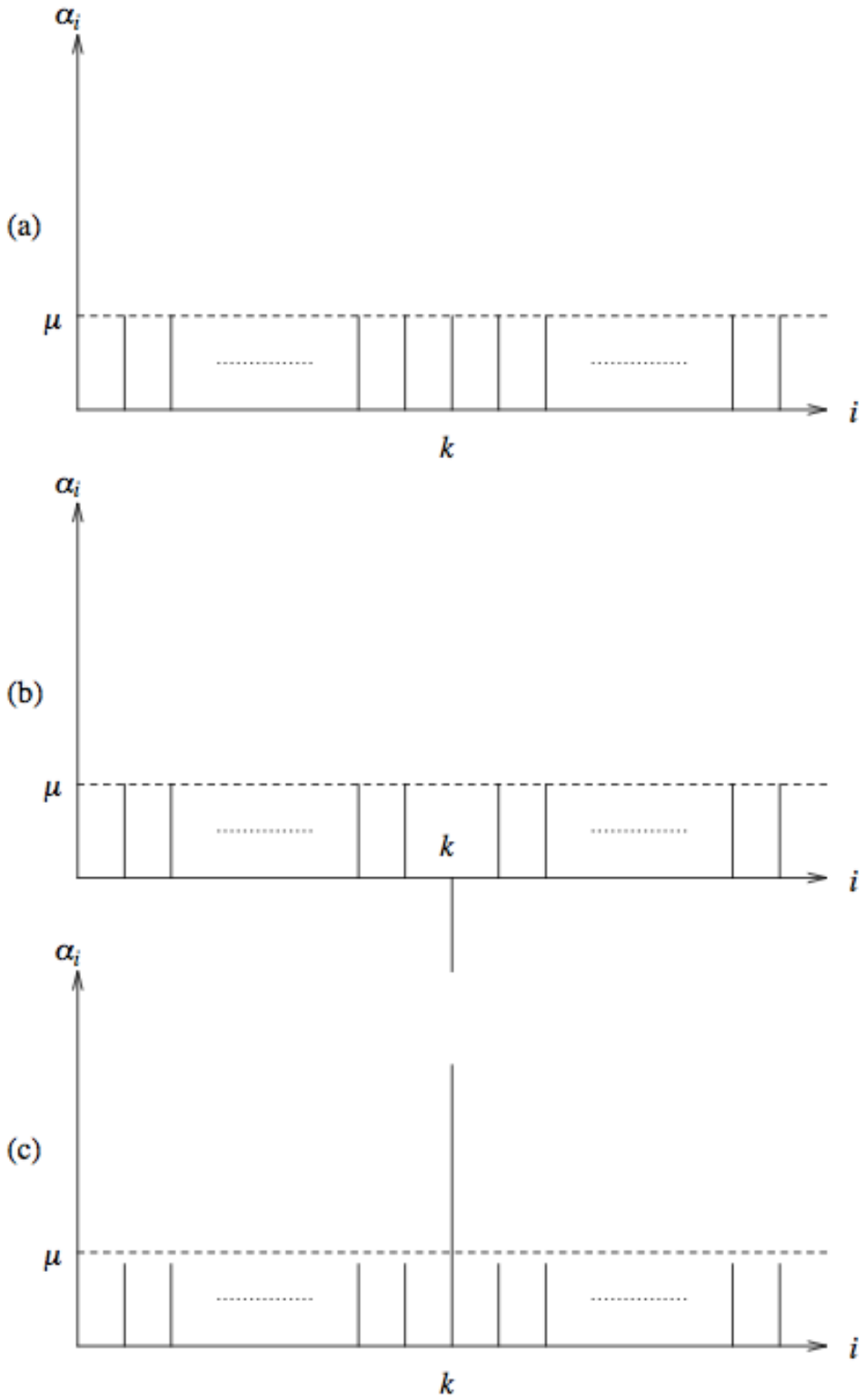



Figure 7.3: The first three steps of Grover's algorithm. We start with a uniform superposition of all basis vectors in (a). In (b), we have used the function $f$ to invert the phase of $\alpha_k$. After running the diffusion operator $D$, we amplify $\alpha_k$ while decreasing all other amplitudes.

# Chapter 8

# Observables

## 8.1 Observables

An observable is an operator that corresponds to a physical quantity, such as energy, spin, or position, that can be measured; think of a measuring device with a pointer from which you can read off a real number which is the outcome of the measurement. For a $k$-state quantum system, observables correspond to $k \times k$ hermitian matrices. Recall that a matrix $M$ is hermitian iff $M^\dagger = M$. Since $M$ is hermitian, it has an orthonormal set of eigenvectors $|\phi_j\rangle$ with real eigenvalues $\lambda_j$. What is the outcome of a measurement of the quantity represented by observable $M$ on a quantum state $|\psi\rangle$? To understand this, let us write $|\psi\rangle = a_0\phi_0 + \cdots + a_{k-1}\phi_{k-1}$ in the $\{|\phi_j\rangle\}$-basis. Now, the result of the measurement must be some $\lambda_j$ (this is the real number we read off our measurement device) with probability $|a_j|^2$. Moreover, the state of the system is collapsed to $|\phi_j\rangle$.

This description of a measurement relates to what we described earlier while explaining the measurement principle: there a measurement was specified by picking an orthonormal basis $\{|\phi_j\rangle\}$, and the measurement outcome was $j$ with probability $|a_j|^2$. The sequence of real numbers $\lambda_j$ simply provide a way of specifying what the pointer of the measurement device indicates for the $j$-th outcome. Moreover, given any orthonormal basis $|\phi_j\rangle$ and the sequence of real numbers $\lambda_j$, we can reconstruct a hermitian matrix $M$ as: $M = \sum_{j=0}^{k-1} \lambda_j |\phi_j\rangle \langle\phi_j|$; in the $\{|\phi_j\rangle\}$-basis this is just a diagonal matrix with the $\lambda_j$'s on the diagonal.

For example, suppose we wish to measure a qubit in the $|+\rangle, |-\rangle$-basis, with measurement results 1 and $-1$ respectively. This corresponds to measuring the observable

$$\begin{aligned} M &= (1)\,|+\rangle\langle+| + (-1)\,|-\rangle\langle-| \\ &= \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} - \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

By construction $M$ has eigenvectors $|+\rangle$ and $|-\rangle$ with eigenvalues 1 and $-1$ respectively.

One important observable of any physical system is its energy; the corresponding hermitian matrix or operator is called the Hamiltonian, and is often denoted by $\hat{H}$. The eigenvectors of this operator

are the states of the system with definite energy, and the eigenvalues are the numerical values of the energies of these eigenstates.

Consider, for example, two states $\psi_1$ and $\psi_2$ such that $\hat{H}\psi_1 = E_1\psi_2$ and $\hat{H}\psi_2 = E_2\psi_2$, where $E_1 \neq E_2$ (in quantum mechanical language this means that the *eigenvalues are non-degenerate*). Suppose we take $10^6$ qubits prepared in state $\psi_1$ and measure the energy of each one and make a histogram. What does the histogram look like? See Figure 1(a).

Now suppose that we prepare $10^6$ qubits in the state $\psi' = \sqrt{\frac{3}{5}}\psi_1 + \sqrt{\frac{2}{5}}\psi_2$, measure each of *their* energies, and make a histogram. How does it look? See Figure 1(b)

Ask yourself, is $\psi'$ a state with well-defined energy? The answer is *NO*. Why? Because $\psi'$ is not an eigenstate of the Hamiltonian operator. Let's check this:

$$\hat{H}\psi' = \hat{H}\left(\sqrt{\frac{3}{5}}\psi_1 + \sqrt{\frac{2}{5}}\psi_2\right) = \sqrt{\frac{3}{5}}E_1\psi_1 + \sqrt{\frac{2}{5}}E_2\psi_2$$

Does this equal (constant)$\times(\psi')$? No, because $E_1$ and $E_2$ are not equal. Therefore $\psi'$ is not an eigenstate of the energy operator and has no well-defined energy.

Even though a given state $|\psi\rangle$ might not have a definite energy, we can still ask the question, "what is the expected energy of this state?" i.e. if we prepare a large number of systems each in the state $|\psi\rangle$, and then measure their energies, what is the average result? In our notation above, this expected value would be $\sum_{j=0}^{k-1} |a_j|^2\lambda_j$. This is exactly the value of the bilinear form $\langle\psi| M |\psi\rangle$. Returning to our example above, where $M = H$, this expected value is $\frac{3}{5}E_1 + \frac{2}{5}E_2$.

How much does the value of the energy of the state $|\psi\rangle$ vary from measurement to measurement? One way of estimating this is to talk about the variance, $\mathrm{var}(X)$ of the measurement outcome. Recall that

$$\mathrm{var}(X) = E(X^2) - E(X)^2.$$

So to compute the variance we must figure out $E(X^2)$, the expected value of the square of the energy. This expected value is

$$\sum_{j=0}^{k-1} |a_j|^2\lambda_j^2.$$

This is exactly the value of the bilinear form $\langle\psi| M^2 |\psi\rangle$. So the variance of the measurement outcome for the state, $|\psi\rangle$ is

$$\mathrm{var}(X) = E(X^2) - E(X)^2 = \langle\psi| M^2 |\psi\rangle - (\langle\psi| M |\psi\rangle)^2.$$

Returning to our example above,

$$\langle\psi'| M^2 |\psi'\rangle = \sum_{j=0}^{k-1} |a_j|^2\lambda_j^2 = \frac{3}{5}E_1^2 + \frac{2}{5}E_2^2.$$

The variance is therefore

$$\mathrm{var}(X) = \frac{3}{5}E_1^2 + \frac{2}{5}E_2^2 - (\frac{3}{5}E_1 + \frac{2}{5}E_2)^2.$$

### Schrödinger's Equation

Schrödinger's equation is the most fundamental equation in quantum mechanics — it is the equation of motion which describes the time evolution of a quantum state.

$$i\hbar \frac{d\,|\psi(t)\rangle}{dt} = H\,|\psi(t)\rangle \quad .$$

Here $H$ is the Hamiltonian or energy operator, and $\hbar$ is a constant (called Planck's constant).

To understand Schrödinger's equation, it is instructive to analyze what it tells us about the time evolution of the eigenstates of the Hamiltonian $H$. Let's assume we are given a quantum system whose state at time $t = 0$ is, $|\psi(0)\rangle = |\phi_j\rangle$, an eigenstate of the Hamiltonian with eigenvalue, $\lambda_j$. Plugging this into Schrödinger's equation,

$$\frac{d\,|\psi(0)\rangle}{dt} = -\frac{i}{\hbar} H\,|\phi_j\rangle = -\frac{i}{\hbar}\lambda_j\,|\phi_j\rangle$$

So let us consider a system that is in the state $|\psi\rangle$ at time $t = 0$ such that that $|\psi(0)\rangle = |\phi_j\rangle$, an eigenvector of $H$ with eigenvalue $\lambda_j$. Now by Schrödinger's equation,

$$\frac{d\,|\psi(0)\rangle}{dt} = -H\,|\phi_j\rangle\,/\hbar = -i\lambda_j/\hbar\,|\phi_j\rangle\,.$$

Thus $|\psi(t)\rangle = a(t)\,|\phi_j\rangle$. Substituting into Schrödinger's equation, we get:

$$i\frac{da(t)\,|\phi_j\rangle}{dt} = H\,|a(t)\phi_j\rangle = a(t)\lambda_j\,|\phi_j\rangle\,.$$

Thus $i\hbar\frac{da(t)}{a(t)} = \lambda_j dt$. Integrating both sides with respect to $t$: $i\hbar \ln a(t) = \lambda_j t$. Therefore $a(t) = e^{-i\lambda_j t/\hbar}$, and $|\psi(t)\rangle = e^{-i\lambda_j t/\hbar}\,|\phi_j\rangle$.

So each energy eigenstate $|\phi_j\rangle$ is invariant over time, but its phase precesses at a rate proportional to its energy $\lambda_j$.

What about a general quantum state $|\psi(0)\rangle = \sum_j a_j\,|\phi_j\rangle$? By linearity, $|\psi(t)\rangle = \sum_j a_j e^{-i\lambda_j t}\,|\phi_j\rangle$.

In the basis of eigenstates of $H$, we can write this as a matrix equation:

$$|\psi(t)\rangle = \begin{pmatrix} e^{-\frac{i}{\hbar}\lambda_1 t} & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & e^{-\frac{i}{\hbar}\lambda_d t} \end{pmatrix} \begin{pmatrix} a_0 \\ \cdot \\ \cdot \\ a_{k-1} \end{pmatrix} = U(t)\,|\psi(0)\rangle$$

We have proved that if the Hamiltonian $H$ is time independent, then Schrödinger's equation implies that the time evolution of the quantum system is unitary. Moreover, the time evolution operator $U(t)$ is diagonal in the basis of eigenvectors of $H$, and can be written as $U(t) = e^{\frac{-iHt}{\hbar}}$.

Returning to our running example, suppose $\psi(x, t = 0) = \psi_1(x)$ where $\hat{H}\psi_1 = E_1\psi_1(x)$. What is $\psi(x, t \neq 0)$? The answer is,

$$\psi(x,t) = \psi_1(x)e^{-iE_1t/\hbar}$$

But what if $\psi(x,t=0) = \psi' = \sqrt{\frac{3}{5}}\psi_1 + \sqrt{\frac{2}{5}}\psi_2$? What's $\psi(x,t\neq 0)$ in this case? The answer then becomes,

$$\psi(x,t) = \sqrt{\frac{3}{5}}\psi_1 e^{-iE_1t/\hbar} + \sqrt{\frac{2}{5}}\psi_2 e^{-iE_2t/\hbar}$$

Each different piece of the wavefunction with differnt well-defined energy dances to its own little drummer. Each piece *spins* at frequency proportional to its energy.

### Conservation Laws and the Hamiltonian

Energy is typically the most important physical observable characterizing any system. You might still wonder, "why is energy so intimately related to the time evolution of a quantum system?" In this section we will try to answer this question. The answer is related to a fundamental physical principle, namely the conservation of energy.

We start by assuming that the time evolution of the state $|\psi\rangle$ in Schrödinger's equation is governed by some arbitrary hermitian operator $M$, or equivalently that the evolution of the system is given by some unitary transformation $U = e^{-iMt}$ (with a little bit of work this can be shown to follow from the third axiom of quantum mechanics in the "time independent situation", where the external conditions the system is subject to do not change over time). So our question reduces to asking, why is the operator $M$ necessarily the energy operator?

To see this, we must first show that if $A$ is any observable corresponding to a physical quantity that is conserved in time, then $A$ commutes with $M$ (as defined above).

Let $|\psi\rangle$ be the initial state of some physical system, and $|\psi'\rangle = U|\psi\rangle = e^{iMt}|\psi\rangle$ be the state after an infinitesimal time interval $t$.
Since $A$ corresponds to a conserved physical quantity, $\langle\psi'|A|\psi'\rangle = \langle\psi|A|\psi\rangle$.
i.e. $\langle\psi|U^\dagger AU|\psi\rangle = \langle\psi|A|\psi\rangle$.
Since this equation holds for every state $|\psi\rangle$, it follows that $U^\dagger AU = A$.
Substituting for $U$, we get
$LHS = e^{-iMt}Ae^{iMt} \approx (1-iMt)A(1+iMt) \approx A - it[M,A]$
where $[M,A] = MA - AM$.
It follows that $[M,A] = 0$.

So any observable corresponding to a conserved quantity must commute with the operator $M$ that describes the time evolution. Now, in addition to energy, there are situations where other physical quantities, such as momentum or angular momentum, are also conserved. These are in a certain sense "accidental" conservation relations — they may or may not hold. Energy however is always conserved. Hence the operator $H$ cannot be just any operator that happens to commute with $M$, but must have some universal property for all physical systems. An intrinsic reason that $H$ might commute with $M$ is that $H = f(M)$. i.e. $H$ is some function of $M$. Since any function of $M$ commutes with $M$ we now assume that $H = f(M)$.

The next critical point to show is that if $H = f(M)$, then $f$ must necessarily be a linear function. Consider a quantum system consisting of two subsystems that do not interact with each other. If $M_1$ and $M_2$ are the time evolution operators corresponding to each subsystem, then $M_1 + M_2$ is the time evolution operator of the system (since the two subsystems do not interact). So the total energy of the system is $f(M_1 + M_2)$. On the other hand, since the two subsystems do not interact, the system hamiltonian, $H = H_1 + H_2 = f(M_1) + f(M_2)$. Hence $f(M_1 + M_2) = f(M_1) + f(M_2)$, and therefore $f$ is a linear function $f(M) = \hbar M$, where $\hbar$ is a constant. So $H = \hbar M$ and $U(t) = e^{iHt/\hbar}$. Since $Ht/\hbar$ must be dimensionless, the constant $\hbar$ must have units of energy x time.

# Chapter 9

# Continuous Quantum Systems

## 9.1 The wavefunction

So far, we have been talking about finite dimensional Hilbert spaces: if our system has $k$ qubits, then our Hilbert space has $2^n$ dimensions, and is equivalent to $\mathbb{C}^{2^n}$. This follows because a set of qubits has a finite number of states: it is only possible to measure each qubit in the state $|1\rangle$ or $|0\rangle$. However, qubits with their finite states are not the only thing that quantum mechanics can deal with. Certainly we could try to measure the position of a quantum particle, and the possible outcomes lie on a continuum. In what follows we describe how to deal with continuous quantum states.

We must now expand our notion of Hilbert space, since the dimension (ie. number of basis states) runs to infinity. A continuous observable, such as position $x$, must be represented by an infinite-dimensional matrix

$$\hat{x} = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_\infty \end{pmatrix}$$

where $\mathbf{x}_j$ denotes all possible positions on a line, in the limit where $j$ becomes a continuous variable. If the particle is sitting at a known position, $x_p$, then its state, $|\psi\rangle$, can be represented in the position-basis by the infinite-dimesional vector

$$|\psi\rangle = |x_p\rangle = (0, 0, \ldots, 0, 1, 0, \ldots, 0, 0),$$

where only the $p^{\text{th}}$ position is nonzero. Of course, the particle's state might alternatively be composed of an arbitrary superposition of position states:

$$|\psi\rangle = a_0 |x_0\rangle + a_1 |x_1\rangle + \cdots$$

where $|a_0|^2 + |a_1|^2 + \cdots = 1$.

The matrix/vector notation becomes extremely awkward as we attempt to cope with an infinite number of infinitesimally-spaced basis states. To deal with this, suppose the particle's state, $|\psi\rangle$ is some arbitrary superposition of infinitesimally-spaced position eigenstates. If we now ask, "What

is the probability-density that the particle will lie at an arbitrary position, $x$, represented by the position eigenstate, $|x\rangle$?" Just as in the finite case, the answer is the inner product $\langle x|\psi\rangle$. Since $x$ is a continuous variable, this inner product is a continuous function of $x$. This leads us to define $\psi(x) = \langle x|\psi\rangle$, and it is called the *wavefunction* of the particle with respect to position.

Now, rather than struggle to tediously write down infinite superpositions of infinitesimally-spaced basis states, we need only specify the continuous function $\psi(x)$. This contains all of the complex information of the infinite-dimensional superposition of states. Since $|\psi\rangle$ is a unit vector in an infinite-dimensional Hilbert space, then $\psi(x)$ must satisfy the condition

$$\langle\psi|\psi\rangle = \lim_{\Delta x_j \to 0} \sum_{x_j=-\infty}^{\infty} \langle\psi|x_j\rangle \langle x_j|\psi\rangle \Delta x_j = \int_{-\infty}^{\infty} \langle\psi|x\rangle \langle x|\psi\rangle \, dx = \int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = 1$$

The operator that represents position, $X$, now operates on the inner product, $\psi(x)$, to yield the eigenvalue equation

$$X\psi(x) = x\psi(x),$$

where $x$ is a scalar.

## 9.2 The Schrödinger Equation

The fundamental equation of quantum mechanics is the Schrödinger Equation, stumbled upon by physicist Erwin Schrödinger in 1925. The Schrödinger equation tells us how a quantum particle in a continuous system should behave. The equation is very difficult to solve, in fact in most real situations it is impossible to solve. For a particle free to move in one dimension, the Schrödinger equation reads

$$H\Psi(x,t) = i\hbar\frac{d}{dt}\Psi(x,t)$$

where $H$ is the Hamiltonian, or energy operator, of a particle that can move in one dimension.

Classically the energy of a particle is simply the sum of its kinetic and potential energy. The total energy of a particle with mass $m$ is well defined in terms of the momentum $p$ and position $x$ of the particle, and is given by

$$E(p,x) = \frac{p^2}{2m} + V(x)$$

Here, $p^2/2m$ is the kinetic energy and $V(x)$ is the classical potential energy of the particle at position $x$. The form of $V(x)$ depends upon what interactions the particle is subjected to (e.g. an electron in a magnetic field, or a free photon). Exactly how to get from the classical energy function, $E(p,x)$ to the quantum mechanical energy operator, $H$, is not totally obvious. We will rely on an axiom of quantum mechanics that we will try to justify (but *not* derive) later on.

Axiom: If the classical energy operator for a system is $E(p,x)$, then the quantum mechanical Hamiltonian can be written as $H = E(\hat{p},\hat{x})$, where $\hat{p}$ and $\hat{x}$ are the quantum mechanical momentum and position operators, respectively. In the position basis, the $\hat{x}$ operator is simply the function $x$, whereas the $\hat{p}$ operator is $\hat{p} = -i\hbar\frac{\partial}{\partial x}$.

To really understand this, we need a few examples.

### The Free Particle

Our first example will be that of a free particle in 1 dimension. Here free means that the particle is subject to no potential interactions with other particles. Schrödinger's equation says

$$H\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t)$$

Because the particle is free, its classical energy is just its kinetic energy $\frac{p^2}{2m}$; the potential energy $V(x) = 0$. Thus,

$$H = \frac{p^2}{2m} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$$

and

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t)$$

At first glance, this equation looks daunting and difficult to solve, it is in fact a second order partial differential equation. However, a mathematical trick called "separation of variables" makes the equation fairly easy to solve. The mathematical formalism of separation of variables is not necessary for this course, so if the following discussion is not helpful to you, feel free to skip it. The important result we derive is that the Schrödinger equation can be separated into into two parts:

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) = E\psi(x) \qquad i\hbar\frac{\partial}{\partial t}\phi(t) = E\phi(t)$$

where $\Psi(x,t) = \psi(x)\phi(t)$.

Here's how it works. Because the derivatives in the Schödinger equation are with respect to different variables, there is an easy way to solve the equation. Because the right hand side and the left hand side of the equation depend on different variables entirely, we can say that the time dependence of $\Psi$ is independent of $x$ dependence: $\Psi(x,t) = \Psi(x,0)\Psi(0,t)$. For if this were not true, then when we change $x$ without changing $t$, the right hand side of the Shrödinger equation changes differently from the left hand side.

In equation form, this translates to:

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,0)\right][\Psi(0,t)] = [\Psi(x,0)]\left[i\hbar\frac{\partial}{\partial t}\Psi(0,t)\right]$$

If we let $\psi(x) = \Psi(x,0)$ and $\phi(t) = \Psi(0,t)$, we see that changing $x$ does not change either $\phi(t)$ or $\frac{\partial}{\partial t}\phi(t)$: these terms are constant with respect to $\psi(x)$. Thus we can separate the above equation into two parts, one that describes the time dependence, and one that describes the position dependence.

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) = E\psi(x) \qquad i\hbar\frac{\partial}{\partial t}\phi(t) = E\phi(t)$$

The first equation is called the *time-independent* Schrödinger equation, and it is fairly easy to solve. The constant $E$ is used because the Hamiltonian on the right hand side is the energy of the particle: $H\psi = E\psi$. The solution to the time independent Schrödinger equation is

$$\psi_k(x) = e^{ikx}, \qquad \psi_k(x) = e^{ikx}$$

where $k = \sqrt{2mE}/\hbar$. If we let $k$ run negative, then we only need to think about the first solution.

The solution to the time portion of the Schrödinger equation is

$$\phi_k(t) = e^{-i\omega t}$$

where $\omega = E/\hbar = \hbar k^2/2m$. Because the time dependent equation is not dependent upon the potential energy, this is *always* the time dependence of a system. To turn a solution of the the time independent Schrödinger euqtion into a time dependent solution, just tack on an $e^{-iE_k/\hbar}$ to each $k^{\text{th}}$ energy eigenstate.

Now, because $\Psi(x,t) = \psi(x)\phi(t)$, the final

$$\boxed{\Psi_k(x,t) = e^{i(kx-\omega t)}}$$

We see that for each $k$ (or $E$) there is a solution, which gives us a continuous set of solutions to the Schrödinger equation. We can think of the set $\Psi_k$ as a basis for the possible states of our particle, since linear combinations of solutions to a linear differential equation are also solutions to the same equation.

### Particle in a Box

Another classic example where Schrödinger's equation is actually solvable is the particle in a box, also known as the infinite square well. In this problem, the particle is free move however it likes within a single line segment (this is its box), but is not allowed to leave.

While this situation is unrealistic and does not occur in nature, it is a half decent way to approximate an atom. You can think of an electron in a hydrogen atom, for example, as being trapped in a box. The potential near the central proton is much lower than far from it. It might be a stretch to say that the this is the same as our particle in a box, but they certainly are similar.

The way to describe the particle in a box is to say that the potential inside the box is 0, while the potential outside is infinite: it would take infinite energy for the particle to exist outside of its little line segment, thus it cannot exist outside of the box.

$$V(x) = \begin{cases} 0 & \text{if } 0 < x < a \\ \infty & \text{otherwise} \end{cases}$$

Inside the box, the Hamiltonian is $H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$. Outside the box we simply mandate that $\Psi(x,t) = 0$.

Furthermore, to make sure that the Schrödinger equation makes sense, we will require that $\Psi(x,t)$ is continuous. Because $\psi(x) = 0$ outside of the box, we require that:

$$\psi(0) = \psi(l) = 0$$

We first solve the time-independent Schrödinger equation without worrying about the above restriction:

$$\frac{d^2}{dx^2}\psi(x) = E\psi(x)$$

We already solved this, and our solution was $\psi_k(x) = Ae^{ikx} + Be^{-ikx}$. But before we can call this done we need to impose our *boundary condition*, i.e. the restriction that $\psi(0) = \psi(a) = 0$.

To get rid of the complex exponentials and make life a little easier, we recall that for some $C$ and $D$,[1]

$$Ae^{ikx} + Be^{-ikx} = C\sin(kx) + D\cos(kx)$$

So $\psi_k(x) = C\sin(kx) + D\cos(kx)$ for some $C$ and $D$. To find the conditions on $C$ and $D$, we impose our boundary conditions:

$$\psi_k(0) = C\sin(0) + D\cos(0) = D$$

But $\psi_k(0) = 0$ so $D = 0$, and we can forget about the cosine solution. The second boundary condition tells us:

$$\psi_k(a) = C\sin(ka) = 0$$

This can only be satisfied when $ka = n\pi$, where $n$ is an integer. And because $k = \sqrt{2mE}/\hbar$, the only allowed energies are $E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}$. Thus, our (almost final) set of solutions is, for each integer $n$,

$$\psi_n(x) = C\sin\left(\frac{n\pi x}{a}\right) \text{with energy} E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}$$

Notice how the *quantization* of energy levels, the fact that there is a discreet set of energy eigenvalues, and quantization of basis states simply falls out of the math. This kind of phenomenon is what gives quantum mechanics its name.

There is one last step before we can call it a day. We need to make sure that $\langle psi_n|\psi_n\rangle = 1$ for all $n$. This is called *normalizing* the wavefunctions.

$$\langle\psi_n|\psi_n\rangle = \int_0^l |\psi_n(x)|^2 dx = 1 \Rightarrow \int_0^l C^2 sin^2\left(\frac{n\pi x}{a}\right) dx = C^2\frac{a}{2} = 1$$

[1] If you have never seen this before, it is not too bad of an exercise to find $C$ and $D$ with the identity $e^{ikx} = \cos(kx) + i\sin(kx)$.

so that $C = \sqrt{2/a}$.

With everything normalized and all boundary conditions accounted for, we finally have our proper set of energy eigenfunctions and eigenvalues:

$$\boxed{\psi_n(x) = C \sin\left(\frac{n\pi x}{a}\right) \text{with energy} E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}}$$

## Qubits

While it is important to have this background in quantum mechanics to better understand quantum information science, there is a direct connection between the solution to the particle in a box problem and qubits. We have discussed several examples of qubits in previous lectures, one of which used the ground state and first excited state of an atom as the two states of a qubit. Because the particle in a box is a simple model for a hydrogen atom, we will discuss hydrogen atom qubits in the context of the square well.

To obtain a qubit from particle in a box system, we can construct our standard basis $|0\rangle$ and $|1\rangle$ by restricting our state space to the bottom two eigenstates:

$$|0\rangle = \sqrt{\frac{2}{a}} \sin\left(\frac{\pi x}{a}\right), \qquad E_1 = \frac{\hbar^2\pi^2}{2ma^2}$$

$$|1\rangle = \sqrt{\frac{2}{a}} \sin\left(\frac{2\pi x}{a}\right), \qquad E_2 = \frac{4\hbar^2\pi^2}{2ma^2}$$

Physically this would mean demanding that the energy in the box be less than or equal to $E_2$, meaning that the particle could never have any overlap with $\psi_n$ for $n > 2$. This can be done in the lab by cooling the atom to very low temperatures (perhaps laser cooling).

In fact, given the right length $a$ of the box, the lowest two states of the hydrogen atom are approximated very well by the infinite square well. Because of this, we can do some calculations to take a look at what a hydrogen atom qubit looks like. An arbitrary qubit superposition of the electron state can be written as

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle = \alpha\sqrt{\frac{2}{a}} \sin\left(\frac{\pi x}{a}\right) + \beta\sqrt{\frac{2}{a}} \sin\left(\frac{2\pi x}{a}\right)$$

The time evolution of this state at some later time $t$ can be tacked on by multiply each eigenstate by $e^{-iE_n/\hbar}$, as noted in the free particle problem:

$$|\psi(t)\rangle = \alpha\,|0\rangle\, e^{-iE_1 t/\hbar} + \beta\,|1\rangle\, e^{-iE_2 t/\hbar}$$

This can be rearranged to become:

$$|\psi(t)\rangle = e^{-iE_1 t/\hbar}\left(\alpha\,|0\rangle + \beta\,|1\rangle\, e^{-i(E_2-E_1)t/\hbar}\right)$$

or

$$\boxed{|\psi(t)\rangle = e^{-iE_1 t/\hbar}\left(\alpha\,|0\rangle + \beta\,|1\rangle\, e^{-i(\Delta E)t/\hbar}\right)}$$

The important point to notice here is that as time passes, the phase difference between the two qubit states differs by a rate that is proportional to $\Delta E$, the energy difference between them. For atomic systems this is a pretty fast rate, since $\Delta E = 10$ eV corresponds to a frequency of $\nu = \frac{\Delta E}{h} = 2.5 \times 10^{15}$ Hz. This is very close to the frequency of optical light, and *thus atomic qubits are controlled optically via interaction with light pulses.*

# Chapter 10

# Spin

## 10.1 Spin $\frac{1}{2}$ as a Qubit

In this chapter we will explore quantum spin, which exhibits behavior that is intrinsically quantum mechanical. For our purposes the most important particles are electrons and protons which are spin 1/2 particles. Their spin state is a qubit, which can be in one of two orthogonal states: spin up denoted by $|\uparrow\rangle$ or spin down $|\downarrow\rangle$. By the superposition principle, the spin state of an electron or proton is thus $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$. The spin state is an excellent way to implement a qubit in real life.

### The Bloch Sphere

Before we can further understand spin, it is useful to digress into a 3-dimensional representation of a qubit via the "Bloch Sphere." This representation allows us to picture the state of a qubit as a point on the unit sphere in 3-dimensional Euclidean space $\mathbb{R}^3$. This convenient mapping between possible single-qubit states and the unit sphere is best explained by picture:

Choosing $\theta$ and $\phi$ as the usual spherical coordinates, every point $(\theta, \phi)$ on the unit sphere represents a possible state of the qubit:

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

And any possible state of a qubit (up to an overall multiplicative phase factor) is represented by a vector on this unit sphere.

What does the action of a single qubit gate correspond to in this Bloch sphere representation? The answer is natural and elegant: a single qubit gate is just a rotation of the Bloch Sphere (see homework). Let's do an example. Consider the Hadamard gate $\mathbf{H}$ that has been discussed in the past. (Note that $\mathbf{H}$ is not equal to the Hamiltonian in this case!)

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

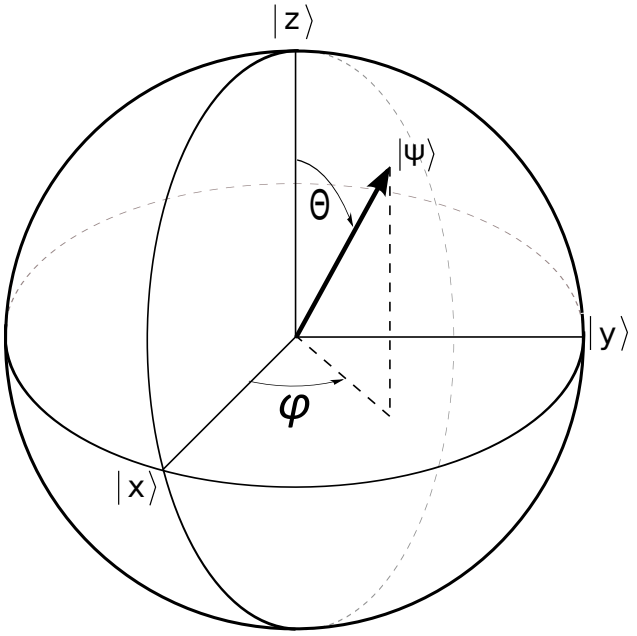



Figure 10.1: Representation of the state of a spin–$\frac{1}{2}$ particle as a point on the surface of the Bloch Sphere.

Given our generalized expression for a quantum state on the Bloch sphere ($|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle$), we see that the action of the Hadamard gate is a $\pi$ rotatation about the $\pi/4$ axis in the $x-z$-plane:y-axis.

### Spatial Interpretation of Spin and Pauli Spin Matrices

Let's return to the question about what we mean spatially when we say that the electron is in the spin up state $|\uparrow\rangle$. We can understand this by referring to our Bloch sphere picture of the spin qubit. Let us align the $z$-axis with the direction "up" associated with spin up (say, as defined by an external B-field). Then the spin up state is identified with the positive $z$ direction or the $|0\rangle$ state on the Bloch sphere. The spin down state is identified with the negative $z$ direction or the $|1\rangle$ state on the Bloch sphere. The positive $x$ direction corresponds to the state $1/\sqrt{2}|\uparrow\rangle + 1/\sqrt{2}|\downarrow\rangle$. Thus the Bloch sphere provides us a way of translating between the abstract vector space in which the state of the spin qubit resides and real 3-dimensional space.

Let us now consider how we would measure whether a spin qubit is in the state spin up $|\uparrow\rangle$ or spin down $|\downarrow\rangle$. We must define an operator whose eigenvectors are $|\uparrow\rangle$ and $|\downarrow\rangle$ (since only through such an operator can we associate a measurable quantity with spin-up or spin-down states). The operator that achieves this is the Pauli matix $\sigma_z$:

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

The eigenvectors of this operator are clearly $|0\rangle$ with eigenvalue 1 and $|1\rangle$ with eigenvalue $-1$.

Similarly if we wish to measure whether the spin points in the plus $x$ direction ($|+\rangle$) or minus $x$ direction ($|-\rangle$) on the Bloch sphere, we would use the Pauli matrix $\sigma_x$:

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

Its eigenvectors are $|+\rangle = 1/\sqrt{2}\,|0\rangle + 1/\sqrt{2}\,|1\rangle$ with eigenvalue 1 and $|-\rangle = 1/\sqrt{2}\,|0\rangle - 1/\sqrt{2}\,|1\rangle$ with eigenvalue $-1$.

Finally if we wish to measure whether the spin points in the plus $y$ direction or the minus $y$ direction we would use the Pauli matrix $\sigma_y$:

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

The magnitude of the spin (intrinsic angular momentum) for an an electon is $\hbar/2$, and so the spin operators are the Pauli operators scaled by $\hbar/2$: $\hat{S_x} = \hbar/2\sigma_x$, $\hat{S_y} = \hbar/2\sigma_y$, $\hat{S_z} = \hbar/2\sigma_z$. These spin operators will play a central role in our study of spin.

Let us now step back and consider what we have learnt so far about spin. Spin is the intrinsic angular momentum carried by elementary particles, and for spin 1/2 particles such as electrons and protons spin is described by a qubit. But how do we reconcile this unit vector in an abstract vector space with the picture of an angular momentum vector pointing in some direction in real Euclidean space? We do this via the Bloch sphere, which maps qubit states onto points on the unit sphere in 3 dimensions.

There are several counter-intuitive consequences of what we have already described. Spin about the three axis are not independent. The $z$ and $y$ component of the spin completely determine the spin (and so do the $z$ and $-z$ component!!). Indeed, we cannot independently measure the spin about the $x$, $y$ and $z$ axes. Another way of saying this is that the operators $\hat{S_x}$, $\hat{S_y}$ and $\hat{S_z}$ do not commute.

We express this via the commutation relations (verify these!) [1]:

$$[\hat{S_x}, \hat{S_y}] = i\hbar\hat{S_z}, [\hat{S_y}, \hat{S_z}] = i\hbar\hat{S_x}, [\hat{S_z}, \hat{S_x}] = i\hbar\hat{S_y}$$

Finally, it will be useful for our later treatment to introduce a fourth operator $\hat{S^2} = \hat{S_x}^2 + \hat{S_y}^2 + \hat{S_z}^2$. It is easy to verify that each of the Pauli spin matrices squares to the identity matrix, and therefore for the electron $\hat{S^2} = 3\hbar^2/4I$. Our final commutation relation simply says that $\hat{S^2}$ commutes with each of the other three spin operators:

$[\hat{S^2}, \hat{S_i}] = 0$

### History and a semi-classical picture

The history of the development of spin is an interesting one. In 1924 Pauli postulated a "two-valued quantum degree of freedom" (a qubit in our notation) associated with the electron in the outermost

[1] recall $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$

shell, which helped him formulate the Pauli exclusion principle. But he sharply criticized Kronig's early suggestions that this degree of freedom was produced by a rotation of the electron. The "discovery" of spin is largely credited to two dutch physicists Uhlenbeck and Goudsmit who, in 1925, introduced it to explain the onset of new energy levels for hydrogen atoms in a magnetic field:

This can be explained if an electron behaves like a little magnet, if it has an *intrinsic* magnetic moment $\vec{\mu}$, since a magnetic moment in a magnetic field $\vec{B}$ has an energy $E = -\vec{\mu} \cdot \vec{B}$. In the context of QM, new energy levels come from $\vec{\mu}$ either parallel or anti-parallel to $\vec{B}$.

But where does $\vec{\mu}$ come from, and how do we explain its QM behavior?

The simplest explanation of the origin of $\vec{\mu}$ is "Classical": Classically, a loop of current gives rise to a magnetic moment $\vec{\mu}$:

The energy $E = -\vec{\mu} \cdot \vec{B}$ comes from $\vec{I} \times \vec{B}$ force of current in a B-field (Lorentz force). The lowest energy, and thereby the place where "the system wants to go", is obtained when the magnetic moment and B-field line up.

So a plausible explanation for "intrinsic" magnetic moment $\vec{\mu}$ of an electron is that the electron spins about some axis (thus effectively creating a loop of current about that axis). We can then express the magnetic moment $\vec{\mu}$ in terms of the angular momentum by a simple calculation:

Consider a charge of mass $m$ moving in a circle with velocity $v$. Then its angular momentum is given by $\vec{L} = \vec{r} \times \vec{p} = \vec{r} \times m\vec{v}$. So in magnitude $L = mvr$. Also, its magnetic moment is given by:

$$\mu = (current)\,(Area) = \frac{e}{\tau} \cdot \pi r^2$$

But $\tau = \frac{2\pi r}{v}$, so $\mu = \frac{e}{2} \cdot vr$.

But now we can express the magnetic moment in terms of the angular momentum:

$$\mu = \frac{e}{2} \cdot vr = \frac{e}{2} \cdot \frac{L}{m} \Rightarrow \vec{\mu} = -\frac{e}{2m}\vec{L}$$

The central hypothesis of electron spin due to Uehlenbeck and Goudsmit states that each electron has an intrinsic angular momentum of spin $\vec{S}$ of magnitude $\hbar/2$ and with an associated magnetic moment $\vec{\mu} = -\frac{ge}{2m}\vec{S}$.

Compare this to the classical equation above: $\vec{\mu} = -\frac{e}{2m}\vec{L}$. What is $g$? $g$ is called the g-factor and it is a unitless correction factor due to QM. For electrons, $g \approx 2$. For protons, $g \approx 5.6$. You should also note that $\frac{m_{proton}}{m_{electron}} \approx 2000$, so we conclude that $\vec{\mu_{proton}} \ll \vec{\mu_{electron}}$.

## 10.2 Stern-Gerlach Apparatus

A Stern-Gerlach device is simply a magnet set up to generate a particular inhomogeneous $\vec{B}$ field. When a particle with spin state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ is shot through the apparatus from the left, its spin-up portion is deflected upward, and its spin-down portion downward. The particle's spin

becomes entangled with its position! Placing detectors to intercept the outgoing paths therefore measures the particle's spin.

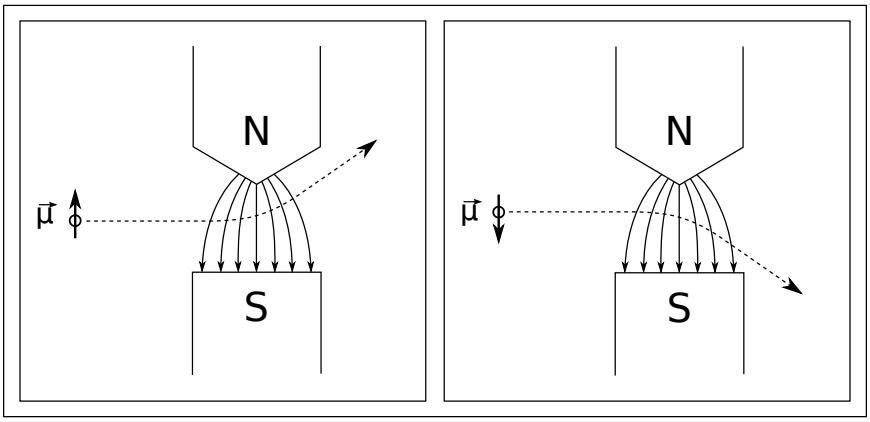



Figure 10.2: Stern-Gerlach device showing the deflection as a function of the orientation of the magnetic moment.

Why does this work? We'll give a semiclassical explanation – mixing the classical $\vec{F} = m\vec{a}$ and the quantum $H\psi = E\psi$ – which is quite wrong, but gives the correct intuition. [See Griffith's § 4.4.2, pp. 162-164 for a more complete argument.] Now the potential energy due to the spin interacting with the field is

$$E = -\vec{\mu} \cdot \vec{B} \ ,$$

so the associated force is

$$\vec{F}_{\text{spin}} = -\vec{\nabla} E = \vec{\nabla}(\vec{\mu} \cdot \vec{B}) \ .$$

At the center $\vec{B} = B(z)\hat{z}$, with $\frac{\partial B}{\partial z} < 0$, so $\vec{F} = \vec{\nabla}(\mu_z B(z)) = \mu_z \frac{\partial B}{\partial z}\hat{z}$. The magnetic moment $\vec{\mu}$ is related to spin $\vec{S}$ by $\vec{\mu} = \frac{gq}{2m}\vec{S} = -\frac{e}{m}\vec{S}$ for an electron. Hence

$$\vec{F} = \tfrac{e}{m}\left|\tfrac{\partial B}{\partial z}\right| S_z \hat{z} \ ;$$

if the electron is spin up, the force is upward, and if the electron is spin down, the force is downward.

## 10.3 Initialize a Qubit

- How can we create a beam of qubits in the state $|\psi\rangle = |0\rangle$? Pass a beam of spin-$\frac{1}{2}$ particles with randomly oriented spins through a Stern-Gerlach apparatus oriented along the $z$ axis. Intercept the downward-pointing beam, leaving the other beam of $|0\rangle$ qubits.

  Note that we *measure* the spin when we intercept an outgoing beam – after this measurement, the experiment is probabilistic and not unitary.

- How can we create a beam of qubits in the state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$? First find the point on the Bloch sphere corresponding to $|\psi\rangle$. That is, write

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle = \cos\tfrac{\theta}{2}\,|0\rangle + e^{i\varphi}\sin\tfrac{\theta}{2}\,|1\rangle$$

(up to a phase), where

$$\tan\frac{\theta}{2} = \left|\frac{\beta}{\alpha}\right| \qquad e^{i\varphi} = \frac{\beta/|\beta|}{\alpha/|\alpha|} \ .$$

The polar coordinates $\theta, \varphi$ determine a unit vector $\hat{n} = \cos\varphi\sin\theta\hat{x} + \sin\varphi\sin\theta\hat{y} + \cos\theta\hat{z}$. Now just point the Stern-Gerlach device in the corresponding direction on the Bloch sphere, and intercept one of the two outgoing beams. That is, a Stern-Gerlach device pointed in direction $\hat{n}$ measures $S_{\hat{n}} = \hat{n} \cdot \hat{\vec{S}}$.

- How can we implement a unitary (deterministic) transformation? We need to evolve the wave function according to a Hamiltonian $\hat{H}$. Then

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}t}\,|\psi(0)\rangle$$

solves the Schrödinger equation (if $\hat{H}$ is time-independent). In the next lecture we will show how to accomplish an arbitrary single-qubit unitary gate (a rotation on the Bloch sphere) by applying a precise magnetic field for some precise amount of time: Larmor precession.

# Chapter 11

# Manipulating Spin

## 11.1 Larmor Precession

Turning on a magnetic field $\vec{B}$, the qubit state rotates. There are two steps to understanding this process, essentially the same steps we make to understand any quantum process:

1. Find $\hat{H}$
2. Solve Schrödinger equation

For the second step, we first solve the "time-independent" Schrödinger equation; that is, we find energy eigenstates

$$\hat{H}\left|\psi_n\right\rangle = E_n\left|\psi_n\right\rangle \quad .$$

The "time-dependent" Schrödinger equation

$$i\hbar\tfrac{d}{dt}\left|\psi(t)\right\rangle = \hat{H}\left|\psi(t)\right\rangle$$

has solution

$$\left|\psi(t)\right\rangle = e^{-i\frac{\hat{H}}{\hbar}t}\left|\psi(t=0)\right\rangle \quad .$$

Expanding $\left|\psi(t=0)\right\rangle = \sum_n c_n\left|\psi_n\right\rangle$, we get

$$\left|\psi(t)\right\rangle = \sum_n c_n e^{-iE_nt/\hbar}\left|\psi_n\right\rangle \quad .$$

(This assumes that $\hat{H}$ is time-independent. If the Hamiltonian is itself a function of $t$, $\hat{H} = \hat{H}(t)$, then we must directly solve the time-dependent Schrödinger equation.)

### Find $\hat{\mathcal{H}}$

Assume there is only potential energy, not kinetic energy. Classically, $E = -\vec{\mu}\cdot\vec{B}$. Quantumly, the magnetic moment is in fact a vector *operator*, $\hat{\vec{\mu}} = \frac{gq}{2m}\hat{\vec{S}} = -\frac{e}{m}\hat{\vec{S}}$. Hence we set the quantum Hamiltonian to be

$$\hat{H} = \tfrac{e}{m}\hat{\vec{S}}\cdot\vec{B} \ .$$

We may choose our coordinate system so $\vec{B} = B\hat{z}$; then

$$\hat{H} = \frac{eB}{m}\hat{S}_z \ .$$

### Solve Schrödinger Equation

Following the recipe we gave above, we start by finding the eigendecomposition of $\hat{H}$. The eigenstates of $\hat{H}$ are just those of $\hat{S}_z$: $|0\rangle$ (up) and $|1\rangle$ (down). The corresponding eigenenergies are $E_0 = \frac{eB}{2m}\hbar$, $E_1 = -\frac{eB}{2m}\hbar$.

Next we solve the time-dependent Schrödinger equation. Write

$$|\psi(t=0)\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \ .$$

Then

$$\begin{aligned} |\psi(t)\rangle &= \alpha e^{-i\frac{eB}{2m}t}\,|0\rangle + \beta e^{i\frac{eB}{2m}t}\,|1\rangle \\ &\propto \alpha\,|0\rangle + \beta e^{i\frac{eB}{m}t}\,|1\rangle \ , \end{aligned}$$

where the proportionality is up to a global phase. On the Bloch sphere,

$$|\psi(t=0)\rangle = \cos\tfrac{\theta}{2}\,|0\rangle + \sin\tfrac{\theta}{2}e^{i\varphi}\,|1\rangle$$

evolves to

$$|\psi(t)\rangle = \cos\tfrac{\theta}{2}\,|0\rangle + \sin\tfrac{\theta}{2}e^{i(\varphi+\frac{eB}{m}t)}\,|1\rangle \ .$$

Thus the state rotates counterclockwise around the $z$ axis, at frequency $\omega_0 \equiv \frac{eB}{m}$ ($\omega_0$ is known as the cyclotron frequency, since it is the same frequency with which a classical $e^-$ cycles in a magnetic field, due to the Lorentz force).

Therefore $\hat{R}_z(\Delta\varphi) = e^{-i\frac{\hat{S}_z}{\hbar}\Delta\varphi}$ is a unitary operation which rotates by $\Delta\varphi$ about the $z$ axis. (Proof: $\hat{R}_z(\Delta\varphi)$ is exactly $e^{-i\frac{\hat{H}}{\hbar}t}$ for $t = \Delta\varphi/\omega_0$.) Being unitary means $\hat{R}_z(\Delta\varphi)^\dagger = \hat{R}_z(\Delta\varphi)^{-1} = \hat{R}_z(-\Delta\varphi)$.

Aligning $\vec{B}$ with the $z$ axis rotates the spin about the $z$ axis. Each state is restricted to the line of latitude it starts on, as illustrated above. For a more general rotation about a different axis, simply point the $\vec{B}$ field in a different direction. For example, the unitary operator

$$\hat{R}_n(\Delta\gamma) = e^{-i\frac{\hat{\vec{S}}\cdot\hat{n}}{\hbar}\Delta\gamma}$$

rotates by $\Delta\gamma$ about the axis $\hat{n}$. To achieve this unitary transformation, set $\vec{B} = B\hat{n}$ for exactly time $t = \Delta\gamma/\omega_0$.

Any unitary transformation on a single qubit, up to a global phase, is a rotation on the Bloch sphere about some axis; mathematically, this is the well-known isomorphism $SU(2)/\pm 1 \cong SO(3)$ between $2\times 2$ unitary matrices up to phase and $3\times 3$ real rotation matrices. Hence Larmor precession, or spin rotation, allows us to achieve any single qubit unitary gate. While theoretically simple, Larmor precession can unfortunately be inconvenient in real life, mostly because of the high frequencies involved and the susceptibility to noise. A more practical method for achieving rotations on the Bloch sphere is spin resonance, which we will describe next.

## 11.2 Spin Resonance

How do we control qubit states in the lab? If $|psi(t)\rangle = \alpha(t)\,|0\rangle + \beta(t)\,|1\rangle$, how do we deterministically change $\alpha$ and $\beta$?

We know that the Hamiltonian evolves things in time, so if we turn on a field then the Hamiltonian will evolve the state via $e^{-i\hat{H}t/\hbar}$.

For a static magnetic field this allows us to rotate qubit state from one point on the Bloch sphere to another via rotations:

$$\hat{R}_i(\Delta\theta) = e^{-i\hat{S}_i\Delta\theta/\hbar}, \Delta\theta = \frac{eB_o}{m}\Delta t, \vec{B} = B_o\hat{x}_i$$

*Question:* How can we maintain energy level splitting between $|0\rangle$ and $|1\rangle$ and *control* the rate at which a qubit rotates between states? (i.e. change it at a rate different from $\omega_o = \frac{eB_o}{m}$.)

*Answer:* Spin Resonance gives us a new level of control (most clearly seen in NMR).

*How it works:* Turn on a big DC field $B_o$ and a little AC field $\vec{B}$ sin( $\omega_o$ t) that is tuned to the resonance $\omega_o = \frac{eB_o}{m}$:

The small AC field induces controlled mixing between $|0\rangle$ and $|1\rangle$... "SPIN FLIPS".

We must solve the Schrodinger equation to understand what is going on:

$$i\hbar\frac{\partial}{\partial t}\,|\psi(t)\rangle = \hat{H}\,|\psi(t)\rangle$$

It is convenient to use column vector notation:

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

What's the Hamiltonian? $\hat{H} = -\vec{\mu}\cdot\vec{B} = \frac{e}{m}\vec{S}\cdot\vec{B}$

We now let the magnetic field be composed of the large bias field and a small oscillating transverse field:

$$\vec{B} = B_o\hat{z} + B_1 cos\omega_o t\hat{x}$$

With this we obtain the Hamiltonian:

$$\hat{H} = \frac{e}{m}B_o\hat{S}_z + \frac{e}{m}B_1 cos\omega_o t\hat{S}_x$$

Now use 2×2 matrix formulation, where the Pauli matrices ($\hat{S}_z = \frac{\hbar}{2}\sigma_z$, etc.) are of course eminently useful:

$$\hat{H} = \frac{e}{m}B_o \cdot \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{e}{m}B_1 cos\omega_o t \cdot \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The two terms sum to give the following $2 \times 2$ Hamiltonian matrix (expressed in the $\hat{S}_z$ basis):

$$\hat{H} = \frac{e\hbar}{2m}\begin{pmatrix} B_o & B_1 cos\omega_o t \\ B_1 cos\omega_o t & -B_o \end{pmatrix}$$

Now we can plug this Hamiltonian into the Schr. equation and solve for $|psi\rangle$.

*A bit of intuition on QM:* If you construct a Hamiltonian matrix out of some basis, then the matrix element $H_{ij}$ tells us how much application of the Hamiltonian tends to send a particle from state $|j\rangle$ to state $|i\rangle$. (The units are of course energy $\Rightarrow$ rate of transitions $\propto$ frequency $\propto \frac{E}{\hbar} \propto \frac{H_{ij}}{\hbar}$.)

So, if we only had $\vec{B} = B_o\hat{z}$ and $\vec{B}_1 = 0$, then what would the rate of spin flip transitions be?

$$rate_{i \leftarrow j} \propto \langle i| \hat{H} |j\rangle = |1\rangle \hat{H} |0\rangle = H_{21} = 0!$$

So, we can conclude that we NEED to have a field perpendicular to the large bias field $\vec{B} = B_o\hat{z}$ to induce "spin flips" or to mix up $|0\rangle$ and $|1\rangle$ states in $|\psi\rangle$. This is perhaps more obvious in case of spin, but not as obvious for other systems. It is important to develop our quantum mechanical intuition which can easily get lost in the math!

Now let's solve the Schr. equation for Spin Resonance.

$$\hat{H} |\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} = \frac{e\hbar}{2m}\begin{pmatrix} B_o & B_1 cos\omega_o t \\ B_1 cos\omega_o t & -B_o \end{pmatrix}\begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix}$$

We get two coupled differential equations. First, we define $\omega_o = \frac{eB_o}{m}$ and $\omega_1 = \frac{eB_1}{2m}$, where the latter quantity is defined with a seemingly annoying factor of 1/2. It'll make sense later, though.

$$i\frac{\partial\alpha(t)}{\partial t} = \frac{\omega_o}{2}\alpha(t) + \omega_1 cos(\omega_o t)\beta(t)$$

$$i\frac{\partial\beta(t)}{\partial t} = \omega_1 cos(\omega_o t)\alpha(t) - \frac{\omega_o}{2}\beta(t)$$

To solve we make a substitution. This may seem weird, but it involves the recognition that the system has a natural rotating frame in which the system should be viewed.

$$a(t) = \alpha(t)e^{i\omega t/2}$$

$$b(t) = \alpha(t)e^{-i\omega t/2}$$

Now we're going to use a dubious approximation, but it involves a recognition that $\omega_o$ is much larger than $\omega_1$ and these fast rotations average to zero on the timescales $1/\omega_1$ (which are the relevant experimental timescales). Anyway, here's the dubious approximation:

$$cos(\omega_o t)e^{i\omega_o t} \approx \frac{1}{2}$$

Using these definitions and dubious approximations and we obtain the following differential equation for a(t) (and correspondingly b(t)):

$$\frac{\partial^2 a(t)}{\partial t^2} + \frac{\omega_1^2}{4}a(t) = 0$$

This is a familiar second order differential equation. Our initial conditions have yet to be specified, but let's say $\alpha(0) = \beta(0) = 0$. This gives the following solution:

$$\begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} = \begin{pmatrix} e^{-i\frac{\omega_o}{2}t}cos\frac{\omega_1}{2}t \\ -e^{+i\frac{\omega_o}{2}t}sin\frac{\omega_1}{2}t \end{pmatrix}$$

What does this mean geometrically? Let's go to the Bloch sphere! Our generalized Bloch vector looks like:

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

Our time-dependent state which is a solution to the Schr. equation looks like:

$$|\psi(t)\rangle = cos\frac{\omega_1 t}{2}|0\rangle + e^{i(\omega_o+\pi)}sin\frac{\omega_1 t}{2}|1\rangle$$

Geometrically we can say that $\phi = \omega_o t + \pi$, so we conclude that the qubit is spinning around $\hat{z}$ at a rate $\omega_o$.

What about $\theta$? $\theta = \omega_1 t$, so we're crawling up the sphere at a rate $\omega_1 = \frac{eB_1}{m}$ at the same time we're spinning rapidly about $\hat{z}$ at the fast $\omega_o$, the *Larmor frequency.* We can control $\omega_1$ precisely by changing the amplitude of $B_1$.

Even though $\omega_0$ is very large, $\omega_1$ can be very small. If we're really good, we can flip spins by applying a "$\pi$-pulse": $\omega_1 \Delta t = \pi$.

*Note:* As spins flip out of ground state they suck energy out of the "RF field" ($B_1 cos\omega_o$). This is easily detected and forms the basis of NMR.