

PHY256 Lecture notes: Introducing Quantum Computing

A. C. Quillen

February 22, 2021

Contents

1	Basis vectors and quantum states	2
2	The qubit	4
2.1	The Bloch Sphere	5
2.2	Single qubit Gates	8
3	Measurements	9
3.1	The Quantum Zeno effect	11
4	Product spaces and 2 qubits	12
4.1	Controlled NOT gate	14
4.2	Other controlled 2 qubit operators	15
4.3	Partial measurement	16
4.4	No cloning Theorem	16
4.5	Thoughts on the EPR paradox	17
4.6	Interpretations	17
4.6.1	Copenhagen interpretation	17
5	Teleportation	18
5.1	The GHZ state	19
6	Universal Gate Sets	20
6.1	3bit gates, the Toffoli gate	20
7	Deutsch Jozca algorithm	21
7.1	Functions and quantum parallelism	23
8	Quantum Random Walks	24

9 Quantum Fourier Transform and Shor algorithm	27
9.1 Discrete Fourier Transforms	27
9.2 Quantum Fourier transforms	28
9.3 Outline of Shor Algorithm	29
10 What's a quantum Computer?	29
10.1 Quantum Compiler	29
10.2 Errors and Error Correction	30
10.3 Entropy and information loss	30
11 The Density matrix	30
11.1 Density operator for a pure state	30
11.2 Mixtures	31
11.3 POVM	32
11.4 Transformations of the density matrix	33

1 Basis vectors and quantum states

A Hilbert space is a complex vector space that has an inner product. We can describe a quantum state $\langle\psi|$ or $|\psi\rangle$ with a vector of unit length in a Hilbert space. With an orthonormal basis for the Hilbert space with basis vectors $\langle a_i|$, (satisfying $\langle a_i|a_j\rangle = \delta_{ij}$) we can write

$$|\psi\rangle = \sum_i c_i |a_i\rangle$$

$$\langle\psi| = \sum_i \langle a_i| c_i^*.$$

Here the c_i are complex numbers. The wave vector or quantum state is normalized;

$$\langle\psi|\psi\rangle = \sum_i c_i c_i^* = 1.$$

The number $c_i c_i^*$ is associated with the probability to be in state $|a_i\rangle$. Quantum mechanics postulates that the world is described with probabilities of knowing something rather than absolutely knowing the position or state an object is in.

The ket or $|\psi\rangle = \sum_i c_i |a_i\rangle$ can also be written as a vector

$$|\psi\rangle = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{N-1} \end{pmatrix}$$

where N is the dimension of the Hilbert space. The bra $\langle\psi|$ can also be written as vector

$$|\psi\rangle = \begin{pmatrix} c_0^* & c_1^* & c_2^* & \dots c_{N-1}^* \end{pmatrix}$$

An operator on $|\psi\rangle$ can be written as a matrix.

A Hamiltonian matrix is a Hermitian or self-adjoint matrix. A Hermitian matrix satisfies $A^\dagger = A$ where the dagger means taking both the transpose and complex conjugate of the matrix. Observables or measurements are also represented by Hermitian matrices. Hamiltonian evolution gives unitary evolution of the wave vector. However the Hamiltonian is not a unitary matrix. A unitary matrix $\mathbf{U}$ satisfies $\mathbf{U}\mathbf{U}^\dagger = \mathbf{I}$ with $\mathbf{I}$ the identity. A Unitary transformation preserves the norm of the wave or state vector ψ . In other words, with $\langle\psi|\psi\rangle = 1$ and $|\psi'\rangle = \mathbf{U}|\psi\rangle$ and $\mathbf{U}$ unitary, then $\langle\psi'|\psi'\rangle = 1$.

Hamiltonian evolution is done using the exponential of the Hamiltonian via

$$|\psi(t)\rangle = e^{-iHt/\hbar} |\psi(t=0)\rangle$$

as this is consistent with Schroedinger's equation

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = H |\psi\rangle$$

In a particular basis, the coefficients of $|\psi\rangle$, (c_i) are directly related to probabilities, with the probability of being in state $|a_i\rangle$ equal to $p_i = c_i c_i^*$. Unitary evolution is not random but deterministic. The coefficients are not randomly varied but updated using a Hamiltonian and evolved using unitary evolution.

The expectation value of an observable $\mathbf{A}$

$$\langle\mathbf{A}\rangle = \langle\psi|\mathbf{A}|\psi\rangle$$

In the basis $|a_i\rangle$, the matrix $\mathbf{A}$ has coefficients

$$A_{ij} = \langle a_i | \mathbf{A} | a_j \rangle$$

and

$$\mathbf{A} = \sum A_{ij} |a_i\rangle \langle a_j|$$

For state $|\psi\rangle = a|0\rangle + b|1\rangle$, aa^* can be interpreted as the probability that the state is in the $|0\rangle$ state and bb^* the probability that the state is in the $|1\rangle$ state. Expectation values are mean values of a system that is measured many times.

We can write a matrix in the form $\sum A_{ij} |a_i\rangle \langle a_j|$ as this gives a linear map taking a vector to another vector in our Hilbert space.

The probability p_i , that a wave vector ψ is in state $|a_i\rangle$ can be computed with the projection operator

$$\mathbf{P}_i = |a_i\rangle \langle a_i|$$

with $\psi = \sum_i c_i |a_i\rangle$,

$$\begin{aligned}
\langle\psi|\mathbf{P}_k|\psi\rangle &= \sum_{jk} \langle a_j| c_j^* \mathbf{P}_{i,jk} c_k |a_k\rangle \\
&= \sum_{ij} \langle a_j| c_j^* |a_i\rangle \langle a_i| c_k |a_k\rangle \\
&= \sum_{ij} c_j^* \delta_{kj} \delta_{ki} c_k \\
&= c_i^* c_i = p_i
\end{aligned}$$

2 The qubit

A two state quantum system can be described with wave-vector

$$\begin{aligned}
|\psi\rangle &= a |0\rangle + b |1\rangle \\
\langle\psi| &= \langle 0| a^* + \langle 1| b^*
\end{aligned}$$

where $aa^* + bb^* = 1$ and $|0\rangle$ and $|1\rangle$ are the two energy states. You can also think of a particle with spin up and spin down as a two energy state object. Here we take $|0\rangle = |\uparrow\rangle$ and $|1\rangle = |\downarrow\rangle$. This is a building block for many quantum computer designs and is called a qubit or q-bit.

In vector notation the first state and second states are

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

and

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix}. \quad (1)$$

Coefficients a, b are both complex numbers, but the vector is normalized so that $aa^* + bb^* = 1$.

A state ψ is described by two complex numbers a, b . Each one has a real and a complex part. This gives a four dimensional space. As $\langle\psi|\psi\rangle = 1$, the sum

$$\text{Re}(a)^2 + \text{Im}(a)^2 + \text{Re}(b)^2 + \text{Im}(b)^2 = 1$$

restricting the space to a 3d spherical surface in this 4d space.

2.1 The Bloch Sphere

It is convient to define

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ |-\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\ |i\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \\ |-i\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \end{aligned}$$

These four and with $|0\rangle, |1\rangle$ are 6 equidistant poles on the Bloch sphere.

Global phase is unobservable whereas *relative phase* is observable. Global means we can multiply the entire state by $e^{i\alpha}$ and we not see any difference in our measurements.

The two states $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ have different *relative* phases. We could rotate the states and then measure their spin. We would measure different spin values in these rotated states. Relative phase is observable.

By adjusting the global phase we can take a wave vector

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

and write it as

$$|\psi'\rangle = a'|0\rangle + b'e^{i\phi}|1\rangle$$

with a', b' both real. We do this by multiplying by $a^*/|a|$.

The map onto the Bloch sphere

$$|\psi\rangle \xrightarrow{\pi} |\psi'\rangle$$

$$\begin{aligned} a' &= |a| \\ b' &= \frac{|ba^*|}{|a|} \\ \phi &= \arctan2(\text{Im}(ba^*), \text{Re}(ba^*)) \end{aligned}$$

The map is a projection as $\pi^2 = \pi$. In other words $\pi(\pi(|\psi\rangle)) = \pi(|\psi\rangle)$.

The wave vector has a norm of 1 so we we can find an angle θ with

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

The factor of 2 within the cosine and sine lets us associate θ with a co-latitude on a sphere. We can describe the wave vector as a point on the sphere, or with unit vector

$$(x, y, z) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta). \quad (3)$$

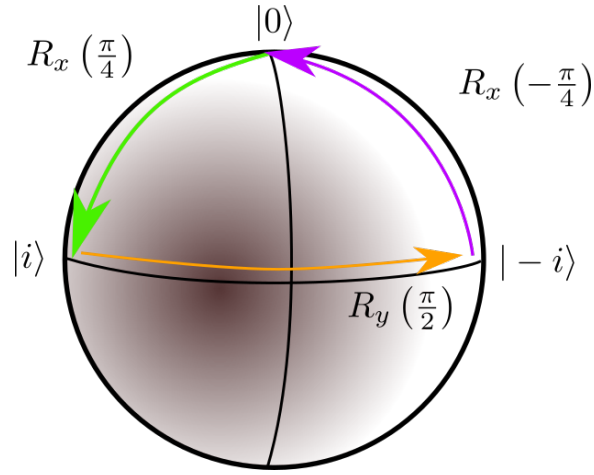


Figure 1: A path on the Bloch sphere.

This is known as the Bloch sphere. The projection π to the Bloch sphere can also be written as a map from complex numbers a, b to angles θ, ϕ ,

$$\begin{aligned}\frac{\theta}{2} &= \arctan2(b', a') = \arctan2(|b|, |a|) \\ \phi &= \arctan2(\text{Im}(ba^*), \text{Re}(ba^*)).\end{aligned}$$

A general qubit state $|\psi\rangle$ depends on two complex numbers, giving 4 degrees of freedom, when counting each real and complex part of the two complex numbers. However $\langle\psi|\psi\rangle = 1$ reduces the degrees of freedom by 1. There is a redundant phase. If we drop it by projecting onto the Bloch sphere, then this reduces the dimension again. This is why the Bloch sphere is a 2-d object, and is a sphere in 3d rather than a 3-sphere embedded in 4-dimensions, like the qubit prior to projection onto the Bloch sphere.

Points 180° apart on the Bloch sphere are orthogonal states. We have given three pairs of them, the spin up and down states, $|0\rangle, |1\rangle$, the states $|+\rangle, |-\rangle$, and the complex states $|i\rangle, |-i\rangle$. We could also rotate into another basis and find a different set of 6 extreme points. The spin up and down states are eigenvectors of σ_z .

Question: What are the θ, ϕ angles and x, y, z coordinates for the 6 states on the Bloch sphere?

Answer: Here is a table:

State	θ	ϕ	(x, y, z)
$ 0\rangle$	0	-	(0,0,1)
$ 1\rangle$	π	0	(0,0,-1)
$ +\rangle$	$\pi/2$	0	(1,0,0)
$ -\rangle$	$-\pi/2$	0	(-1,0,0)
$ i\rangle$	$\pi/2$	$\pi/2$	(0,1,0)
$ -i\rangle$	$\pi/2$	$-\pi/2$	(0,-1,0)

Question: How do the Pauli spin matrices, $\sigma_x, \sigma_y, \sigma_z$ operate on the Bloch states?

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

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

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

Answer: here is a table:

$ \psi\rangle$	$\sigma_x \psi\rangle$	$\sigma_y \psi\rangle$	$\sigma_z \psi\rangle$
$ 0\rangle$	$ 1\rangle$	$i 1\rangle$	$ 0\rangle$
$ 1\rangle$	$ 0\rangle$	$-i 0\rangle$	$- 1\rangle$
$ +\rangle$	$ +\rangle$	$-i -\rangle$	$ -\rangle$
$ -\rangle$	$- -\rangle$	$i +\rangle$	$ +\rangle$
$ i\rangle$	$i -i\rangle$	$ i\rangle$	$ -i\rangle$
$ -i\rangle$	$-i i\rangle$	$- -i\rangle$	$ i\rangle$

Question: Find a matrix for which the complex states $|i\rangle$ and $|-i\rangle$ are eigenvectors.
The answer is the Pauli-Y or σ_y .

Question: Find a matrix for which the states $|+\rangle$ and $|-\rangle$ are eigenvectors.
Answer is the Pauli-X or σ_x .

Question: Is there a way to figure this out without guessing?

Answer: Yes. Make a unitary transformation U to transfer from $|0\rangle, |1\rangle$ basis to $|+\rangle, |-\rangle$. This can be done with the states themselves as these states are orthonormal and rows of unitary matrices are orthonormal. Then compute $U\sigma_z U^\dagger$.

2.2 Single qubit Gates

Quantum gates are unitary transformations. I give some examples of common gates that operate on a single qubit.

Hadamard	$\text{---}\boxed{\text{H}}\text{---}$	$\mathbf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$
NOT or Pauli-X	$\text{---}\boxed{\text{X}}\text{---}$	$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
Pauli-Y	$\text{---}\boxed{\text{Y}}\text{---}$	$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$
Pauli-Z	$\text{---}\boxed{\text{Z}}\text{---}$	$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
Phase = $\sqrt{Z}$	$\text{---}\boxed{\text{S}}\text{---}$	$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$
$\pi/8 = \sqrt{S}$	$\text{---}\boxed{\frac{\pi}{8}}\text{---}$	$\frac{\pi}{8} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$

I have also include the quantum circuit symbols. The Hadamard gate takes a $|0\rangle$ and returns a $|+\rangle$ and takes a $|1\rangle$ and returns a $|-\rangle$. The Hadamard gate obeys $\mathbf{H}^2 = \mathbf{I}$. The NOT interchanges the $|0\rangle, |1\rangle$ states. The $\pi/8$ is the square root of the Phase Gate which itself is the square root of the Pauli-Z gate.

Unitary matrices satisfy $\mathbf{U}\mathbf{U}^\dagger = \mathbf{I}$. The rows and columns of a unitary matrix are orthonormal. These gates are unitary matrices.

The exponentials of the Pauli matrices are sometimes described like rotations $R_x(\alpha)$, $R_y(\alpha)$, $R_z(\alpha)$,

$$R_x(\alpha) = e^{i\alpha\sigma_x} = \cos \alpha \mathbf{I} + i \sin \alpha \sigma_x = \begin{pmatrix} \cos \alpha & i \sin \alpha \\ i \sin \alpha & \cos \alpha \end{pmatrix} \quad (4)$$

$$R_y(\alpha) = e^{i\alpha\sigma_y} = \cos \alpha \mathbf{I} + i \sin \alpha \sigma_y = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \quad (5)$$

$$R_z(\alpha) = e^{i\alpha\sigma_z} = \cos \alpha \mathbf{I} + i \sin \alpha \sigma_z = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \quad (6)$$

and they are functions of an angle α . They can be considered rotations in $U(2)$ that are generated exponentially from infinitesimal operators, the Pauli spin matrices.

The matrices $R_x(\alpha)$, $R_y(\alpha)$, $R_z(\alpha)$ are unitary transformations.

Some properties of *unitary transformations*:

- Reversible. They have an inverse.
- Deterministic. No random choices are required.

- They can be applied in a time continuous way with an exponential matrix.
- They preserve the wave vector norm $\langle\psi|\psi\rangle = 1$.

A unitary matrix has columns that are orthonormal and rows that are orthonormal. The determinant of a unitary matrix $|\det(U)| = 1$. A unitary matrix can be written as $U = e^{iH}$ where H is a Hermitian matrix. A Hermitian matrix is equal to the complex conjugate of its transpose, it is diagonalizable and it has real eigenvalues.

3 Measurements

Measurements are

- Irreversible.
- Information is lost.
- Probabilistic.
- Involve collapse of wave-function. They can be discontinuous. (Though not necessarily in the setting of weak measurements).

Measurements can be described as a projection of the wave-function with projection operator chosen using a probability.

In contrast, unitary transformations always preserve the norm, are reversible, and involve no random choices.

Measurements involve a sum of probabilities that equal 1.

The Pauli spin matrix

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

The expectation

$$\langle\sigma_z\rangle = \langle\psi|\sigma_z|\psi\rangle$$

gives us the expectation value of the spin value with a measured 1 being spin up and a measured -1 being spin down. (For a spin 1/2 particle the spin operator is $\mathbf{J} = \hbar\boldsymbol{\sigma}/2$ so a measurement of spin in the z direction actually gives $\pm\hbar/2$). With $|\psi\rangle = |0\rangle$ the spin is up and with $|\psi\rangle = |1\rangle$ the spin is down. With $|\psi\rangle = a|0\rangle + b|1\rangle$, the probability of measuring spin up is aa^* and the probability of measure spin down is $bb^* = 1 - aa^*$. After measuring a spin up, the wave vector collapses and becomes a complex number of magnitude 1 times $|0\rangle$. After measuring a spin down, the wave vector collapses and becomes $|1\rangle$ times a complex number of magnitude 1. Is the phase important? If you are later on carrying out an interference experiment, the phase could be important.

We can collapse the wave function with two projection operators

$$\mathbf{P}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \mathbf{P}_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

The projection operators can also be written

$$\mathbf{P}_0 = |0\rangle \langle 0| \quad \mathbf{P}_1 = |1\rangle \langle 1|$$

The projection operators satisfy $\mathbf{P}_0^2 = \mathbf{P}_0$, $\mathbf{P}_1^2 = \mathbf{P}_1$. After a measurement giving spin up, the wave-vector must be normalized

$$|\psi'\rangle = \frac{\mathbf{P}_0 |\psi\rangle}{\langle \psi | \mathbf{P}_0 | \psi \rangle}$$

The probability that a wave vector $|\psi\rangle$ is in the spin up state is $\langle \psi | \mathbf{P}_0 | \psi \rangle$ and that it is in the spin down state is $\langle \psi | \mathbf{P}_1 | \psi \rangle$.

To simulate measurement of a single state, you need a set of projection operators, each associated with each possible measurement value. To simulate a measurement, project the state with each projection operator. For example if $|\psi\rangle = a|0\rangle + b|1\rangle$ then

$$\begin{aligned} P_0 |\psi\rangle &= a|0\rangle \\ P_1 |\psi\rangle &= b|1\rangle \end{aligned}$$

The probability of getting spin up is aa^* and the probability of getting spin down is bb^* . Now you would choose a possible measurement value randomly, using these two probabilities. After you know what state resulted from the measurement you then would normalize the state vector. If spin up was measured then the resulting statevector after measurement is $\frac{a}{|a|} |0\rangle$, otherwise it is $\frac{b}{|b|} |1\rangle$. Here I am keeping track of the phase. If you don't want to keep track of the phase then after measurement you get $|0\rangle$ or $|1\rangle$.

Simulation of measurement on a single state involves making a random choice. A series of measurements would then be a series of random choices each giving you a new statevector. Mimicking a series of measurements is a Markov Chain Monte Carlo model.

If you have an experiment that you run many times, then the expectation value of the spin would be

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

with 1 corresponding to spin up and -1 corresponding measuring spin down. This expectation value would be the average over many possible measurements of the state vector $|\psi\rangle$.

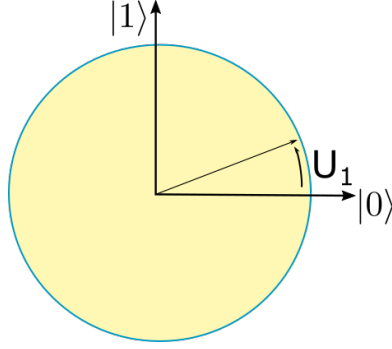


Figure 2: A qubit drifts in the counter clockwise due to continuous unitary evolution. Measurement by σ_z projects the state either into $|0\rangle$ or $|1\rangle$. If the state originates at $|0\rangle$ repeated measurements tend to keep it near $|0\rangle$. This is called the Quantum Zeno effect.

3.1 The Quantum Zeno effect

The **Quantum Zeno** effect describes what happens if you measure a state over and over again. It is nearly frozen into a single measured state. With a slowly drifting system, the process of repeated measurement and collapse keeps the probability low that the system can evolve into a different state and that a subsequent measurement will give a different value.

Consider a single qubit that is initially in the spin up state $|\psi\rangle = |0\rangle$. We operate on it with a smoothly varying unitary operation $U(t) = R_x(\beta t) = e^{i\beta\sigma_x t}$ where t is time. We consider the state at time intervals of $dt = \frac{2\pi}{N\beta}$. At time interval jdt , the unitary transformation $U_j = e^{\frac{i2\pi j\sigma_x}{N}}$ giving rotation angle $\alpha = \frac{2\pi j}{N}$. We evaluate this matrix using equation 4 to find

$$U_j = \begin{pmatrix} \cos \frac{2\pi j}{N} & i \sin \frac{2\pi j}{N} \\ i \sin \frac{2\pi j}{N} & \cos \frac{2\pi j}{N} \end{pmatrix} \quad (7)$$

With a single time interval dt the unitary transformation is

$$U_1 = \begin{pmatrix} \cos \frac{2\pi}{N} & i \sin \frac{2\pi}{N} \\ i \sin \frac{2\pi}{N} & \cos \frac{2\pi}{N} \end{pmatrix} \quad (8)$$

We evolve with U_1 then measure the state then evolve it again and so on alternating between continuous unitary evolution and measurement.

After evolving via U_1 the wave vector becomes

$$|\psi\rangle = \cos \frac{2\pi}{N} |0\rangle + i \sin \frac{2\pi}{N} |1\rangle$$

We measure it with σ_z . The probability that the state has a spin up is $\cos^2 \frac{2\pi}{N} = 1 - \sin^2 \frac{2\pi}{N}$ and that the state has spin down is $\sin^2 \frac{2\pi}{N}$. Let $\sin \frac{2\pi}{N} = \epsilon$ with ϵ small which is equivalent to N large. The probability that a spin up is measured is $1 - \epsilon^2$.

The probability that a spin up is measured M times in a series of M continuous unitary evolution and evenly spaced measurements by σ_z is

$$P_{UM} = (1 - \epsilon^2)^M \sim 1 - M\epsilon^2.$$

The probability that a spin up is measured after M unitary evolutions by U_1 and measurements by σ_z is actually higher than this since one of the intermediate measurements could have been spin down.

Suppose instead the system is allowed to evolve without measurement during the M time intervals. The wave-vector becomes

$$|\psi\rangle = \cos \frac{2\pi M}{N} |0\rangle + i \sin \frac{2\pi M}{N} |1\rangle$$

The probability that spin up is measured after Mdt is

$$P_U \sim \cos^2 \frac{2\pi M}{N}.$$

Suppose we chose $M = N/4$. This gives $P_U = 0$ because the state evolves to $|1\rangle$. In contrast

$$P_{UM} \sim 1 - M \frac{4\pi^2}{N^2} \sim 1 - \frac{\pi^2}{N}.$$

The probability can be high that the spin up is measured after the evenly spaced M measurements.

4 Product spaces and 2 qubits

With two qubits we have a Hilbert space that is a product of two Hilbert spaces.

A state looks like this

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

with four complex coefficients. Our Hilbert space has 4 elements in its basis and must be normalized so that $|a_{00}|^2 + |a_{01}|^2 + |a_{10}|^2 + |a_{11}|^2 = 1$.

We could also write $|00\rangle$ as $|0\rangle |0\rangle$ making it clearer that our Hilbert space is a product of two complex vectors spaces. The product space contains elements like $a_{ij} |i\rangle |j\rangle$ or $a_{ij} |ij\rangle$.

A single qubit looks like $a|0\rangle + b|1\rangle$. We can make a tensor product with two single qubits (each in their own 2d Hilbert space) with

$$\begin{aligned} (a|0\rangle + b|1\rangle) \otimes (c|0\rangle + d|1\rangle) &= ac|0\rangle|0\rangle + bd|1\rangle|1\rangle + ad|0\rangle|1\rangle + bc|1\rangle|0\rangle \\ &= ac|00\rangle + bd|11\rangle + ad|01\rangle + bc|10\rangle \end{aligned}$$

Not all states in the full 4d Hilbert space can be written as a tensor product. For example $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ gives $ac = 1/\sqrt{2} = bd$ and $ad = bc = 0$ that has no solution. We

can call such states *entangled*. The state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is called a **Bell pair** or EPR pair state.

We consider a state-vector $|\psi\rangle = \sum_{ij} a_{ij} |i\rangle |j\rangle$. There might be a basis in which we can write the state-vector as

$$|\psi\rangle = \sum_{ij} a_{ij} |i\rangle |j\rangle = \sum_{ij} a_i |i\rangle a_j |j\rangle$$

If there is no-such basis, then the state is described as **entangled**.

A qutrit is a three state system with $|0\rangle, |1\rangle, |2\rangle$. We could make a Hilbert space that is a product of a qubit and a qutrit.

For the product of two qubits, we can order matrices and vectors using the order above so

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (9)$$

The identity for the 2 qubit system can be written as $\mathbf{I} \otimes \mathbf{I}$. Many operators can be written as direct products, like $\mathbf{A} \otimes \mathbf{I}$ where $\mathbf{I}$ is the identity for a single qubit and $\mathbf{A}$ is an operator for a single qubit.

We can construct operators (or gates) in the full space that are products of gates in the subspaces. For example $\mathbf{H} \otimes \mathbf{H}$ where $\mathbf{H}$ is the Hadamard gate or $\mathbf{H} \otimes \mathbf{I}$.

$$\mathbf{I} \otimes \mathbf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

$$\mathbf{H} \otimes \mathbf{I} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

The product of $\mathbf{I} \otimes \mathbf{H}$ and $\mathbf{H} \otimes \mathbf{I}$ is

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

It may help to know that unitary matrices have rows and columns that have norm 1 and are orthogonal. It is also useful to check that the conjugate transpose of the matrix times itself gives the identity.

4.1 Controlled NOT gate

An interesting new gate that operates on 2 qubits is the **Controlled NOT** or **cNOT** where the second bit is flipped only if the first bit is 1.

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

With basis as defined in equation 9.

$$\text{cNOT} |00\rangle = |00\rangle$$

$$\text{cNOT} |01\rangle = |01\rangle$$

$$\text{cNOT} |10\rangle = |11\rangle$$

$$\text{cNOT} |11\rangle = |10\rangle$$

The cNOT *cannot* be written as a tensor product.

The cNOT can also be written as

$$\text{cNOT} |xy\rangle = |x, x + y\rangle$$

where the sum $x + y$ is mod 2 and x, y are either 0 or 1. Treating the states as classical bits, $x + y \text{ mod } 2$ is also **XOR** applied to bit x, y , which gives 0 if both bits are the same but 1 otherwise.

I can start with a state that is a tensor product and apply the cNOT

$$\begin{aligned} \text{cNOT} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle &= \text{cNOT} \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) \\ &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \end{aligned}$$

The result is an entangled state (as we showed above that this state could not be written as a tensor product). So the cNOT takes a tensor product state, that is not entangled, and turns it into an entangled state.

The cNOT is said to induce **correlations**. With $|+0\rangle$ the original state, there is 50% chance for the first qubit to be 0, 50% chance of it being 1 and 100% chance that the second qubit is 0. The final state has a 50% chance of both qubits being zero and 50% chance of both qubits being 1.

Example: Starting from $|00\rangle$ construct a Bell pair state using simple gates. See Figure 3 for an illustration.

Here's how we do it.

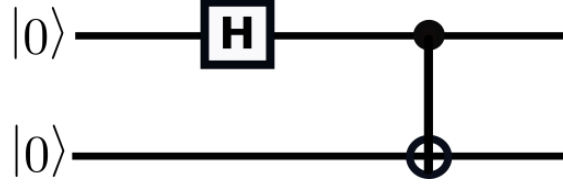


Figure 3: The first quantum operation is a Hadamard on the first qubit. This corresponds to the unitary transformation $\mathbf{H} \otimes \mathbf{I}$. The second operation is the cNOT but with first bit as control and second bit as target. The cNOT looks like a plus on the second qubit because the cNOT can be written as $|x, x + y\rangle$. Starting with state $|00\rangle$ on the left, the result, on the right is the Bell pair state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$.

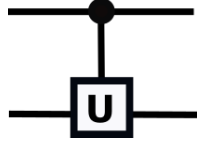


Figure 4: A controlled gate. The control bit is the top one and if it is 1 then the gate U operates on the second bit

- Apply $\mathbf{H} \otimes \mathbf{I}$, the Hadamard gate to the first qubit.

$$\mathbf{H} \otimes \mathbf{I} |00\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle)$$

- Then apply the cNOT where the control bit is the first qubit and the 2nd qubit is the target

$$\text{cNOT} \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Quantum circuit notation for the creating the Bell pair state is shown in Figure 3

4.2 Other controlled 2 qubit operators

We can consider a 2 qubit gate with the first bit a control bit and the second bit a target bit. Instead of flipping the target bit if the first bit is 1, we can execute a gate on the target bit if the first bit is 1 and not change it if the first bit is 0. Any one bit gate can be controlled. We write the controlled cNOT gate as $\Lambda(X)$, where X is the NOT gate and equivalent to the Pauli-X gate. We can use the Λ symbol to denote other types of controlled gates. For example a controlled phase gate $\Lambda(\mathbf{S})$ gate operates on the second bit with the phase gate if the first bit is 1. Recall that the phase gate looks like

$$\mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

The controlled phase gate

$$\Lambda(\mathbf{S}) = \mathbf{P}_0 \otimes \mathbf{I} + \mathbf{P}_1 \otimes \mathbf{S} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & i \end{pmatrix}$$

where $\mathbf{P}_0 = |0\rangle\langle 0|$ is the projection operator to $|0\rangle$ for the first qubit and $\mathbf{P}_1 = |1\rangle\langle 1|$ is the projection operator to $|1\rangle$ for the first qubit.

4.3 Partial measurement

Suppose we have

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$$

Question: We measure the first qubit. What is the probability to get a 0 (or spin up)?

Answer: The probability to get 0 in a measurement of the first qubit is $|\alpha|^2 + |\beta|^2$.

Question: We measure a 0. What does the wave vector look like now?

Answer:

$$|\psi'\rangle = \alpha'|00\rangle + \beta'|01\rangle$$

with coefficients

$$\alpha' = \frac{\alpha}{|\alpha|^2 + |\beta|^2}$$

$$\beta' = \frac{\beta}{|\alpha|^2 + |\beta|^2}$$

What are the projection operators associated with measuring the first bit? They look like $\mathbf{P}_0 \otimes \mathbf{I}$ and $\mathbf{P}_1 \otimes \mathbf{I}$ or

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

The projection operators can also be written as tensor products

$$|0\rangle\langle 0| \otimes \mathbf{I} \quad |1\rangle\langle 1| \otimes \mathbf{I}$$

4.4 No cloning Theorem

Cloning means taking an arbitrary state in the first qubit, a specific state in the second one and making the second qubit the same as the first qubit and keeping the state unentangled. In other words finding a transformation

$$(\alpha|0\rangle + \beta|1\rangle) \otimes |0\rangle \rightarrow (\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha|0\rangle + \beta|1\rangle)$$

with a unitary transformation and for **any** α, β .

In matrix form, the transformation would satisfy

$$U \begin{pmatrix} \alpha \\ 0 \\ \beta \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha^2 \\ \alpha\beta \\ \alpha\beta \\ \beta^2 \end{pmatrix}$$

The transformation is non-linear. There is no way to solve for a matrix U that does not depend on α, β . So, it is not possible to do this with a unitary transformation.

You can clone a specific state with a cNOT. But that's a specific state, not any state.

4.5 Thoughts on the EPR paradox

We start with the Bell pair state $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Then the two qubits are spit up and Al (short for Alice or Alan) given one of them and Bobby (short for Bob or Barbara) given the other. Al measures 0 then Bobby must also measure 0. This could be interpreted as transfer of information in an infinitely small period of time. What if both Al and Bobby are given a newspaper. They both can simultaneously know the same information. In both cases the information is not actually transferred instantaneously.

4.6 Interpretations

So far we have divided up our operations on quantum states into two categories

1. Unitary evolution. No body is watching. No information loss.
2. Measurement. The wave-function is collapsed to a single state which is chosen based on a probability described by the wave-function.

This division presents some paradoxes. If physical laws are based on quantum mechanics why can't everything be described via unitary evolution only? When can we approximate coupling between systems as a measurement?

4.6.1 Copenhagen interpretation

It's not all that easy to pin down the Copenhagen interpretation. It is a set of guiding principles.

In the appropriate limit, quantum theory should resemble classical physics and reproduces the classical predictions. Quantum mechanics obeys different rules than classical physics. The results provided by measuring devices are essentially classical. Measurement involves an interaction between the system and a laboratory device and this interaction collapses the wave function. A wave function is a mathematical entity that provides a probability distribution for the outcomes of each possible measurement on a system.

There are numerous other interpretations.

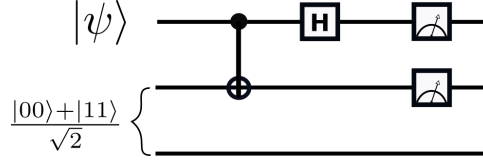


Figure 5: A recipe for teleporting the qubit $|\psi\rangle$ using two entangled qubits, a cNOT, a Hadamard operation and two measurements. The receiver must apply a transformation on the third bit that is based on the measurements of the first two.

5 Teleportation

We describe how an unknown state can be transferred from one qubit to another with a series of measurements and two qubits in an entangle state.

We have three qubits. Two of them are in an entangled Bell pair state. Al has a single qubit $|\psi\rangle = a|0\rangle + b|1\rangle$. She wants to teleport its state to Bobby. Al can measure 1 qubit of the Bell pair and Bobby can measure the other qubit in the Bell pair.

- Al performs a transformation on $|\psi\rangle$ that entangles it with the Bell pair.
- Al measures $|\psi\rangle$ and her qubit that is part of the entangled pair.
- Al tells Bobby what she has measured.
- Bobby performs a transformation on his half of the entangled qubit pair.
- Bobby now holds the information that was in $|\psi\rangle$.
- Alice no longer holds the information that was in $|\psi\rangle$.
- The information that was in $|\psi\rangle$ (the values of a, b) has teleported to Bobby.

We start with total state

$$a|0\rangle + b|1\rangle \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

This can be written as

$$\frac{1}{\sqrt{2}}(a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle)$$

Al performs the cNOT gate on the total state with target the second qubit. The cNOT gate flips the second qubit if the first one is 1. The total state becomes

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

Al performs a Hadamard operation on the first qubit. The total state becomes

$$\frac{1}{2} (a |000\rangle + a |100\rangle + a |011\rangle + a |111\rangle - b |110\rangle + b |010\rangle - b |101\rangle + b |001\rangle)$$

Al now measures the first two bits. She can measure 0,0 with a probability of $1/4(aa^* + bb^*) = 1/4$. If this happens the state becomes

$$a |000\rangle + b |001\rangle$$

Let's make a table

Al measures	00	01	10	11
The state becomes	$a 000\rangle + b 001\rangle$	$a 011\rangle + b 010\rangle$	$a 100\rangle - b 011\rangle$	$a 111\rangle - b 010\rangle$

Al sends her measurements to Bobby who then performs the following transformations on the last qubit.

If the first bit that Al measures is 1, Bob applies

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

This only flips the sign of $|1\rangle$. If the second bit that Al measures is 1, Bob applies

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

This sends $|1\rangle \rightarrow |0\rangle$ and vice versa.

Let's apply these transformations to the above table. If Al measured 00 then the third qubit is $a |0\rangle + b |1\rangle$ and it already is equal to the original state of the first qubit. If Al measured 01 then the states are reversed and again we recover $a |0\rangle + b |1\rangle$. And so on for the last two possibilities.

Note: Al never knew what a, b were. At the end, there is nothing left of interest in the first two bits as both Al and Bob know what they are. The Bell pair is 'used up'.

With more Bell pairs, more qubit states can be teleported. If two qubits are teleported and they are initially entangled, the resulting teleported state will also be entangled.

5.1 The GHZ state

Consider $\frac{1}{\sqrt{2}} (|000\rangle + |111\rangle)$. Each qubit in the entangled state is hosted by a different person. Alice gets the first one, Bobby gets the second and Charlie gets the third.

If Alice has a qubit that is maximally entangled with one that Bob has, then Alice's qubit cannot be maximally entangled with Charlie's.

Let's look at some density matrices.



Figure 6: The circuit notation for the 3 bit Toffoli gate. The third bit is flipped if both the first and second bits are up.

6 Universal Gate Sets

Question: Will we be able to solve any problem on a quantum computer that can't be solved on a classical computer, regardless of resources?

Answer: No. However, the number of states in an n qubit machine is 2^n . This number gets large quickly as n increases. So a few qubits system could be a powerful computer.

Unitary matrices form a continuous manifold that has dimension $2^n \times 2^n$.

It's possible to implement any unitary transformation, using many simple operators or gates. We can ask what is a minimal set for control? Such a set is called a **universal gate set**.

6.1 3bit gates, the Toffoli gate

The Toffoli gate is a controlled-controlled NOT or a ccNOT. It is a 3-qubit gate that flips the third bit if and only if the first two bits are both 1. Another way to write the Toffoli gate is

$$|x, y, z\rangle \rightarrow |x, y, xy + z\rangle$$

The Toffoli gate is special in that it has a product in the output. Controlled two bit gates are affine.

The 3 qubit Toffoli gate can be used to construct a NAND. The third bit is actually the output of a NAND given the inputs of the other two and setting the z bit to 0.

In a classical setting any logical operation involving N bits can be constructed of NAND operations involving 2 bits. Thus NAND is a universal gate set for classical computing. The NAND gate is universal as all other logic computations can be constructed from NANDS. The set (AND, OR) is not universal as it is linear or monotone or affine mod 2. For the same reason NAND is needed for machine learning applications.

For a quantum computer, if we can construct a NAND, then we can use the classical universality of the NAND to show that our quantum computer has a universal gate set.

In the quantum qubit setting cNOT (involving each pair of qubits) along with sets of single qubit gates (and covering the unitary operations for each individual qubit) gives a universal gate set.

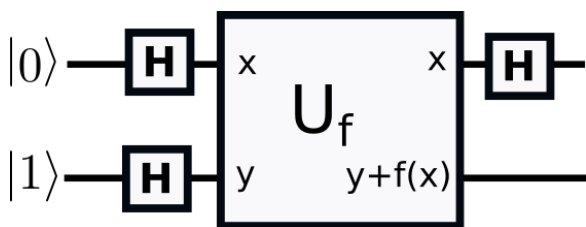


Figure 7: The Deutsch algorithm is the following quantum circuit applied to an initial state of $|01\rangle$. The U_f transformation is $|xy\rangle \rightarrow |x, y + f(x)\rangle$ where the $+$ is mod 2. The goal is to determine if $f(x)$ is constant or balanced with a single use of the function $f()$.

Unitary transformations are reversible. Any classical reversible gate can be used as a quantum gate.

ANDs and ORs are not reversible! We can construct ANDs and ORs using reversible gates but the extra information (that is not lost) goes into other bits that keep the remaining information.

Any set of transformations that acts on a state and only takes that into a subspace of the entire Hilbert space would not be a universal set. The group of unitary transformations does contain subgroups, so if you are trying to determine if your gate set is universal you need to make sure your set does not restrict you to one of these subgroups.

Two bit gates suffice for universal irreversible computation. However three-bit gates are needed for universal reversible computation.

What is the role of interference? The goal, in quantum computing, is to choreograph a pattern of interference so that the wrong answers somehow cancel out but the right ones are reinforced.

Question: is it possible to construct a 3-bit ccNOT (or Toffoli gate) using 2-qubit gates?

Answer: Yes. But it requires at least 5 2-qubit gates. Apparently, the Toffoli gate can be constructed from single qubit gates and six cNOTs.

7 Deutsch Jozsa algorithm

Consider a function $f(x)$ on a single bit. There are 4 possibilities for the function

$$f(x) = 0 \quad f(x) = 1 \quad f(x) = x \quad f(x) = \text{NOT } x$$

The left two are constant functions as they always give the same out bit. The right two possibilities are called balanced functions as they give 0,1 as outputs. Classically you would have to use the function twice to figure out if the function is constant or balanced. However it is possible to design a quantum circuit that uses interference and a single function call to determine whether the function is constant or balanced. The circuit is illustrated in Figure 7. The unitary operation U_f takes $|xy\rangle \rightarrow |x, y + f(x)\rangle$ and is only called one time.

To illustrate this computationally rather than analytically we need to implement the unitary operation U_f . We compute the following for U_f

$$\begin{aligned} f(x) = 0 & \quad |00\rangle \rightarrow |00\rangle, |01\rangle \rightarrow |01\rangle, |10\rangle \rightarrow |10\rangle, |11\rangle \rightarrow |11\rangle \\ f(x) = 1 & \quad |00\rangle \rightarrow |01\rangle, |01\rangle \rightarrow |00\rangle, |10\rangle \rightarrow |11\rangle, |11\rangle \rightarrow |10\rangle \\ f(x) = x & \quad |00\rangle \rightarrow |00\rangle, |01\rangle \rightarrow |01\rangle, |10\rangle \rightarrow |11\rangle, |11\rangle \rightarrow |10\rangle \\ f(x) = \bar{x} & \quad |00\rangle \rightarrow |01\rangle, |01\rangle \rightarrow |00\rangle, |10\rangle \rightarrow |10\rangle, |11\rangle \rightarrow |11\rangle \end{aligned}$$

$$U_{f(x)=0} = \mathbf{I}$$

$$U_{f(x)=1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \mathbf{I} \otimes \sigma_x$$

This one is equivalent to $\mathbf{I} \otimes \sigma_x$.

$$U_{f(x)=x} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \text{cNOT}(c=0, t=1)$$

This one is equivalent to the cNOT gate with control bit the first one and target bit the second one. It is also $|xy\rangle \rightarrow |x, y+x\rangle$.

$$U_{f(x)=\bar{x}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \text{cNOT}(c=0, t=1) * (\mathbf{I} \otimes \sigma_x)$$

This one flips the second bit only if the first bit is 0. It is also $|xy\rangle \rightarrow |x, y + \text{NOT } x\rangle$.

Let's mimic the operation of the circuit. Starting with initial state $|01\rangle$ we apply $\mathbf{H} \otimes \mathbf{H}$

$$\begin{aligned} \mathbf{H} \otimes \mathbf{H} |01\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\ &= \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle) \end{aligned}$$

Now apply $U_f = |xy\rangle \rightarrow |x, y + f(x)\rangle$

$$\begin{aligned} |00\rangle &\rightarrow |0, 0 + f(0)\rangle = |0, f(0)\rangle \\ |01\rangle &\rightarrow |0, 1 + f(0)\rangle = |0, \text{NOT } f(0)\rangle \\ |10\rangle &\rightarrow |1, 0 + f(1)\rangle = |1, f(1)\rangle \\ |11\rangle &\rightarrow |1, 1 + f(1)\rangle = |1, \text{NOT } f(1)\rangle \end{aligned}$$

The result is

$$\mathbf{U}_f \mathbf{H} \otimes \mathbf{H} |01\rangle = \frac{1}{2}(|0, f(0)\rangle - |0, \text{NOT } f(0)\rangle + |1, f(1)\rangle - |1, \text{NOT } f(1)\rangle). \quad (10)$$

Now we apply the Hadamard operation to the first bit. The Hadamard operator takes

$$\begin{aligned} |0, f(0)\rangle &\rightarrow \frac{1}{\sqrt{2}}(|0, f(0)\rangle + |1, f(0)\rangle) \\ |0, \text{NOT } f(0)\rangle &\rightarrow \frac{1}{\sqrt{2}}(|0, \text{NOT } f(0)\rangle + |1, \text{NOT } f(0)\rangle) \\ |1, f(1)\rangle &\rightarrow \frac{1}{\sqrt{2}}(|0, f(1)\rangle - |1, f(1)\rangle) \\ |1, \text{NOT } f(1)\rangle &\rightarrow \frac{1}{\sqrt{2}}(|0, \text{NOT } f(1)\rangle - |1, \text{NOT } f(1)\rangle) \end{aligned}$$

The operation of the Hadamard to the first qubit on the state in equation 10 gives a final or output state

$$\begin{aligned} |\text{final}\rangle &= (\mathbf{H} \otimes \mathbf{I}) * \mathbf{U}_f * (\mathbf{H} \otimes \mathbf{H}) |01\rangle \\ &= \frac{1}{\sqrt{8}}(|0, f(0)\rangle - |0, \text{NOT } f(0)\rangle + |0, f(1)\rangle - |0, \text{NOT } f(1)\rangle + \\ &\quad |1, f(0)\rangle - |1, \text{NOT } f(0)\rangle - |1, f(1)\rangle + |1, \text{NOT } f(1)\rangle) \end{aligned}$$

If $f(0)$ is equal to $f(1)$ then the final state evaluates to

$$|\text{final}\rangle = \frac{1}{\sqrt{2}}(|0, f(0)\rangle - |0, \text{NOT } f(0)\rangle)$$

We notice that the first bit is now exclusively in the 0 state!

If $f(0)$ is opposite to $f(1)$ then the final state is

$$|\text{final}\rangle = \frac{1}{\sqrt{2}}(|1, f(0)\rangle - |1, \text{NOT } f(0)\rangle)$$

We notice that the first bit is now exclusively in the 1 state! Measurement of the first bit can determine whether the function is constant or balanced.

This is the Deutsch algorithm. A more general version with function $f(x)$ that depends on a series of bits is known as the Deutsch-Jozsa algorithm.

7.1 Functions and quantum parallelism

For a classical computer one evaluates a function of a series of bits. For a quantum computer one evaluates a function on a series of qubits that can be in quantum superpositions of states. Calling the function a single time gives what is called ‘quantum parallelism’.

Quantum gates are unitary operations which are invertible. How do we evaluate functions that are not necessarily invertible? How do we write irreversible functions in terms of unitary or invertible operations? Consider a function $f(\mathbf{x})$ of N bits gives 0 or 1.

$$f(\mathbf{x}, y) = (\mathbf{x}, y + f(\mathbf{x}))$$

is invertible. Here the plus is mod 2 or equivalent to the XOR.

8 Quantum Random Walks

A popular way to do a random walk is to start with an initial state vector $|\psi\rangle$ in a product Hilbert space and then repetitively perform a series of unitary transformations. This is the type introduced by Aharonov, Davidovich, and Zagury 1993. Quantum random walks can also be called *quantum cellular automata*. Unlike a classical and stochastic random walk, all transformations are unitary and hence reversible. There is no actual stochasticity. The walk does not lose its recollection of the initial state and it cannot converge to a stationary distribution.

We explore a discrete quantum random walk of a spin on a circle.

Take a state space that is a tensor product of a space with two states (a qubit), $\mathcal{H}_2$, and a space with N states, $\mathcal{H}_N$. A basis for this space is

$$|jn\rangle$$

where $j \in [0, 1]$ is spin up or down and $n \in [0, N - 1]$. The N dimensional Hilbert space can be described as N possible particle positions.

The entire Hilbert space is a tensor product space $\mathcal{H} = \mathcal{H}_2 \otimes \mathcal{H}_N$.

We start with $|\psi\rangle = |00\rangle$ and alternate applying a spin mixing operator

$$\mathbf{H} \otimes \mathbf{I}$$

and a position change that depends on the spin

$$\mathbf{C} = \mathbf{P}_0 \otimes \mathbf{U}_+ + \mathbf{P}_1 \otimes \mathbf{U}_-$$

$\mathbf{H}$ is the Hadamard operator on the spin state (in $\mathcal{H}_2$) and it takes

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

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

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

The transformation $\mathbf{H} \otimes \mathbf{I}$ shifts the spin state but operates on a state in $\mathcal{H}$.

$\mathbf{P}_0$ projects the spin state onto $|0\rangle$ and $\mathbf{P}_1$ projects the spin state onto $|1\rangle$,

$$\begin{aligned}\mathbf{P}_0|0\rangle &= |0\rangle \\ \mathbf{P}_0|1\rangle &= 0 \\ \mathbf{P}_1|0\rangle &= 0 \\ \mathbf{P}_1|1\rangle &= |1\rangle\end{aligned}$$

These two operate on spin states, or those in $\mathcal{H}_2$. We can also write

$$\begin{aligned}P_0 &= |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ P_1 &= |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\end{aligned}$$

The operators $\mathbf{U}_+$ and $\mathbf{U}_-$ raise and lower n

$$\begin{aligned}\mathbf{U}_+|n\rangle &= |n+1 \bmod N\rangle \\ \mathbf{U}_-|n\rangle &= |n-1 \bmod N\rangle\end{aligned}$$

These two operate on states in $\mathcal{H}_N$. We can also write

$$\begin{aligned}\mathbf{U}_+ &= \sum_{j=0}^{N-2} |j+1\rangle\langle j| + |0\rangle\langle N-1| = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & \dots \\ \dots & & & & & \\ 1 & 0 & 0 & 0 & 0 & \dots \end{pmatrix} \\ \mathbf{U}_- &= \sum_{j=1}^{N-1} |j-1\rangle\langle j| + |N-1\rangle\langle 0| = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & 0 & \dots & \dots \\ 0 & 1 & 0 & 0 & \dots & \dots \\ 0 & 0 & 1 & 0 & \dots & \dots \\ \dots & & & & & \end{pmatrix}\end{aligned}$$

The operator $\mathbf{C}$ moves the particle to the right if the spin is up and moves the particle to left if the spin is down. It simulates a coin flip. This is why the procedure can be called a random walk. Because $\mathbf{U}_+|N-1\rangle = |0\rangle$ and $\mathbf{U}_-|0\rangle = |N-1\rangle$, the end of the N state space is connected with the beginning of it, so our position space is equivalent to N equidistant points on a circle.

We can think of operator $\mathbf{C}$ as defining a network of connections between states in the N -dimensional Hilbert space $\mathcal{H}_N$. The $\mathbf{C}$ operator entangles the spin with the particle position.

Combining the two operators

$$\mathbf{V} = \mathbf{C} * (\mathbf{H} \otimes \mathbf{I})$$

One can plot the probability of being in each state $|n\rangle$ at each iteration.

The resulting vector has even/odd parity and spreads out ballistically (rather than diffusively) with more iterations. The shape of the probability distribution differs from that of a classical random walk.

Let's evaluate the first few states. Starting with $|\psi_0\rangle = |00\rangle$ we first apply the Hadamard to the spin state and we get

$$\mathbf{H} \otimes \mathbf{I} |\psi_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle)$$

We now apply the controlled raising and lower operator, $\mathbf{C}$. $|00\rangle \rightarrow |01\rangle$ and $|10\rangle \rightarrow |1, N-1\rangle$.

$$\mathbf{V} |\psi_0\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |1, N-1\rangle)$$

Let's operate again with the Hadamard

$$(\mathbf{H} \otimes \mathbf{I}) \mathbf{V} |\psi_0\rangle = \frac{1}{2}(|01\rangle + |11\rangle + |0, N-1\rangle - |1, N-1\rangle)$$

Now we operate again with $\mathbf{C}$.

$$\mathbf{V}^2 |\psi_0\rangle = \frac{1}{2}(|02\rangle + 2|00\rangle - |0, N-2\rangle)$$

Notice the even/odd parity developing! Apply the Hadamard again

$$(\mathbf{H} \otimes \mathbf{I}) \mathbf{V}^2 |\psi_0\rangle = \frac{1}{\sqrt{8}}(|02\rangle + |12\rangle + 2|00\rangle + 2|10\rangle - |0, N-2\rangle - |1, N-2\rangle)$$

Apply $\mathbf{C}$ again,

$$\mathbf{V}^3 |\psi_0\rangle = \frac{1}{\sqrt{8}}(|03\rangle + |11\rangle + 2|01\rangle + 2|1, N-1\rangle - |0, N-1\rangle - |1, N-3\rangle)$$

And so on.

A quantum random walk might be used to find a short path on a complicated network. They also are interesting as a class of cellular automata. One can study the influence of imperfections and external perturbations on the behavior of a quantum random walk. While static spatial random changes of the coin may lead to Anderson localization temporal randomness in the coin operator can cause decoherence resulting in a transition to classical random walking behavior. Their transport and percolation behavior is dependent on the network or graph of the shift operator. Some times people approximate continuous (in time) Hamiltonian evolution with a series of discretely (in time) applied operators. Quantum random walks can be considered relevant for numerical techniques aiming to approximate continuous systems.

9 Quantum Fourier Transform and Shor algorithm

9.1 Discrete Fourier Transforms

The discrete Fourier transform is used on an evenly spaced set of data. A particularly efficient (in both numbers of operations and with memory usage) is the Fast Fourier Transform (FFT) which is done on $N = 2^n$ data samples, where the number of samples is a power of 2. We take x_j to be our N data samples with index j going from 0 to $N - 1$. The numbers x_j can be complex. The result of the Discrete Fourier Transform is another sequence of complex numbers

$$X_k = \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N} k j}$$

with index k also ranging from 0 to $N - 1$. The inverse transform is

$$x_j = \frac{1}{N} \sum_{l=0}^{N-1} X_l e^{\frac{2\pi i}{N} l j}$$

The complex numbers $e^{\frac{2\pi i}{N} j}$ with $j \in [0, N - 1]$, on the complex plane, are N equidistant points on the unit circle. The points are symmetrically distributed, so the sum

$$\sum_{j=0}^{N-1} e^{\frac{2\pi i}{N} j} = 0$$

With integer $q \in [0, N - 1]$, the sum

$$\sum_{j=0}^{N-1} e^{\frac{2\pi i}{N} q j} = \begin{cases} 0 & \text{if } q \neq 0 \\ N & \text{if } q = 0 \end{cases}$$

We can check that inverse transform is an inverse transform by inserting the sum for

x_j inside that for X_k

$$\begin{aligned}
X_k &= \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N} k j} \\
&= \frac{1}{N} \sum_{j=0}^{N-1} \sum_{l=0}^{N-1} X_l e^{\frac{2\pi i}{N} l j} e^{-\frac{2\pi i}{N} k j} \\
&= \frac{1}{N} \sum_{l=0}^{N-1} X_l \sum_{j=0}^{N-1} e^{\frac{2\pi i}{N} (l-k) j} \\
&= \frac{1}{N} \sum_{l=0}^{N-1} X_l N \delta_{lk} \\
&= X_k
\end{aligned}$$

9.2 Quantum Fourier transforms

The Quantum Fourier transform is a map of a wavefunction $|\psi\rangle = f(x) |x\rangle$ assuming that there are $N = 2^n$ states.

$$\begin{aligned}
g(x) |x\rangle &\rightarrow G(y) |y\rangle \\
G(y) &= \text{Discrete Fourier Transform } g(x)
\end{aligned}$$

Here x, y is are integers from 0 to $N-1$. We assume that $N = 2^n$ is a multiple of 2. The quantum Fourier transform is a unitary transformation.

We express x, y as binary expansions or in terms of bits

$$\begin{aligned}
x &= x_{n-1}2^{n-1} + x_{n-2}2^{n-2} + \dots + x_12 + x_0 \\
y &= y_{n-1}2^{n-1} + y_{n-2}2^{n-2} + \dots + y_12 + y_0
\end{aligned}$$

In base 2 the integer x would have the bits $x_{n-1}x_{n-2}..x_1x_0$ and likewise for y .

Let's compute the product of xy modulo 2^n

$$\frac{xy}{2^n} = y_{n-1}(x_0) + y_{n-2}(x_1x_0) + y_{n-3}(x_2x_1x_0) \dots + y_1(\prod_{j=0}^{n-2} x_j) + y_0 \prod_{j=0}^{n-1} x_j$$

$$G(y) = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{y=N-1} e^{2\pi i xy/N} |y\rangle$$

where x, y are integers. The Discrete Fourier transform is a unitary operation and can be constructed from 2-bit Hadamard gates and phase shifts.

9.3 Outline of Shor Algorithm

Aim: Find prime factors of given integer $N = pq$ efficiently using the Quantum Fourier transform.

10 What's a quantum Computer?

What are the requirements for a quantum computer?

- A set of quantum bits or few state quantum systems.
- Long coherence times in each quantum system. You need to be able to perform operations on the states before coherence is lost.
- The ability to perform some operations on individual bits.
- The ability to perform operations on pairs of bit that entangle them.
- The ability to prepare quantum states.
- The ability to make measurements.

All bits need not be localized to the same region. (Teleportation!)

Much of modern physics strives to be formulated in geometric or coordinate free ways. However in this set of notes, and as commonly done in introductions to quantum information theory, we started with a preferred basis. The preferred bases depends on how the entire system is decomposed into subsystems. For example we used basis vectors for a 2 qubit systems that look like $|01\rangle$ and $|00\rangle$. A 2 qubit system is equivalent to a 4 dimensional Hilbert space. Why do we do our calculations assuming a specific basis? Why not just work with generators of $U(4)$?

Realizations of quantum computers are probably going to be strongly tied to particular basis states. Feasible quantum gates are likely to operate in physically bounded regions in space. Likewise errors, decoherence and connections or interactions between qubits are going to be specifically implemented. Eventually more abstract versions of quantum computers might be realizable.

10.1 Quantum Compiler

A quantum algorithm is described by a single unitary matrix W . Your goal is to find optimal implementation of W using elementary Quantum gates. Such an implementation could be called a compiler.

The meaning of *optimal* may depend on mechanisms for error correction in the presence of decoherence or noise and the number of available qubits.

Unitary transformations are reversible and no information is lost. However not all desired computations are reversible. With extra qubits you can always describe an irreversible function with a reversible one.

A NAND serves as a classical universal gate set, and the 3 qubit Toffoli gate, gives you a NAND operation but it takes an extra bit to do so. Each time a NAND is executed a qubit is wasted. Erasing bits costs energy (increases entropy) and is not reversible (or unitary). A way around this problem is to do the calculation, copy the output to a set of qubits, then reverse the computation, so that original used bits are reset. This saves the output and does not require any erasing. This procedure of computing and uncomputing saves space (needs fewer qubits) and maintains the entire calculation as a unitary operation, at the cost of using extra gates.

10.2 Errors and Error Correction

Using extra bits to correct errors defeat decoherence. I would love to add here!

10.3 Entropy and information loss

All quantum gates are Unitary. Many calculations are not. This means extra bits are used to store lost information within a quantum computation. Loss of information is equivalent to entropy increase and dissipation. One possible way to design good algorithms is to minimize use of gates that are busy because of information loss in a calculation. Partly reversing the calculation can recover the use of these bits?

11 The Density matrix

11.1 Density operator for a pure state

The density operator is a matrix

$$\begin{aligned}\rho &= |\psi\rangle \langle\psi| \\ \rho &= \sum_{ij} c_i^* c_j |a_i\rangle \langle a_j|\end{aligned}$$

Since ψ has length 1,

$$\text{tr } \rho = \sum_i c_i^* c_i = 1.$$

The trace of a matrix is independent of basis. In other words $\text{tr } UAU^\dagger = \text{tr } A$, where U is a unitary matrix that allows you to change basis from one orthonormal basis to another. Eigenvalues are also basis independent and this means that $\text{tr } \lambda_i = 1$ for the density matrix, where λ_i are its eigenvalues.

A unitary transformation transforms from one orthonormal basis to another. We can always find an orthonormal basis that has our state vector as a basis vector. In this basis

$|\psi\rangle = |0\rangle$ where $|0\rangle$ is that basis vector. In this basis the density matrix has a single 1 on its diagonal and the rest of the matrix contains zeros. That means that $\text{tr} \rho^2 = 1$ and the density matrix has a single non-zero eigenvalue that is 1. This is only true for density matrices that are derived from a single pure state.

We can write the expectation of an observable as

$$\langle \mathbf{A} \rangle = \text{tr}(\rho \mathbf{A})$$

While there is a complex phase ambiguity in a state vector, there is none in the density matrix.

11.2 Mixtures

We can consider a system which has probability of being in various pure states,

$$|\psi\rangle = \sum_i p_i |\psi_i\rangle$$

where p_i are probabilities with $\sum p_i = 1$. Call ρ_i the density matrix for each pure state.

$$\rho = \sum_i p_i \rho_i$$

and

$$\langle \mathbf{A} \rangle = \sum_i p_i \text{tr}(\rho_i \mathbf{A}) = \sum_i \text{tr}(p_i \rho_i \mathbf{A}) = \text{tr}(\rho \mathbf{A})$$

The trace $\text{tr} \rho = 1$, as before for a pure state. Using the Schwartz inequality one can show that $\text{tr} \rho^2 \leq 1$. If $\text{tr} \rho^2 = 1$ then the system is said to be in a *pure* state.

For a non-pure state, the off-diagonal elements are called *coherences*. Because ρ is Hermitian one can always find a basis that makes it diagonal. Eigenvalues are real. If the system is a pure state, then in that basis, the density matrix has a single 1 on the diagonal.

The amount of information can be described with an entropy (sometimes called Shannon entropy)

$$s \equiv -\text{tr}(\rho \ln \rho)$$

The exponential of a matrix can be defined in terms of a series. The log of a matrix can be defined as coming from that definition. If we diagonalize and write ρ as a diagonal matrix with eigenvalues λ_i on its diagonal then

$$s \equiv -\sum_i \lambda_i \ln \lambda_i$$

A pure state has a single eigenvalue of 1 so its entropy $s = 0$. A mixed state has a positive entropy as the eigenvalues are positive and they are all smaller than 1.

Summary.

- **Classical states** An object is either in one or not. We can have a linear vector space describing all the possible states $|a_i\rangle$.
- **Pure quantum states** These are complex superpositions of classical states; $|\psi\rangle = \sum c_i |a_i\rangle$. The probability p_i that an object is in a state $|a_i\rangle$ depends on the amplitude; $p_i = c_i c_i^*$.
- **Mixed states** Classical probability distribution over pure states.

$$\rho = \sum p_j |\psi_j\rangle \langle \psi_j|$$

11.3 POVM

Positive operator value measurement is a collection of operators that sum to the identity

$$I = \sum_m M_m M_m^\dagger$$

We don't worry about the value measured, only its probabilities. The $M_m M_m^\dagger$ are the POVM elements associated with measurement and they are sufficient to determine probabilities of outcomes.

Measurements can be described by a collection of measurement operators M_m that satisfy completeness

$$\sum_m M_m M_m^\dagger = I$$

The state of a system after measurement finding that it is in the m state is

$$\frac{M_m |\psi\rangle}{\sqrt{\langle \psi | M_m^\dagger M_m | \psi \rangle}}.$$

A projective measurement is a Hermitian matrix that has spectral decomposition

$$M = \sum_m m P_m$$

where P_m is a projector with eigenvalue m . Measurement of m has probability

$$p(m) = \langle \psi | P_m | \psi \rangle$$

and if the state was in the m state, the state after measurement is

$$\frac{P_m |\psi\rangle}{\sqrt{p(m)}}.$$

Projective measurements satisfy $P_i P_j = \delta_{ij} P_i$ but a general measurement might not.

11.4 Transformations of the density matrix

If the wave-function $|\psi'\rangle = U |\psi\rangle$ where U is a unitary operator, then the density matrix transforms as

$$\rho' = U \rho U^\dagger$$

This is true even for mixed states.

If we have a measurement operator M_m the probability of getting value m is

$$p(m) = \text{tr}(M_m^\dagger M_m \rho)$$

and the density matrix after measurement becomes

$$\rho'_m = \frac{M_m \rho M_m^\dagger}{\text{tr}(M_m^\dagger M_m \rho)}$$

If we don't record the result of the measurement then the wavefunction is not collapsed and we sum over all the probabilities for each possible measurement

$$\begin{aligned} \rho' &= \sum_m p(m) \rho'_m \\ &= \sum_m \text{tr}(M_m^\dagger M_m \rho) \frac{M_m \rho M_m^\dagger}{\text{tr}(M_m^\dagger M_m \rho)} \\ &= \sum_m M_m \rho M_m^\dagger \\ \rho' &= \sum_m M_m \rho M_m^\dagger \end{aligned} \tag{11}$$

This type of transformation is usually not unitary. However it is deterministic as a measurement value is not chosen with a probability.

This gives two types of types of quantum simulations involving a series of measurements.

- We can start with a density matrix and transform it via each measurement, conserving probability as we go and using equation 11 each time a measurement is done.
- Or we can start with a wave-vector, and simulate the measurements by randomly choosing measured values consistent with the probability of measurement. After each measurement the wave function is collapsed. Each measurement requires a monte-Carlo simulated random event.

In the first case we evolve probability. In the second case we follow a path of possible measurements. An ensemble of paths should be consistent with the probability distribution predicted via the first method. In other words, if we carry out the second method a number of times and average the density matrices of the results at the end, we should get the density matrix estimated from the first method.