Contents

1 Introduction

In this paper I will recapitulate much of APMA 1930W: *Probabilities in Quantum Mechanics*, taught at Brown University by Professor Stuart Geman, and build upon the lessons of the course to investigate the basics of quantum computing. Starting with the assumption that the reader is familiar with linear algebra, statistics, and conventional computers, I will explain the mathematical tools necessary for quantum computing. I will then cover the main findings from the course while highlighting important experimental discoveries in the field of quantum mechanics. Quantum computers will be briefly discussed, and the final section includes an explanation of Grover's quantum search algorithm along with a demonstration of its implementation.

2 Basic Mathematical Tools

2.1 Hilbert Spaces

An inner product space is a vector space V over a field F that has an inner product operator such that $\langle \cdot, \cdot \rangle : V \times V \to F$. F is often either $\mathbb{R}$ for a real inner product space or $\mathbb{C}$ for a complex inner product space. The inner product operator fulfils three properties,[17] $\forall x, y, z \in V$ and $\forall a, b \in F$:

1. *Conjugate symmetry*: $\langle x, y \rangle = \overline{\langle y, x \rangle}$ where for some complex number $c = a + bi$, its complex conjugate is defined as $\bar{c} = a - bi$; other notation for $\bar{c}$ is c^*. Note that this implies $\langle x, x \rangle \in \mathbb{R}$ regardless of whether V is $\mathbb{R}$ or $\mathbb{C}$.

2. *Linearity in the first argument*: $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$.

3. *Positive-definiteness*: $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.

Intuitively, the inner product of vectors is just the product of vectors that results in a scalar and where a vectors inner product with itself is a positive real number (unless that vector is 0, in which case the inner product is 0). The above paragraphs formalize this.

Having an inner product operator $\langle \cdot, \cdot \rangle$ allows us to define a norm $\| \cdot \|$ such that $\|x\| = \sqrt{\langle x, x \rangle}$. The norm is a generalization of absolute value that exists for scalars. In the real coordinate space $\mathbb{R}^n$ of basic Euclidean geometry, the inner product operator is the dot product, where for $x, y \in \mathbb{R}^n, \langle x, y \rangle = x \cdot y = x^T y = \|x\|\|y\| \cos\theta$. This definition of the norm ($\|x\| = \sqrt{x \cdot x}$) is called the Euclidean norm, alternatively the ℓ_2 norm which is sometimes noted as $\| \cdot \|_2$. While there are many other definitions for the norm,[12] the ℓ_2 norm is what we will consider for the purpose of this paper, and the next section will show that the norm is largely irrelevant due to the Hilbert space that we choose.

A Hilbert Space H is an inner product space that is also complete. Completeness is the property such that every Cauchy sequence of elements of the space converges to an element of that same space. Intuitively, a Cauchy sequence is a sequence of elements—with some *distance* defined over the set which the sequence is a subset of—where for all positive real numbers, there is an index into the sequence such that all pairs of elements past it have distance between them smaller than that number. $\mathbb{R}$ and $\mathbb{C}$ are complete,[6] as is any finite-dimensional inner product space.[14] For this course we only consider finite-dimensional inner product spaces, and thus the spaces we use for quantum mechanics are Hilbert spaces.

2.2 The $\mathbb{C}_1^2$ Space

The most common Hilbert space to consider for quantum computing is $\mathbb{C}_1^2$. The superscript 2 means that an element of the set is a vector composed of 2 complex numbers. The subscript 1 means that elements of the set are normalized, i.e., for any $x \in \mathbb{C}_1^2$, $\|x\| = 1$. Informally, this is saying that the length of all the vectors in the space is 1.

The absolute value of a complex number $z = a + bi$ is defined as $|z| = \sqrt{a^2 + b^2}$. Furthermore, we see that $|z|^2 = a^2 + b^2 = z^* z$. Also, $z^* z = z z^*$ and $z^{**} = z$. If z and w are both complex, $(z+w)^* = z^* + w^*$.

2.2.1 Dirac Notation, Adjoints, & Hermitian Matrices

In Dirac Notation, complex vectors are represented with a *ket* symbolized as $|x\rangle$. If $x \in \mathbb{C}_1^2$, then

$$|x\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$ (this is a consequence of $\|x\| = 1$ that is proven in subsubsection 2.2.2). The conjugate transpose (a.k.a. the *adjoint*) of a ket $|x\rangle$ is the *bra* $\langle x|$ where

$$\langle x| = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix}.$$

Notationally this can be written out as $|x\rangle^{*T} = \langle x|$, for the conjugate transpose. But this is clunky, and so dagger notation is used for adjoints such that $|x\rangle^\dagger = \langle x|$.

A matrix A that is self-adjoint is called a *Hermitian* matrix. That is, A is Hermitian iff $A = A^\dagger$ (i.e., A is its own adjoint). It is easy to see that any Hermitian matrix must be a square matrix. Finally, it is proven that the eigenvalues of a Hermitian matrix must be real. Consider Hermitian matrix A, ket $|x\rangle$, and eigenvalue α. By definitions, $A|x\rangle = \alpha|x\rangle$, so $\langle x|A|x\rangle = \langle x|\alpha|x\rangle = \alpha\langle x|x\rangle = \alpha\|x\| = \alpha$. Taking the adjoint of both sides, we see

$$\begin{aligned} [\langle x|A|x\rangle]^\dagger &= \alpha^\dagger \\ |x\rangle^\dagger A^\dagger \langle x|^\dagger &= \alpha^* \\ \langle x|A|x\rangle &= \alpha^* \end{aligned}$$

and thus $\alpha = \alpha^*$, indicating that any eigenvalue α of a Hermitian is real.

2.2.2 Inner Product & Outer Product

The inner product of a bra and ket $\langle\langle x|, |y\rangle\rangle$ is more simply written as $\langle x|y\rangle$.

Since the ket is on the right, and kets are column vectors, taking an inner product results in a scalar. This aligns with the formal definition of the inner product laid out in subsection 2.1.

If $\alpha = a + bi$ and $\beta = c + di$, we see that $\||x\rangle\|^2 = \langle x|x\rangle = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha^*\alpha + \beta^*\beta = |\alpha|^2 + |\beta|^2 = a^2 + b^2 + c^2 + d^2$. And since $x \in \mathbb{C}_1^2$, $\|x\| = 1$, so $\|x\|^2 = 1$ and thus $|\alpha|^2 + |\beta|^2 = 1$.

Taking an outer product is like the reverse of the inner product. It is vector multiplication with a ket on the left and a bra on the right, e.g., $|x\rangle\langle y|$. The product is an $n \times n$ matrix, where $|x\rangle$ and $|y\rangle$ are n-dimensional vectors.

2.3 Unitary Operators

A square matrix is unitary if it is invertible and its inverse is its adjoint. That is, U is unitary iff $UU^\dagger = U^\dagger U = UU^{-1} = U^{-1}U = 1$. Because unitary matrices have an inverse (its adjoint), multiplying one by a vector is a reversible operation—you can simply multiply it by its adjoint. We multiply unitary matrices by vectors to apply a linear transformation to them, and use the terminology "unitary operator" for these matrices that operate on vectors/kets/bras.

One property of unitary operators is that they preserve inner product. $\forall x, y \in \mathbb{C}$ and a unitary operator U, $\langle Ux|Uy\rangle = \langle x|y\rangle$. We will return to the importance of this in subsection 4.2. Also, the notation

used here shows U being applied to x within the bra. Since $\langle x|$ is the adjoint of $|x\rangle$, to extract U from within the bra we need to take its adjoint. That is, $\langle Ux| = \langle x|U^\dagger$, and so $\langle Ux|Uy\rangle = \langle x|U^\dagger U|y\rangle$. Lastly, unitary operators preserve norm, so $\||x\rangle\| = 1 \iff \||Ux\rangle\| = 1$.

3 Spin State

In quantum mechanics we are interested in studying the "spin states" of particles. We represent such states as vectors in a Hilbert Space; in particular, we use $\mathbb{C}_1^2$. It is important at this point to distinguish between complex vectors and kets/bras. A complex vector $x \in \mathbb{C}_1^2$ is just that, a complex vector. $|x\rangle$ is a spin state that a particle can have.

3.1 Physical Understanding of Spin

The "spin" of a particle is its intrinsic angular momentum,[30] notationally s.[4] This property causes particles to have their trajectory deflected when moving through an inhomogenous magnetic field.[10] Fermions are particles with half integer spin (e.g., $\frac{1}{2}$, $\frac{3}{2}$). Bosons are particles with integer spin (e.g., 0, 1, 2). Fundamental particles of the Standard Model (which currently explains the strong, weak, and electromagnetic forces, but not gravity) that are fermions include electrons, neutrinos, and quarks, all with spin $\frac{1}{2}$; photons are an example of a fundamental particle boson with spin 1. Composite particles may have higher spin. The Higgs boson—discovered in 2012—has spin 0, and the graviton (the hypothesized force-carrying particle behind the gravitational force) has spin 2.[7]

3.2 Stern-Gerlach Experiment

In classical physics, angular momentum is a continuous variable. In 1922, the Stern-Gerlach experiment demonstrated that atoms have quantized spin. The results of the experiment have been more recently reproduced and shown to extend to individual electrons.[10]

The experimental setup of the Stern-Gerlach experiment starts with an oven that heats up silver atoms. Silver atoms were easier to deal with than individual electrons and they have a single electron in their outermost (5s) shell, meaning the atoms' magnetic momentum is the same as that of the one electron in the 5s shell.[32][10] The only exit through this oven is a collimating slit that results in a parallel beam of vaporized silver atoms. Assume this beam is propogating in the positive y-direction. The beam then passes through a pair of magnets that are set up to create a magnetic field in the z-direction perpendicular to the atoms' motion. Further along the y-direction there is a screen set up (perpendicular to the y-direction and expanding along both the x- and z-axes) that detects where the atoms strike the screen after passing through the magnetic field.

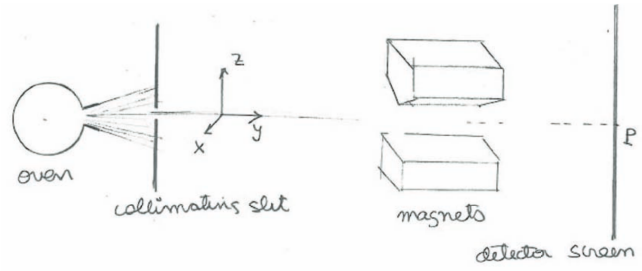

Figure 1: Sketch of the Stern-Gerlach experiment[32]

Because angular momentum is a continuous variable in classical mechanics and a particle traveling through a magnetic field is deflected to a degree proportional to the strength of its spin, the expected outcome of the Stern-Gerlach experiment was that the atoms would have all different spins (randomized in the oven) and their trajectories would thus be bent in a continuous spectrum of angles. The result would be the screen detecting a line of where the different atoms hit. However, the result was that all the atoms hit the screen at one of two points.

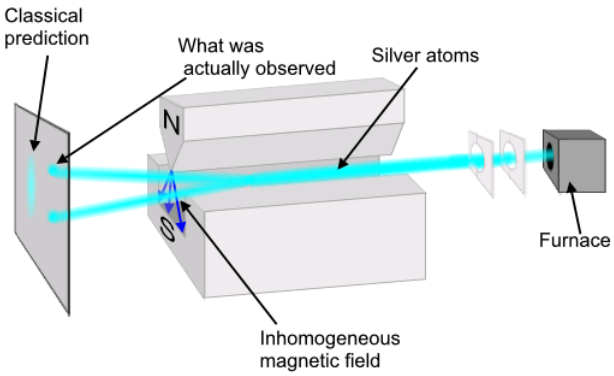

Figure 2: Rendering of the Stern-Gerlach experiment demonstrating the expected result based on the classical assumption of spin versus the actual results.[20] Since the particles are deflected up or down in the z-direction, the particles that are deflected up are labelled as moving in the $\lambda_z = 1$ direction and those deflected down in the $\lambda_z = -1$ direction.

Generalizing this result, it is now known that a beam of atoms passing through an inhomogenous magnetic field will split into $2s + 1$ beams (with s still being the spin number of the particle).[28][4] So for silver atoms, which have the same spin as an electron ($\frac{1}{2}$), the stream split into $2 \cdot \frac{1}{2} + 1 = 2$ beams.

3.3 Classical vs. Quantum Mechanics

What happened is that spin states at the quantum level are, well, quantized. Hence "quantum mechanics" and how it differs from classical mechanics. The spin magnetic quantum number, m_s, takes values from $-s$ to s in steps of 1. For an arbitrary direction z (recall the direction of the magnetic field in the Stern-Gerlach experiment is along the z-axis) the z-component of spin s_z, which causes the trajectory change when passing through in an inhomogenous magnetic field in the z-direction, can take on any value $m_s\hbar$.[4]

This means that the s_z of the silver atoms (or pure electrons, which has been confirmed in experimental settings more recently[10]) could either be $\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$. All of the silver atoms coming out of the oven had one of these spins and changed direction when passing between the magnets creating the inhomongenous magnetic field (the Stern-Gerlach machine) at a degree corresponding to their spin, resulting in the two discrete beams exiting the Stern-Gerlach machine.

3.4 Feynman Stern-Gerlach Machines

Richard Feynman proposed as a thought experiment a modified type of Stern-Gerlach machine. It uses two additional sets of magnets after the first to recombine the split beam back into one.[13][10]

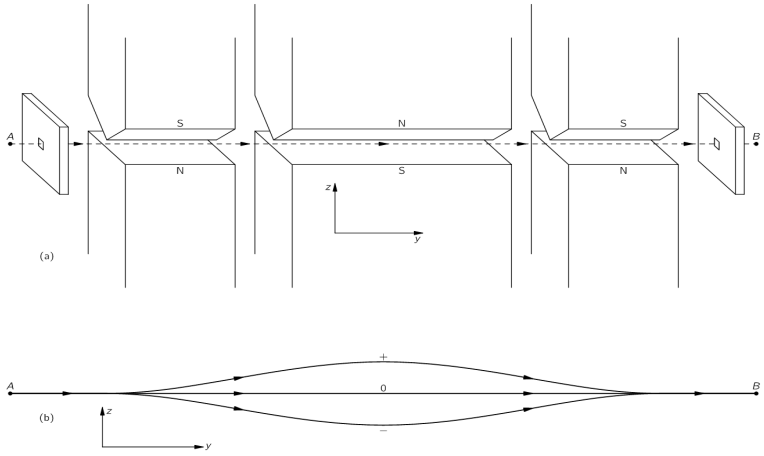

Figure 3: Example of the modified Feynman Stern-Gerlach machine; note that this shows its input beam splitting into three, which would correspond to a beam of particles with $s = 1$.[13]

The second part to these Feynman Stern-Gerlach machines is that you can selectively put blockers between the second pair of magnets so as to block the path of only some of the split beams. This results in the final beam that exits having the same orientation as the input beam but that it has been filtered to only include atoms in the spin states not blocked.

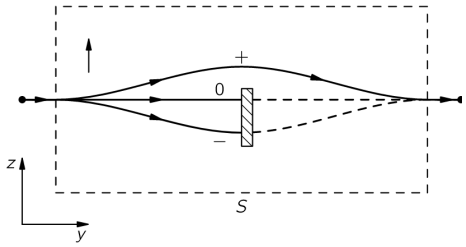

Figure 4: Example of the blocking that can be used in a Feynman Stern-Gerlach machine to filter particles by spin state[13]

This new instrumentation aids in constructing more complicated experiments that further elucidate the nature of spin states. We do so by chaining together multiple Feynman Stern-Gerlach machines.

3.4.1 Repeating the Same Filter

Looking back at Figure 4, we see only the beam in the + direction was allowed to exist. Since this Feynman Stern-Gerlach machine measured spin in the z-direction, the particles in that beam are said to be in the $|z^+\rangle$ state. Now, suppose we feed this beam into another Feynman Stern-Gerlach machine with its magnetic fields in the same directions as those in the first. We see that the beam in the second machine does not split at all. This is because all the particles were prepared by the first machine in the $|z^+\rangle$ state, and so when they are measured again in the same direction (the z-axis) all of the particles bend upwards.

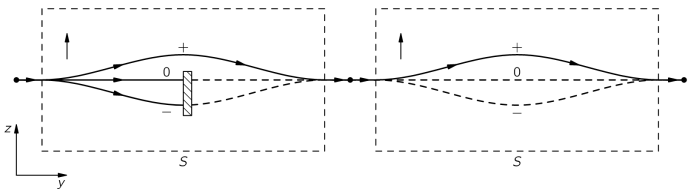

Figure 5: Two Feynman Stern-Gerlach machines with the same orientation[13]

3.4.2 Changing Orientation in Space

Now suppose we take the second Feynman Stern-Gerlach machine in Figure 5 and rotate it (along with reflecting the beam so that it travels into the reoriented machine) to measure the particles exiting the first machine—prepared in the $|z^+\rangle$ state—along the x-axis instead. What happens is that the beam splits again, with half of the particles moving in the $\lambda_x = 1$ direction and the other half in the $\lambda_x = -1$ direction.

In the next section we will discuss how to use Dirac notation to probabilistically model the amount of particles that move in the $\lambda_n = 1$ direction versus the $\lambda_n = -1$ direction for any arbitrary second direction n. As it turns out, it is dependent on what state the particles are prepared in prior to their measurement in direction n. We can start by explaining what happens with more traditional notation.

3.4.3 The Conditional Nature of State

Generalizing our two measurement setup, we will say that the first Feynman Stern-Gerlach machine measures along the n-axis and the second machine measures along the m-axis. What has been found experimentally[10] is that for any particle, the direction it is bent in inside the second machine has the conditional probability as follows

$$\begin{aligned} P(\lambda_m = 1 | \lambda_n = 1) &= \frac{1 + m \cdot n}{2} \\ &= \frac{1 + \|m\|\|n\| \cos\theta}{2} && \text{where } \theta \text{ is the angle between } m \text{ and } n \\ &= \frac{1 + \cos\theta}{2} && \text{since we are in the } \mathbb{C}_1^2 \text{ space, } \|m\| = \|n\| = 1 \\ &= \cos^2 \frac{\theta}{2} \end{aligned}$$

Thus we see that the probability of a particle's state when measured in the m-direction is dependent entirely on the angle between the this direction and the n-direction which it was previously measured in.

This explains the outcome of the experiment in subsubsection 3.4.2 where exactly half the particles bent upwards in the x-direction machine after being measured in the z-direction machine. The x- and z-axes are separated by an angle of $\theta = \frac{\tau}{4}$ (where $\tau = 2\pi$, a convention that will be used to make visualization of angles easier[16]), and $\cos^2 \frac{\frac{\tau}{4}}{2} = \cos^2 \frac{\tau}{8} = \sqrt{\frac{1}{2}}^2 = \frac{1}{2}$. This is contextually interpreted as saying that half the particles will spin $\lambda_x = 1$ given they were all prepared with spin $\lambda_z = 1$, which is that same as what is found experimentally.

3.4.4 Erasing State

Consider again the same experiment from subsubsection 3.4.2. A beam of particles goes into a z-directional Feynman Stern-Gerlach machine with a filter on the inside so that all the particles that

exit are prepared in the $|z^+\rangle$ state. The beam is then passed through a x-directional Feynman Stern-Gerlach machine where it splits in half once again. Now, take the beam exiting that second machine and pass it through a third machine which is oriented to measure along the z-axis just like the first. You may expect that the beam is uniformly bent up, since all the particles were previously filtered so as to be in the $|z^+\rangle$ state. However, this is not what happens! The beam splits in half once again, with half of the particles previously prepared in the $|z^+\rangle$ state now being bent in the $\lambda_z = -1$ direction, indicating they are in the $|z^-\rangle$ state now. This seems to directly contradict the results of the experiment in subsubsection 3.4.1 wherein the particles retained their $|z^+\rangle$ state.

What is happening is that measurements at the quantum level are not gentle. Measuring the state of the particles along x-axis disturbed them so much that their state in the z-axis was erased. As Susskind and Friedman put it,

> *Any* interaction that is strong enough to measure some aspect of a [quantum] system is necessarily strong enough to disrupt some other aspect of the same system. Thus, you can learn nothing about a quantum system without changing something else... The intermediate measurement along the x axis will leave the spin in a completely random configuration as far as the next measurement is concerned. There is no way to make the intermediate determination of the spin without completely disrupting the final measurement. One might say that measuring one component of the spin destroys the information about another component. In fact, one simply cannot simultaneously know the components of the spin along two different axes, not in a reproducible way in any case.[31]

Indeed, this aligns with the probabilistic method outlined in subsubsection 3.4.3. We see that the measurement in the x-direction erased the state in the z-direction, and the third measurement (the second one in the z-direction) depended entirely on the state that the immediately previous measurement prepared the particles in. Just as $P(\lambda_x = 1|\lambda_z = 1) = \frac{1}{2}$, $P(\lambda_z = 1|\lambda_x = 1) = \frac{1}{2}$ as the angle between the two axes remains as $\frac{\tau}{4}$.

4 Mathematically Explaining Spin State

We can use the tools introduced in section 2 to model quantum systems in a manner that aligns with experimental findings such as those found by Stern and Gerlach. We will start by formalizing the probability discussed in subsubsection 3.4.3 using Dirac notation.

4.1 Probability

We can use the Dirac notation for inner products to model probability regarding quantum states. Revisiting the example from subsubsection 3.4.3, we can rewrite $P(\lambda_m = 1|\lambda_n = 1)$ as $|\langle m^+|n^+\rangle|^2$. Through this, we see that the interpretation of the inner product is that it represents the positive square root of the probability of a particle being in the $|m^+\rangle$ state when measured in the m-direction given that it was prepared in the $|n^+\rangle$ state.[9]

The findings of subsubsection 2.2.2 show that for any x, $\langle x|x\rangle = 1$, and so $|\langle x|x\rangle|^2 = 1$. This makes sense with the probability interpretation of inner product, because if a particle is measured to be in the $|x\rangle$ state then it has a 100% chance of being found in the $|x\rangle$ state again with no other intervening measurements. This result is found in the experiment of subsubsection 3.4.1.

4.2 Superposition

So far we have been looking at particles with states like $|z^+\rangle$ or $|z^-\rangle$. However, particles can also exist in a superposition of states. A superposition of states is essentially a linear combination. That is, a particle could be in the state $\frac{1}{\sqrt{2}}|z^+\rangle + \frac{1}{\sqrt{2}}|z^-\rangle$. The physical interpretation of this is

that a particle with this superposition is simultaneously in the $|z^+\rangle$ state and the $|z^-\rangle$ state at the same time. The coefficients on the states in the linear combination are called the *probability amplitudes*. They are called this because, for a particle in a general superposition state $\alpha|x\rangle + \beta|y\rangle$, the particle has a $|\alpha|^2$ probability of being in the $|x\rangle$ state and a $|\beta|^2$ probability of being in the $|y\rangle$ state when measured along the $\{x, y\}$ basis. This follows from subsection 4.1 because, for example, $P(|\alpha\rangle \,|\, \alpha|x\rangle + \beta|y\rangle) = |\langle x|\alpha|x\rangle + \langle x|\beta|y\rangle|^2 = |\alpha\langle x|x\rangle + \beta\langle x|y\rangle|^2 = |\alpha \cdot 1 + \beta \cdot 0|^2 = |\alpha|^2$. (Note: α and β are scalar multiples on the inner product and were factored out, also $\langle x|y\rangle = 0$ because they are distinct bases of the vector space). That means a particle in the $\frac{1}{\sqrt{2}}|z^+\rangle + \frac{1}{\sqrt{2}}|z^-\rangle$ state has a $\frac{1}{2}$ probability of bending upwards (the $\lambda_z = 1$ direction) and a $\frac{1}{2}$ probability of bending downwards (the $\lambda_z = -1$ direction) when put through a Stern-Gerlach machine oriented to measure spin in the z-axis. Note that for any superposition, whether it is in the generic $|x\rangle$ or $|y\rangle$ state is a well defined probability function since $|\alpha|^2 + |\beta|^2 = 1$ as a result of states being elements of the $\mathbb{C}_1^2$ space (see subsubsection 2.2.2). Since probability amplitudes are found by taking the inner product between two states, they are also called the *overlap* between those states.

As found in subsubsection 3.4.4, measuring a quantum state is not gentle. This extends to superpositions as well. When we measure a particle in a superposition, it "collapses" (see: subsection 4.7) and fully assumes the state it was found to be in. To go back to the previous example, when we measure a particle in the $\frac{1}{\sqrt{2}}|z^+\rangle + \frac{1}{\sqrt{2}}|z^-\rangle$ state, if we see it is $|z^+\rangle$ then its state has actually changed and is no longer a superposition but rather is just plainly $|z^+\rangle$.

4.3 The Computational Basis

The spin states relevant to quantum computing will be superpositions of the states in the *computational basis*. The computational basis is the set $\{|0\rangle, |1\rangle\}$ where $|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$. It is called this because this set is an orthonormal basis of the $\mathbb{C}_1^2$ space and the "0" and "1" labels are reminiscent of the traditional binary values of conventional computing. In physics, $|0\rangle$ and $|1\rangle$ are sometimes referred to as $|u\rangle$ and $|d\rangle$, respectively, where u stands for "up" and d for "down." "Up" comes from the arbitrary convention of assigning the positive direction along the z-axis such that $|z^+\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|z^-\rangle = \begin{pmatrix}0\\1\end{pmatrix}$).[10] Note that while these vectors are $\frac{\tau}{2}$ radians apart, which would make them collinear in the traditional geometric sense, they are orthogonal in the $\mathbb{C}_1^2$ Hilbert space. This is easily confirmed by checking that $\langle 0|1\rangle = \begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = 0$. This aligns with the experimental findings of subsubsection 3.4.1, because when you measure a particle repeatedly in the same direction (in that example, measuring it along the z-axis) you will continue to find a particle in the same state. You cannot prepare a particle in the $|0\rangle$ state and then measure it immediately after and find it in the $|1\rangle$ state; this is what $\langle 0|1\rangle = 0$ is saying. Also note that any pure (see: subsection 4.5) state in the computational basis is a superposition of $|0\rangle$ and/or $|1\rangle$.

4.4 The Bloch Sphere

To visualize the computational basis and superpositions of it we use the Bloch Sphere. Points on the Bloch Sphere can be represented using the two values θ and ϕ, where $|\psi\rangle$ (which is used to represent a generic unknown state) is given by $|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle$. θ is the angle between the vector and $|0\rangle$ (the z-axis in the positive direction as explained in subsection 4.3) and ϕ is the angle between the vector and $|+\rangle$ measured counterclockwise. $|+\rangle$ is the vector that is on the x-axis in the positive direction (what is referred to previously as $|x^+\rangle$).

Thus for $|+\rangle$, $\theta = \frac{\tau}{4}$ and $\phi = 0$, so $|+\rangle = \cos\frac{\tau}{8}|0\rangle + e^{i0}\sin\frac{\tau}{8}|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$. This is the superposition of the bases $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$, which can be easily verified by remembering the vectors these kets represent. In the negative direction along the x-axis is the vector called $|-\rangle$. When we look at the superposition it represents we see that $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. As a heuristic for remembering these, note that in $|+\rangle$ the basis vectors $|0\rangle$ and $|1\rangle$ are added together whereas their difference is found in $|-\rangle$. In the positive direction along the y-axis is $|i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ (refer to Euler's Formula if you are confused about the i coefficient on $|1\rangle$), and in the negative direction is $|-i\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)$.

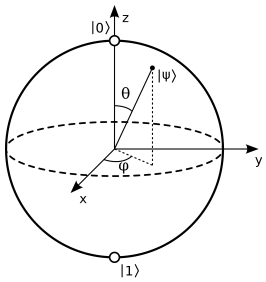

Figure 6: The Bloch Sphere

4.5 Pure vs. Mixed States & Density Matrices

All the states we've considered thus far—which have been superpositions of the computational basis that are normalized—are *pure* states. Any point on the surface of the Bloch Sphere is a pure state. Consider the pure state $|+\rangle$, a particle in this state is in the state $|+\rangle$ with probability 1 ($\langle+|+\rangle = 1$). While this state is in a superposition of $|0\rangle$ and $|1\rangle$, a particle prepared in it is not in both of those other states or in one or the other with some probability, it is in the pure state $|+\rangle$. Again, if you then measure a particle with state $|+\rangle$ along the z-axis it has a $\frac{1}{2}$ probability of being found to be $|0\rangle$ and a $\frac{1}{2}$ probability of being found to be $|1\rangle$ (owing the fact that the $\frac{1}{\sqrt{2}}$ coefficients are not probabilities but probability amplitudes), though that's because the measurement *changed* the state of the particle. It is no longer in $|+\rangle$, it is in either $|0\rangle$ or $|1\rangle$, but the angles between the state the particle previously had ($|+\rangle$) and the axis of measurement did affect the probability for what state the particle would be changed to and found in when measured. But overall, a particle in the state $|+\rangle$ is in the pure state $|+\rangle$. You can, however, have a state that is a collection of multiple different pure states where each is assigned some probability for what the actual state is. These are *mixed* states. Formally, mixed states are probability distribution over pure states,[1] where we have a set of states $|\psi_1\rangle, |\psi_2\rangle, ..., |\psi_n\rangle$ and probabilities $p_1, p_2, ..., p_n$ such that $\sum_{i=1}^{n} p_i = 1$. Mixed states are thus a generalized version of state, since if $n = 1$ then the mixed state is a pure state.[10]

While different pure states are distinguishable experimentally, there are different mixed states (different probability distributions of different pure states) that cannot be distinguished experimentally. So to represent mixed states in a way where each unique representation is experimentally distinguishable we use density matrices. The density matrix ρ for a mixed state of pure states $|\psi_1\rangle, |\psi_2\rangle, ..., |\psi_n\rangle$ and corresponding probabilities $p_1, p_2, ..., p_n$ is given by

$$\rho = \sum_{i=1}^{n} p_i |\psi_i\rangle\langle\psi_i|.$$

All density matrices are Hermitian. This can be seen by the looking at the expansion of a generic density matrix

$$\begin{aligned}
\rho &= p_i \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \begin{pmatrix} \alpha_0^* & \alpha_1^* \end{pmatrix} + ... + p_j \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \begin{pmatrix} \beta_0^* & \beta_1^* \end{pmatrix} \\
&= \begin{pmatrix} p_i\alpha_0\alpha_0^* & p_i\alpha_0\alpha_1^* \\ p_i\alpha_1\alpha_0^* & p_i\alpha_1\alpha_1^* \end{pmatrix} + ... + \begin{pmatrix} p_j\beta_0\beta_0^* & p_j\beta_0\beta_1^* \\ p_j\beta_1\beta_0^* & p_j\beta_1\beta_1^* \end{pmatrix} \\
&= \begin{pmatrix} p_i\alpha_0\alpha_0^* + ... + p_j\beta_0\beta_0^* & p_i\alpha_0\alpha_1^* + ... + p_j\beta_0\beta_1^* \\ p_i\alpha_1\alpha_0^* + ... + p_j\beta_1\beta_0^* & p_i\alpha_1\alpha_1^* + ... + p_j\beta_1\beta_1^* \end{pmatrix}
\end{aligned}$$

and seeing that

$$\begin{aligned}\rho^\dagger &= \begin{pmatrix} p_i\alpha_0\alpha_0^* + ... + p_j\beta_0\beta_0^* & p_i\alpha_0\alpha_1^* + ... + p_j\beta_0\beta_1^* \\ p_i\alpha_1\alpha_0^* + ... + p_j\beta_1\beta_0^* & p_i\alpha_1\alpha_1^* + ... + p_j\beta_1\beta_1^* \end{pmatrix}^\dagger \\ &= \begin{pmatrix} (p_i\alpha_0\alpha_0^* + ... + p_j\beta_0\beta_0^*)^* & (p_i\alpha_1\alpha_0^* + ... + p_j\beta_1\beta_0^*)^* \\ (p_i\alpha_0\alpha_1^* + ... + p_j\beta_0\beta_1^*)^* & (p_i\alpha_1\alpha_1^* + ... + p_j\beta_1\beta_1^*)^* \end{pmatrix} \\ &= \begin{pmatrix} p_i\alpha_0\alpha_0^* + ... + p_j\beta_0\beta_0^* & p_i\alpha_0\alpha_1^* + ... + p_j\beta_0\beta_1^* \\ p_i\alpha_1\alpha_0^* + ... + p_j\beta_1\beta_0^* & p_i\alpha_1\alpha_1^* + ... + p_j\beta_1\beta_1^* \end{pmatrix}\end{aligned}$$

and thus $\rho = \rho^\dagger$. From subsubsection 2.2.1 this gives us that the eigenvalues of a density matrix are real.

We also see that for any state $|\psi_j\rangle$, $\langle\psi_j|\rho|\psi_j\rangle = \sum_{i=1}^n p_i|\langle\psi_j|\psi_i\rangle|^2$. Since inner products are positive-definite, and $p_1, ..., p_n$ are all probabilities and thus ≥ 0, we have that $\langle\psi_j|\rho|\psi_j\rangle \geq 0$, making density matrices *positive-semidefinite*. This means that all the eigenvalues of any density matrix ρ are non-negative.

4.6 Qubits

Conventional computers utilize bits, which are physical components (e.g., capacitors, transistors, magnetic disk, etc. depending on the type of memory) that can be electromagnetically charged to be in one of two binary states: 0 or 1. However, quantum computers use qubits (*qu*antum *bits*). Qubits are particles that can be in any superposition of the computational basis. There are many different methods for physically implementing qubits. The most common one today utilizes superconducting; the electronic circuits used for this must be cooled to incredibly low temperatures (< 0.1Kelvin).[19] Because our computational basis is $\{|0\rangle, |1\rangle\}$, any qubit will have its state represented as a superposition of these states (since you will recall any element of a vector space can be defined as a linear combination of the elements in the space's basis). So the general state $|\psi\rangle$ of a qubit is given by $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ (where of course $|\alpha|^2 + |\beta|^2 = 1$). Measuring the value of a qubit is done at the very end of a quantum algorithm, because again, measurement of quantum systems is not gentle (see subsubsection 3.4.4 and subsection 4.2). Since we measure qubits in the computational basis, a qubit in the general $|\psi\rangle$ state has a $|\langle 0|\psi\rangle|^2 = |\alpha|^2$ probability of being found in the $|0\rangle$ state and a $|\langle 1|\psi\rangle|^2 = |\beta|^2$ probability of being found in the $|1\rangle$ state. To reiterate the terminology used earlier, α is the probability amplitude for the qubit being in the $|0\rangle$ state and β is the probability amplitude for the qubit being in the $|1\rangle$ state. Through this, and through our defintion of elements of the $\mathbb{C}_1^2$ space in subsection 2.2, we notice that the vector representation of states in the computation basis—$|\psi\rangle = \alpha\begin{pmatrix}1\\0\end{pmatrix} + \beta\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}\alpha\\0\end{pmatrix} + \begin{pmatrix}0\\\beta\end{pmatrix} = \begin{pmatrix}\alpha\\\beta\end{pmatrix}$—is an easy way to see the probability amplitudes of a qubit in that state.

Because unitary operators preserve inner product, they also preserve probability amplitudes. That is, for each possible state, the output vector has the same probability of being measured in that state transformed by the same operator as the input vector did for the original state. As it's formally written in subsection 2.3, $\langle Ux|Uy\rangle = \langle x|y\rangle$. Quantum systems can only evolve across time from one state to another through the application of unitary operators.[31] This principle is derived from the solution to the Schrödinger Equation, a partial differential equation that explains the way waves—which are how all matter behaves at the quantum level (see: subsection 4.7)—evolve over time. This means that the state of quantum computers has to be manipulated through the application of unitary operators. So a unitary operator is to a quantum computer as a logic gate is to a conventional computer. Thus, unitary operators are referred to as quantum logic gates in the context of quantum computing. For a single-qubit system, a quantum logic gate is a unitary operator with the shape 2×2. A table of common quantum logic gates are found in subsection 8.1.

With n qubits, a quantum computer can be in any superposition of 2^n states. For example, $n = 1$ qubit can be in a superposition of the $2^1 = 2$ states $|0\rangle$ and $|1\rangle$. A quantum computer with $n > 1$ qubits represents its state as a superposition of tensor products of n states in the computational basis. Compare this to a classical computer, whose binary bits can collectively exist in only one of the 2^n possible states at a time. This ability to store exponentially more information than conventional computers is useful in many quantum algorithms. However, a common misconception is that this ability allows quantum computers to try multiple different solutions to a problem in parallel—this is not quite the case. Quantum parallelism allows gates to perform operations on multiple potential solution states at the same time when they are applied on superpositions, and then, as Dr. Scott Aaronson puts it, quantum algorithms "choreograph an interference pattern, where the unwanted [outcome states] cancel."[2][23] What this means is that in a quantum algorithm you start with the quantum computer in some initial state (usually with all the qubits in the $|0\rangle$ state) and apply a series of quantum logic gates in order to increase the probability that the computer (really the measurement made on the tensor product of all its qubits—see subsection 4.8) is in the state that solves the problem at hand. Measurement will be covered in subsection 4.7 and interference in subsection 4.9.

4.7 Wave-Particle Duality & Observables

The double-slit experiment was first performed by Thomas Young in 1801 and demonstrated the wave behavior of light. The setup involves a coherent light source (e.g., a laser beam) emitting at a screen which has two slits in it and a detector placed a distance behind that screen. The detector observes the intensity of light at each location (the relevant dimension of the location being the offset from the slits in the direction perpendicular to that the light is emmited in) where the intensity measures how much of the light is hitting a given location. If light were to behave only as a particle then it would pass through one slit or the other—possibly bouncing off the walls within the slit to redirect—and we would expect the greatest intensities to be directly behind the slits (offset of 0). However, the greatest intensity is in the middle of the two slits, and the overall manner with which the light illuminates the detector is consistent with the interference pattern that would be generated from two waves originating at the slits. This indicates that the light is actually traveling as a wave, goes through both slits, and diffraction at the two slits creates two new wavefronts that interfere with each other as they travel towards the detector. Where the two waves happen to meet in phase there is constructive interference causing the light to become more intense, and where the waves meet out of phase there is destructive interference such that the waves cancel out and the light is dimmer.

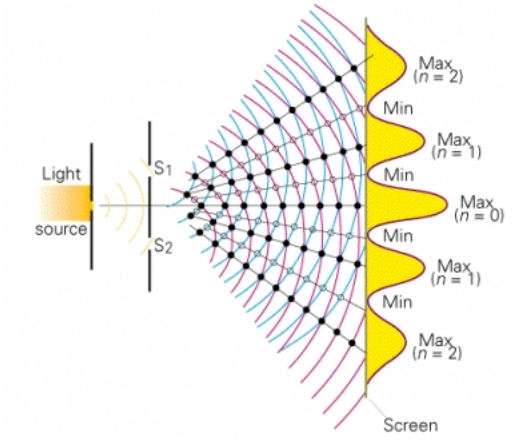

Figure 7: Interference pattern of the two light waves in the double slit experiment. Solid circles represent points of maximal constructive interference, circle outlines represent points of maximal destructive interference and diminished light.

About a hundred years later light was shown both theoretically and experimentally (notably with Einstein's famous $E = mc^2$ equation and Arthur Compton's discovery of the Compton Effect, respectively) to also behave as a particle.[29] These particles of light are discrete quanta that carry the electromagnetic force known as photons. This led to the wave-particle duality understanding of light, in which it exhibits properties of both. Louis de Broglie extended this duality to all matter in 1924 with his wave equation $\lambda = \frac{h}{mv}$ that gives a formula for any mass's wavelength λ. It was first demonstrated in the Davisson–Germer experiment three years later. In this equation, h is the incredibly small Planck constant, and so when there's mass such that $m \gg 0$, then the wavelength, and thus wave properties, become imperceptible. This is what occurs at the scale of classical mechanics and the everyday world. The wave-particle duality that is observable at the quantum level due to the the small mass of particles studied is essential to the way in which quantum mechanics differs from classical mechanics.

Quantum systems evolve over time as a wave—this is what the wavefronts of light are doing in the double slit experiment. Describing this evolution is the Schrödinger Equation that was mentioned in subsection 4.6. The classical counterpart for this is Newton's second law that can be used to predict how a system of particles will move over time. When the light is measured, however, it is observed to behave like a particle; the photon strikes the detector at exactly one point. Modern experimental design also enabled recreation of the double-slit experiment where single photons are emitted at a time and then their location when hitting the detector is measured. Repeating this single shot process many times yields a distribution of where the photons strike the detector, and the same interference pattern occurs! This indicates that individual photons are going through both slits simultaneously and then interfering with themselves, further proving their wave nature prior to observation. The role of the observer here is key as was discussed in subsubsection 3.4.4 and we will now formalize the phenomenon that occurred therein.

In physics, physical properties that can be measured (e.g., position, momentum, spin) are known as observables. In quantum mechanics we apply operators—specifically Hermitian matrices—to a system (e.g., qubits in quantum computing or spinning particles when using the spin operators of the Stern-Gerlach experiment) to measure these observables. For each different observable we may be interested in there is an operator to use. There are also other types of operators that do not measure observables when applied. The set of potential values that an observable can be measured as is the set of eigenvalues of the operator used, which are all necessarily real since the operators for measuring observables are Hermitian matrices. When measuring a particle the application of the operator leaves it in a new superposition of eigenvectors of the operator. The outcome of the measurement is probabilistic, where the observed value will be the eigenvalue associated with a given eigenvector in the superposition with probability equal to the square of the norm of the coeffecient—the probability amplitude—on that given eigenvector. Once we have the measurement, the wave is said to have "collapsed" and is no longer in the superposition but has the state of the eigenvector associated with the measurement; this is known as the projection postulate of quantum mechanics.[21][11] Applying this to the example of subsubsection 3.4.4, consider the Pauli-Z operator $\sigma_z = |z^+\rangle\langle z^+| - |z^-\rangle\langle z^-| = |1\rangle\langle 1| - |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. This is the spin operator for the z-axis, and so applying it allows us to observe a particle's spin in the z-direction as is done in the Stern-Gerlach experiment; note that the operator is unitary and can thus be used as a quantum logic gate. It in fact often is, called the Pauli-Z Gate, and you can learn more about it subsection 8.1.

We will go through some examples of applying operators to mathematically formalize what is occurring in the experiments described in subsection 3.4. Assume we have a particle in the $|0\rangle$ state, as covered in subsection 4.3 we then know that it has $\lambda_z = 1$ spin. Now say we want to measure this particle's spin in the z-direction, to do so we apply the σ_z operator:

$$\sigma_z|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (1)|0\rangle.$$

We thus see that $|0\rangle$ is an eigenvector of σ_z with the associated eigenvalue of 1, so the observable (spin in the z-direction) is measured with value of 1 as we expect.

Next assume we have a particle with state $|1\rangle$ and thus $\lambda_z = -1$ spin. We see:

$$\sigma_z|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = (-1)|1\rangle.$$

So $|1\rangle$ is also an eigenvector with associated eigenvalue -1. The spin of the particle in the z-direction, as measured via application of the σ_z operator, is again as expected from our construction of the computational basis.

Finally consider a particle in the $|+\rangle$ state:

$$\sigma_z|+\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}[(1)|0\rangle + (-1)|1\rangle].$$

This is the first example we have seen where applying the operator results in a superposition of eigenvectors rather than a lone eigenvector. The eigenvector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ of the operator has eigenvalue 1 and probability for being found (as the resulting spin state of the particle after the measurement) of $|\frac{1}{\sqrt{2}}|^2 = \frac{1}{2}$; the eigenvector $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$ has eigenvalue -1 with probability also $\frac{1}{2}$. This means, for a particle in the $|+\rangle$ state which has its spin measured in the z-direction by the σ_z operator, there is a $\frac{1}{2}$ chance it will collapse into the $|0\rangle$ state with $\lambda_z = 1$ spin and a $\frac{1}{2}$ chance it will collapse into the $|1\rangle$ state with $\lambda_z = -1$ spin. The particle now is in the state is has collapsed to, and a repeated measurement of spin in the z-direction (i.e., applying the σ_z operator) will result in the same spin again as shown mathematically directly above and demonstrated experimentally in subsubsection 3.4.1.

4.8 Tensor Product Spaces & Multiple Qubits

So far we've been discussed the states of one qubit at a time. However, it is necessary to think about the state of many qubits simultaneously. Doing so requires the use of tensor product spaces which allow us to consider the quantum state of systems with multiple particles. The tensor product of two vectors, say $|x\rangle \in V$ and $|y\rangle \in W$ (where V and W are both vector spaces with n and m dimensions, respectively), is written $|x\rangle \otimes |y\rangle$ and is an nm-dimensional vector. This vector is an element of the $V \otimes W$ tensor product space. To construct this space we take the bases of V and W, sets B_V and B_W respectively, and use their tensor product $(B_V \otimes B_W)$ as the basis of the new space. The tensor product symbol $\otimes$ is also often elided, and $|x\rangle \otimes |y\rangle$ is rewritten as $|xy\rangle$.

But what actually is $|xy\rangle$? It is an ordered list of the product of every combination of the elements in the $|x\rangle$ and $|y\rangle$ vectors. As a concrete example, see $|0\rangle \otimes |1\rangle = |01\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0 \\ 1 \cdot 1 \\ 0 \cdot 0 \\ 0 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

Like with a single qubit, this vector gives us the probability amplitudes for finding the two-qubit system in each of the elements of the basis of the tensor product of the computational basis with itself. That is, a two-qubit system with the state $\begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix}$ has a $|\alpha|^2$ probability of being found in the $|00\rangle$ state, a $|\beta|^2$ probability of being found in the $|01\rangle$ state, a $|\gamma|^2$ probability of being found in the $|10\rangle$ state, and a $|\delta|^2$ probability of being found in the $|11\rangle$ state. So we see in the example above that a two qubit system in the $|01\rangle$ state has a probability of $|1|^2 = 1$ for being found in the $|01\rangle$ state when measured,

which is of course what we would expect. Naturally, this representation of the probability amplitudes is more useful when the two-qubit system is in a superposition and thus the vector is not a one-hot vector.

4.9 Interference & the Hadamard Gate

As was mentioned in subsection 4.6, a crucially important part of quantum computing is cancelling out certain states through interference. Suppose we have a qubit in the $|+\rangle$ state which we have seen is an even (probabilistically) superposition of the basis vectors of the computational basis. We can use the Hadamard gate to apply intereference to this superposition such that we are left with the basis state $|0\rangle$.[26] The Hadamard gate is a unitary operator represented as H, more details can be found in Appendix subsection 8.1.

$$H|+\rangle = H\begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix} = H\tfrac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \tfrac{1}{\sqrt{2}}([H|0\rangle] + [H|1\rangle]) =$$
$$\tfrac{1}{\sqrt{2}}\left(\left[\tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}|0\rangle\right] + \left[\tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}|1\rangle\right]\right) = \tfrac{1}{2}\left(\left[\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix}\right] + \left[\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix}\right]\right) =$$
$$\tfrac{1}{2}\left(\left[\begin{pmatrix}1\\ 1\end{pmatrix}\right] + \left[\begin{pmatrix}1\\ -1\end{pmatrix}\right]\right) = \tfrac{1}{2}([|0\rangle + |1\rangle] + [|0\rangle - |1\rangle]) = |0\rangle.$$

As $|+\rangle$ is in a superposition of $|0\rangle$ and $|1\rangle$, applying H creates destructive interference between the two states $H|0\rangle$ and $H|1\rangle$ such that $|1\rangle$ as a potential outcome (i.e., part of the superposition with probability amplitude > 0) is canceled out and there is constructive interference with the parts of the two states that yield the $|0\rangle$ outcome.[26] Another way to think about this is by inspecting the third to last state where we see $\frac{1}{2}\left(\left[\begin{pmatrix}1\\ 1\end{pmatrix}\right] + \left[\begin{pmatrix}1\\ -1\end{pmatrix}\right]\right)$. In this form we see that the $|1\rangle = \begin{pmatrix}0\\ 1\end{pmatrix}$ portions of the superposition are perfectly out of phase causing them to cancel out.[15] This results in the deterministic $|0\rangle$ state for $H|+\rangle$.

We can see this same type of interference occur when we have multiple qubits:

$$H|++\rangle = H|+\rangle \otimes H|+\rangle = |0\rangle \otimes |0\rangle = |00\rangle.$$

Note that $|++\rangle = |+\rangle \otimes |+\rangle = \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix} \otimes \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix} = \begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\\ \frac{1}{2}\\ \frac{1}{2}\end{pmatrix} = \begin{pmatrix}\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2}\end{pmatrix}^T$. Recall, the first entry in this ket means that the probability amplitude for the state $|00\rangle$ is $\frac{1}{2}$ and thus the probability for finding $|++\rangle$ in the $|00\rangle$ state when measured in the computational basis is $|\frac{1}{2}|^2 = \frac{1}{4}$.

5 Quantum Computer Hardware

Because the Hadamard gate is its own inverse, $H|00\rangle$ gets us back to $|++\rangle$. Qubits in a quantum computer at the start of an algorithm are often all in the $|0\rangle$ state; for a general quantum computer with n bits this is notationally written as being in the $|0\rangle^{\otimes n}$ state. For example, if there are 3 qubits then the computer is in the $|000\rangle = |0\rangle \otimes |0\rangle \otimes |0\rangle = |0\rangle^{\otimes 3}$ state. Multiple qubits are kept together in a register; the qubits can all be organized into one or multiple different registers depending on the needs of the algorithm being run.[23][10] Registers and qubits make up the *quantum data plane*.[3][27]

The *control processor plane* implements the quantum algorithm by triggering the series of quantum gates and measurements to be applied. The *control and measurement plane* takes the digital signals that are representing these gates and measurements and converts them to analog signals which are applied to the quantum data plane; these signals carry out the operation on the qubits.[3][27]

The *host processor* is a classical computer that allows a developer to interface with the quantum control processor. Running software on a host processor such as IBM's Qiskit allows one to write quantum algorithms. This algorithm is then sent to the control processor plane that executes the algorithm via the control and measurement plane operating on the data plane.[3][27]

6 Grover's Algorithm

Grover's algorithm is used to solve the so-called *unstructured search* problem. The setup for this problem is that we have N elements $\{x_0, x_1, ..., x_{N-1}\} \subseteq \{0,1\}^n$, where in the largest case $N = 2^n$ for n bits as that is how many unique elements are in $\{0,1\}^n$. We have a function $f : \{0,1\}^n \to \{0,1\}$ where there is exactly one x_i such that $f(x_i) = 1$ and for all other x_j such that $j \neq i$ we have $f(x_j) = 0$; we want to find x_i.[15][8] Because the elements $\{x_0, x_1, ..., x_{N-1}\}$ are binary strings with length n we use a register with n qubits; this makes it so that when the final superposition after running the algorithm is measured it collapses to a tensor product of the computational basis. This tensor product is an n-length binary string that represents x_i. There are variants of the problem where there can either be multiple $x_i \in \{x_0, x_1, ..., x_{N-1}\}$ such that $f(x_i) = 1$ or no solution x_i at all. In this version the goal is to find any x_i such that $f(x_i) = 1$, and if there are no such x_i we must determine that.[24] There is no inherent structure, organization, or sorting of the elements that can be leveraged to accelerate finding x_i, hence *unstructured*. The function f is called the *oracle* as is tells us whether we have the right answer or not and it is assumed that consulting the oracle with $f(x)$ takes a single operation. Algorithms that use classical computers (so-called *classical algorithms*) to solve unstructured search have runtime $O(N)$ because they must brute-force search through the elements in a linear fashion.[18] Alternatively, Grover's Algorithm has runtime $O(\sqrt{N})$ which provides a quadratic speedup.

6.1 Theory

A good high level overview of the theory behind the algorithm is given by IBM,

> Grover's algorithm takes an iterative approach: it evaluates f on superpositions of [multiple inputs at a time] and intersperses these evaluations with other operations that have the effect of creating interference patterns, leading to a solution with high probability (if one exists) after $O(\sqrt{N})$ iterations.[24]

As Grover himself puts it,

> By properly adjusting the phases of various operations, successful computations reinforce each other while others interfere randomly. As a result, the desired $[x_i]$ can be obtained in only $O(\sqrt{N})$ steps.[15]

The initial step of the algorithm is applying the Hadamard gate H to the register (assuming it starts out in the $|0\rangle^{\otimes n}$ state) rendering the qubits in a superposition with probability amplitudes such that it is a uniform distribution across all possible $N = 2^n$ states of the computational basis. Recall that as mentioned previously, $H|0\rangle^{\otimes n} = |+\rangle^{\otimes n}$.

Next is the core routine of the algorithm which involves a phase query gate followed by the diffusion operator.[15] The phase query gate applies a phase inversion on the probability amplitude that corresponds to the desired state x_i. The diffusion operator inverts all of the amplitudes around the mean of all amplitudes. We repeat this two step routine $\sqrt{N}$ times.

6.1.1 Phase Inversion

The algorithm uses a phase query gate which is an operator that negates the amplitude on the entry corresponding to the solution binary string x_i and leaves the other amplitudes untouched. We refer to

the gate as U_f; in the context of this algorithm it is also called the Grover oracle as it depends on the oracle function f. For any $x \in \{0,1\}^n$,

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

If $x = x_i$, then $(-1)^{f(x)} = (-1)^1 = -1$ and $|x\rangle$ has its phase inverted such that $U_f|x\rangle = -|x\rangle$. If $x \neq x_i$, then $(-1)^{f(x)} = (-1)^0 = 1$ and $U_f|x\rangle = |x\rangle$.

This gate provides a good example for understanding quantum parallelism as mentioned in subsection 4.6.[23] Suppose $x \notin \{0,1\}^n$, but rather $n = 2$ and $|x\rangle = |++\rangle$ with $x_i = 01$. In this case,

$$\begin{aligned}
U_f|x\rangle &= U_f|++\rangle \\
&= U_f \tfrac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \\
&= \tfrac{1}{2}(U_f|00\rangle + U_f|01\rangle + U_f|10\rangle + U_f|11\rangle) \\
&= \tfrac{1}{2}((-1)^{f(00)}|00\rangle + (-1)^{f(01)}|01\rangle + (-1)^{f(10)}|10\rangle + (-1)^{f(11)}|11\rangle) \\
&= \tfrac{1}{2}(1|00\rangle + -1|01\rangle + 1|10\rangle + 1|11\rangle) \\
&= \tfrac{1}{2}(|00\rangle - |01\rangle + |10\rangle + |11\rangle) \\
&= \begin{pmatrix}\frac{1}{2} & \frac{-1}{2} & \frac{1}{2} & \frac{1}{2}\end{pmatrix}^T.
\end{aligned}$$

We see that for the solution state x_i we are interested in finding the phase has been inverted.

6.1.2 Inversion About the Mean

Inversion about the mean is performed with the diffusion operator D. $D = -I + 2P$ where P is a $N \times N$ matrix whose entries are all $\frac{1}{N}$. For an N-dimensional ket $|x\rangle$, $P|x\rangle$ is an N-dimensional ket $|\overline{x}\rangle$ (using $\overline{x}$ in this context for the mean and not complex conjugate) with all identical entries which are the average $\overline{x}$ of the entries in $|x\rangle$. Thus, $D|x\rangle = 2P|x\rangle - I|x\rangle = 2|\overline{x}\rangle - I|x\rangle = |\hat{x}\rangle$ where for each entry $\hat{x}_k$ in $\hat{x}$, $\hat{x}_k = 2\overline{x} - x_k$. Prior to applying this transformation D, the entries of $|x\rangle$ differ from their collective mean by $x_k - \overline{x}$. The entries of $D|x\rangle = |\hat{x}\rangle$ differ by $(2\overline{x} - x_k) - \overline{x} = \overline{x} - x_k$. Thus, we see that applying D means that the absolute value of the difference between the entries of $D|x\rangle$ and the mean of $|x\rangle$ are the same, but the values have been inverted about the mean such that the difference is multiplied by -1; note that the mean of $D|x\rangle$ is the same as the mean of $|x\rangle$.

When it comes to actually implementing D as a gate, we use the fact that $D = HRH$ where $R = -I + 2|0\rangle\langle 0|$ and is a rotational transform which flips the phase of every bit except the first one which represents the $|0\rangle^{\otimes n}$ state.[24] Furthermore,

$$\begin{aligned}
R &= -I + 2|0\rangle\langle 0| \\
&= \begin{pmatrix} -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} + \begin{pmatrix} 2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
&= -1 \begin{pmatrix} -1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

This last step factors out a global phase flip. In quantum computing, only the relative phase between qubits matters so we can ignore this.[10] Without this coefficient, the transform that need be applied is one that flips the phase of just the $|0\rangle^{\otimes n}$ state. If we apply a NOT gate (the Pauli-X operator) to all the qubits then the phase we want to flip becomes the $|1\rangle^{\otimes n}$ state, which is exactly what the Pauli-Z gate (see: subsection 8.1) does. We then apply another NOT gate to go back so that it is the $|0\rangle^{\otimes n}$ state that has had its phase flipped. This succeeds in achieving the transform defined by R using X and Z gates. Summarily, we can implement the inversion about the mean with Hadamard (H), NOT (X), and Pauli-Z (Z) gates.

6.1.3 Runtime

Grover's algorithm has been abstracted to a general technique called *amplitude amplification* which is used when for problems that can be solved by amplifying the amplitude of desired "good" states (such as x_i in unstructured search) and decreasing the amplitude of unwanted "bad states". There is a geometric understanding of these problems that I will not get into at any depth, but suffice it to say that you start with some initial state ($H|0\rangle^{\otimes n}$ in Grover's) and each iteration up to a point projects the state closer to a subspace of good states by amplifying the amplitudes of these good states and decreasing other amplitudes. It has been shown that the optimal amount of iterations for running an amplitude amplification algorithm, when there are N possible states and the initial state has a uniform distribution for being found in any given one of these states (as is the case with Grover's), is given by $\left\lfloor \frac{\pi}{4 \sin^{-1} \sqrt{\frac{1}{N}}} \right\rfloor$ (the bracket-like objects around the equation are the floor function).[5] The amplitude amplifying piece in Grover's is the application of the two inversions DU_f. Note that $\frac{1}{\sin^{-1} \sqrt{\frac{1}{N}}} \approx \sqrt{N}$, hence why the runtime of Grover's algorithm is $O(\sqrt{N})$.

6.2 Example

We will look at an example to better understand how the combined power of phase inversion and inversion about the mean enable unstructured search. Consider the case where we have $n = 3$ qubits so $N = 2^3 = 8$; this example is borrowed from Dr. Anastasios Kyrillidis.[22] The 8 elements are $\{000, 001, 010, 011, 100, 101, 110, 111\}$, assume $x_i = 100$. The first step of the algorithm yields $H|000\rangle = |{+}{+}{+}\rangle = (\frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}})^T$.

Next, we invert the phase of the desired state with U_f which yields $(\frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{-1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}}\ \frac{1}{\sqrt{8}})^T$. The mean across the entries of this ket is $\frac{3}{4\sqrt{8}}$, so applying D begets $(\frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{5}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}})^T$.

With $N = 8$ the optimal number of iterations is $\lfloor \frac{\pi}{4}\sqrt{8} \rfloor = 2$, so we will apply the inversion operators once more. $DU_f(\frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{5}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}})^T = D(\frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{-5}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}}\ \frac{1}{2\sqrt{8}})^T = (\frac{-1}{4\sqrt{8}}\ \frac{-1}{4\sqrt{8}}\ \frac{-1}{4\sqrt{8}}\ \frac{-1}{4\sqrt{8}}\ \frac{11}{4\sqrt{8}}\ \frac{-1}{4\sqrt{8}}\ \frac{-1}{4\sqrt{8}}\ \frac{-1}{4\sqrt{8}})^T$. At this point, we would measure the state of the quantum computer; $|\frac{11}{4\sqrt{8}}|^2 \approx 0.9453$ which means there is a 94.53% chance of observing the computer in the solution state $|100\rangle$.

We see that jointly applying phase inversion and inversion about the mean increases the probability amplitude of the desired state x_i and dampens all other probability amplitudes of bad states.

6.3 Implementation

Login at https://quantum-computing.ibm.com creating an IBMid if you do not already have one. Then, while logged in, note that your API token is available to copy in the upper right corner of this page (the IBM Quantum Platform dashboard). If it is not there, you can click on the user icon

in the top right corner, select Manage Account, and then the token will be available under Account Overview > API Token. You need this token to access IBM's quantum computers and simulators. The token can be copied to clipboard by clicking the copy icon ⧉.

Qiskit is an open-source software development kit (SDK) that enables writing code for quantum computers in python. I have implemented Grover's Algorithm using it in a Colab notebook available at https://colab.research.google.com/drive/1pjp4IYr9yRQqdHSF53U3fJIaZzqBvpro?usp=sharing. This work includes code from learning materials created by IBM Quantum Learning.[25] You can open this and go to File > Save a copy in Drive to get your own copy of this notebook. Then, paste in your API token in the setup cell and click Runtime > Run all to run the algorithm on an IBM quantum computer. Comments throughout the notebook explain details relevant to implementation.

7 References

[1] Scott Aaronson. UT Austin, Lecture Notes: Introduction to Quantum Information Science. 2018. URL: https://www.scottaaronson.com/qclec.pdf.

[2] Scott Aaronson. When Exactly Do Quantum Computers Provide A Speedup? Yale Quantum Institute Seminar. 2014. URL: https://www.scottaaronson.com/talks/speedup.ppt.

[3] Amazon Web Services. What Is Quantum Computing? 2023. URL: https://aws.amazon.com/what-is/quantum-computing/.

[4] Peter Atkins and Julio de Paula. Atkins' Physical Chemistry. 8th ed. 2006, pp. 308–309. URL: https://www.rnlkwc.ac.in/pdf/study-material/chemistry/Peter_Atkins__Julio_de_Paula__Physical_Chemistry__1_.pdf.

[5] Gilles Brassard et al. "Quantum Amplitude Amplification and Estimation". In: Contemporary Mathematics 305 (2002): Quantum Computation and Information. Ed. by Samuel J. Lomonaco Jr. and Howard E. Brandt. URL: https://arxiv.org/abs/quant-ph/0005055.

[6] Daniel Bump. Group Representation Theory. 2010. URL: http://sporadic.stanford.edu/bump/group/gind2_1.html.

[7] CERN. The Standard Model. 2023. URL: https://home.web.cern.ch/science/physics/standard-model.

[8] christianb93. Grover's algorithm – unstructured search with a quantum computer. LeftAsExercise. Oct. 29, 2018. URL: https://leftasexercise.com/2018/10/29/grovers-algorithm-unstructured-search-with-a-quantum-computer/.

[9] Charles Francis. "The Hilbert space of conditional clauses". In: International Journal of Quantum Information 13.03 (2015). URL: https://arxiv.org/abs/1205.4607v2.

[10] Stuart Geman et al. Brown University APMA 1930W, Lecture Notes: Probabilities in Quantum Mechanics. 2022.

[11] Giancarlo Ghirardi and Angelo Bassi. Collapse Theories. The Stanford Encyclopedia of Philosophy. May 15, 2020. URL: https://plato.stanford.edu/archives/sum2020/entries/qm-collapse/.

[12] Jonathan Goodman. NYU, Lecture Notes: Scientific Computing. 2022. URL: https://math.nyu.edu/~goodman/teaching/ScientificComputing2022/materials/LinearAlgebra2.pdf.

[13] Michael A. Gottlieb and Rudolf Pfeiffer. The Feynman Lectures. California Institute of Technology. 2013. URL: https://www.feynmanlectures.caltech.edu/.

[14] Robert B. Griffiths. CMU 33-658, Lecture Notes: Hilbert Space Quantum Mechanics. 16 January 2014. URL: https://quantum.phys.cmu.edu/QCQI/qitd114.pdf.

[15] Lov K. Grover. "A fast quantum mechanical algorithm for database search". In: STOC. 1996. URL: https://arxiv.org/pdf/quant-ph/9605043.pdf.

[16] Micael Hartl. The Tau Manifesto. 2022. URL: https://tauday.com/tau-manifesto.

[17] John K. Hunter and Bruno Nachtergaele. Applied Analysis. 2005. URL: https://www.math.ucdavis.edu/~hunter/book/pdfbook.html.

[18] Ataç İmamoğlu and Renato Renner. Quantum Information Processing Solutions 6. ETH Zurich. 2015. URL: https://edu.itp.phys.ethz.ch/15FS/qip/sol06.pdf.

[19] The Quantum Insider. A Detailed Review of Qubit Implementations for Quantum Computing. May 21, 2020. URL: https://thequantuminsider.com/2020/05/21/tqd-exclusive-a-detailed-review-of-qubit-implementations-for-quantum-computing/.

[20] Theresa Knott. Stern-Gerlach Experiment. URL: https://upload.wikimedia.org/wikipedia/commons/2/29/Stern-Gerlach_experiment.PNG.

[21] Pieter Kok. 2: The Postulates of Quantum Mechanics. 2023. URL: https://phys.libretexts.org/Bookshelves/Quantum_Mechanics/Advanced_Quantum_Mechanics_(Kok)/02%3A_The_Postulates_of_Quantum_Mechanics.

[22] Anastasios Kyrillidis. Introduction to quantum computing: Grover's algorithm. Sept. 11, 2023. URL: https://akyrillidis.github.io/notes/quant_post_9.

[23] Ryan LaRose. QuIC Seminar 9: Introduction to the Query Model of Quantum Computing. Michigan State University. URL: https://www.ryanlarose.com/uploads/1/1/5/8/115879647/quic09-dj-algorithm.pdf.

[24] IBM Quantum Learning. Grover's Algorithm. 2023. URL: https://learning.quantum-computing.ibm.com/course/fundamentals-of-quantum-algorithms/grovers-algorithm.

[25] IBM Quantum Learning. Grover's algorithm using the Sampler primitive. 2023. URL: https://learning.quantum-computing.ibm.com/tutorial/grovers-algorithm-using-the-sampler-primitive.

[26] MITRE. Quantum Interference. July 1, 2022. URL: https://stem.mitre.org/quantum/quantum-concepts/quantum-interference.html.

[27] National Academies of Sciences, Engineering, and Medicine. Quantum Computing: Progress and Prospects. Ed. by Emily Grumbling and Mark Horowitz. 2019. Chap. 5: Essential Hardware Components of a Quantum Computer. URL: https://nap.nationalacademies.org/read/25196/chapter/7.

[28] MIT Department of Physics. The Stern-Gerlach Experiment. Oct. 13, 2004. URL: https://web.mit.edu/8.13/www/JLExperiments/JLExp18.pdf.

[29] Glenn Stark. Quantum Theory of Light. Encyclopaedia Britannica. 2023. URL: https://www.britannica.com/science/light/Quantum-theory-of-light.

[30] Victor J. Stenger. What exactly is the 'spin' of subatomic particles such as electrons and protons? Scientific American. 1999. URL: https://www.scientificamerican.com/article/what-exactly-is-the-spin/.

[31] Leonard Susskind and Art Friedman. Quantum Mechanics: The Theoretical Minimum. Basic Books, 2014.

[32] Barton Zwiebach. MIT 8.05, Lecture Notes. 2013. URL: https://ocw.mit.edu/courses/8-05-quantum-physics-ii-fall-2013/pages/lecture-notes/.

8 Appendix

8.1 Common Quantum Logic Gates

Recall that a 2×2 matrix which is a unitary operator can be applied to manipulate the state of single qubit.

Unitary Operator (Gate)	Symbol	Icon in IBM Quantum Composer	Matrix	Notes
NOT	$\boxed{X}$		$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	This is the Pauli-X operator, also known as σ_x. This gate rotates $\frac{\tau}{2}$ radians around the x-axis, so it swaps the computational basis elements (i.e., $\sigma_x\|0\rangle = \|1\rangle$ and $\sigma_x\|1\rangle = \|0\rangle$) hence its name.
Hadamard	$\boxed{H}$		$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$	This gate rotates $\frac{\tau}{4}$ radians around the y-axis and $\frac{\tau}{2}$ radians around the x-axis. That means $H\|0\rangle = \|+\rangle$ where $\|+\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ which correlates to $\|x^+\rangle$ (positive along the x-axis) and $H\|1\rangle = \|-\rangle$ where $\|-\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix}$ which correlates to $\|x^-\rangle$.
Pauli-Z	$\boxed{Z}$		$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	The Pauli-Z operator, also seen as σ_z, measures spin in the z-direction. Because of this, it also can be used to invert/flip the phase of $\|1\rangle$ qubits, since $\sigma_z\|1\rangle = -\|1\rangle$ whereas $\sigma_z\|0\rangle = \|0\rangle$.