Introduction to Quantum Information Theory

Joshua Parmenter

19 Jan, 2021

Abstract

Quantum Information is a field that is becoming more and more relevant as computers get faster and more intelligent. This document is designed to be a fast paced, rapid introduction to the concepts necessary to comprehend the math and theories behind the creation, editing, and analysis of quantum bits, especially with a focus on sending and receiving information, or processing large or mathematically intensive computations.

Contents

1 Vectors and Vector Spaces

1.1 Vector Spaces

In modern mathematics, an area of importance is first defining what field the problems are being asked in. When speaking on vectors and vector spaces, it is easiest to work in two main fields of numbers.

1.1.1 Relevant Vector Spaces

1. $\mathbb{R}$: **All Real Numbers** $(\infty, -\infty)$

   This includes all real numbers from zero to infinity, something of note is that this number field is **not algebraically closed**, meaning not all numbers can be produced from combinations of other numbers in the field. An example of this is $x^2 + 1 = 0$, meaning $x = 1 * i$

2. $\mathbb{C}$: **All Complex Numbers**

   This number field **is algebraically closed**. This means that every possible number in the field can be produced from combinations of other numbers in the field. This makes sense as two complex numbers multiplied together can produce 1. A completely real number 2. A complex number with a real and imaginary component 3. A complex number with only an imaginary component.

1.2 Complex Number Operations

In the following sections, we will make the assumption $a, b, c, d, e \in \mathbb{R}$

1.2.1 Addition

$$(a + bi) + (c + di) = (a + c) + (b + d)i \tag{1.1}$$

1.2.2 Multiplication

$$(a + bi) * (c + di) = (ac + adi) + (bci - bd) \tag{1.2}$$

1.2.3 Conjugation

$$\overline{a + bi} = (a + bi)^* = a - bi \tag{1.3}$$

1.2.4 Modulus

$$|z| = |a + bi| = \sqrt{a^2 + b^2} \tag{1.4}$$

1.3 Common Vector Space Notation and Definitions

This section is to lay out some of the most important pieces of notation and definitions for the complex mathematics required for this course.

1.3.1 Linear Algebra Terms

Definition 1.1 An element $|v\rangle \in \mathbb{C}^n$ is called a **(ket) vector** and is expressed as a **column** of n complex numbers. The integer n is called the **dimension** of the vector space $\mathbb{C}^n$. ⧫

Definition 1.2 A **linear combination** of $|v_1\rangle, ..., |v_k\rangle \subset \mathbb{C}^n$ is just a single vector in the form $\lambda_1 |v_1\rangle + \lambda_2 |v_2\rangle + ... + \lambda_k |v_k\rangle$ for some $\lambda_1, \lambda_2, ..., \lambda_k \in \mathbb{C}$. ⧫

Example 1.3 Create a **linear combination** of $|v_1\rangle = \begin{bmatrix} i \\ 1 \end{bmatrix}, |v_2\rangle = \begin{bmatrix} -1 \\ i+1 \end{bmatrix}$

To create a linear combination, we simply must add two complex number constants λ_1, λ_2 in the form shown in Definition 1.2, $\lambda_{\mathbf{1}} |\mathbf{v_1}\rangle + \lambda_{\mathbf{2}} |\mathbf{v_2}\rangle$.

$$\lambda_{final} |v_{final}\rangle = \lambda_1 |v_1\rangle + \lambda_2 |v_2\rangle \qquad \begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= 3 + 2i \end{aligned}$$

$$\begin{aligned} \lambda_{final} |v_{final}\rangle &= 1 \begin{bmatrix} i \\ 1 \end{bmatrix} + (3+2i) \begin{bmatrix} -1 \\ i+1 \end{bmatrix} \\ &= \begin{bmatrix} i \\ 1 \end{bmatrix} + \begin{bmatrix} -3-2i \\ 1+5i \end{bmatrix} \\ &= \begin{bmatrix} -3-i \\ 3+5i \end{bmatrix} \end{aligned} \tag{1.5}$$

⧫

Definition 1.4 Two vectors $|v_1\rangle, |v_2\rangle$ are **linearly independent** if the only way to create the 0 vector from a linear combination is through setting all complex constants $\lambda_1, \lambda_2, ..., \lambda_k$ to zero in the linear combination formula described in Definition 1.2. ⧫

Definition 1.5 A **subspace** of complex plane $\mathbb{C}$ is a set $\mathbf{M} \in \mathbb{C}^n$ that satisfies three properties:

- $|0\rangle \in \mathbf{M}$
- $|v\rangle + |u\rangle \in \mathbf{M}$ for all $|v\rangle, |u\rangle$
- $c|v\rangle \in \mathbf{M} \;\; \forall \;\; |v\rangle \in \mathbf{M},\; c \in \mathbb{C}$ ⧫

Definition 1.6 The **span** of a set $\mathbf{S}$ is the smallest linear subspace $\in \mathbf{S}$. ⧫

Theorem 1.7 ***(Span of a set is always a subspace theorem)***
Let $\mathbf{s} = \{|v_1\rangle, |v_2\rangle, ..., |v_k\rangle\} \subset \mathbb{C}^n$*, then* $span(\mathbf{s})$ *is a subspace of* $\mathbb{C}^n$

Proof. 1. Note that $|0\rangle = 0\,|v_1\rangle + 0\,|v_2\rangle + ... + 0\,|v_k\rangle \in span(\mathbf{s})$.

2. Let $|v\rangle, |u\rangle \in span(\mathbf{s})$ be arbitrary.

   Since $|v\rangle, |u\rangle \in span(\mathbf{s})$, there exists scalars $c_1, c_2, ..., c_k$ and $d_1, d_2, ..., d_k$ such that

$$\begin{aligned}|v\rangle =& c_1\,|v_1\rangle, ..., c_k\,|v_k\rangle\\ |u\rangle =& d_1\,|v_1\rangle, ..., d_k\,|v_k\rangle\end{aligned}$$

   then,

$$\begin{aligned}|v\rangle + |u\rangle &= (c_1\,|v_1\rangle, ..., c_k\,|v_k\rangle) + d_1\,|u_1\rangle, ..., d_k\,|u_k\rangle\\ &= (c_1 + d_1)\,|v_1\rangle + ... + (c_k + d_k)\,|v_k\rangle \;\in\; span(\mathbf{s})\end{aligned}$$

   This is the true because all of these are linear combinations consisting of a complex constant and a ket vector, which coincides with the Linear Combination Definition 1.2.

3. Let $|w\rangle \in span(\mathbf{s})$ and $c \in \mathbb{C}$. If this is true, then there exists $\lambda_1, \lambda_2, ..., \lambda_k$ such that $|v\rangle = \lambda_1\,|v_1\rangle, ..., \lambda_k\,|v_k\rangle$. Consider:

$$\begin{aligned}c\,|w\rangle &= c\,|v\rangle\\ &= C(\lambda_1\,|v_1\rangle, ..., \lambda_k\,|v_k\rangle)\\ &= (C\lambda_1)\,|v_1\rangle + ... + (C\lambda_k)\,|v_k\rangle\end{aligned} \tag{1.6}$$

   Equation 1.6 shows that after distributing **C** in, we have another linear combination, which must exist inside of $span(\mathbf{s})$, and since $span(\mathbf{s})$ satisfies 1,2,3, $span(\mathbf{s})$ must be a subspace. ■

Theorem 1.8 ***Basis Theorem***
If $\boldsymbol{S}$ *is a set of* $\mathbf{n}$ *linearly independent vectors in* $\mathbb{C}^{\mathbf{n}}$*, then* $span(\mathbf{S}) = \mathbb{C}$.

Definition 1.9 A **basis** of a subspace $M \subseteq \mathbb{C}^n$ is a set of vectors **S** such that

1. S is linearly independent.
2. $span(\mathbf{S}) = M$, or in plain English ”S spans M” ♦

Definition 1.10 The **dimension** of a space is the number of vectors in it. ♦

Example 1.11 Standard basis of different complex vector spaces.

- Standard Basis for $\mathbb{C}^2$: $\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}$

- Standard Basis for $\mathbb{C}^n$: $\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix},...,\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}$ ♦

Exercise 1.12 (Vector Space Exercise) Below are some exercises designed to put the concepts in this section to work, in order to further reinforce learning.

1. Consider S $= \{|v_1\rangle, |v_2\rangle, |v_3\rangle\}$, where $|v_1\rangle = \begin{bmatrix}1\\-1\end{bmatrix}, |v_2\rangle = \begin{bmatrix}i\\1\end{bmatrix}, |v_3\rangle = \begin{bmatrix}0\\i\end{bmatrix}$

   (a) Give a linear combination of the vectors in S.

$$\lambda|v\rangle = 1\begin{bmatrix}1\\-1\end{bmatrix} + 1\begin{bmatrix}i\\1\end{bmatrix} - (1+200i)\begin{bmatrix}0\\i\end{bmatrix} \tag{1.7}$$

   (b) Determine if $\begin{bmatrix}1+i\\200-i\end{bmatrix}$ is in Span(S).

$$\begin{aligned}\alpha|v1\rangle + \beta|v_2\rangle + \gamma|v_3\rangle &= \begin{bmatrix}1+i\\200-i\end{bmatrix}, \qquad \alpha = 1, \beta = 1, \gamma = ?\\ 1\begin{bmatrix}1\\-1\end{bmatrix} + 1\begin{bmatrix}i\\1\end{bmatrix} - \gamma\begin{bmatrix}0\\i\end{bmatrix} &= \begin{bmatrix}1+i\\200-i\end{bmatrix} \text{Choose alpha and beta to obtain} \begin{bmatrix}1+i\\0\end{bmatrix}\\ \begin{bmatrix}1+i\\0\end{bmatrix} - \gamma\begin{bmatrix}0\\i\end{bmatrix} &= \begin{bmatrix}1+i\\200-i\end{bmatrix} \qquad \text{Now we must isolate } \gamma\\ \begin{bmatrix}0\\\gamma * i\end{bmatrix} &= \begin{bmatrix}0\\-200+i\end{bmatrix} \qquad \text{We now set the matrix as an equality}\\ \gamma * i &= -200 + i\\ \gamma &= -(1+200i)\\ 1\begin{bmatrix}1\\-1\end{bmatrix} + 1\begin{bmatrix}i\\1\end{bmatrix} - (1+200i)\begin{bmatrix}0\\i\end{bmatrix} &= \begin{bmatrix}1+i\\200-i\end{bmatrix}\end{aligned}$$

   Since we were able to create a linear combination of the vectors that satisfy the given vector, it is in Span(S).

   (c) Describe Span(S) geometrically.

   Span(S) is every possible vector $\in \mathbb{C}^2$ ♦

Exercise 1.13 (Linear Independence Exercise) Find the condition under which the following two vectors are linearly independent.

$$|v_1\rangle = \begin{bmatrix} x \\ y \\ 3 \end{bmatrix} \in \mathbb{R}, \qquad |v_2\rangle = \begin{bmatrix} 2 \\ x-y \\ 1 \end{bmatrix} \in \mathbb{R}$$

$$\begin{bmatrix} x \\ y \\ 3 \end{bmatrix} = \alpha \begin{bmatrix} 2 \\ x-y \\ 1 \end{bmatrix} \tag{1.8}$$

First, we must make this linearly dependent so we know when x and y fail linear independence

$$\begin{aligned} \alpha = 3 \quad \rightarrow \quad & \begin{bmatrix} x \\ y \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 2 \\ x-y \\ 1 \end{bmatrix} \\ & \begin{bmatrix} x \\ y \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 3(x-y) \\ 3 \end{bmatrix} \end{aligned}$$

Now we need to make this an equality to solve for x and y.

$$\begin{aligned} x = 6 \quad \rightarrow \quad y &= 3(x-y) \\ y &= 3x - 3y \\ 4y &= 18 \\ y &= 9/2 (\text{or } 4.5) \end{aligned}$$

♦

Exercise 1.14 (Basis Exercise) Show that the set formed by the following vectors is a basis for C^3

$$|V_1\rangle = \begin{bmatrix}1\\1\\1\end{bmatrix}, \ |V_1\rangle = \begin{bmatrix}1\\0\\1\end{bmatrix}, \ |V_1\rangle = \begin{bmatrix}1\\-1\\-1\end{bmatrix}$$

There are two things we must prove in order for a set of vectors to form a basis: linear independence and an equal number of vectors as the dimension of the space.
First we will prove linear independence:

1. Linear Independence Test

$$\begin{gathered}
\begin{bmatrix}1&1&1&0\\1&0&-1&0\\1&1&-1&0\end{bmatrix}\text{First, we build augmented matrix}\\
\xrightarrow{R_2-R_3}\begin{bmatrix}1&1&1&0\\0&-1&0&0\\1&1&-1&0\end{bmatrix}\\
\xrightarrow{R_3-R_1}\begin{bmatrix}1&1&1&0\\0&-1&0&0\\0&0&-2&0\end{bmatrix}\\
\xrightarrow{-1*R_2,\,-\frac{1}{2}*R_3}\begin{bmatrix}1&1&1&0\\0&1&0&0\\0&0&1&0\end{bmatrix}\\
\xrightarrow[R_1-R_3]{R_1-R_2}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\end{bmatrix}
\end{gathered} \tag{1.9}$$

Thus, we have proved that if the only way to obtain the [0] matrix is through $\lambda_1, ..., \lambda_k = 0$ as mentioned in Definition 1.4, we know that the vectors are **linearly independent**.

2. Dimension Test

Now we must prove that the dimension of the space is equal to the number of the vectors, and since we are working with C^3 and we have $|V_1\rangle, |V_2\rangle, |V_3\rangle$ we know that our space has three dimensions and we have three vectors, therefore we have just proven that **the given set of vectors forms a basis for $\mathbb{C}^3$.** ♦

1.4 Inner/Outer Products, and Norms of Vectors

Definition 1.15 The **dual space of $\mathbb{C}^n$** ($\mathbb{C}^{n^*}$) is a space of row vectors where there are n entries from $\mathbb{C}$. ⧫

Definition 1.16 A **transpose** is an operation that every column into a row vector and vice versa without changing the order. ⧫

Example 1.17 (Example of transpose operation)

$$|V\rangle^\dagger = \begin{bmatrix} 7 \\ 8i \\ \pi + 3i \\ 0 \end{bmatrix}^\dagger \tag{1.10}$$

$$\langle V| = \begin{bmatrix} 7 & 8i & \pi + 3i & 0 \end{bmatrix} \tag{1.11}$$

It can be seen that the only difference between **1.10** and **1.11** is that we have converted the column to a row, this is accomplished through the **dagger (†)** being applied to a matrix or vector. ⧫

Definition 1.18 Given $|V\rangle, |W\rangle$, the **inner product** of $|V\rangle, |W\rangle$ is $\langle W|V\rangle$ which can also be written as $\langle W| * |V\rangle$.
It should be noted that the inner product of a bra and ket vector will always yield a constant.⧫

Example 1.19 (Example of an Inner Product on $\mathbb{R}^n$)

$$|V\rangle = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, |W\rangle = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} \tag{1.12}$$

$$\langle W|V\rangle = \begin{bmatrix} 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} = 5 - 2 + 3 = 6 \tag{1.13}$$

It should be noted that the inner product on $\mathbb{R}^n$ is just the usual dot product. ⧫

Example 1.20 (Example of an Inner Product on $\mathbb{C}^n$)

$$|V\rangle = \begin{bmatrix} i \\ 1 \end{bmatrix}, |W\rangle = \begin{bmatrix} -i \\ -1 \end{bmatrix} \tag{1.14}$$

$$\langle W|V\rangle = \begin{bmatrix} -i & -1 \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = 1 - 1 = \mathbf{0} \tag{1.15}$$

*The inner product yielding 0 means that the two vectors are **orthogonal**.* ⧫

Definition 1.21 The **norm (or magnitude)** of $|V\rangle \in \mathbb{C}^n$ is

$$\||V\rangle\| \coloneqq \sqrt{\langle V|V\rangle} \tag{1.16}$$

♦

Example 1.22 (Example of taking the norm of a complex vector)

$$v = \begin{bmatrix}1, 1\end{bmatrix} \in \mathbb{R}^2$$
$$\||V\rangle\| \coloneqq \sqrt{\begin{bmatrix}1 & 1\end{bmatrix} | \begin{bmatrix}1\\1\end{bmatrix}} = \sqrt{1+1} = \sqrt{2}$$

Two points of note:
1. Can $\langle W|V\rangle$ be complex? **YES**
2. Can $\langle V|V\rangle = \||V\rangle\|^2 \in \mathbb{C}$? **YES** ♦

Definition 1.23 A set of nonzero vectors $\mathrm{S} = \{|V_1\rangle, |V_2\rangle, ..., |V_k\rangle\} \leq \mathbb{C}^n$ is called **an orthogonal set** if $\langle \mathbf{v_i}|\mathbf{v_j}\rangle = \mathbf{0}$ if $\mathbf{i} \neq \mathbf{j}$. ♦

Theorem 1.24 (Orthogonality and Linear Independence Theorem) *In this theorem, we will use what we have learned so far in order to gain a deeper understanding into the properties of orthogonality in a vector space.*

If $S = |V_1\rangle, ..., |V_1\rangle$ is an orthogonal vector set of nonzero vectors in $\mathbb{C}^n$, then S is linearly independent.

Proof. Let $c_1, ..., c_k \in \mathbb{C}$ s.t. $c_1|v_1\rangle, c_2|v_2\rangle, ..., c_k|v_k\rangle = |0\rangle$.
Goal: Show $c_1, ..., c_k = 0$ using assumption S is orthogonal set.

Let $j \in \{1, ..., k\}$, then

$$\begin{aligned}&\langle v_j|(c_1|v_1\rangle, ..., c_k|v_k\rangle) = \langle v_j|0\rangle\\ &c_1\langle v_j|v_1\rangle + ... + c_k\langle v_j|v_k\rangle\\ &c_j\langle v_j|v_j\rangle = \text{Since S is an orthogonal set}\end{aligned} \tag{1.17}$$

Since $\langle v_j|v_j\rangle$ can't possibly be 0 because we know that $j \neq 0$, it means that c_k must be zero, which as stated in the Linear Independence Definition, means that S is linearly independent.■

Definition 1.25 A basis is **orthonormal** if two conditions are satisfied:

1. S is an orthogonal set of n unit vectors.
2. Span(S) $\in \mathbb{C}^n$

In plain English, an orthonormal basis is a basis in which there are exactly the number of vectors as the dimension of the space (from the basis definition) as well as well as all of those vectors being orthogonal to each other. ♦

Exercise 1.26 (Basis Practice) In this example, there will be multiple transformations completed to gain a better understanding of what a basis is, and what it means to be orthogonal.

1.
$$|b_1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} \qquad |b_2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} \tag{1.18}$$

   (a) Show that $\mathcal{B}$ is an orthonormal basis for $\mathbb{C}^2$.

      i. Check orthogonality:

$$\langle b_1|b_2\rangle = \frac{1}{\sqrt{2}}(\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} = 1 - 1 = \mathbf{0}$$

      These two vectors are orthogonal since their inner product is 0.

      ii. Check that they are unit vectors: (is the norm one?)

$$\||b_1\rangle\| = \sqrt{\langle b_1|b_1\rangle} = \frac{1}{\sqrt{2}}\sqrt{1+1} = \frac{\sqrt{2}}{\sqrt{2}} = \mathbf{1}$$
$$\||b_2\rangle\| = \sqrt{\langle b_2|b_2\rangle} = \frac{1}{\sqrt{2}}\sqrt{1^2+(-1)^2} = \frac{\sqrt{2}}{\sqrt{2}} = \mathbf{1}$$

   This means that, since the vectors are orthogonal and unit vectors, they are orthonormal, forming an orthonormal basis for $\mathbb{C}^2$.

(b) Find the coordinates of $|x\rangle = \begin{bmatrix}-21\end{bmatrix}$ relative to basis $\mathcal{B}$ (i.e. find the scalars such that...)

$$c_1, c_2 \in \mathbb{C} \text{ such that } c_1 |b_1\rangle + c_2 |b_2\rangle = |x\rangle$$

$$\langle b_1| \left(c_1 |b_1\rangle + c_2 |b_2\rangle\right) = \langle b_1|x\rangle$$
$$c_1 \langle b_1|b_1\rangle + c_2 \langle b_1|b_2\rangle = \langle b_1|x\rangle$$

$$b_1, b_2 \text{ are orthogonal so } \langle b_1|b_2\rangle = 0 \text{ so}$$

$$c_1 = \langle b_1|x\rangle$$

$$\boxed{c_1 = \frac{1}{\sqrt{2}} \begin{bmatrix}1 & 1\end{bmatrix} \begin{bmatrix}-2 \\ 1\end{bmatrix} = \frac{\mathbf{-1}}{\sqrt{\mathbf{2}}}}$$

$$\langle b_2| \left(c_1 |b_1\rangle + c_2 |b_2\rangle\right) = \langle b_2|x\rangle$$
$$c_1 \langle b_2|b_1\rangle + c_2 \langle b_2|b_2\rangle = \langle b_2|x\rangle$$

$$b_1, b_2 \text{ are orthogonal so } \langle b_2|b_1\rangle = 0 \text{ so}$$

♦

$$c_2 = \langle b_2|x\rangle$$

$$\boxed{c_2 = \frac{1}{\sqrt{2}} \begin{bmatrix}1 & -1\end{bmatrix} \begin{bmatrix}-2 \\ 1\end{bmatrix} = \frac{\mathbf{-3}}{\sqrt{\mathbf{2}}}}$$

Definition 1.27 Define $P_i \coloneqq |f_i\rangle \langle f_i|$ to be the **projection operator** onto span($|k_i\rangle$) ♦

Proposition 1.28 *Let $\mathcal{B} = \{|f_1\rangle, ..., |f_n\rangle\}$ be an orthonormal basis for $\mathbb{C}^n$ then...*

1. $P_i(|v\rangle) \in span(|f_1\rangle)$
2. $|v\rangle - P_i|v\rangle$ *is orthogonal to* $|f_i\rangle$
3. ${P_i}^2 = P_i * P_i = P_i$ *Projection will still be the same.*
4. $P_iP_j = 0$ *shadow on one plane will be "nothing" in the perspective of another plane.*
5. $P_i^\dagger = P_i$ *self-adjoint*
6. $\sum_{i=1}^n P_i = I_n$

Example 1.29 (Outer Product Example) Given $P_2 = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$, what is $P_1 + P_2$

$$\begin{aligned}
P_1 &= |f_1\rangle \langle f_1| \\
&= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} * \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\
P_1 + P_2 &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}$$

This is an interesting result, it shows that if you add all of the projections in a space you will get the identity matrix, proof available elsewhere as I am too lazy to type it out. ♦

Why is projection useful?: It allows us to find $c_1, ..., c_k$ for $|x\rangle = c_1|f_1\rangle + ... + c_k|f_k\rangle$ when all $|f\rangle$ are orthogonal to each other.

Theorem 1.30 (The Gram Schmidt Process) *Let $|b_1\rangle, ..., |b_k\rangle$ be the basis for a subspace M in $\mathbb{C}^n$. Define the set of vectors $|f_1\rangle, ..., |f_k\rangle$ as follows:*

Step 1:
Define $|f_1\rangle \coloneqq |b_1\rangle$.

Step i+1:
for i = 1:k-1

$$|f_{i+1}\rangle \coloneqq |b_{i+1}\rangle - (\sum_{j=1}^{i} \frac{|f_j\rangle \langle f_j|}{\langle f_j|f_j\rangle}) |b_{i+1}\rangle = |b_{i+1}\rangle - (\sum_{j=1}^{i} \frac{\langle f_j|b_{i+1}\rangle}{\langle f_j|f_j\rangle}) |f_j\rangle \tag{1.19}$$

Done

Now normalization step
for i = 1 : k

$$|f_i\rangle \coloneqq \frac{1}{\||f_i\rangle\|} |f_i\rangle \tag{1.20}$$

Done, now we know that $|f_1\rangle, ... |f_k\rangle$ is a orthonormal basis for M

A special thank you for Dr. Hamidi and Dr. Ismert for giving this psuedocode function for Gram Schmidt, I have pulled from it heavily in this page as it is the best way to explain this code.

2 Matrices and Linear Transformations

2.1 Matrices

This section assumes a basic knowledge of matrix operations, and should be a basic overview.

Definition 2.1 $\mathbf{M}_{\mathbf{mn(}}\mathbb{C}) =$ set of all mxn matrices with complex entries. ♦

Definition 2.2 $\mathbf{M}_{\mathbf{n}}(\mathbb{C}) \equiv \mathbf{M}_{\mathbf{nn}}(\mathbb{C})$ meaning that it is just a square matrix with nxn size. ♦

Example 2.3 (Pauli Matrices)

$$\begin{aligned} \sigma_x &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \sigma_y &= \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \qquad \in M_2 \\ \sigma_z &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{aligned} \tag{2.21}$$

These three matrices are called the ***Pauli matrices*** *the the basic matrices we will use for quantum computations in the future.* ♦

Definition 2.4 The **Identity Matrix** (I_n) is the matrix M_n whose columns are the standard basis $|e_1\rangle, ..., |e_n\rangle$ for $\mathbb{C}_n$.

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.22}$$

♦

Theorem 2.5 *M_n has a well defined matrix multiplication, in general, if $A \in M_{mn}, B \in M_{nk}$.*

The matrix $AB = [A|b_1\rangle \ \ ... \ \ A|b_k\rangle]$

Example 2.6 (Pauli Matrices in action) For this example, we will be scaling the σ_y Pauli matrix.

$$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & -3i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 1|a_1\rangle & 3|a_2\rangle \end{bmatrix} \tag{2.23}$$

This shows that in general, multiplying a row matrix by a diagonal matrix will scale the row matrix by the respective diagonal matrix value. ♦

2.2 Eigenvectors and Eigenvalues

2.2.1 Eigenvalues/Eigenvectors

Definition 2.7 An **eigenvalue** for $A \in M_n$ is a complex number $\lambda \in \mathbb{C}$ such that there is a nonzero vector $|x\rangle \in \mathbb{C}^n$ satisfying

$$A\,|x\rangle = \lambda\,|x\rangle \tag{2.24}$$

♦

Definition 2.8 A mmatrix $A \in M_n$ is **diagnolizable** if and only if:

1. There exists a diagonal matrix D and an invertable matrix P such that $A = PDP^-1$ if and only if
2. there exists a basis for C^n consisting of eigenvalues for A. ♦

Example 2.9 Consider:

$$A = \begin{bmatrix} I_2 & 0 \\ 0 & \sigma_y \end{bmatrix} \equiv \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \end{bmatrix} \tag{2.25}$$

1. Get eigenvalues and some normalized eigenvectors:

$$\begin{aligned} &\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix} \\ = &\begin{bmatrix} 1-\lambda & 0 & 0 & 0 \\ 0 & 1-\lambda & 0 & 0 \\ 0 & 0 & \lambda & -i \\ 0 & 0 & i & -\lambda \end{bmatrix} \\ = &(1-\lambda)(1-\lambda)[\lambda^2 - 1] = 0 \\ &\lambda = 1, 1, 1, -1 \end{aligned} \tag{2.26}$$

2. Now plug in eigenvalue and find eigenvector:

$$A\left|e_1\right\rangle = \left|e_1\right\rangle$$
$$A\left|e_2\right\rangle = \left|e_2\right\rangle$$
$$A - \lambda * I = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -i \\ 0 & 0 & i & -1 \end{bmatrix}$$
$$\xrightarrow{R_4 + iR_3} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & i \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$$\xrightarrow[R_1 \leftrightarrow R_3]{-1R_1} \begin{bmatrix} 0 & 0 & 1 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Now we must use this and transfer it into an equation to solve for x_3, x_4

$x_3 = ix_4$

Choose $x_4 = 1$

$$\left|x\right\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ i \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ i \end{bmatrix} \tag{2.27}$$

These first two arrays are just chosen because there are no x_1, x_2 values in the eigenvector array, so we choose two unit vectors.

3. Finally, normalize the vector.

$$\left|x\right\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 0 \\ i \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 0 \\ 1 \\ i \end{bmatrix}$$

4. Final answer

$$\text{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} * i \\ 0 & 0 & \frac{1}{\sqrt{2}} * i & \frac{1}{\sqrt{2}} \end{bmatrix}, \text{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$ ♦

2.2.2 Eigenvectors and Eigenvalues Big Picture

1. Self-Adjoint (Hermitian)

$$\begin{aligned} &A \in M_n \text{ is Hermitian if} \\ &A^\dagger = A \\ &{\sigma_x}^\dagger = \sigma_x \\ &{\sigma_y}^\dagger = \sigma_y \\ &{\sigma_z}^\dagger = \sigma_z \end{aligned} \tag{2.28}$$

   Additionally, all Hermitian matrices are diagnolizable, meaning their eigenvalues form a orthonormal basis.

2. Normal Matrices

$$A * A^\dagger = A^\dagger * A \leftrightarrow \text{eigenvectors form orthonormal basis} \tag{2.29}$$

3. Positive semi-definite matrix M_n is:

$$\begin{aligned} &\forall \left|x\right\rangle \in \mathbb{C}^n, \text{we have} \\ &\langle x|A|x\rangle \geq 0 \text{ where } A\left|x\right\rangle = y \end{aligned} \tag{2.30}$$

4. (Unitary Matrices) The following are all equivalent:

   - $U \in M_n$ is unitary
   - $U^\dagger = U^{-1}$ or $(U^{-1} * U = U * U^{-1} = I_n)$
   - $U * U^\dagger = I_n$ or $U^\dagger * U = I_n$ *note: $U^\dagger$ not always equal to U, but they are normal.*
   - The column of U form an orthonormal basis for $\mathbb{C}^n$.
   - $\forall \left|x\right\rangle, \left|y\right\rangle \in \mathbb{C}^n$ and $\langle U_x|U_y\rangle = \langle x|y\rangle$ meaning unitary matrices are rotation matrices.

3 Introduction to Quantum Theory

Now that we have laid out the mathematical framework necessary to understand this subject, it is time to lay our the physics necessary to complete the basis (no pun intended) for Quantum Information Theory.

3.1 Quantum Theory Axioms

These axioms are humans attempts at creating rules and methods of abstractions to describe and understand

3.1.1 Axiom 1.1: A vector state x is a unit vector in a complex Hilbert space.

We want to figure out the photon's (particle) polarization, to figure this out we are going to shoot it at a vertically polarized filter. Classically, $|x\rangle$ should be polarized ↑ or →, if this is the case we expect two outcomes:

1. If $|x\rangle$ goes through filter, then $|x\rangle$ is ↑.
2. If $|x\rangle$ is deflected, then $|x\rangle$ is →.

3.1.2 Axiom 1.2: Linear combinations (or superposition) of the physical states are allowed to act as x vectors.

Physical States (Choices made by humans to describe quantum mechanics):

$$\uparrow = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \rightarrow = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Something of note, these two vectors form a basis for $\mathbb{C}^2$ and all combinations are:

$$|x\rangle = \alpha \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ where } \|\alpha\|^2 + \|\beta\|^2 = 1$$

To keep track of this choice, we make the matrix $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ representing vertically polarized filter.

Another note, the e-vals of that matrix are $|v\rangle, |h\rangle$.

Example 3.1 $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ might be a nice to represent a horizontally polarized filter. ⧫

3.1.3 Axiom 2: Observable

An observable of a state $|x\rangle$ corresponds to a Hermitian matrix A. $|x\rangle$ is in state $|u\rangle$, an e-vect for A with probability $|\langle u|x\rangle|^2$

Definition 3.2 Expectation value of A (or mean value) of observable associated to A after measurements with respect to many copies of $|\psi\rangle$ is **the weighted average of the expected outcomes**.

$$\langle A\rangle_{\psi} = \sum_{i=1}^{n} |\langle u_i|\psi\rangle|^2$$
$$|\psi\rangle = \sum_{i=1}^{n} c_i\,|u_i\rangle \text{ where } u_i \text{ is an e-basis from A}$$

♦

3.1.4 Axiom 3: The time dependence of a state is governed by the Schrödinger equation

$$i\bar{h}\frac{\delta|\psi|}{\delta t} = H\,|\psi\rangle \tag{3.31}$$

$\bar{h}$ is reduced Planck's constant
H is Hermitian matrix corresponding to energy of the system **Hamiltonian**.

When H is time invariant (constant), the Schrödinger equation becomes:

$$|\psi(t)\rangle = e^{\frac{-itH}{\bar{h}}}\,|\psi(0)\rangle \tag{3.32}$$

3.2 More Introductory Quantum Concepts

Now that the basic framework of what a quantum state is, it is now time to apply these concepts and build on them to further characterize what a quantum system is, and why it is important to us. Prepare yourself, because we are jumping right into an example applying the previous axioms to characterize a *real world* quantum system.

Example 3.3 Consider a physical system with Hamiltonian $H = \frac{\hbar}{2}\omega\sigma_x$ and suppose $|\psi(0)\rangle = \begin{bmatrix}1\\0\end{bmatrix}$ (The first column of σ_z.)

- Find the wave function $|\psi(t)\rangle, t > 0$, which by the Shrödinger equation is:

$$\begin{aligned}
|\psi(t)\rangle &= \exp\frac{-itH}{\hbar} * |\psi\rangle \\
&= \exp\left(i(\frac{-t}{\hbar})(\frac{\hbar}{2}\omega\sigma_x)\right)\begin{bmatrix}1\\0\end{bmatrix} \\
&= \exp i(\frac{t}{2}\omega)\sigma_x)\begin{bmatrix}1\\0\end{bmatrix} \qquad \text{where } \frac{t}{2}\omega \text{ is } \alpha \\
&= \exp i(\alpha)\sigma_x)\begin{bmatrix}1\\0\end{bmatrix} \\
&= [\cos(\frac{t}{2}\omega)I + i\sin(\frac{t}{2}\omega)\sigma_x]\begin{bmatrix}1\\0\end{bmatrix} \\
&= \begin{bmatrix}\cos(\frac{t}{2}\omega) & i\sin(\frac{t}{2}\omega)\\ i\sin(\frac{t}{2}\omega) & \cos(\frac{t}{2}\omega)\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} \\
|\psi(t)\rangle &= \begin{bmatrix}\cos(\frac{\omega t}{2})\\ i\sin\frac{\omega t}{2}\end{bmatrix}
\end{aligned} \tag{3.33}$$

- Find the probability for the system to have outcome +1 upon measurement of σ_z

$$\begin{aligned}
P_2|\psi(t)\rangle = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}\cos(\frac{\omega t}{2})\\ i\sin\frac{\omega t}{2}\end{bmatrix} = \begin{bmatrix}\cos(\frac{\omega t}{2})\\ 0\end{bmatrix} \\
P_\downarrow(t) = |\cos\frac{\omega t}{2}|^2 = \cos^2\frac{\omega t}{2}
\end{aligned} \tag{3.34}$$

As you can see, we have picked off the cos out of ψ and in order to compute the probability distribution of the wave, we square the final answer, this will look familiar to a probability and statistics class topic of expected values and distributions (because it is... Wow! never thought it would be useful again)

- Find the probability for the system to have outcome -1 upon measurement of σ_z

$$\begin{aligned} P_2 \left|\psi(t)\right\rangle = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\frac{\omega t}{2}) \\ i \sin \frac{\omega t}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ i \sin \frac{\omega t}{2} \end{bmatrix} \\ P_{\downarrow}(t) = |i \sin \frac{\omega t}{2}|^2 = \sin^2 \frac{\omega t}{2} \end{aligned} \tag{3.35}$$

Same thing as the one above, but notice that this one is picking out the sin in ψ instead using the other diagonal term in the P matrix.

- Find the expectation value under many measurements of σ_z:

$$\begin{aligned} < \sigma_x >_{\psi} &= \langle \psi(t) | \sigma_z \psi(t) \rangle = (+1)(\cos^2 \frac{\omega t}{2}) + (-1)(\sin^2 \frac{\omega t}{2}) \\ &= (\cos^2 \frac{\omega t}{2}) - (\sin^2 \frac{\omega t}{2}) \end{aligned} \tag{3.36}$$

As you can see, we have set matrix σ_z against vector $\psi(t)$, giving us the overall expected outcome if we measured wave function $\psi(t)$ in the σ_z ”direction”. ⧫

3.3 Multipartite Physical States

3.3.1 Tensor Products

Definition 3.4 Tensor Product Consider vector space H with an inner product $H = H_1 \otimes H_2$ where $\otimes$ is the **tensor product** combining H_1, H_2 in such a way that things in H_1 effect H_2 and vice versa. ⧫

A general vector of H is a linear combination of vectors $\{|v_1\rangle \otimes |v_2\rangle : |v_1\rangle \in H_1, |v_2\rangle \in H_2\}$

Now to take the tensor of two matrices where $A \in M_{mn}, B \in M_{pq}$:

$$A \otimes B = A = \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & & \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & \dots & & A_{KK} \end{bmatrix} \in M_{(mp)(nq)} \tag{3.37}$$

This equation may look overwhelming at first glance, but it can be seen that all it is doing is applying the *elements* of A to the *matrix* B. This means that you will get a matrix that any change in A affects B, and vice versa just like we hoped.

Example 3.5

$$\sigma_x \otimes i\sigma_y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{bmatrix} \tag{3.38}$$

Understand that this is actually just a 4X4 matrix, and was only written with the extra brakets in order to provide clarity. ⧫

Rules of tensor products Now that we have learned a little about tensor products, lets learn some basic rules that apply to this operation:

1. $(\lambda A) \otimes B = A \otimes (\lambda B), \quad \lambda \in \mathbb{C}$ *(Scaling one matrix affects the other and vise verse)*
2. In general $A \otimes B \neq B \otimes A$ *(Generally noncommutative, not always true though)*
3. $(A \otimes B)(C \otimes D) = AC \otimes BD$ *(Distributivity law)*
4. $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ *(Another distributivity law)*
5. $(A \otimes B) + (C \otimes D) \neq (A + C) \otimes (B + D)$ *(Nonassociative)*
6. $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$ *(Daggers are distributive)*
7. If A and B are invertible, $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ *(Inverses are distributive)*

Example 3.6 (Tensor product of two ket vectors)

$$|1\rangle \otimes |1\rangle =: |11\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$|1\rangle \otimes |2\rangle =: |12\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$|2\rangle \otimes |1\rangle =: |21\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$|2\rangle \otimes |2\rangle =: |22\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Note: $\{|11\rangle, |12\rangle, |21\rangle, |22\rangle\}$ *forms a basis for* $\mathbb{C}^4$ *but these vectors live in* $\mathbb{C}^2 \otimes \mathbb{C}^2 \approx \mathbb{C}^4$ ♦

This example showed a property of tensor products in which if $|v\rangle \in \mathbb{C}^m, |w\rangle \in \mathbb{C}^n$, then $|wv\rangle := |v\rangle \otimes |w\rangle \in \mathbb{C}^m \otimes \mathbb{C}^n \approx \mathbb{C}^{mn}$, so the result is the two dimensions multiplied.

Since we have done so much fun stuff this chapter, lets finish it off with something fun, a proof!

Theorem 3.7 (Tensor Product Basis)
If H_1 has ONB $\varepsilon_1 = \{|e_{1,1}\rangle, |e_{1,2}\rangle, \dots, |e_{1,m}\rangle\}$ and H_2 has ONB $\varepsilon_1 = \{|e_{2,1}\rangle, |e_{2,2}\rangle, \dots, |e_{2,n}\rangle\}$

Note: First subscript is the basis where the vector is formed.

Then $H \coloneqq H_1 \otimes H_2$ has an orthonormal basis $\{|e_{1,i}\rangle \otimes |e_{2,i}\rangle : 1 \leq i \leq m, 1 \leq j \leq n\}$

Proof. Let $\sum_p \lambda_p |v_p\rangle \otimes |w_p\rangle \in H$
It suffices to show that just one elementary tensor $ketv_p \otimes |w_p\rangle$ belongs to Span(ε), so since ε_1 is a basis for H_1 and $ketv_p \in H_1$, we can write

$$|v_p\rangle = c_1 |e_{1,1}\rangle + c_2 |e_{1,2}\rangle + \cdots + c_m |e_{1,m}\rangle \text{ for some } c_1, \dots, c_m \in \mathbb{C}$$

$$|w_p\rangle = d_1 |e_{2,1}\rangle + d_2 |e_{2,2}\rangle + \cdots + d_n |e_{2,n}\rangle \text{ for some } d_1, \dots, d_n \in \mathbb{C}$$

$$\begin{aligned}
\text{So } |v_p\rangle \otimes |w_p\rangle &= (c_1 |e_{1,1}\rangle + c_2 |e_{1,2}\rangle + \cdots + c_m |e_{1,m}\rangle) \otimes (d_1 |e_{2,1}\rangle + d_2 |e_{2,2}\rangle + \cdots + d_n |e_{2,n}\rangle) \\
&= (\sum_{i=1}^{m} c_i |e_{1,i}\rangle) \otimes (\sum_{j=1}^{n} d_j |e_{2,j}\rangle) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} c_i |e_{1,i}\rangle \otimes d_j |e_{2,j}\rangle \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} c_i d_j (|e_{1,i}\rangle \otimes |e_{2,j}\rangle) \in Span(\varepsilon) \textbf{ by tensor product fact 1}
\end{aligned}$$

What is the inner product of H?

$$\begin{aligned}
&\text{let } |v_1\rangle, |w_1\rangle \in H_1 \text{ and} |v_2\rangle, |w_2\rangle \in H_2 \\
\langle |v_1\rangle \otimes |v_2\rangle \| w_1\rangle \otimes |w_2\rangle\rangle &\coloneqq \langle v_1 v_2 | w_1 w_2\rangle \\
&= \langle v_1 | w_1\rangle_{H_1} \langle v_2 | w_2\rangle_{H_2} \textit{ Inner products of first factors form first set and same for seco}
\end{aligned}$$

Finally we must check that ε is an orthonormal set.

Let $|e_{1,i}\rangle \otimes |e_{2,j}\rangle, |e_{1,k}\rangle \otimes |e_{2,l}\rangle \in \varepsilon$
Then $\langle e_{1,i} e_{2,j} | e_{1,k} e_{2,l}\rangle = \langle e_{1,i} | e_{1,k}\rangle_{H_1} * \langle e_{2,j} | e_{2,l}\rangle_{H_2}$
That means this breaks down to when i=k and j=l otherwise they are orthogonal so
$= \| |e_{1,i} e_{2,j}\rangle \|$ which must be 1 since they are parallel so ε is an orthonormal set.
(3.39)

■

3.3.2 Quantum Multipartite Systems

$$H = H_1 \otimes H_2 \otimes \cdots \otimes H_n \text{ is a multipartite system} \tag{3.40}$$

*Note: if n is two then the system is **bipartite***

Definition 3.8 (Bipartite state) $|\psi\rangle$ state in $H_1 \otimes H_2$ is called a **bipartite state** ♦

Definition 3.9 (Separability) A vector $|\psi\rangle \in H = H_1 \otimes H_2$ is **separable** (or an elementary tensor) if $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$ for some $|\psi_1\rangle \in H_1, |\psi_2\rangle \in H_2$ ♦

In plain English, separability states that there are vectors that can be tensored together to create the final tensor, an example below will illustrate this better.

Example 3.10 Is $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{C}^4 (\approx \mathbb{C}^2 \otimes \mathbb{C}^2)$ an elementary tensor?

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 * \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ 1 * \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \tag{3.41}$$

It *is* separable. ♦

Definition 3.11 A vector that is *not* separable is called **entangled**. ♦

Remark 3.12 A state $\psi = C_{11}|11\rangle + C_{12}|12\rangle + C_{21}|21\rangle + C_{22}|22\rangle$ is separable only if (in C^2) rank is one (i.e. only if its rows are scalar multiples of each other), where the coeff matrix is $\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$. ♦

Most of the time, finding if $|\psi\rangle \in H_1 \otimes H_2$ is separable is not as easy as the exercises as above, this is because most of the time we cannot look at a matrix and infer the vectors being tensored together, this is where the next section comes into play, and we use a singular value decomposition.

3.3.3 Singular Value Decomposition

First things first, we must know that everything that is about to be shown is only applicable in bipartite systems, anything larger than that is an impossible, and yes that really means impossible, task in most situations because of the properties of quantum systems and their probabilistic nature. If you can figure out an algorithmic way to solve the decomposition of multipartite systems, you will be very famous and more than likely have a decomposition named after you. Now that we have laid that out there, lets continue with singular value decompositions.

The Singular Value Decomposition takes on the form $C = U\Sigma V^\dagger$, this will mean very little, but keep this in mind as you move through the example as this is the motivation to the steps below.

Definition 3.13 If $A \in M_{mn}$, then the **singular values** for A are the square roots of the e-vals for $A^\dagger A \in M_{nn}$. ⧫

Fact: For any $A \in M_{mn}$, $A^\dagger A$ is positive semi-definite, meaning it is Hermetian and its eigenvalues are non-negative.

Example 3.14 Find the singular value decomposition (SVD) for A = $\begin{bmatrix} 1 & 1 \\ 0 & 0 \\ i & i \end{bmatrix}$

1. **Compute eigenvalues and eigenvectors**

$$A^\dagger A = \begin{bmatrix} 1 & 0 & -i \\ 1 & 0 & -i \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ -i & -i \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \tag{3.42}$$

Find the eigenvalues

$$det(A^\dagger A - \lambda I) = 0 \to \boxed{\lambda = 0, 4}$$

Find the eigenvectors

$$\lambda = 0 \to A^\dagger A |x\rangle = |0\rangle \rightsquigarrow |x\rangle = \boxed{\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}}$$

$$\lambda = 4 \to A^\dagger A |x\rangle = 4|x\rangle \rightsquigarrow |x\rangle = \boxed{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

Now we must take the spectral decomposition to get the e-vec matrix (see below). We are allowed to take the spectral decomposition because we know that $A^\dagger A$ is positive semi definite.

$$A^\dagger A = \boxed{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}} \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\right)^\dagger \tag{3.43}$$

2. **Make sure V's columns are in decreasing order with respect to eigenvalues.**

$$V = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \leftarrow \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \\ 0 & 0 \end{bmatrix} \text{(This is generally how it is)}$$

3. **Construct more e-vectors** $A\left|\lambda_1\right\rangle, A\left|\lambda_2\right\rangle$

$$A\left|\lambda_1\right\rangle = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ i & i \end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} \frac{2}{\sqrt{2}} \\ 0 \\ \frac{2i}{\sqrt{2}} \end{bmatrix} = \sqrt{2}\begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix}$$

$$A\left|\lambda_2\right\rangle = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ i & i \end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{This one is useless as there is no new info}$$

4. **Now use** $\left|u_1\right\rangle = \sqrt{2}\begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix}$ **to create an ONB for** $\mathbb{C}^3$**.**

$\{\left|u_1\right\rangle, \left|u_2\right\rangle, \left|u_3\right\rangle\} \rightsquigarrow \{\left|u_1\right\rangle, \left|2\right\rangle, \left|3\right\rangle\}$ where $\left|2\right\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\left|3\right\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

This is linearly independent but not orthogonal, we must use Gram-Schmidt to change that:

5. **Apply the Gram-Schmidt Process to the basis** My poor fingers tire of this, therefore I implore you to go look at the Gram-Schmidt example above to understand how this process works. If you don't believe me, feel free to try it for yourself and email me telling me of my stupidity if it is indeed wrong. Thank you ☺.

Final Orthonormal Basis:

$$\left|u_1\right\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{i}{\sqrt{2}} \end{bmatrix}$$

$$\left|u_2\right\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\left|u_3\right\rangle = \begin{bmatrix} \frac{i}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

6. **Form U vector with orthonormal basis calculated above**

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{i}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{i}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ is } \textbf{unitary} \tag{3.44}$$

7. **Finally, assemble the SVD!**

$$C = U\Sigma V^\dagger = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{i}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{i}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^\dagger \tag{3.45}$$

Whew, that was a hard process.. but now that we have this set of matrices, you can see that we can recover A through these, super useful! ⧫

3.3.4 Schmidt Decomposition

Theorem 3.15 *Let $H = H_1 \otimes H_2$ where dim $H_1 = m < infinity$ and dim $H_2 = n < infinity$, every vector $\psi \in H$ admits a Schmidt Decomposition:*

$$\psi = \sum_{j=1}^{r} S_j \left|u_j\right\rangle \otimes \left|v_j\right\rangle \tag{3.46}$$

where $s_j > 0$ are the ***Schmidt Coefficients*** *satisfying $\sum_{j=1}^{r} {S_j}^2 = 1$, $\{\left|u_j\right\rangle\} \leq H_1$ and $\{\left|v_j\right\rangle\} \leq H_2$ are orthonormal, and $r \leq min\{m, n\}$ is called the* ***Schmidt number****.*

Note: r is just the rank of the coeff matrix
Remark: $\left|\psi\right\rangle \in H_1 \otimes H_2$ is separable iff it's Schmidt number is 1.

3.4 Mixed States as Density Matrices

In this section, we will learn how to describe quantum systems with density matrices, which we will find has many advantages compared to the ways we have learned to describe systems thus far, so before we get into any of the math, let us discuss the motivation more specifically.

3.4.1 Motivation for density matrices

Quantum particles are waves, and as such, there is some specificity when it comes to describing those particles with respect to time. This would be a problem if two of the same particles were measured, there would be no guarantee that the particle would behave in the same way because of this phase shift with respect to time. With a density matrix, we encompass the phase shift within the projection (math for this shown below) in creating one, meaning that we can describe particles regardless of their phase shift. Now lets get into how to make one.

3.4.2 Creating a density matrix

Example 3.16 Suppose $|\psi\rangle \in \mathbb{C}^2$ is the state of some quantum system. That means that $\||\psi\rangle\| = 1$ and if we were hoping to measure vertical or horizontal polarization.

Let being vertically polarized by represented by:

$$|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{and} \qquad |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

These values should come from the eigenstates from chosen Hermitian matrix
which representsthe apparatus.

$$A = 0\,|0\rangle\langle 0| + 1\,|1\rangle\langle 1| = |1\rangle\langle 1|$$

$$= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \textit{ This can be assigned as the "apparatus"}$$ ♦

To glean information about ψ regarding its polarization, we use the equation

$$\psi = \alpha\,|0\rangle + \beta\,|1\rangle$$

$$\text{where } \|\psi\| = \sqrt{\langle\psi|\psi\rangle} = \sqrt{\|\alpha\|^2 + \|\beta\|^2} = 1$$

We interpret this expression of ψ in the $\{|0\rangle, |1\rangle\}$-base as:

$$|\psi\rangle \text{ is in } |0\rangle \text{ with } \|\alpha\|^2 \text{ probability and } |\psi\rangle \text{ is in } |1\rangle \text{ with } \|\beta\|^2 \text{ probability}$$

Recall: $|\psi_1\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ when we take this, and let it develop over time, the Schrödinger equation states that time development creates a phase shift in the particle vector.
This leads to $|\psi_2\rangle = e^{i\theta}\alpha\,|0\rangle + e^{i\theta}\beta\,|1\rangle$ and it is in $|0\rangle$ with $\|e^{i\theta}\alpha\|^2$ probability and is in $|1\rangle$ with

$\|e^{i\theta}\beta\|^2$ probability.

Note: Something interesting is that based on this is that $|\psi_1\rangle \neq |\psi_2\rangle$ as ***vectors*** *in $\mathbb{C}^2$ but $|\psi_1\rangle = |\psi_2\rangle$ as* ***states*** *of the system in "quantum land".*

As we can see, it is going to be really hard to understand what state a particle is in this format, but this is where density matrices come in.
Consider:

$$\begin{aligned}\rho_{\psi_1} &= |\psi_1\rangle\langle\psi_1| = (\alpha|0\rangle + \beta|1\rangle)(\alpha^*\langle 0| + \beta^*\langle 1|)\\ &= |\alpha|^2|0\rangle\langle 0| + \alpha\beta^*|0\rangle\langle 1| + \alpha^*\beta|1\rangle\langle 0| + |\beta|^2|1\rangle\langle 1|\\ \rho_{\psi_1} &= \begin{bmatrix}|\alpha|^2 & \alpha\beta^*\\ \alpha^*\beta & |\beta|^2\end{bmatrix}\end{aligned}$$

Now let's try the same thing with $|\psi_2\rangle$

$$\begin{aligned}\rho_{\psi_2} &= |\psi_2\rangle\langle\psi_2| = e^{i\theta}\alpha(e^{i\theta}\alpha)^* = e^{i\theta}\alpha * e^{-i\theta}\alpha^* = \alpha\alpha^2 = |\alpha|^2...\\ &= \begin{bmatrix}|\alpha|^2 & \alpha\beta^*\\ \alpha^*\beta & |\beta|^2\end{bmatrix}\checkmark\end{aligned}$$

If you continue this you will see that they are the same matrix, meaning that using these matrices we can eliminate the time dependence of the vectors, and still keep the same data about the states.

3.4.3 Density Matrix Observations

Example 3.17 Given $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ or $|\psi\rangle = \begin{bmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}$

Find ρ_ψ (w.r.t. $\{|0\rangle, |1\rangle\}$):

$$\rho_\psi = \begin{bmatrix}\frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & \frac{1}{2}\end{bmatrix} \tag{3.47}$$

♦

Now, a couple observations about density matrices:

- $\rho_\psi \in M_n\ (\mathbb{C})$
- ρ_ψ has trace 1 (i.e. $\text{Tr}(\rho_\psi) = 1$ where Tr() is the sum of the diagonal matrices)
- ρ_ψ is **Hermetian**, and in fact furthermore **positive semi-definite**.

Definition 3.18 (Trace) The **trace** of an nxn matrixis the sum of its diagonal entries. ♦

Definition 3.19 (Density Matrix) A **density matrix** $\rho \in M_n$ is a Hermetian, positive semi-definite matrix such that $\text{tr}(\rho) = 1$. ♦

3.4.4 Density Matrix Axioms

- Axiom 1': A physical state of a system, whose Hilburt Space $\mathbb{H}$, is completely determined by it's associated density matrix.
- Axiom 2': The mean value of an observable A is $< A >_\psi = tr(\rho A)$
- Axiom 3': The time evolution of density matrix is given by the Louisville - von Nuemann equation
$$i\hbar \frac{d}{dt}\rho = [H, \rho] \tag{3.48}$$

3.4.5 What about bipartite?

I can already hear the cries from the judgmental people reading this document, "we will never have a single particle system". You are completely right, but we have to walk before we run young padawan, be patient. In this next section, we will discuss systems of tensor products of multiple states, and how they are represented in density matrices.

Definition 3.20 Let $\rho \in M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$ be a density matrix.

Recall: $\rho = \sum_{i=1}^N \ c_i \ \underbrace{A_{1,i}}_{\in M_n(\mathbb{C})} \ \otimes \ \underbrace{A_{2,i}}_{\in M_n(\mathbb{C})}$

1. ρ is called "uncorrelated" if:

   $\rho = \rho_1 \otimes \rho_2$ where ρ_i is a density matrix. (i.e. simple factorization)

2. ρ is "separable" if:

   $\rho = \sum_{i=1}^N p_i \rho_{1,i} \otimes \rho_{2,i}$. (i.e. factorable, but not actually two elementary tensors, actually a sum of multiple)

3. If ρ is not "separable", so we call ρ inseparable. (i.e. NO factorization)

Example 3.21 Find the tensor product of the quantum vector.

$$\begin{aligned} |\psi\rangle &= |02\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^3 \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \in \mathbb{C}^6 \end{aligned} \tag{3.49}$$

$$\rho = |02\rangle\langle 02| = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\text{"Tells the row"}} \underbrace{\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}}_{\text{"Tells the column"}} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{\text{"3rd row, 3rd col"}} \tag{3.50}$$

♦

Definition 3.22 Partial Trace The partial trace of A over $\mathbb{C}^m$ is

$$A_1 = \sum_{i=1}^{N} c_i \underbrace{A_{1,i}}_{\text{elementary tensor}} \otimes \underbrace{A_{2,i}}_{\text{elementary tensors}} \tag{3.51}$$

♦

What is the point of the partial trace?
The partial trace gives you information about the specific particle instead of the overall system as you are looking at the elementary tensors of the system. ♦

4 Qubits and Information Theory

Finally! I know, there has been large amounts of math up until this point, and there will continue to be, but with all of this math I hope that you have not lost sight of the purpose of this document: quantum computing! All of this math was designed to give you some perspective on what is happening in these computers, and this is the section in which we will connect these math concepts to the quantum concepts integral to quantum computing, but first lets start with a quick review of classical computers.

4.1 Classical and Quantum Computer Comparisons

As a computer engineer, this is where I start to get *super* interested, no offense to the mathematicians out there, but I find that placing these concepts to a real world parallel makes everything more interesting, so this first section will be a slight history lesson on how computers have worked since their inception, and how quantum computers solve many of the problems of classical computers, as well as create a whole new set of problems.

4.1.1 Classical Computer Basics

- Boolean bit is either a 1 or 0.
- Data is transmit in strings of bits.

Example 4.1 A bit of length 8 is a byte. 00100100 in utf-8 is "$". ☺ ♦

4.1.2 Quantum Computer Basics

Definition 4.2 Qubit A **qubit** is a unit vector in $\mathbb{C}^2$

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle\, suchthat \left\|a^2\right\| + \left\|b^2\right\| = 1 \tag{4.52}$$

♦

Remark 4.3 Some other facts about qubits

- To extract info from a qubit, make a measurement which results in **collapse** to a $|0\rangle$ or $|1\rangle$.
- Qutrits ($\mathbb{C}^3$) and qudits ($\mathbb{C}^n$) also exist, but are not discussed in this paper. ♦

Definition 4.4 A group or system of n qubits is called a quantum register. ♦

4.1.3 What problems does a quantum computer solve?

As you can see, for a quantum computer, there is little crossover, but the main point is that while a classical computer can represent information as a 0 or 1, a quantum computer can represent the same information with a 0, 1, or anything in between, with the particle in ”superposition”. An example of this contrast of the two computers will follow.

Example 4.5 Classical bit stream Suppose we have a bit stream, xyzw, of length 4. Also suppose that each bit in this bitstream is determined completely independent of each other. This would mean that there is a maximum of 2^4 possibilities, and we are choosing 4 out of those to represent our bitstream. This is quite a bit of data, that can be stored in those four bits, but this pales in comparison when looking at a quantum register. The example below illustrates that situation. ⧫

Example 4.6 Suppose we can describe a quantum register of n qubits analogous to classical framework... meaning we can write each qubit as

$$\begin{aligned} & a_i \left|0\right\rangle + b_i \left|1\right\rangle fori = 1, ..., n. \\ & (a_1 \left|0\right\rangle + b_1 \left|1\right\rangle) \otimes (a_2 \left|0\right\rangle + b_2 \left|1\right\rangle) \otimes ... \otimes (a_n \left|0\right\rangle + b_n \left|1\right\rangle) \quad \in \mathbb{C}^{2^n} \end{aligned} \tag{4.53}$$

The main problem here is that we have superposition, we need to include the tangled vectors along with the pure states. this means that a state of the system should be

$$\left|\psi\right\rangle = \sum a \left|i_1\right\rangle \otimes \left|i_2\right\rangle \otimes \left|i_3\right\rangle \otimes ... \otimes \left|i_n\right\rangle \tag{4.54}$$

Written differently, this is a linear combination of basis vectors:

$$\{\left|0000\right\rangle, \left|0001\right\rangle, \left|0010\right\rangle, ..., \left|1111\right\rangle\} \tag{4.55}$$

As you can see, in the same four quantum bits, we can store infinitely more data in the same four bits because we can have some linear combination of the same bits, all because of superposition. This is only one of the powerful aspects of quantum computing, more will be discussed in future paragraphs. ⧫

4.1.4 More common quantum states

Lets look at a different basis, far different from our standard basis we use.

Example 4.7 Suppose we have two qubits ($C^2 \otimes \mathbb{C}^2$), the set

$$\begin{aligned} \{ \left|\phi^+\right\rangle = \frac{1}{\sqrt{2}}(\left|00\right\rangle + \left|11\right\rangle), \qquad & \left|\phi^-\right\rangle = \frac{1}{\sqrt{2}}(\left|00\right\rangle - \left|11\right\rangle), \\ \left|\psi^+\right\rangle = \frac{1}{\sqrt{2}}(\left|01\right\rangle + \left|10\right\rangle), \qquad & \left|\psi^-\right\rangle = \frac{1}{\sqrt{2}}(\left|01\right\rangle + \left|10\right\rangle)\} \end{aligned} \tag{4.56}$$

This is called the Bell Basis, another orthonormal basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$. Something about this basis is very different from other basis we have seen in the past, can you see it? None of these states are separable, *each of the Bell states is* ***entangled!*** This will become more important when we further understand why superposition is so powerful, but I will save that definition when we start to talk about cryptography. ⧫

4.1.5 Error recovery and avoidance

If there is one thing that we all know, it is that there are two things that are actually perfect: sweet baby Jesus and Grandma's pecan pie. Quantum bits are neither of these, so they are in fact not perfect. Therefore we need avenues to safeguarding and detecting against imperfections in the line, whether that be noise, some delay, or a software reception error. Let us first look at how this is solveed in a classical manner then we can approach how it is solved in the quantum realm.

Classical repetition codes Suppose a user sends 10**1**1 over a channel, but Sally, the person receiving the message, gets 10**0**1.

We must be able to identify that there has been an error in the third bit. One of the simplest ways to avoid this is by sending four of each bit, like this: 1111 0000 1111 1111.

Now suppose that Sally receives 1111 0100 1110 1110. She can take the bit that shows up most commonly in these four bit strings, and choose that as the bit to be used in that slot. The more bits there are, the lower the chance that there will be an error in the message transmission.

Woohoo, that is awesome, surely it is that simple to do the same in the quantum realm right? **WRONG, SO SO WRONG, HOW NIAVE OF YOU.** This fails for a couple reasons:

1. Measurements cause collapse, meaning that as soon as we measure that bit, it is useless to us, at least in a quantum sense.
2. All bits are distinct to one another, so sending "the same bit" carries a different weight when thinking about quantum bits compared to classical bits.

So now we know that there is no way to send redundant information because of the properties of particles in superposition, but there must be a way to interact with quantum bits without collapsing them right? There is, it is through quantum gates, which simply put are unitary transformations applied to qubits. Just because we can use quantum gates to interact with qubits does not mean that we are able to freely clone systems with these gates to accomplish this redundancy, in fact, this is the opposite of what we can do, a theory called the no-cloning theorem disproves this. I may add it if I have time, if not, sorry. We will see how my fingers feel after I finish this chapter.

4.2 Qubits and Quantum Gates

Looking back, it seems that we jumped straight into this wild world of qubits and quantum gates, but I do not think there was every a simple definition of what a qubit is, so lets review that before jumping into another section.

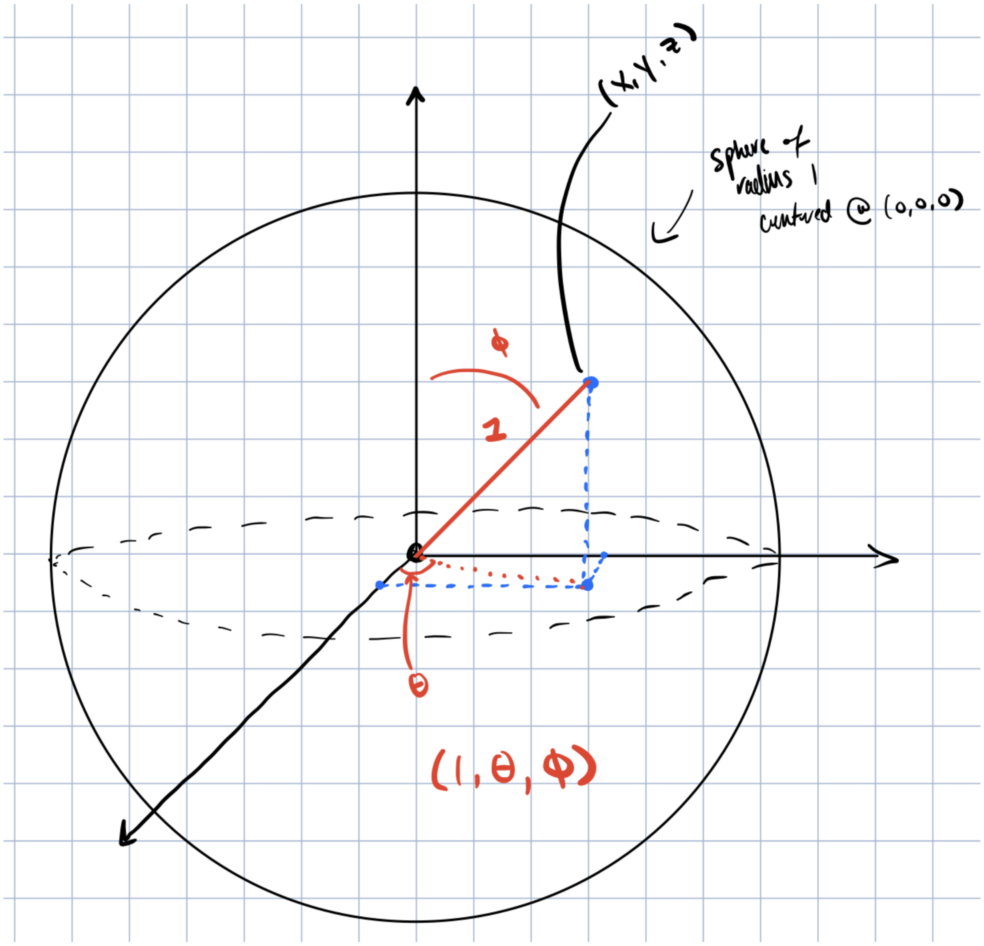
Qubit Review A qubit is simply a linear combination of two of our standard basis vectors, written in math it is

$$|\psi\rangle = \alpha \underbrace{|0\rangle}_{\begin{bmatrix}1\\0\end{bmatrix}} + \beta \underbrace{|1\rangle}_{\begin{bmatrix}0\\1\end{bmatrix}} \text{ where } \alpha = \alpha_1 + \alpha_2 i, \;\; \beta = \beta_1 + \beta_2 i \tag{4.57}$$

Well this is all well and good, but we really need another way to visualize this that can get rid of these pesky complex numbers. This should ring some bells for any of you math enthusiasts out there, it is time to break out some 3d polar coordinates! In this case, we plan to use the Bloch Sphere.

4.2.1 Bloch Sphere Introduction

As mentioned above, the Bloch Sphere is a polar coordinate system used to describe quantum particles, while minimizing the complex plane interaction. I think the easiest way to understand it is to see a picture of it then dive into the math, so check this out:



As you can see, the coordinate system we use for this is $(\underbrace{1}_{\text{Magnitude}}, \underbrace{\theta}_{\text{y-axis rotation}}, \underbrace{\phi}_{\text{Downward z-direction}})$, as you can see, the highest magnitude of a point is 1, meaning that it lies on the outer rim of the graph, with 100% of the system "accounted for". The following two entries in the coordinates will be described in depth as they are much more complicated.

What is ψ when in polar form?

$$\begin{aligned}
|\psi\rangle &= \mathbb{C}^2 \\
|\psi\rangle &= \alpha\,|0\rangle + \beta\,|1\rangle \\
|\psi\rangle &= \underbrace{|\alpha| e^{i\gamma_1}}_{\text{Polar form of } \alpha} |0\rangle + \underbrace{|\beta| e^{i\gamma_2}}_{\text{Polar form of } \beta} |1\rangle \\
|\psi\rangle &= \boxed{\underbrace{e^{i\gamma_1}}_{\text{Global phase shift}} (|\alpha|\,|0\rangle + |\beta| \underbrace{e^{i(\gamma_2-\gamma_1})}_{\text{Relative phase shift}}}
\end{aligned} \tag{4.58}$$

What is θ, ϕ? Now that we have the polar form of the state vector ψ, we can derive what θ and ϕ are. Let us first define

$$\underbrace{\theta}_{[0,2\pi]} = \gamma_2 - \gamma_1 \tag{4.59}$$

Simplifying what this math is saying is take the global phase shift, set this as the "perspective", then we can look at how far out of phase our β is and choose that as our relative phase shift. This means we can completely eliminate the imaginary component from *alpha* as we are setting that vector as the point of view. Now lets find ϕ.

Find $\underbrace{\phi}_{[0,\pi]}$ such that $(cos(\frac{\phi}{2}), sin(\frac{theta}{2}) = (|\alpha|, |\beta|)$, that means that

$$|\psi\rangle \underbrace{\approx}_{\text{"Up to a global phase shift"}} cos(\frac{\phi}{2})\,|0\rangle + sin(\frac{\theta}{2}) e^{i\theta}\,|1\rangle \tag{4.60}$$

Lets look at an example of what this equation looks like with actual numbers.

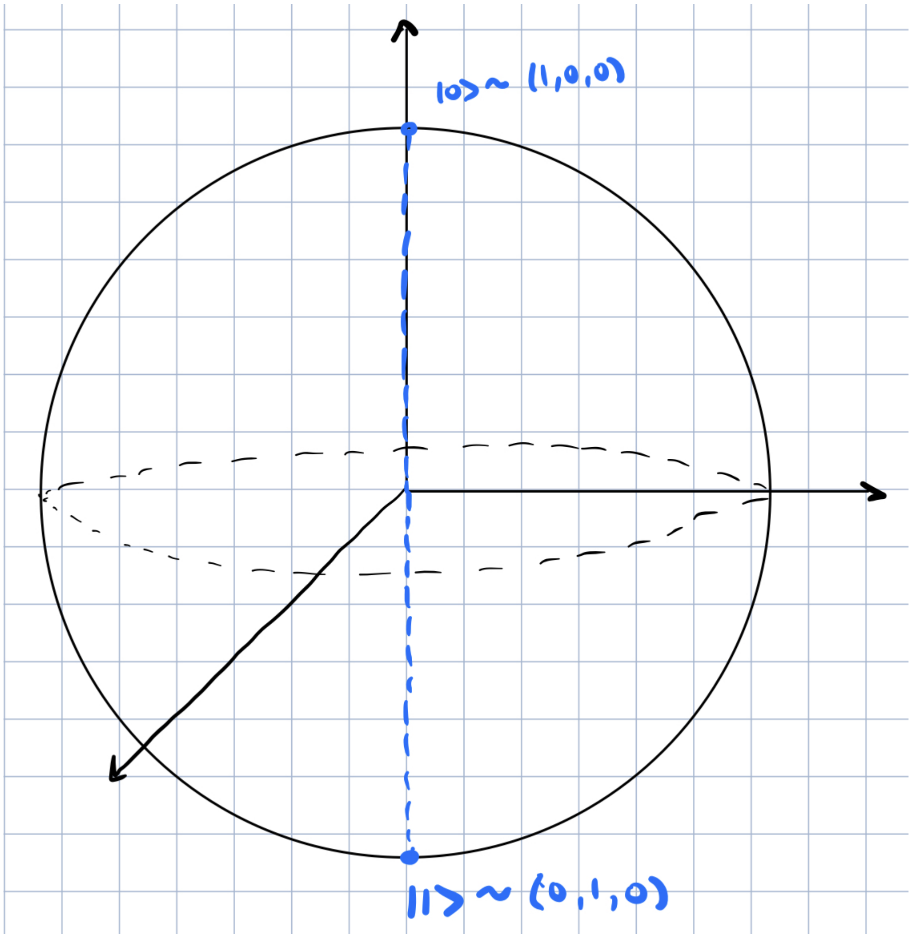
Example 4.8

$$\begin{gathered}
|0\rangle \rightarrow \alpha = 1, \quad \beta = 0 \\
cos(\frac{\phi}{2}) = 1 \quad sin(\frac{\phi}{2}) = 0 \\
\rightarrow \phi = 0 \\
\rightarrow \theta = 0 \\
|0\rangle \approx (1, 0, 0)
\end{gathered} \tag{4.61}$$

$$\begin{gathered}
|1\rangle \rightarrow \alpha = 0, \quad \beta = 1 \\
cos(\frac{\phi}{2}) = 0 \quad sin(\frac{\phi}{2}) = 1 \\
\rightarrow \phi = \pi \\
\rightarrow \theta = 0 \\
|0\rangle \approx (1, 0, \pi)
\end{gathered} \tag{4.62}$$

As you can see in both of these we do not have any theta, that is because these are both pure states, meaning there is no relative phase shift, or more so it does not matter because we are able to get 0 or 1 from the sin and cos, this will change with states that are in some state of superposition, meaning that relative phase shift matters. ♦

Before we move on lets look at what that looks like on the Bloch Sphere:



As you can see, they are on the two "north/south" poles on the sphere, further reinforcing why that theta does not matter: because that spin on the y axis would not change how the vector looks! Very interesting!

4.3 Quantum Gates

Well we know what these particles look like, and we know how they are represented, but we still don't know how to do anything with them! Good thing we are about to learn about quantum gates. As we mentioned before, these gates are very important with doing things with the particles, more importantly without collapsing them. Lets get into it.

4.3.1 Classical Gate Overview

I have little desire to completely rehash classical logic gates, but in general it is something that changes a 0 to a 1, or swap, or clone. A couple quick examples are below.

Example 4.9

$$\begin{aligned} I &: 0 \to 0 \\ NOT &: 0 \to 1 \end{aligned} \tag{4.63}$$

As you can see, the identity passess the value of the bit through, while the not gate flips the value. This may seem elementary at first but combining these into more complex systems can lead to large and very complex operations. ♦

4.3.2 Quantum Gate Introduction

So real quick let us recall what a quantum gate is, it is just a unitary transformation (i.e. unitary matrices applied to our qubits). Something interesting is because they are unitary transformations, this means that all of these operations can be reversed. Lets look at a couple common ones.

Example 4.10

$$\begin{aligned} U &: |0\rangle \to \alpha\,|0\rangle + \beta\,|1\rangle \\ \underbrace{\sigma_x = X}_{\text{QNOT}} &: |0\rangle\,, |1\rangle \to 1 \\ -i\sigma_Y = Y &= |0\rangle\,, |1\rangle \to |{-1}\rangle\,, |0\rangle \\ \sigma_z = Z &= |0\rangle\,, |1\rangle \to |0\rangle\,, |{-1}\rangle \\ H &= |0\rangle\,, |1\rangle \to |\sigma_x^+\rangle\,, |\sigma_x^-\rangle \end{aligned} \tag{4.64}$$

As you can see we are using the Pauli spin matrices to make these quantum gates, but one of these gates may seem foreign, and rightfully so. "H" is the Hadamard gate, which if you can see, puts the particle into a state of complete superposition. This is an important gate that is used constantly for this exact purpose. ♦

4.3.3 Common Quantum Gates

Below is a list of some of the most common quantum gates that are useful in information transmission.

QNOT:

$$\begin{gathered} QNOT : \sigma_x \left(X\right) \begin{bmatrix} 0 & 12 \\ 1 & 0 \end{bmatrix} \\ |0\rangle \rightarrow |1\rangle \\ |1\rangle \rightarrow |0\rangle \end{gathered} \tag{4.65}$$

The next couple need three bits, two bits that are the inputs, then one bit that is the output based off of the two bits.

QAND:

$$|xyz\rangle \rightarrow \begin{bmatrix} x = y = 1, \ |xy\neg z\rangle \\ \text{else } \ |xyz\rangle \end{bmatrix} \tag{4.66}$$

QOR:

$$\begin{gathered} (I \otimes I \otimes X) {*} CNOT * (X \otimes X \otimes I) \\ |00\rangle \rightarrow |00\rangle \\ |01\rangle \rightarrow |01\rangle \\ |10\rangle \rightarrow |11\rangle \\ |11\rangle \rightarrow |10\rangle \end{gathered} \tag{4.67}$$

Hadamard:

$$\begin{gathered} |0\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ |1\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \end{gathered} \tag{4.68}$$

4.3.4 Measuring Qubits

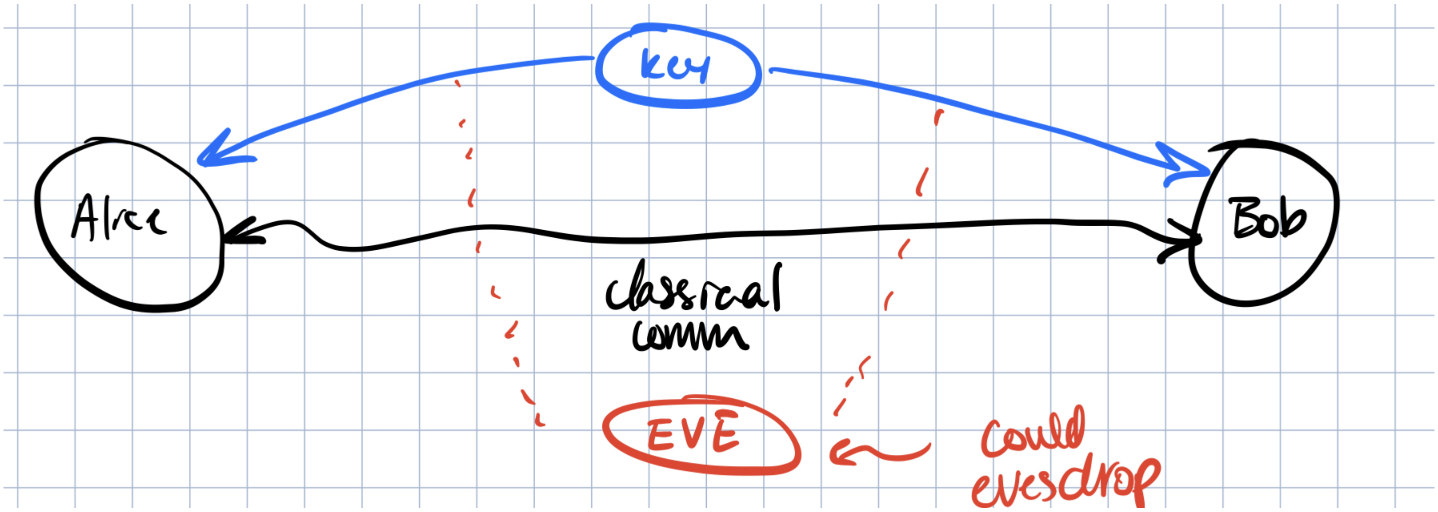
Recall: Suppose the single qubit system is in state $|x\rangle = a|0\rangle + b|1\rangle$, find the probability the state is $|0\rangle$ (+1 outcome) or $|1\rangle$ (-1 outcome) upon measurement of σ_z

Idea: Let's try and use the spectral projection associated to $\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$:

$$\begin{aligned} &= 1\,\underbrace{|0\rangle\langle 0|}_{M_0} + (-1)\,\underbrace{|1\rangle\langle 1|}_{M_1} \\ p(0) &= a^2 = \langle x|\,M_0\,|x\rangle \\ p(1) &= b^2 = \langle x|\,M_1\,|x\rangle \end{aligned} \tag{4.69}$$

4.3.5 Applying quantum measurements

BB84 Protocol - Quantum Key Distributions One of the most prevalent aspects of quantum information currently is through quantum encryption. The aspect of collapse makes this it impossible for someone to eavesdrop, more on that in the chapter. For a quick introduction into quantum key distribution, lets look at a protocol developed in 1984, muich before there was a single quantum computer even beginning to be built. Anyways, lets get into it, let us visualize what a typical communication looks like.



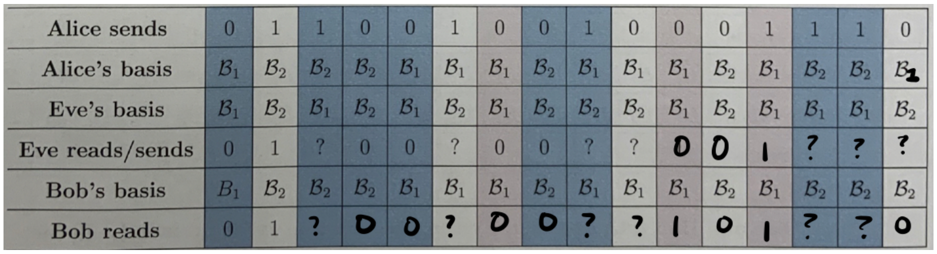
We will be using a **one time pad**, meaning when we get the key we will apply it to our data, flipping bits if it is a one and leaving it the same if not. Now we can set up the problem.

- First, over the classical channel, Alice will communicate to Bob what basis they will be preparing the particles with.
- Bob and Alice then take the stream of qubits (in superposition) and measure each measure them with random basis.

- Next, Bob tells Alice which basis he used to measure all of the different qubits, and if they did not choose the same, Alice secretly throws out those qubits because no information can be gleaned from them.
- In the end we should have ≈ $N/2$ good qubits (meaning w.r.t. the same basis, 50% chance.)
- If no one was listening, this is it, Alice chooses some number of those good qubits, as they both have the same values as those qubits are entangled, and use that as the one time pad key.

This can seem easy enough, and as long as there is no outside person listening this seems fine. Even if someone was listening, lets say for example on the classical channel, they would not be able to do anything with the information because they never dicuss the outcome of any of the particles. Eve actually can't even listen on the quantum channel where the information is being sent because if she made a measurement, the particle would collapse! That means that to make this impossible, like actually mathematically impossible to crack, we need to detect that someone is collapsing the particles.

Detecting collapse, tampering with information To check that someone is listening we will have to sacrifice a certain numbers of our above "good" qubits to check the outcome value to make sure no one collapsed the value and sent the particle on again. If eve collapsed the particle, the value that Bob reads will differ from what Alice reads. Alice will choose bits at random and check values, and if there are more than $\frac{N}{4}$ bits, we must retransmit because someone was listening in. A quick matrix of what this would look like is below.



Alice sends	0	1	1	0	0	1	0	0	1	0	0	0	1	1	1	0
Alice's basis	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_1$
Eve's basis	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$
Eve reads/sends	0	1	?	0	0	?	0	0	?	?	0	0	1	?	?	?
Bob's basis	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_2$
Bob reads	0	1	?	0	0	?	0	0	?	?	1	0	1	?	?	0

As you can see, when Eve measures, Bob receives a 0 even though Alice received a one and vice versa. These are the bits that are detected in tamper investigation.

Apologies for the slightly abbreviated version of this, this is very in depth and many mathematic concepts at play, but I realistically do not have time to spend writing all of those equations down. Please show me mercy.

4.4 Quantum Gates as Density Matrices

Something that is always a problem when it comes to quantum information transmission is reliability. When you are working at the atomic and sometimes subatomic particles, you have to be extremely precise. What happens if there is any level of noise or inaccuracy in the bits. Will the quantum logic still apply to the particles? This is where quantum gates being represented as density martices comes in.

Example 4.11 Suppose we want to send $|11\rangle$ as $|11\rangle$ by applying a CNOT. 80% of the time $|11\rangle$ is sent, 15% of the time $(H \otimes I)|10\rangle$ is sent, and 5% of the time $(H \otimes H)CNOT(H \otimes H)$ is sent. This means that the vector received is

$$0.80CNOT|10\rangle + .15(H \otimes I)|10\rangle + 0.05CNOT'|10\rangle \tag{4.70}$$

This can be represented in a much simpler manner using density matrices, to do this we must apply some quantum gate U to a vector. This gives us $\psi \to U\psi$.

Theorem 4.12 *If ρ is a pure state density matrix $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in \mathbb{C}^n$,*

$$\rho = |\psi\rangle\langle\psi| \to (u|\psi\rangle)(u|psi\rangle)^\dagger = U|\psi\rangle\langle\psi|U^\dagger = U\rho U^\dagger \tag{4.71}$$

5 Conclusion and Parting Words

Thank you for spending the time to read through this document, I realize that this may not be the best document by any stretch, and there is still lots of work to make this actually useful to someone who wants to learn, but taking into account that this was completed in tandem with two jobs and developing a satellite, I am actually somewhat satisfied. Overall, this is a field that is just barely blooming, and this is the math and concepts that are necessary to even begin to comprehend the new research being done in this field. Please take as much as possible from this, and have as much fun with this as I did in the semester I took this class. Goodbye!