MA 399 Intro to Quantum Information Theory

Mitch Hamidi and Lara Ismert

February 5, 2024

Abstract

The following are notes from a direct study at Embry-Riddle Aeronautical University, Prescott, AZ during Spring 2022.

Course webpage: https://cklixx.people.wm.edu/teaching/QC-invitation.html

"Quantum information science and quantum computing are rapidly growing areas. The study concerns the use of quantum properties to store, transmit, and manipulate data. Recent study has connected the topic to other research areas such as image processing, machine learning, neural network, etc. The study requires knowledge from a wide spectrum of different disciplines including mathematics, physics, computer science, chemistry, engineering, material science, etc.

The goal of this lecture series is to use elementary matrix theory approach to introduce the subject to beginners, and also provide a platform for people from different background to exchange experiences and idea about quantum properties and how they can be use to solve problems in quantum information, quantum computing, and other topics.

We will not assume any quantum mechanics background from the audience, and require only basic courses in calculus and linear algebra. Background in group theory and differential equations will be useful, but not necessary. The discussion will focus on three components and their expected outcome are listed below."

Contents

1 Vector Spaces

History: "Matrix Mechanics" led by Heisenberg vs. "Wave Mechanics" led by Schrödinger. It has been proven that the two conventions are equivalent.

A photon has two (classical) states: vertical and horizontal polarization – think Schrödinger's cat. The two states would be represented in binary – 0 or 1.

In quantum physics, we use the unit vectors "$|0\rangle$" $= \begin{bmatrix}1\\0\end{bmatrix}$ and "$|1\rangle$" $= \begin{bmatrix}0\\1\end{bmatrix}$ in $\mathbb{C}^2$ to describe these states.

1.1 Vectors

A vector space is an additive group equipped with a compatible scalar multiplication from $\mathbb{R}$ or $\mathbb{C}$. We will care about two vector spaces $\mathbb{C}^n$ over $\mathbb{C}$ and $\mathbb{R}^n$ over $\mathbb{R}$. Recall that $\mathbb{C}$ is the scalar field of complex numbers $a+bi$, where multiplication is defined by the rule $i^2=-1$, equipped with:

(a) a complex conjugation operation – $\overline{a+bi} = (a+bi)^* = a-bi$ and

(b) a size function called the **modulus** – $|a+bi| = \sqrt{a^2+b^2}$.

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 $\mathbb{C}^n$. ♦

Example 1.2 $|0\rangle, |1\rangle$ are vectors in the 2-dimensional vector space, $\mathbb{C}^2$. ♦

Remark 1.3 Usually, $|0\rangle$ will denote the **zero vector** in $\mathbb{C}^n$. ♦

Definition 1.4 A **linear combination** of the vectors $|v_1\rangle, \ldots, |v_k\rangle$ is a vector $|v\rangle$ such that $|v\rangle = \sum_{i=1}^{k} c_i |v_i\rangle$ for some scalars $c_1, \ldots, c_k$. If $S = \{|v_1\rangle, \ldots, |v_k\rangle\}$, then $\text{Span}(S)$ is the set of all linear combinations of $|v_1\rangle, \ldots, |v_k\rangle$. ♦

Exercise 1.5 Consider $S = \{|v_1\rangle, |v_2\rangle, |v_3\rangle\}$, where

$$|v_1\rangle = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad |v_2\rangle = \begin{bmatrix} i \\ 1 \end{bmatrix}, \quad |v_3\rangle = \begin{bmatrix} 0 \\ i \end{bmatrix}$$

1. Give a linear combination of the vectors in S.
2. Determine if $\begin{bmatrix} 1+i \\ 200-i \end{bmatrix}$ is in $\text{Span}(S)$. ♦

Definition 1.6 A **subspace** of $\mathbb{C}^n$ is a set of vectors M in $\mathbb{C}^n$ satisfying:

(i) $|0\rangle \in M$,

(ii) $|u\rangle + |v\rangle \in M$ for all $|u\rangle, |v\rangle \in M$, and

(iii) $c|v\rangle \in M$ for all $|v\rangle \in M, c \in \mathbb{C}$. ♦

Proposition 1.7 *Let $S = \{|v_1\rangle, \ldots, |v_k\rangle\}$ be a set of vectors in $\mathbb{C}^n$. Then $\text{Span}(S)$ is a subspace of $\mathbb{C}^n$.*

Exercise 1.8 Describe $\text{Span}(S)$ in Exercise 1.5 "geometrically." ♦

1.2 Linear Independence

Definition 1.9 A set of vectors $\{|v_1\rangle, \ldots, |v_k\rangle\}$ is **linearly dependent** if

$$\sum_{i=1}^{k} c_i |v_k\rangle = |0\rangle \tag{1.1}$$

has a non-trivial solution, i.e., there exist scalars $c_1, \ldots, c_k$ not all zero that satisfy (1.1). If there exists no non-trivial solution to (1.1), then $\{|v_1\rangle, \ldots, |v_k\rangle\}$ is **linearly independent**.♦

Example 1.10 Show that S in Example 1.5 is linearly dependent. ♦

Exercise 1.11 (Exercise 1.1 in [NO08]) Find the condition under which the following two vectors are linearly independent

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

♦

Theorem 1.12 (Theorem 1.1 in [NO08], "Theorem Too Many Vectors") *If a set of k vectors in $\mathbb{C}^n$ is linearly independent, then the number k satisfies $k \leq n$. Equivalently, any set of k vectors in $\mathbb{C}^n$ with $k > n$ is linearly dependent.*

Proof. (Matricial proof.) Let $\{|v_1\rangle, \ldots, |v_k\rangle\}$ be a set of k vectors in $\mathbb{C}^n$. Suppose $k > n$. Then $A = \begin{bmatrix} |v_1\rangle & \cdots & |v_k\rangle \end{bmatrix}$ is an $n \times k$ matrix with at most n pivot positions since $n < k$. Hence, A cannot have a pivot position in each column. Thus, the columns of A are linearly dependent. ■

Theorem 1.13 (Basis Theorem) *If S is a set of n linearly independent vectors in $\mathbb{C}^n$, then S must **span** $\mathbb{C}^n$, i.e., $Span(S) = \mathbb{C}^n$.*

Proof. (Matricial proof.) Let $S = \{|v_1\rangle, \ldots, |v_n\rangle\} \subset \mathbb{C}^n$ be linearly independent and suppose $|v\rangle \in \mathbb{C}^n$. Since S is linearly independent, $\begin{bmatrix} |v_1\rangle & \cdots & |v_n\rangle \end{bmatrix}$ is an invertible $n \times n$ matrix, which implies the matrix equation $\begin{bmatrix} |v_1\rangle & \cdots & |v_n\rangle \end{bmatrix} |c\rangle = |v\rangle$ must have a solution.■

1.2.1 Bases and Dimension

Definition 1.14 Let M be a subspace of $\mathbb{C}^n$. A **basis** for M is a set of linearly independent vectors S in M that span M, i.e., S is a linearly independent **spanning set** for M. The **dimension** of M is the number of vectors in a basis for M. – (which is well-defined!) ♦

Example 1.15 The set of vectors $\mathcal{E} = \{|e_1\rangle, \ldots, |e_n\rangle\}$ in $\mathbb{C}^n$, where $|e_k\rangle$ is the vector with 1 in the k^{th} entry and 0 elsewhere, is called the **standard basis** for $\mathbb{C}^n$. ♦

Remark 1.16 A photon (**state**) $|\psi\rangle$ in a quantum environment is a unit vector in **superposition** of "$|0\rangle$" $= |e_1\rangle$ and "$|1\rangle$" $= |e_2\rangle$. ♦

Example 1.17 Find a basis for Span (S) from Example 1.5.

Exercise 1.18 (Exercise 1.2 in [NO08]) Show that the set formed by the following vectors is a basis for $\mathbb{C}^3$.

$$|v_1\rangle = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad |v_2\rangle = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad |v_3\rangle = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}.$$

♦

End of Day 1: Jan 13, 2022

1.3 Inner Products

Every column ("ket") vector $|\psi\rangle \in \mathbb{C}^n$ has an associated row ("bra") vector $\langle\psi|$. We call the vector space of row vectors with n entries from $\mathbb{C}$ the **dual space** of $\mathbb{C}^n$ and denote it by $\mathbb{C}^{n*}$.

Definition 1.19 We equip $\mathbb{C}^n$ with an **involution** (*conjugate transpose*) operation $\dagger : \mathbb{C}^n \to \mathbb{C}^{n*}$ defined by

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}^\dagger = \begin{bmatrix} x_1^* & \cdots & x_n^* \end{bmatrix}.$$

Given a vector $|v\rangle \in \mathbb{C}^n$, we define the **bra vector** $\langle v|$ to be $\langle v| = |v\rangle^\dagger$. ♦

Example 1.20 If $|v\rangle = \begin{bmatrix} 1 \\ -\pi i \\ 5 - \frac{5}{6}i \end{bmatrix}$, then $\langle v| = \begin{bmatrix} 1 & \pi i & 5 + \frac{5}{6}i \end{bmatrix}$. ♦

Definition 1.21 Given vectors $|u\rangle, |v\rangle \in \mathbb{C}^n$, the **inner product** of $|u\rangle$ and $|v\rangle$ is

$$\langle u|v\rangle := (|u\rangle)^\dagger |v\rangle = \sum_{i=1}^{n} u_i^* v_i.$$

♦

Example 1.22 Let $|x\rangle = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, |y\rangle = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, and $|z\rangle = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$. Then $\langle x|y\rangle = 1, \langle x|z\rangle = 0$, and $\langle y|z\rangle = -2$. Since $\langle x|z\rangle = 0$, we say $|x\rangle$ and $|z\rangle$ are **orthogonal**. ♦

On $\mathbb{R}^n$, the inner product is just the usual dot product. An n-dimensional vector space over $\mathbb{C}$ equipped with an inner product is called a **Hilbert space**. Given an inner product, we can define a norm (size function) on $\mathbb{C}^n$.

Definition 1.23 Let $|v\rangle \in \mathbb{C}^n$. We define the **norm** of $|v\rangle$, denoted $\||v\rangle\|$ to be

$$\||v\rangle\| := \sqrt{\langle v|v\rangle} = \left(\sum_{i=1}^{n} |v_i|^2\right)^{1/2}.$$

If $\||v\rangle\| = 1$, we say $|v\rangle$ is a **unit vector**. In general, unit vectors in $\mathbb{C}^n$ represent quantum states with n physical states. ♦

Remark 1.24 The defining properties of a norm is that for all vectors $|u\rangle, |v\rangle \in \mathbb{C}^n$ and scalars $\alpha \in \mathbb{C}$ it satisfies:

(i) the **triangle inequality**, i.e. $\||u\rangle + |v\rangle\| \leq \||u\rangle\| + \||v\rangle\|$,

(ii) a **homogeneity property**, i.e., $\|\alpha\,|u\rangle\| = |\alpha|\,\||u\rangle\|$, and

(iii) a **positive definite property**, i.e., $\|u\| \geq 0$ with equality if and only if $|u\rangle = |0\rangle$.

Convince yourself that each of these properties holds true for the norm on $\mathbb{C}^n$.

The **Cauchy-Schwarz inequality** is a useful tool for estimates in an inner product space. It says for all vectors $|u\rangle, |v\rangle \in \mathbb{C}^n$, we have

$$|\langle u|v\rangle| \leq \||u\rangle\| \cdot \||v\rangle\|.$$ ♦

Exercise 1.25 (Exercise 1.3 in [NO08]) Let

$$|x\rangle = \begin{bmatrix} 1 \\ i \\ 2+i \end{bmatrix}, \quad |y\rangle = \begin{bmatrix} 2-i \\ 1 \\ 2+i \end{bmatrix}, \quad |z\rangle = \frac{\sqrt{2+\sqrt{5}}}{2+\sqrt{5}}\,|x\rangle.$$

Find $\||x\rangle\|, \langle x|y\rangle, \langle y|x\rangle$, and $\||z\rangle\|$. ♦

Exercise 1.26 (Exercise 1.4 in [NO08]) Prove that for all $|u\rangle, |v\rangle \in \mathbb{C}^n$

$$\langle u|v\rangle = \langle v|u\rangle^*.$$ ♦

Exercise 1.27 Prove that the inner product on $\mathbb{C}^n$ is a **sesquilinear form**. That is, show that the inner product is linear in the "ket component" and conjugate linear in the "bra component." ♦

1.3.1 Orthonormal Bases

Definition 1.28 A set of vectors $S = \{|v_1\rangle, \ldots, |v_k\rangle\}$ in $\mathbb{C}^n$ is called an **orthogonal set** if $\langle v_i|v_j\rangle = 0$ when $i \neq j$. ♦

Theorem 1.29 *If $S = \{|v_1\rangle, \ldots, |v_k\rangle\}$ in $\mathbb{C}^n$ is an orthogonal set of nonzero vectors, then S is linearly independent.*

Corollary 1.30 *Every orthogonal set of n vectors in $\mathbb{C}^n$ is a basis for $\mathbb{C}^n$.*

Proof. This follows from the previous theorem and Theorem 1.13. ■

Definition 1.31 An orthogonal basis of unit vectors in $\mathbb{C}^n$ is called an **orthonormal basis**. ♦

Example 1.32 The standard basis $\mathcal{E} = \{|e_1\rangle, \ldots, |e_n\rangle\}$ is an orthonormal basis for $\mathbb{C}^n$. ♦

Exercise 1.33 Show that $\mathcal{B} = \{|b_1\rangle, |b_2\rangle\}$, where

$$|b_1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, |b_2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$$

is an orthonormal basis for $\mathbb{C}^2$. This basis is the basis obtained by putting the screen at a 45° angle for the incoming photon in the polarization experiment. ♦

Theorem 1.34 *Let $\mathcal{B} = \{|f_1\rangle, \ldots, |f_n\rangle\}$ be an orthonormal basis for $\mathbb{C}^n$. Then each $|x\rangle \in \mathbb{C}^n$ can be expressed as*

$$|x\rangle = \sum_{i=1}^{n} \langle f_i|x\rangle |f_i\rangle .$$

It follows that

$$\sum_{i=1}^{n} |f_i\rangle \langle f_i| = I_n \tag{1.2}$$

where I_n is the $n \times n$ identity matrix.

Equation 1.2 is called the **completeness relation**. An orthogonal basis is sometimes called a *complete orthogonal set.*

1.3.2 (Vector) Projections

Example 1.35 Let $\mathcal{B} = \{|b_1\rangle, |b_2\rangle\}$ from Exercise 1.33. Find the **coordinates** (or components) of $|x\rangle = \begin{bmatrix}-2\\1\end{bmatrix}$ relative to basis $\mathcal{B}$, i.e., find the scalars $c_1, c_2 \in \mathbb{C}$ such that $c_1|b_1\rangle + c_2|b_2\rangle = |x\rangle$. ♦

Proposition 1.36 *Let $\mathcal{B} = \{|f_1\rangle, \ldots, |f_n\rangle\}$ be an orthonormal basis for $\mathbb{C}^n$ and for each $i = 1, \ldots, n$, define $P_i := |f_i\rangle\langle f_i|$. Then $\{P_1, \ldots, P_n\}$ is a set of $n \times n$ matrices satisfying for all $|v\rangle \in \mathbb{C}^n$ and each $i = 1, \ldots, n$:*

(i) $P_i |v\rangle \in Span(|f_i\rangle)$, *i.e.,* $P_i |v\rangle$ *is on the line spanned by* $|f_i\rangle$

(ii) $|v\rangle - P |v\rangle$ *is orthogonal to* $|f_i\rangle$

(iii) $P_i^2 = P_i$

(iv) $P_i P_j = 0$ *when* $i \neq j$

(v) $\sum_{i=1}^{n} P_i = I_n$

Definition 1.37 Let $\mathcal{B} = \{|f_1\rangle, \ldots, |f_n\rangle\}$ be an orthonormal basis for $\mathbb{C}^n$ and for each $i = 1, \ldots, n$, define $P_i := |f_i\rangle \langle f_i|$. Each P_i is an $n \times n$ matrix called a **projection** onto Span $(|b_i\rangle)$. ♦

Example 1.38 Let $\mathcal{B} = \{|b_1\rangle, |b_2\rangle\}$ be as in Exercise 1.33. Compute the projections P_1 and P_2 relative to $\mathcal{B}$. ♦

Exercise 1.39 (Exercise 1.5 in [NO08]) Let $\mathcal{B} = \{|b_1\rangle, |b_2\rangle\}$ be as in Exercise 1.33. Find the coordinates of $|v\rangle = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ relative to $\mathcal{B}$. ♦

1.3.3 The Gram-Schmidt Process

The **Gram-Schmidt Process** is an algorithm to produce orthonormal bases/sets from bases/linearly independent sets.

Theorem 1.40 (The Gram Schmidt Process) *Let* $\{|b_1\rangle, \ldots, |b_k\rangle\}$ *be a basis for a subspace* M *in* $\mathbb{C}^n$. *Define the set of vectors* $\{|f_1\rangle, \ldots, |f_k\rangle\}$ *as follows:*

Step 1:
  Define $|f_1\rangle := |b_1\rangle$.
Step $i+1$*:*
  for $i = 1 : k-1$

$$|f_{i+1}\rangle := \underbrace{|b_{i+1}\rangle - \left(\sum_{j=1}^{i} \frac{|f_j\rangle \langle f_j|}{\langle f_j | f_j \rangle} \right) |b_{i+1}\rangle}_{\text{projection operator notation}} = \underbrace{|b_{i+1}\rangle - \left(\sum_{j=1}^{i} \frac{\langle f_j | b_{i+1} \rangle}{\langle f_j | f_j \rangle} |f_j\rangle \right)}_{\text{vector notation}}$$

  end

Step normalization:
  for $i = 1 : k$

$$|f_i\rangle := \frac{1}{\| |f_i\rangle \|} |f_i\rangle$$

  end

Then $\{|f_1\rangle, \ldots, |f_k\rangle\}$ *is an orthonormal basis for* M.

Proof. ($k = 2$ case... not a full proof.) First, note that $|b_i\rangle$ is a nonzero vector for all $i = 1, \ldots, k$ since $\{|b_1\rangle, \ldots, |b_k\rangle\}$ is linearly independent. Thus, the definition of $|f_{i+1}\rangle$ in the algorithm is well-defined. Second, the vectors are clearly unit vectors after the normalization step so we need only show that $\{|f_1\rangle, \ldots, |f_k\rangle\}$ as defined in the first loop is an orthogonal set that spans M. The formal proof is completed by induction. Let's just think about the $k = 2$ case.

Define $|f_1\rangle$ and $|f_2\rangle$ according to the above algorithm. We need to show that $|f_1\rangle$ and $|f_2\rangle$ are orthogonal and $\text{Span}(\{|f_1\rangle, |f_2\rangle\}) = \text{Span}(\{|b_1\rangle, |b_2\rangle\})$.

Note that $|f_2\rangle$ is nonzero since $|b_2\rangle \notin \text{Span}(\{|b_1\rangle\}) = \text{Span}(\{|f_1\rangle\})$ as $\{|b_1\rangle, |b_2\rangle\}$ is linearly independent. To see that $|f_1\rangle$ and $|f_2\rangle$ are orthogonal, we express the inner product of $|f_1\rangle$ and $|f_2\rangle$ using the vector notation in the algorithm to get

$$\langle f_1|f_2\rangle = \langle f_1| \left(|b_2\rangle - \left(\frac{\langle f_1|b_2\rangle}{\langle f_1|f_1\rangle} \right) |f_1\rangle \right).$$

Distributing $\langle f_1|$, we get

$$\langle f_1|f_2\rangle = \langle f_1|b_2\rangle - \left(\frac{\langle f_1|b_2\rangle}{\langle f_1|f_1\rangle} \right) \langle f_1|f_1\rangle.$$

Cancelling the factor of $\langle f_1|f_1\rangle \neq 0$, we get $\langle f_1|f_2\rangle = \langle f_1|b_2\rangle - \langle f_1|b_2\rangle = 0$. Thus, $|f_1\rangle$ and $|f_2\rangle$ are orthogonal.

By definition, we have $\text{Span}(\{|f_1\rangle, |f_2\rangle\})$ is a subspace of $\text{Span}(\{|b_1\rangle, |b_2\rangle\})$. Since $|b_1\rangle = |f_1\rangle \in \text{Span}(\{|f_1\rangle, |f_2\rangle\})$, we get that $|b_2\rangle = |f_2\rangle + \left(\frac{\langle f_1|b_2\rangle}{\langle f_1|f_1\rangle} \right) |f_1\rangle \in \text{Span}(\{|f_1\rangle, |f_2\rangle\})$. Hence, $\text{Span}(\{|b_1\rangle, |b_2\rangle\})$ must be a subspace of $\text{Span}(\{|f_1\rangle, |f_2\rangle\})$, and it follows that $\text{Span}(\{|f_1\rangle, |f_2\rangle\}) = \text{Span}(\{|b_1\rangle, |b_2\rangle\})$. This proves the $k = 2$ case. ■

Example 1.41 Use the Gram-Schmidt Process to produce an orthonormal basis $\mathcal{B}$ for $M = \text{Span}(\{|v_1\rangle, |v_2\rangle\})$, where

$$|v_1\rangle = \begin{bmatrix} 1 \\ 1 \\ i \end{bmatrix}, \quad |v_2\rangle = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Extend $\mathcal{B}$ to an orthonormal basis for $\mathbb{C}^3$. ♦

Exercise 1.42 (Exercise 1.6 in [NO08]) Use the Gram-Schmidt Process on $\{|v_1\rangle, |v_2\rangle, |v_3\rangle\}$ to produce an orthonormal basis $\mathcal{B}$ for $\mathbb{C}^3$, where

$$|v_1\rangle = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \quad |v_2\rangle = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \quad |v_3\rangle = \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}.$$

Find the coordinates of $|u\rangle = \begin{bmatrix} 1 \\ -2 \\ 7 \end{bmatrix}$ relative to $\mathcal{B}$. ♦

Exercise 1.43 (Exercise 1.7 in [NO08]) Use the Gram-Schmidt Process to produce an orthonormal basis $\mathcal{B}$ for $M = \text{Span}(\{|v_1\rangle, |v_2\rangle\})$, where

$$|v_1\rangle = \begin{bmatrix} 1 \\ i \\ 1 \end{bmatrix}, \quad |v_2\rangle = \begin{bmatrix} 3 \\ 1 \\ i \end{bmatrix}$$

♦

End of Day 2: Jan 18, 2022

2 Linear Transformations & Matrices

Let $M_{m,n}(\mathbb{C})$ denote the set of all $m \times n$ complex matrices. If $n = m$, just use $M_n(\mathbb{C})$. Since we'll always be considering matrices with entries from $\mathbb{C}$, we will simply write $M_{m,n}$ or M_n.

Example 2.1 The **Pauli matrices**, also known as the **spin matrices**, are

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \in M_2.$$

♦

Example 2.2 The **identity matrix** in M_n is the matrix I_n whose columns are the standard basis $\{|e_1\rangle, \ldots, |e_n\rangle\}$ in $\mathbb{C}^n$. ♦

2.1 Algebra of Matrices

$M_{m,n}$ is a vector space under matrix addition and scalar multiplication. When $n = m$, this vector space *also* has multiplication. Given $A \in M_{m,n}$ and $B \in M_{n,k}$ where $A = (a_{ij})$ and $B = (b_{rs})$, we can define $C := AB \in M_{m,k}$ to be the $m \times k$ matrix with pq-entry given by the dot product of the (row) vector $(a_{p1}, ..., a_{pn})$ against the (column) vector $(b_{1q}, ..., b_{nq})$:

$$c_{pq} = \begin{bmatrix} a_{p1} & \cdots & a_{pn} \end{bmatrix} \begin{bmatrix} b_{1q} \\ \vdots \\ b_{nq} \end{bmatrix} = \sum_{i=1}^{n} a_{pi} b_{iq}$$

Alternatively, if A has rows $\langle A_1|, ... \langle A_m|$ and B has columns $|B_1\rangle, ..., |B_k\rangle$, then

$$AB = \begin{bmatrix} A|B_1\rangle \ldots A|B_k\rangle \end{bmatrix} = \begin{bmatrix} \langle A_1| B \\ \vdots \\ \langle A_m| B \end{bmatrix}$$

Note that the second equality is the dual.

We can think of matrix multiplication as an *outer product* of two vectors, as opposed to an *inner product* that we like in Calculus 3. In particular, if A has columns $|A_1\rangle, \ldots, |A_n\rangle$ and B has rows $\{\langle B_1|, \ldots, \langle B_n|\}$, then we have

$$AB = \sum_{j=1}^{n} |A_j\rangle \langle B_j| \tag{2.3}$$

When D is an $n \times n$ diagonal matrix, multiplication is super easy:

$$AD = \begin{bmatrix} d_1 |A_1\rangle & \dots & d_n |A_n\rangle \end{bmatrix} \qquad \text{and} \qquad DB = \begin{bmatrix} d_1 \langle B_1| \\ \vdots \\ d_n \langle B_n| \end{bmatrix}$$

Sometimes we don't just want to look at scalar coefficients of a matrix (like "pixels"), but at a general block inside the matrix (as a submatrix). Suppose $D = D_1 \oplus \mathbf{0}_{n-\ell}$, i.e., D is a diagonal matrix with zero diagonal entries after the ℓ^{th}. Then $ADB = A_1 D_1 B_1$–the 0's in D's lower diagonal zero out. In particular, it sees only the first ℓ columns of A and the first ℓ rows of B.

2.1.1 Eigenvalues and Eigenvectors

We can find eigenvalues and eigenvectors for *square* matrices. Let $A \in M_n(\mathbb{C})$.

Definition 2.3 A vector $|x\rangle \in \mathbb{C}^n$ is an **eigenvector for** A if $|x\rangle \neq \mathbf{0}$ and there exists $\lambda \in \mathbb{C}$ such that $A|x\rangle = \lambda |x\rangle$. ♦

The application is that, when you can find a linearly independent set of n eigenvectors for A, and you define S to be the matrix whose columns are those eigenvectors, then $AS = SD$, where D is a diagonal matrix with corresponding eigenvalues as its entries.

Example 2.4 Consider

$$A = \begin{bmatrix} I_2 & 0 \\ 0 & \sigma_y \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix} \in M_4.$$

(a) Find the eigenvalues and corresponding normalized eigenvectors for A.

The eigenvalues are $1, 1, 1, -1$ with eigenvectors $|e_1\rangle, |e_2\rangle, \frac{1}{\sqrt{2}}(0,0,1,i)^t, \frac{1}{\sqrt{2}}(0,0,i,1)^t$.

(b) Show that the eigenvectors are mutually orthogonal.

(c) Show the projection operators associated to the mutually orthogonal eigenvectors satisfy the *completeness relation.*

(d) Find an invertible matrix which diagonalizes A.

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

Example 2.5 Suppose A is diagonalizable. Prove that $\mathrm{tr}(A)$ is the sum of the eigenvalues for A and $\det(A)$ is the product of the eigenvalues for A. ♦

This statement is true in more generality, but is harder to show.

2.1.2 Special Classes of Matrices

Definition 2.6 Given a matrix $A = (a_{ij}) \in M_{mn}$, the **Hermitian adjoint (or Hermitian conjugate)** of A is the matrix $A^\dagger = (a^*_{ji})$. ♦

It's important to notice that $A^\dagger$ is the **unique** matrix such that $\langle u|A|v\rangle = \langle v|A^\dagger|u\rangle^*$ for all $|u\rangle \in \mathbb{C}^m$ and $|v\rangle \in \mathbb{C}^n$.

Definition 2.7 $A \in M_n$ is **Hermitian** if $A = A^\dagger$, where $(A^\dagger)_{ij} = a^*_{ji}$. ♦

Example 2.8 The Pauli matrices are all Hermitian. Check σ_y. The matrix $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ is Hermitian but **not** unitary and **not** positive. ♦

Definition 2.9 The following are equivalent:

1. $A \in M_n$ is **positive semidefinite**.
2. $\langle x|A|x\rangle \geq 0$ for all $|x\rangle \in \mathbb{C}^n$.
3. A is Hermitian with nonnegative eigenvalues. ♦

Example 2.10 The Pauli matrices are Hermitian but **not** positive. The matrix $P_2 = |e_2\rangle\langle e_2|$ is positive semidefinite by **not** unitary. ♦

Definition 2.11 The following are equivalent.

1. $U \in M_n$ is **unitary**.
2. $U^\dagger = U^{-1}$
3. $UU^\dagger = I_n$
4. $U^\dagger U = I_n$
5. the columns of U form an orthonormal basis for $\mathbb{C}^n$ ♦

Example 2.12 The invertible matrix U in Example 2.4 is unitary but **not** Hermitian. ♦

Definition 2.13 $N \in M_n$ is **normal** if $NN^\dagger = N^\dagger N$. We say N **commutes** with its Hermitian adjoint. ♦

Example 2.14 The matrix $A = \begin{bmatrix} 2 & -i \\ -i & 2 \end{bmatrix}$ is **not** Hermitian or unitary, but it is normal since $A^\dagger A = AA^\dagger = 5I_2$. ♦

Note that Hermitian and unitary matrices are special cases of normal matrices. The **Spectral Theorem** says that a **non-Hermitian**, **non-unitary** normal matrix must have complex non-unimodular eigenvalues.

2.2 Decomposition of Matrices

2.2.1 Spectral Decomposition for Normal Matrices

Theorem 2.15 (Spectral Theorem for Normal Matrices) *A matrix $N \in M_n$ is normal if and only if there is a unitary matrix $U = \begin{bmatrix} |u_1\rangle & \dots & |u_n\rangle \end{bmatrix}$ and a diagonal matrix $D = diag(\lambda_1, ..., \lambda_n)$ such that*

$$N = \underbrace{UDU^\dagger}_{\text{unitary diagonalization of } N} = \underbrace{\sum_{j=1}^{n} \lambda_j |u_j\rangle \langle u_j|}_{\text{spectral decomposition of } N} .$$

End of Day 3: Jan 25, 2022

The matrices $P_j := |u_j\rangle \langle u_j|$ for $j = 1, \dots, n$ are called the spectral projections of N.

Example 2.16 Show that $A = \begin{bmatrix} 2 & -i \\ -i & 2 \end{bmatrix}$ has the unitary diagonalization

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2-i & 0 \\ 0 & 2+i \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^\dagger$$

and the spectral decomposition

$$A = (2-i) \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} + (2+i) \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} .$$

♦

Remark 2.17 That is, N has an orthonormal set of eigenvectors (that form a basis for $\mathbb{C}^n$) coming from the columns of U with corresponding eigenvalues coming from D. Note that there **are** diagonalizable matrices which are NOT normal. ♦

Example 2.18 Find a diagonalizable matrix A which is not normal. Prove both claims. [Hint: Your invertible matrix S had better not have orthogonal columns :)]

Try $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$. ♦

The following corollary distinguishes classes of normal matrices.

Corollary 2.19 *Let $A \in M_n$.*

- *A is Hermitian if and only if A is normal with real eigenvalues.*
- *U is unitary if and only if U is normal with eigenvalues of modulus 1 (we say "on $\mathbb{T}$").*
- *A is positive semidefinite if and only if A is normal with nonnegative eigenvalues.*

Example 2.20 Draw a Venn diagram that shows the relationships between normal, unitary, Hermitian, and positive semidefinite matrices. ♦

Theorem 2.21 *Suppose $N \in M_n$ is normal in the form*

$$N = UDU^\dagger = \sum_{j=1}^{n} \lambda_j |u_j\rangle \langle u_j| .$$

Then you get some really convenient, pretty [insert your favorite positive adjective] "functional" results.

- *If $k \in \mathbb{Z}^+$, then $N^k = \sum_{j=1}^n \lambda_j^k |u_j\rangle \langle u_j|$.*
- *If N is invertible and $k \in \mathbb{Z}^+$, then $N^{-k} = \sum_{j=1}^n \lambda_j^{-k} |u_j\rangle \langle u_j|$.*
- *If N has positive eigenvalues, then $N^r = \sum_{j=1}^n \lambda_j^r |u_j\rangle \langle u_j|$ for all $r \in \mathbb{R}$.*
- *If f is an* **analytic function on** $\mathbb{C}$*, then $f(N) = \sum_{j=1}^n f(\lambda_j) |u_j\rangle \langle u_j|$*

Definition 2.22 An **analytic function** f on a domain (open subset) $\Omega \subseteq \mathbb{C}$ is a function $f : \Omega \to \mathbb{C}$ which can be represented by a power series expansion $f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n$ for some $a \in \Omega$. ♦

Example 2.23 Consider $\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

(a) Find a unitary diagonalization for σ_x.

$$\sigma_x = \underbrace{\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}}_{UDU^\dagger}$$

(b) Find the spectral decomposition for σ_x.

$$\sigma_x = \underbrace{(1)\left(\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right) + (-1)\left(\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right)}_{(1)|\lambda_1\rangle\langle\lambda_1| + (-1)|\lambda_2\rangle\langle\lambda_2|} .$$

(c) Compute the matrix exponential Compute $\exp(i\alpha\sigma_x)$.

Using the functional calculus, $f(z) = e^{i\alpha z}$ is analytic everywhere so

$$\exp(i\alpha\sigma_x) = f(\sigma_x) = f(1)P_1 + f(-1)P_2.$$

By Euler's identity, we have $f(1) = e^{i\alpha} = \cos(\alpha) + i\sin(\alpha)$ and $f(-1) = e^{-i\alpha} = \cos(\alpha) - i\sin(\alpha)$. Hence, we have

$$\exp(i\alpha\sigma_x) = \cos(\alpha)(P_1 + P_2) + i\sin(\alpha)(P_1 - P_2).$$

Since the *completeness relation* guarantees $P_1 + P_2 = I$, we get

$$\exp(i\alpha\sigma_x) = \cos(\alpha)I + i\sin(\alpha)(P_1 - P_2) = \begin{bmatrix} \cos(\alpha) & i\sin(\alpha) \\ i\sin(\alpha) & \cos(\alpha) \end{bmatrix}. \tag{2.4}$$

♦

Remark 2.24 It's important to note that the first equality in Equation (2.4) generalizes to σ_y and σ_z by finding the appropriate spectral projections P_1 corresponding to $\lambda_1 = 1$ and P_2 corresponding to $\lambda_2 = -1$ to get ♦

Example 2.25 Compute $\exp(A)$ using A and the spectral decomposition found in Example 2.16. ♦

Example 2.26 Suppose H is Hermitian. Show that e^{iH} exists, determine what special type of matrix it is, and give its spectral decomposition. ♦

2.3 Hilbert Space Formalism for Quantum Mechanics

Quantum information theory uses quantum properties to help store, process, and transmit information. Let's discuss the basic background in quantum mechanics.

2.3.1 Copenhagen Interpretation of Quantum Mechanics

A single interpretation of the wave function is not agreed upon. We adopt the most popular interpretation, the **Copenhagen interpretation**. We'll explore the axioms through an example of a photon and its polarization.

A photon is prepared in the lab. Axiom 1 says we should represent the state of this photon by $|x\rangle$, but we've not yet explained why. You cannot definitively know the properties this photon possesses, despite having prepared it. This is because it behaves as a wave. This is what's meant by being a "quantum state."

> **Axiom 1.1.** A *vector state* $|\psi\rangle$ is a unit vector in a complex inner product space $\mathcal{H}$ (usually $\mathbb{C}^n$), which we call a *Hilbert space*.

You want to know the photon's polarization, so you decide to shoot the photon through a screen (generally called an "apparatus"). Classically, you would expect that, if $|x\rangle$ is vertically polarized, it will pass through the screen, but if $|x\rangle$ is horizontally polarized, it will be deflected by the screen.

But experiment shows that, even if you prepare a ton of copies of $|x\rangle$ in the exact same way, when you try to measure their polarization by passing them through a screen, *you don't get the same polarization every time.* Instead, experiment shows that $|x\rangle$ passes through the screen witht some probability $|\beta|^2$ and is deflected with some probability $|\alpha|^2$. Laws of probability require

$$|\alpha|^2 + |\beta|^2 = 1.$$

> **Axiom 1.2:** Linear combinations (or *superposition*) of the physical states are allowed in the state space.

A photon can be in a superposition of "horizontal" and "vertical" polarization:

$$|x\rangle = \alpha\,|h\rangle + \beta\,|v\rangle$$

such that $\|x\| = 1$. We associate to the apparatus used to make this measurement a Hermitian operator A given above. The possible states are orthogonal unit vectors and, in particular, these will be eigenvectors for A with real eigenvalues. We associate to this vertical screen apparatus a Hermitian operator $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, which has eigenvalues 1 and 0 with eigenvectors

$$|v\rangle := \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ and } |h\rangle := \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

respectively.

Axiom 2: An *observable* of a state $|\psi\rangle$ (or a physical quantity) corresponds to a Hermitian operator A such that a measurement will change $|\psi\rangle$ to an *eigenstate* (unit eigenvector) $|u\rangle$ of A with probability $|\langle u|\psi\rangle|^2$.

The act of shooting $|x\rangle$ through a screen is "observing its polarization," so we call the polarization of the photon an "observable." The probability $|x\rangle$ is vertically polarized is

$$|\langle v|x\rangle|^2 = |\beta \langle v|v\rangle|^2 = |\beta|^2$$

and the probability $|x\rangle$ is horizontally polarized is

$$|\langle h|x\rangle|^2 = |\alpha \langle h|h\rangle|^2 = |\alpha|^2.$$

Remark 2.27 If the photon passes through the screen, it will exit the other side as $|v\rangle$– all trace of α and β are lost. Both collapse AND renormalization occur. If the photon is deflected, it is bounces back as $|h\rangle$, and again, all information about α and β is completely gone. ♦

Example 2.28 In general, consider the finite dimensional case $\mathcal{H} = \mathbb{C}^n$. We can represent the observable A and the state $|\psi\rangle$ according to a spectral decomposition for A given by

$$A = \sum_{j=1}^{n} \lambda_j |u_j\rangle \langle u_j| = \sum_{j=1}^{n} \lambda_j P_j$$

and

$$|\psi\rangle = \sum_{j=1}^{n} c_j |u_j\rangle.$$

Note that each $c_j = \langle u_j|\psi\rangle$ (Why?). Apply A to ψ makes a measurement of the system. When the measurement is applied, the state $|\psi\rangle$ (the wave function) becomes (**collapses** to) $|u_j\rangle$ with a probability of $|c_j|^2 = |\langle u_j|\psi\rangle|^2$. The **expectation value** (or **mean value**) of the observable associated to A after measurements with respect to many copies of $|\psi\rangle$ is the weighted average of the expected outcomes

$$\langle A\rangle_\psi = \sum_{j=1}^{n} \lambda_j |c_j|^2 = \langle \psi|A|\psi\rangle.$$

♦

Remark 2.29 Note that a measurement produces only one outcome. The probability of obtaining an outcome λ_i is experimentally evaluated only after repeating measurements with many copies of the same state.

The coefficients $c_1, \ldots, c_n$ are called the **probability amplitude** of $|\psi\rangle$ with respect to the observable associated with A. ♦

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

$$i\hbar \frac{\partial |\psi\rangle}{\partial t} = H |\psi\rangle , \tag{2.5}$$

where $\hbar = 6.6260700410 \times 10^{-34}/2\pi$ m^{2}kg/s is the (reduced) **Planck constant** and H is a Hermitian matrix corresponding to the energy of the system and is called the **Hamiltonian**.

When the Hamiltonian H is time-independent, the Schrödinger equation's (2.5) has the solution

$$|\psi(t)\rangle = \exp\left(-itH/\hbar\right) |\psi(0)\rangle .$$

When H is time-dependent,

$$|\psi(t)\rangle = \exp\left(-\frac{i}{\hbar}\int_0^t H(s)\, ds\right) |x(0)\rangle .$$

Compare this to the 1×1 case, where $\psi'(t) = \alpha\psi(t)$ has the general solution $\psi(t) = e^{\alpha t}\psi(0)$.

End of Day 4: Feb 1, 2022

Example 2.30 Consider a physical system with Hamiltonian

$$H = -\frac{\hbar}{2}\omega\sigma_x = \text{“spin under magnetic field along } x\text{-axis”}$$

and suppose the initial state of the system is

$$|\psi(0)\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \text{“spin points in the } z\text{-direction at } t = 0\text{”}$$
$$= \text{eigenstate of } \sigma_z \text{ with eigenvalue } +1 \text{ at } t = 0.$$

(a) Find the wave function $|\psi(t)\rangle$ for $t > 0$.

The wave function $|\psi(t)\rangle$ $(t > 0)$ is found by solving the *Schrödinger equation* to get

$$|\psi(t)\rangle = \exp\left(i\frac{\omega}{2}\sigma_x t\right) |\psi(0)\rangle .$$

Thus, Equation (2.4) with $\alpha = \omega t/2$ yields

$$|\psi(t)\rangle = \begin{bmatrix} \cos(\omega t/2) & i\sin(\omega t/2) \\ i\sin(\omega t/2) & \cos(\omega t/2) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos(\omega t/2) \\ i\sin(\omega t/2) \end{bmatrix} .$$

(b) Find the probability for the system to have the outcome +1 and -1 upon measurement of σ_z at $t > 0$.

We find the probabilities by first finding the coordinates of $|\psi(t)\rangle$ relative to an orthonormal eigenbasis for σ_z, i.e., write $|\psi(t)\rangle = c_1 |\lambda_1\rangle + c_2 |\lambda_2\rangle$ where $|\lambda_1\rangle, |\lambda_2\rangle$ are unit eigenvectors for σ_z corresponding to the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$, respectively. Then we have

$$P_\uparrow(t) = |c_1|^2 = \cos^2\left(\frac{\omega}{2}t\right) \quad \text{and} \quad P_\downarrow(t) = |c_2|^2 = \sin^2\left(\frac{\omega}{2}t\right)$$

(c) Find the expected value after many measurements of σ_z.

$$\langle A\rangle_\psi = \langle\psi(t)|A|\psi(t)\rangle = (1)P_\uparrow(t) + (-1)P_\downarrow(t) = \cos^2\left(\frac{\omega}{2}t\right) - \sin^2\left(\frac{\omega}{2}t\right)$$

(d) Find the probability for the system to have the outcome +1 upon measurement of σ_x at $t > 0$.

$$P_\uparrow(t) = \frac{\left[\cos\left(\frac{\omega}{2}t\right) + \sin\left(\frac{\omega}{2}t\right)\right]^2}{2} \quad \text{and} \quad P_\downarrow(t) = \frac{\left[\cos\left(\frac{\omega}{2}t\right) - \sin\left(\frac{\omega}{2}t\right)\right]^2}{2}$$

(e) Find the expected value after many measurements of σ_y.

$$\langle A\rangle_\psi = \langle\psi(t)|A|\psi(t)\rangle = (1)P_\uparrow(t) + (-1)P_\downarrow(t) = 2\cos\left(\frac{\omega}{2}t\right)\sin\left(\frac{\omega}{2}t\right) = \sin(\omega t) \quad \blacklozenge$$

2.4 Heisenberg Uncertainty Principle

Definition 2.31 If the expectation value of the observable associated to A is $\mu = \langle A\rangle_\psi$, the **variance** of the observable associated to A after measurements with respect to many copies of $|\psi\rangle$ is the expectation value of the observable associated to $(A - \mu I)^2$

$$\mathrm{Var}_\psi(A) = \langle (A-\mu I)^2\rangle_\psi = \langle\psi|(A-\mu I)^2|\psi\rangle = \|(A-\mu I)|\psi\rangle\|^2.$$

The **standard deviation** of the observable associate to A is $\Delta(A) = \sqrt{\mathrm{Var}_\psi(A)}$. $\blacklozenge$

In a deterministic model, the variance of measurements should approach zero as the experiment apparatus is made more accurate.

2.5 Tensor Products

We will define a new way to combine two matrices to form a larger one. When the two matrices are both *vectors* from vector spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, we can actually build a *new* vector space $\mathcal{H}_1 \otimes \mathcal{H}_2$ from them.

Definition 2.32 Let $A \in M_{m,n}$ and $B \in M_{p,q}$, where $m, n, p, q \in \mathbb{N}$ could be entirely distinct. We define *the tensor product of A with B to be* the $mp \times nq$-matrix

$$A \otimes B := \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & \dots & \dots & A_{2n}B \\ \vdots & \ddots & & \vdots \\ A_{m1}B & \dots & \dots & A_{mn}B \end{bmatrix}$$

Example 2.33 Let $\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, and $\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Compute

1. $\sigma_x \otimes i\sigma_y$
2. $i\sigma_x \otimes \sigma_y$
3. $\sigma_y \otimes i\sigma_x$
4. $(\sigma_x \otimes \sigma_y) + (\sigma_y \otimes \sigma_z)$
5. $(\sigma_x + \sigma_y) \otimes (\sigma_y + \sigma_z)$
6. $(\sigma_z \otimes \sigma_z)^\dagger$

Note that, like matrix multiplication, the tensor product is not a commutative operation. ♦

Exercise 2.34 Let $\lambda \in \mathbb{C}$, $A \in M_{mn}$, $B \in M_{pq}$, $C \in M_{nr}$, and $D \in M_{qs}$. Show that

1. $\lambda(A \otimes B) = \lambda A \otimes B = A \otimes \lambda B$
2. $\lambda(A \otimes B) \neq \lambda A \otimes \lambda B$
3. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.
4. $A \otimes (B + C) = A \otimes B + A \otimes C$.
5. $(A \otimes B) + (C \otimes D) \neq (A + C) \otimes (B + D)$
6. $A \otimes B \neq B \otimes A$
7. $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$
8. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$

whenever the matrix operations are well-defined. ♦

2.5.1 Tensor Product Hilbert Spaces

Given two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, we define $\mathcal{H}_1 \otimes \mathcal{H}_2$ to be the Hilbert space generated by elements of the form $|v\rangle \otimes |w\rangle$ where $|v\rangle \in \mathcal{H}_1$ and $|w\rangle \in \mathcal{H}_2$. Generally speaking, they are finite linear combinations of the form

$$\sum_p \lambda_p |v_p\rangle \otimes |w_p\rangle$$

where $\lambda_p \in \mathbb{C}$, $|v_p\rangle \in \mathcal{H}_1$, $|w_p\rangle \in \mathcal{H}_2$, and p is any natural number.

Example 2.35 We sometimes denote $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ by $|1\rangle$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ by $|2\rangle$. Furthermore, (physics) folks like to write

- $|11\rangle := |1\rangle \otimes |1\rangle$
- $|21\rangle := |2\rangle \otimes |1\rangle$
- $|12\rangle := |1\rangle \otimes |2\rangle$
- $|22\rangle := |2\rangle \otimes |2\rangle$.

Compute these four vectors and their lengths. What would you conjecture $\||v\rangle \otimes |w\rangle\|$ to be in terms of $\|v\|$ and $\|w\|$? Prove that $\{|11\rangle, |12\rangle, |21\rangle, |22\rangle\}$ forms a basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$. Can you identify $\mathbb{C}^2 \otimes \mathbb{C}^2$ as $\mathbb{C}^n$ for some n? ⧫

Notation 2.36 The previous example makes the notational convention look a bit too simple. This is because the exact same 2 vectors coming from $\mathbb{C}^2$ are tensored together. Let's see what that looks like in general. Since we tend to subscript our Hilbert spaces with numbers, vectors that come from $\mathcal{H}_1$ will also have a subscript of 1 to help us remember where they came from: $|v_1\rangle \in \mathcal{H}_1, |v_2\rangle \in \mathcal{H}_2$. When you want to talk about *multiple* vectors from $\mathcal{H}_1$ and *multiple* vectors from $\mathcal{H}_2$, we now need two subscripts. So $\{|v_{1,1}\rangle, |v_{1,2}\rangle, ..., |v_{1,m}\rangle\}$ are m vectors coming from $\mathcal{H}_1$, while $\{|v_{2,1}\rangle, |v_{2,2}\rangle, ..., |v_{2,n}\rangle\}$ are n vectors coming from $\mathcal{H}_2$.

Because tensor products are yucky in bra-ket notation, we will go ahead and use the physicists' convention.

- $|v_1\rangle \otimes |v_2\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ will be written as $|v_1 v_2\rangle$. This makes sense because a state that has two particles in states $|v_1\rangle$ and $|v_2\rangle$ would be described as $|v_1 v_2\rangle$. Moral:

$$|v_1\rangle \otimes |v_2\rangle =: |v_1 v_2\rangle .$$

- Using this new convention and algebraic rules of tensor products, determine what $\langle v_1 v_2|$ naturally represents. ⧫

Definition 2.37 Just like there is an inner product on $\mathcal{H}_1$ and $\mathcal{H}_2$, there is an inner product on their tensor product, $\mathcal{H}_1 \otimes \mathcal{H}_2$. Given $|v_1\rangle, |w_1\rangle \in \mathcal{H}_1$ and $|v_2\rangle, |w_2\rangle \in \mathcal{H}_2$, define

$$\langle v_1 v_2 | w_1 w_2 \rangle_{\otimes} := \langle v_1 | w_1 \rangle_{\mathcal{H}_1} \cdot \langle v_2 | w_2 \rangle_{\mathcal{H}_2} .$$

Proposition 2.38 *If* $\mathcal{E}_1 = \{|e_{1,1}\rangle, |e_{1,2}\rangle, ..., |e_{1,m}\rangle\}$ *and* $\mathcal{E}_2 = \{|e_{2,1}\rangle, |e_{2,2}\rangle, ..., |e_{2,n}\rangle\}$ *are orthonormal bases for* $\mathcal{H}_1$*, and* $\mathcal{H}_2$*, respectively then*

$$\{|e_{1,i}\rangle \otimes |e_{2,j}\rangle : i \in [m], j \in [n]\}$$

forms an orthonormal basis for $\mathcal{H}_1 \otimes \mathcal{H}_2$*.*

Proof. Let $|e_{1,i}\rangle, |e_{1,k}\rangle \in \mathcal{H}_1$ and $|e_{2,j}\rangle, |e_{2,\ell}\rangle \in \mathcal{H}_2$. Then

$$\langle e_{1,i} e_{2,j} | e_{1,k} e_{2,\ell} \rangle_{\otimes} = \langle e_{1,i} | e_{1,k} \rangle_{\mathcal{H}_1} \cdot \langle e_{2,j} | e_{2,\ell} \rangle_{\mathcal{H}_2} = \delta_{i=k} \cdot \delta_{j=\ell}.$$

Therefore, $\{|e_{1,i} e_{2,j}\rangle : i \in [m], j \in [n]\}$ forms an orthonormal set, which is necessarily linearly independent.

Now, let $\sum_p \lambda_p |v_p\rangle \otimes |w_q\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Since $|v_p\rangle \in \mathcal{H}_1$ and $|w_p\rangle \in \mathcal{H}_2$, they can be expressed as

$$|v_p\rangle = \langle e_{1,1}|v_p\rangle |e_{1,1}\rangle + \langle e_{1,2}|v_p\rangle |e_{1,2}\rangle + ... + \langle e_{1,m}|v_p\rangle |e_{1,m}\rangle = \sum_{i=1}^{m} \langle e_{1,i}|v_p\rangle |e_{1,i}\rangle$$

and

$$|w_p\rangle = \langle e_{2,1}|w_p\rangle |e_{2,1}\rangle + \langle e_{2,2}|w_p\rangle |e_{w,2}\rangle + ... + \langle e_{2,n}|w_p\rangle |e_{2,n}\rangle = \sum_{j=1}^{n} \langle e_{2,j}|w_p\rangle |e_{2,j}\rangle .$$

Hence,

$$\begin{aligned}\sum_p \lambda_p[(|v_p\rangle) \otimes (|w_p\rangle)] &= \sum_p \lambda_p \left[\left(\sum_{i=1}^{m} \langle e_{1,i}|v_p\rangle |e_{1,i}\rangle \right) \otimes \left(\sum_{j=1}^{n} \langle e_{2,j}|w_p\rangle |e_{2,j}\rangle \right) \right] \\ &= \sum_p \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_p \langle e_{1,i}|v_p\rangle \langle e_{2,j}|w_p\rangle (|e_{1,i}\rangle \otimes |e_{2,j}\rangle).\end{aligned}$$

We have shown that an arbitrary element of $\mathcal{H}_1 \otimes \mathcal{H}_2$ can be written as a linear combination of the set $\{|e_{1,i}\rangle \otimes |e_{2,j}\rangle : i \in [m], j \in [n]\}$. Therefore, $\{|e_{1,i}\rangle \otimes |e_{2,j}\rangle : i \in [m], j \in [n]\}$ is an orthonormal basis for $\mathcal{H}_1 \otimes \mathcal{H}_2$. ■

In convenient notation, we'd say $\{|e_{1,i}e_{2,j}\rangle : i \in [m], j \in [n]\}$ forms an orthonormal basis, so

$$\sum_p \lambda_p |v_p w_p\rangle = \sum_p \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_p \langle e_{1,i}|v_p\rangle \langle e_{2,j}|w_p\rangle (|e_{1,i}e_{2,j}\rangle).$$

Corollary 2.39 *If* $\dim(\mathcal{H}_1) < \infty$ *and* $\dim(\mathcal{H}_2) < \infty$, *then*

$$\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = \dim(\mathcal{H}_1) \cdot \dim(\mathcal{H}_2).$$

Corollary 2.40 $\mathbb{C}^m \otimes \mathbb{C}^n \cong \mathbb{C}^{mn}$

Example 2.41 (2.3) Consider a *bipartite state* $|\psi\rangle$ written in terms of an orthonormal basis as

$$|\psi\rangle = \frac{1}{2} (|e_{1,1}e_{2,1}\rangle + |e_{1,1}e_{2,2}\rangle + i\,|e_{1,3}e_{2,1}\rangle + i\,|e_{1,3}e_{2,2}\rangle)$$

whose coefficients form a matrix $C = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ i & i \end{bmatrix}$.

1. How many Hilbert spaces comprise the system, and what are their dimensions?
2. What Hilbert space does $|\psi\rangle$ belong to? What is the dimension of this Hilbert space?
3. Find the probability that after measurement $|\psi\rangle$ is found in state $|e_{1,1}e_{2,1}\rangle$ OR $|e_{1,3}e_{2,2}\rangle$.

♦

–End class 2/8/22–

Exercise 2.42 Let $|a\rangle, |b\rangle, |c\rangle, |d\rangle \in \mathbb{C}^n$. Show that

$$(|a\rangle\langle b|) \otimes (|c\rangle\langle d|) = (|a\rangle \otimes |c\rangle)(\langle b| \otimes \langle d|) = |ac\rangle\langle bd|.$$

Exercise 2.43 Suppose

$$e_{1,1} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ i\end{bmatrix}, \qquad e_{1,2} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ -i\end{bmatrix}$$

and

$$e_{2,1} = \begin{bmatrix}1\\ 0\\ 0\end{bmatrix}, \quad e_{2,2} = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\ 1\\ -1\end{bmatrix}, \quad e_{2,3} = \begin{bmatrix}0\\ 0\\ i\end{bmatrix}.$$

Prove or disprove: $\{|e_{1,i}e_{2,j}\rangle : i \in [2], j \in [3]\}$ forms an orthonormal basis for $\mathbb{C}^2 \otimes \mathbb{C}^3$. ♦

2.6 Multipartite Systems

A quantum system could be comprised of multiple subsystems. For example, the system could contain two particles – each having its own Hilbert space $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. The total system containing both particles is $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. We call this a *bipartite system*. When there are more than two subsystems, we call the system *multipartite*.

We will want to consider multiple systems together as we march towards quantum computing.

Exercise 2.44 Show that if $|\psi_1\rangle \in \mathbb{C}^n$ and $|\psi_2\rangle \in \mathbb{C}^m$ are states (unit vectors), then $|\psi_1\rangle \otimes |\psi_2\rangle \in \mathbb{C}^n \otimes \mathbb{C}^m$ is a state (unit vector). ♦

Definition 2.45 A vector $|\psi\rangle \in \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ is *separable* (*or a tensor product state*) if it can be written as $|\psi_1\rangle \otimes |\psi_2\rangle$ for some $|\psi_1\rangle \in \mathcal{H}_1$ and $|\psi_2\rangle \in \mathcal{H}_2$. ♦

Example 2.46 Let $|\psi\rangle$ represent a bipartite state in $\mathbb{C}^2 \otimes \mathbb{C}^2$ which can be described as

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|1\rangle \otimes |1\rangle + |2\rangle \otimes |2\rangle) = \frac{1}{\sqrt{2}}(|11\rangle + |22\rangle).$$

Determine if $|\psi\rangle$ is separable.

If $|\psi\rangle$ were separable, then it should be able to be written as $|\psi_1\rangle \otimes |\psi_2\rangle$, and moreover, using $|1\rangle, |2\rangle$ as a basis for $\mathbb{C}^2$, we could write $|\psi_1\rangle = c_1|1\rangle + c_2|2\rangle$ and $|\psi_2\rangle = d_1|1\rangle + d_2|2\rangle$ for some $c_1, c_2, d_1, d_2 \in \mathbb{C}$. Thus, if $|\psi\rangle$ is separable, we should be able to write

$$\begin{aligned}|\psi\rangle &= |\psi_1\rangle \otimes |\psi_2\rangle \\ &= (c_1|1\rangle + c_2|2\rangle) \otimes (d_1|1\rangle + d_2|2\rangle) \\ &= c_1d_1|1\rangle \otimes |1\rangle + c_1d_2|1\rangle \otimes |2\rangle + c_2d_1|2\rangle \otimes |1\rangle + c_2d_2|2\rangle \otimes |2\rangle \\ &= c_1d_1|11\rangle + c_1d_2|12\rangle + c_2d_1|21\rangle + c_2d_2|22\rangle.\end{aligned}$$

Then $c_1d_2 = c_2d_1 = 0$ and $c_1d_1 = c_2d_2 = \frac{1}{\sqrt{2}}$. There are, however, no solutions for $c_1, c_2, d_1, d_2 \in \mathbb{C}$ subject to these conditions. Therefore, $|\psi\rangle$ is not separable. ♦

Remark 2.47 One can verify that a state $|\psi\rangle = c_{11}|11\rangle + c_{12}|12\rangle + c_{21}|21\rangle + c_{22}|22\rangle \in \mathbb{C}^2$ is separable if and only if the rows of the coefficient matrix $C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$ are scalar multiples. Indeed, $|\psi\rangle$ is separable if and only if there exist scalars $a, b, c, d \in \mathbb{C}$ such that

$$|\psi\rangle = \begin{bmatrix} a \\ b \end{bmatrix} \otimes \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} a\begin{bmatrix} c \\ d \end{bmatrix} \\ b\begin{bmatrix} c \\ d \end{bmatrix} \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{12} \\ c_{21} \\ c_{22} \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \begin{bmatrix} c & d \end{bmatrix}. \tag{2.6}$$

Equivalently, Equation (2.6) says $|\psi\rangle$ is separable if and only if the coefficient matrix C is *rank one*. ♦

Definition 2.48 A non-separable state is called *entangled*. ♦

It is clear that there are significantly more entangled states than separable states. It would be really nice to know when a state is separable – this is called the **separability problem** in quantum information theory. It has been shown to be an **NP-Hard** problem in many cases, meaning it is suspected that there cannot exist an algorithm to solve this problem in polynomial time.

This might seems extremely disappointing, but it turns out that entangled states are a super powerful tool in quantum computing and quantum information theory (for secure information encryption and decryption).

Despite the difficulty of the **separability problem** in general, there is a nice criterion for bipartite systems called the *Schmidt decomposition*, which follows from the well-known *singular value decomposition* for matrices.

2.6.1 Singular Value Decomposition

Even if a matrix is not normal, we can still decompose it in a useful way.

Definition 2.49 Let $A \in M_{mn}$. The **singular values** for A are the square roots of the eigenvalues for $A^\dagger A$, i.e., if λ is an eigenvalue of $A^\dagger A$, $\sqrt{\lambda}$ is called a *singular value* for A. ♦

It should be noted that the singular values of a matrix are well-defined. Indeed, if $A \in M_{mn}$, then $A^\dagger A$ is $n \times n$ and will necessarily have real, non-negative eigenvalues, i.e., $A^\dagger A$ is positive semidefinite.

Exercise 2.50 Let $A \in M_{mn}$ be given. Show that $A^\dagger A$ is positive semidefinite. ♦

Theorem 2.51 (Singular Value Decomposition) *Let $A \in M_{mn}$ with rank$(A) = r$. Define Σ to be the $m \times n$ matrix whose n diagonal entries are the singular values for A in descending order, and 0 elsewhere. Then there exists a unitary matrix $U \in M_m$ and a unitary matrix $V \in M_n$ such that*

$$A = U\Sigma V^\dagger.$$

*Such a factorization is called a **singular value decomposition** or (**SVD**) for A.*

Proof. (Sketch) It can be shown that, of the n total eigenvalues for $A^\dagger A$, $r = \text{rank}(A)$ of the eigenvalues for $A^\dagger A$ are strictly positive and the other $n-r$ are 0. Order them in descending order $\lambda_1, ..., \lambda_r, ..., \lambda_n$ and let $|v_1\rangle, ..., |v_r\rangle, ..., |v_n\rangle$ be their corresponding eigenvectors.. Note that the eigenvectors for $A^\dagger A$ form an orthonormal basis for $\mathbb{C}^n$ by the Spectral Theorem because $A^\dagger A$ is normal (why?). Define

$$V := \begin{bmatrix} |v_1\rangle & |v_2\rangle & \dots & |v_n\rangle \end{bmatrix}.$$

Since the columns of V are orthonormal, V is a unitary matrix and $V(A^\dagger A)V^\dagger$ is a diagonal matrix with entries $\lambda_1, ..., \lambda_n$. Define Σ to be the $m \times n$ matrix with $\sqrt{\lambda_i}$ in the i^{th} diagonal entry, for $1 \leq i \leq n$, and 0's elsewhere. Note that

$$AV = \begin{bmatrix} A|v_1\rangle & \dots & A|v_n\rangle \end{bmatrix} = \begin{bmatrix} A|v_1\rangle & \dots & A|v_r\rangle & |0\rangle & \dots & |0\rangle \end{bmatrix}$$

by one of our versions of matrix multiplication, and because $A|v_j\rangle = |0\rangle$ whenever $j > r$ (check that $\|A|v_j\rangle\| = \lambda_j = 0$). Last, define $U \in M_m$ to be the unitary matrix with columns

$$|u_j\rangle := \frac{1}{\sqrt{\lambda_j}} A|v_j\rangle \qquad \text{for } 0 \leq j \leq r$$

(the eigenvalues which are nonzero) and $|u_{r+1}\rangle, ..., |u_m\rangle$ are normal vectors which are "taken to be orthogonal" to the first r.

Notice that

$$U\Sigma = \begin{bmatrix} \sqrt{\lambda_1}|u_1\rangle & \dots & \sqrt{\lambda_r}|u_r\rangle & 0 \dots & 0 \end{bmatrix} = \begin{bmatrix} A|v_1\rangle & \dots & A|v_r\rangle & 0 \dots & 0 \end{bmatrix} = AV.$$

Therefore, $A = U\Sigma V^\dagger$. ■

Example 2.52 Find the singular value decomposition (SVD) for

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

.

Step 1: Find a unitary diagonalization for $A^\dagger A$, i.e. find an $n \times n = 2 \times 2$ unitary matrix V and a $n \times n = 2 \times 2$ diagonal matrix D such that $A^\dagger A = VDV^\dagger$: Consider $A^\dagger A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ with eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 0$ and corresponding normalized eigenvectors $|\lambda_1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $|\lambda_2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, respectively. Thus, $A^\dagger A = VDV^\dagger$ where $V = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$.

Step 2: Construct V **and** Σ: If necessary, redefine V by reordering the columns in descending order according to the corresponding eigenvalues. In our case, we already have

$\lambda_1 \geq \lambda_2$ and $V = \begin{bmatrix} |\lambda_1\rangle & |\lambda_2\rangle \end{bmatrix}$. Define Σ to be the 3×2 matrix with the singular values for A down its diagonal, i.e.,

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Step 3: Extend the set of (nonzero) vectors of the form $|\mu_i\rangle = A|\lambda_i\rangle$ to an orthonormal basis for $\mathbb{C}^m$:

We have $|\mu_1\rangle = A|\lambda_1\rangle = \sqrt{2}\begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix}$ and $A|\lambda_2\rangle = |0\rangle$. Thus, we extend the set $\{|\mu_1\rangle\}$ to an orthonormal basis for $\mathbb{C}^3$. Note that the set $\{|\mu\rangle_1, |2\rangle, |3\rangle\}$ is a linearly independent set...

We use the Gram-Schmidt process to refine the set to an orthonormal basis.

GS Step 1:

Set

$$|u_1\rangle = |\mu_1\rangle = \sqrt{2}\begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix}.$$

GS Step 2:

Define

$$|u_2\rangle = |2\rangle - \frac{\langle u_1|2\rangle}{\langle u_1|u_1\rangle}|u_1\rangle = |2\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

GS Step 3:

Define

$$\begin{aligned} |u_3\rangle &= |3\rangle - \frac{\langle u_1|3\rangle}{\langle u_1|u_1\rangle}|u_1\rangle - \frac{\langle u_2|3\rangle}{\langle u_2|u_2\rangle}|u_2\rangle = |3\rangle - \frac{\langle u_1|3\rangle}{\langle u_1|u_1\rangle}|u_1\rangle \\ &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{(-\sqrt{2}i)}{4}\left(\sqrt{2}\begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix} i \\ 0 \\ 1 \end{bmatrix}. \end{aligned}$$

Thus, $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$ is an orthogonal set, hence an orthogonal basis for $\mathbb{C}^3$ by the Basis Theorem.

Step Normalization:

Normalizing each vector gives $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$ is an orthonormal basis, where

$$|u_1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix} \qquad |u_2\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad |u_3\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} i \\ 0 \\ 1 \end{bmatrix}.$$

♦

Step 4: Construct U: Define $U = \begin{bmatrix} |u_1\rangle & |u_2\rangle & |u_3\rangle \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & i \\ 0 & \sqrt{2} & 0 \\ i & 0 & 1 \end{bmatrix}$.

Step 5: Verify that $U\Sigma = AV$:

Observe that

$$U\Sigma = \begin{bmatrix} \sqrt{\lambda_1}u_1 & \sqrt{\lambda_2}u_2 \end{bmatrix} = \begin{bmatrix} A|v_1\rangle & A|v_2\rangle \end{bmatrix} = AV.$$

Exercise 2.53 Find the SVD of $A = \begin{bmatrix} 1 & 0 & i \\ i & 0 & 1 \end{bmatrix}$. ♦

2.6.2 Schmidt Decomposition

A consequence of the SVD is a characterization of separable states in bipartite systems.

Theorem 2.54 (Schmidt Decomposition) *Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ be the Hilbert space of a bipartite system, where $\dim \mathcal{H}_1 = m < \infty$ and $\dim \mathcal{H}_2 = n < \infty$. Then every vector $|x\rangle \in \mathcal{H}$ admits a **Schmidt decomposition***

$$|x\rangle = \sum_{j=1}^{r} s_j |u_j\rangle \otimes |v_j\rangle,$$

*where $s_j > 0$ are the **Schmidt coefficients** satisfying $\sum_{j=1}^{r} s_j^2 = 1$, $\{|u_1\rangle, \ldots, |u_r\rangle\}$ and $\{|v_1\rangle, \ldots, |v_r\rangle\}$ are orthonormal sets of $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, and $r \leq \min\{m, n\}$ is the **Schmidt number** of $|x\rangle$.*

Proof. (Sketch) Let $|x\rangle \in \mathcal{H}$ be given. Proposition 2.38 says we can write

$$|x\rangle = \sum_{i,j} c_{ij} |e_{1,i}\rangle \otimes |e_{2,j}\rangle,$$

where $\{|e_{a,i}\rangle\}$ $(a = 1, 2)$ is an orthonormal basis for $\mathcal{H}_a$ and $\sum_{i,j} |c_{ij}|^2 = 1$. Let $C = \begin{bmatrix} c_{ij} \end{bmatrix}$ be the coefficient matrix for $|x\rangle$. If C is rank one, then $|x\rangle$ is separable by an argument similar to Remark 2.47. If $r = \text{rank}(R) > 1$, then C has a SVD, which (using Equation 2.3 we can write

$$C = U\Sigma V^\dagger = \sum_{j=1}^{n} s_j |U_j\rangle \langle V_j^\dagger| = \sum_{j=1}^{r} s_j |U_j\rangle \langle V_j^\dagger|,$$

where $U \in M_m, V \in M_n$ are unitaries and $s_1 \geq \cdots \geq s_{\min\{m,n\}}$ are the singular values of C. Moreover, note that $\sum_{j=1}^{r} s_j^2 = \text{tr}(A^\dagger A) = \|C\|_F^2 = \sum_{p,q} |c_{pq}|^2 = 1$, where $\|\cdot\|_F$ is the Frobenius norm. Since the coefficient matrix C decomposes as the sum of rank one matrices of the form $|U_j\rangle \langle V_j^\dagger|$ $(j = 1, \ldots, r)$, it follows from an argument similar to Remark 2.47 that

$$|\psi\rangle = \sum_{j=1}^{r} s_j |U_j\rangle \otimes |V_j^\dagger\rangle,$$

where $s_1 \geq \cdots \geq s_{\min\{m,n\}}$ are the singular values of C, $\{|U_1\rangle, \ldots, |U_r\rangle\}$ and $\left\{\langle V_1^\dagger|^t, \ldots, \langle V_r^\dagger|^t\right\}$ are the first r columns of U and the first r **rows of** $V^\dagger$ written as column vectors, respectively. ■

Remark 2.55 The *Schmidt number* of a vector $|x\rangle$ from a bipartite system, introduced in Theorem 2.54, is unique and is the minimal number of vectors needed to write $|\psi\rangle$ as a linear combination of separable states. (Prove this? Homework?) It follows that a bipartite state is separable if and only if its Schmidt number is 1. One should observe that **the Schmidt number of a vector is the rank of its coefficient matrix**.

Example 2.56 Consider a *bipartite state* $|\psi\rangle$ written in terms of an orthonormal basis as

$$|\psi\rangle = \frac{1}{2}\left(|e_{1,1}e_{2,1}\rangle + |e_{1,1}e_{2,2}\rangle + i\,|e_{1,3}e_{2,1}\rangle + i\,|e_{1,3}e_{2,2}\rangle\right) = \frac{1}{2}\begin{bmatrix}1\\1\\0\\0\\i\\i\end{bmatrix} \in \mathbb{C}^6.$$

whose coefficients form a matrix $C = \frac{1}{2}\begin{bmatrix}1 & 1\\ 0 & 0\\ i & i\end{bmatrix}$. Find a Schmidt decomposition for $|\psi\rangle$.

Step 1: Find the SVD for the coefficient matrix C:

Using the SVD in Example 2.52, we get $C = U\Sigma V^\dagger$, where

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 0 & i\\ 0 & \sqrt{2} & 0\\ i & 0 & 1\end{bmatrix}, \quad \Sigma = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \quad \text{and} \quad V = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}.$$

♦

Step 2: Construct the Schmidt decomposition for $|\psi\rangle$, i.e., write $|\psi\rangle$ as the linear combination of the outer product of columns of U and the **rows of** $V^\dagger$ (written as column vectors) with weights given by the singular values of C, in descending order:

As the rows of $V^\dagger$ written as column vectors are the columns of V, we have

$$|\psi\rangle = s_1\,|U_1\rangle \otimes \langle V_1^\dagger|^t = \left(\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\i\end{bmatrix}\right) \otimes \left(\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix}1\\0\\i\end{bmatrix} \otimes \begin{bmatrix}1\\1\end{bmatrix}.$$

In particular, we have that $|\psi\rangle$ is separable.

Exercise 2.57 Consider the state $|\psi\rangle = \frac{1}{2}(1,0,i,i,0,1)^t \in \mathbb{C}^6$.

(a) Find a Schmidt decomposition for $|\psi\rangle$ in $\mathbb{C}^2 \otimes \mathbb{C}^3$.

HINT: Note that $|\psi\rangle = \frac{1}{2}\left(|e_{1,1}\rangle \otimes |e_{2,1}\rangle + i\,|e_{1,1}\rangle \otimes |e_{2,3}\rangle + i\,|e_{1,2}\rangle \otimes |e_{2,1}\rangle + |e_{1,2}\rangle \otimes |e_{2,3}\rangle\right) \in \mathbb{C}^2 \otimes \mathbb{C}^3$. You might be able to use work from a previous exercise...

(b) Find a Schmidt decomposition for $|\psi\rangle$ in $\mathbb{C}^3 \otimes \mathbb{C}^2$.

Remark: This problem illustrates that the components of the bipartite system must be specified in order to write out a Schmidt decomposition. ♦

–End class 2/22/22–

2.6.3 Lara Pontification

Mixed States as Density Matrices

Single particle system

Example 2.58 Suppose $|\psi\rangle \in \mathbb{C}^2$ is the state of some quantum system. That means...

- $\||\psi\rangle\| = 1$, and v. polarization, h. polarization...

Let being vertically polarized be represented by $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and horizontal polarization is represented by $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Note that $|0\rangle, |1\rangle$ are eigenstates of a chosen Hermitian matrix which represents the apparatus. To glean probabilistic information about $|\psi_1\rangle = a|0\rangle + b|1\rangle$ ($|a|^2 + |b|^2 = 1$) regarding its polarization, we make measurements to find $|\psi_1\rangle$ is in state $|0\rangle$ with probability $|a|^2$ and in state $|1\rangle$ with probability $|b|^2$.

Suppose $|\psi_2\rangle = e^{i\theta}|\psi_1\rangle$. Then $|\psi_1\rangle \neq |\psi_2\rangle$ as vectors in $\mathbb{C}^2$, but they represent the same state in our quantum system since $\left|e^{i\theta}a\right|^2 = |a|^2$ and $\left|e^{i\theta}b\right|^2 = |b|^2$ and those probabilities are the only information we can gather about our states.

Instead of vectors, we going to use *density matrices* of vectors. Consider the matrix $\rho_{\psi_1} = |\psi_1\rangle\langle\psi_1| = \begin{bmatrix} |a|^2 & ab^* \\ ba^* & |b|^2 \end{bmatrix}$. Observe that $\rho_{\psi_2} = \rho_{\psi_1}$. So in some sense, the matrix ρ_{ψ_1} better represents the state since it ignores *phase shifts*. ♦

Remark 2.59 Observe that if $|\psi\rangle \in \mathbb{C}^2$ is a state, then

- $\rho_\psi \in M_2$
- ρ_ψ has trace 1
- ρ_ψ is Hermitian and positive definite – note that $\det(\rho_\psi - \lambda I) = \lambda(\lambda - (|a|^2 + |b|^2)) = \lambda(\lambda - 1)$. ♦

Definition 2.60 A *density matrix* $\rho \in M_n$ is a positive semi-definite Hermitian matrix such that $\mathrm{tr}(\rho) = 1$. ♦

Discuss Pure vs. Mixed States... we need mixed states due to noise.

3 Building Blocks of Quantum Information

3.1 Qubits as Density Matrices

A classical (Boolean) bit is an element $\mathtt{x} \in \{0, 1\}$.

Definition 3.1 A *qubit* is a (unit) vector in $\mathbb{C}^2$ written in terms of the standard basis $\{|1\rangle, |2\rangle\}$, which we now write as $\{|0\rangle, |1\rangle\}$. In other words, a *qubit* is a vector $|\psi\rangle \in \mathbb{C}^2$ such that

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \text{where } |a|^2 + |b|^2 = 1.$$

♦

Even though a qubit can assume infinitely many states, our method of extracting information from a qubit – namely, measurement – will still only give $|0\rangle$ or $|1\rangle$ with some probability due to collapse. For example, suppose a qubit is in the state $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$. Even by making measurements on many identical copies of the system, we are not able to exactly determine the coefficients a and b, i.e., we cannot know the exact state of the system. At best, we know that the probability that $|\psi\rangle$ is state $|0\rangle$ is given by the expectation value $\langle\psi|P_1|\psi\rangle = |a|^2$ and the probability that $|\psi\rangle$ is state $|1\rangle$ is given by the expectation value $\langle\psi|P_2|\psi\rangle = |b|^2$, where $P_1 = |0\rangle\langle 0|$ and $P_2 = |1\rangle\langle 1|$.

Remark 3.2 The states $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ represent the same quantum state – we cannot distinguish them. So often we will work with the rank one orthogonal projection $\rho = |\psi\rangle\langle\psi|$, called a *pure state*, to represent a quantum state. Note that ρ is a **matrix** and $|\psi\rangle$ is a **vector**. Moreover, $|a|^2$ and $|b|^2$ are the diagonal entries of ρ. ♦

Definition 3.3 A *pure state* is the density matrix of the state of a quantum system that **can** be described by a single ket vector, i.e., $\rho = |\psi\rangle\langle\psi|$. A *mixed state* is the density matrix of a quantum system that cannot be described by a single ket vector, i.e., $\rho = \sum_{i=1}^{N} p_i\,|\psi_i\rangle\langle\psi_i|$ where $\sum_{i=1}^{N} p_i = 1$.

Example 3.4 (a) Suppose the state of the system is a uniform mixture of $|0\rangle$ and $|1\rangle$. Find the density matrix associated to the mixed state.

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}I.$$

(b) Suppose the state of the system is in a pure state of $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Find the density matrix associated to $|\psi\rangle$.

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2}(|0\rangle - |1\rangle)(\langle 0| - \langle 1|) = \frac{1}{2}(|0\rangle\langle 0| - |0\rangle\langle 1| - |1\rangle\langle 0| + |1\rangle\langle 1| = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

♦

Remark 3.5 If the preparation of the system is not fully known, we can at best consider the state as a convex/"probablistic" linear combination of the possible preparations, i.e., the state of the system is mixed. In the perspective of density matrices and why we need not only pure states, but also mixed states, is to consider a pure state $\rho = |\psi\rangle\langle\psi|$ that is to be transmitted across a channel, but the channel has noise. Then the message that's observed may be ρ half of the time Φ_1 25% of the time, and Φ_2 the other 25% of the time. We'd describe this observed state as a *mixed state*, denoted $\frac{1}{2}\rho + \frac{1}{4}\Phi_1 + \frac{1}{4}\Phi_2$. As mentioned in Remark 3.25, the density matrix associated to the state of the system is often referred to as a state itself... but why? ♦

Compare the following properties with our original axioms for quantum mechanics. But first a definition...

Definition 3.6 The *trace* of a matrix $A \in M_n$, denoted $\mathrm{tr}(A)$, is the sum of its diagonal entries.

> **Axiom 1'.** A physical state of a system, whose Hilbert space is $\mathcal{H}$, is completely determined by its associated density matrix $\rho : \mathcal{H} \to \mathcal{H}$. A density matrix is a positive semi-definite matrix with $\mathrm{tr}(\rho) = 1$.

Exercise 3.7 Show that density matrices are positive semi-definite, i.e., show that a matrix written as $\rho = \sum_{i=1}^{N} p_i^2 |\psi_i\rangle\langle\psi_i|$ such that each $p_i \in \mathbb{R}$ and $\||\psi_i\rangle\| = 1$ is positive semi-definite. ♦

Example 3.8 Show that density matrices have trace 1.

Let $\rho = \sum_{i=1}^{N} p_i^2 |\psi_i\rangle\langle\psi_i|$ be a density matrix and let $\{|e_1\rangle, \ldots, |e_n\rangle\}$ be an orthonormal basis for $\mathcal{H}$. Then the trace of ρ can be written as

$$\mathrm{tr}(\rho) = \sum_{k=1}^{n} \langle e_k|\rho|e_k\rangle = \sum_{i=1}^{N} \left(\sum_{k=1}^{n} \langle e_k|p_i^2|\psi_i\rangle\langle\psi_i|e_k\rangle \right) = \sum_{i=1}^{N} p_i^2 \left(\sum_{k=1}^{n} \langle\psi_i|e_k\rangle\langle e_k|\psi_i\rangle \right) = 1$$

by the completeness relation $\sum_{k=1}^{n} |e_k\rangle\langle e_k| = I$. ♦

Exercise 3.9 Show that the following are equivalent:

(a) A state is pure, i.e., the density matrix associated to the state of the system is a rank-one projection $\rho = |\psi\rangle\langle\psi|$, where $\||\psi\rangle\| = 1$.

(b) $\rho^2 = \rho$.

(c) $\mathrm{tr}(\rho^2) = 1$. ♦

> **Axiom 2'.** The mean value of an observable a is $\langle A\rangle = \mathrm{tr}(\rho A)$.

Example 3.10 Show that $\langle A \rangle = \mathrm{tr}(\rho A)$.

Let $\{|e_1\rangle, \ldots, |e_n\rangle\}$ be an orthonormal basis for $\mathcal{H}$. If $|\psi_i\rangle$ occurs with probability p_i^2, then the expectation value of a is $\langle A \rangle_{\psi_i} = \langle \psi_i | A | \psi_i \rangle$. Hence, the weighted mean of a occurring after measurements with respect to many copies of the mixed/pure state $|\psi\rangle = \sum_{i=1}^N p_i |\psi_i\rangle$ is

$$\langle A \rangle_\psi = \sum_{i=1}^N p_i^2 \langle \psi_i | A | \psi_i \rangle .$$

On the other hand, observe that we have

$$\begin{aligned}
\mathrm{tr}(\rho A) &= \sum_{k=1}^n \langle e_k | \rho A | e_k \rangle \\
&= \sum_{i=1}^N p_i^2 \left(\sum_{k=1}^n \langle e_k | \psi_i \rangle \langle \psi_i | A | e_k \rangle \right) \\
&= \sum_{i=1}^N p_i^2 \left(\sum_{k=1}^n \langle \psi_i | A | e_k \rangle \langle e_k | \psi_i \rangle \right) \\
&= \sum_{i=1}^N p_i^2 \langle \psi_i | A | \psi_i \rangle .
\end{aligned}$$

♦

Axiom 3'. The time evolution of the density matrix is given by the **Liouville-von Neumann equation**

$$i\hbar \frac{d}{dt} \rho = [H, \rho]$$

where H is the Hamiltonian of the system and $[H, \rho] = H\rho - \rho H$ is the *commutator* of H and ρ.

End of March 1, 2022

3.1.1 How do we measure probabilities of states as density matrices?

Theorem 3.11 *Suppose an observable A of a quantum system represented by $\mathbb{C}^n$ has spectral decomposition $A = \sum_j \lambda_j |\lambda_j\rangle \langle \lambda_j| =: \sum_j \lambda_j P_j$, where $\lambda_j \in \mathbb{R}$ for all j. If the system is in state $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$, then the probability of obtaining outcome λ_m for A is*

$$\langle \lambda_m | \rho | \lambda_m \rangle = \sum_i p_i |\langle \lambda_m | \psi_i \rangle|^2 = tr(P_m \rho).$$

Proof. Observe

$$\begin{aligned}
\langle\lambda_m|\,\rho\,|\lambda_m\rangle &= \langle\lambda_m|\left(\sum_i p_i\,|\psi_i\rangle\,\langle\psi_i|\right)|\lambda_m\rangle \\
&= \sum_i p_i\,\langle\lambda_m|\psi_i\rangle\,\langle\psi_i|\lambda_m\rangle \\
&= \sum_i p_i\,|\langle\lambda_m|\psi_i\rangle|^2 \\
&= \sum_i p_i\,\langle\psi_i|\lambda_m\rangle\,\langle\lambda_m|\psi_i\rangle \\
&= \sum_i p_i\,\langle\psi_i|\,P_m\,|\psi_i\rangle \\
&= \mathrm{tr}(P_m\rho)
\end{aligned}$$

■

3.2 Classes of Density Matrices

So far we've seen that, if quantum system can be represented on a Hilbert space $\mathbb{C}^n$ in state $|\psi\rangle \in \mathbb{C}^n$, then we can translate all the data $|\psi\rangle$ and its phase shifts $\{e^{i\theta}\,|\psi\rangle : \theta \in [0, 2\pi)\}$ contain to a density matrix $\rho := |\psi\rangle\,\langle\psi| \in M_n(\mathbb{C})$. The range of the transform consists of *pure states*. Moreover, we can form new density matrices by taking convex combinations of pure states, and we call these *mixed states*. If $\rho = \sum_i p_i\,|\psi_i\rangle\,\langle\psi_i|$ is a density matrix such that $p_i = p_j$ for all i, j ("ρ consists of uniformly weighted (equal probability) pure states"), we say ρ *maximally mixed*. In the rest of this section, we classify mixed states into three classes, analogous to the classification of pure states into separable v. entangled.

Definition 3.12 Let $\rho \in M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$ be a density matrix.

- ρ is *uncorrelated* if it **can be written as** $\rho = \rho_1 \otimes \rho_2$ for some density matrices $\rho_1 \in M_n(\mathbb{C})$ and $\rho_2 \in M_m(\mathbb{C})$.
- ρ is *separable* if it **can be written in the form** $\rho = \sum_i p_i \rho_{1,i} \otimes \rho_{2,i}$ for some probabilities $0 < p_i \leq 1$, $\sum_i p_i = 1$, and density matrices $\rho_{1,i} \in M_n(\mathbb{C})$, $\rho_{2,i} \in M_m(\mathbb{C})$.
- ρ is *inseparable* if ρ **cannot be written** in any separable form above. ⧫

Throughout this section, we will be considering states $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$, where we consider the previously-defined basis vectors $|0\rangle = \begin{bmatrix}1\\0\end{bmatrix}$, $|1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$. Recall the tensor product notation $|v\rangle \otimes |w\rangle =: |vw\rangle$. Then

$$|00\rangle = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} \quad |01\rangle = \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \quad |10\rangle = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \quad |11\rangle = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

Also, some projection operators in $M_4(\mathbb{C})$ are:

$$|00\rangle\langle 00| = E_{11} \quad |00\rangle\langle 01| = E_{12} \quad |00\rangle\langle 10| = E_{13} \quad |00\rangle\langle 11| = E_{14}$$

$$|01\rangle\langle 00| = E_{21} \quad |01\rangle\langle 01| = E_{22} \quad |01\rangle\langle 10| = E_{23} \quad |01\rangle\langle 11| = E_{24}$$

$$|10\rangle\langle 00| = E_{31} \quad |10\rangle\langle 01| = E_{32} \quad |10\rangle\langle 10| = E_{33} \quad |10\rangle\langle 11| = E_{34}$$

$$|11\rangle\langle 00| = E_{41} \quad |11\rangle\langle 01| = E_{42} \quad |11\rangle\langle 10| = E_{43} \quad |11\rangle\langle 11| = E_{44}$$

In $\mathbb{C}^3$, we denote the standard basis by

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad |2\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Remark 3.13 The language of density matrices is a little confusing, especially compared to the language of vector states in multipartite systems. For example, why should we use *uncorrelated* for density matrices as the analogue of *separable* for vector states?

A **classical correlation** is statistical information about one system that allows us to infer info about another system. (We should flush this out and get some examples of classical correlations. A table here would be nice.) A **quantum correlation** is "a correlation that does not occur classically." **For example, entanglement is a quantum correlation.** With this in mind, an uncorrelated density matrix $\rho = \rho_1 \otimes \rho_2$ is clearly not entangled, i.e., is not (quantum) correlated. ♦

3.2.1 Classifying Pure States

Example 3.14 Let $|\psi\rangle = |02\rangle = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}^T \in \mathbb{C}^2 \otimes \mathbb{C}^3$. By definition, $|\psi\rangle$ is separable because it is the elementary tensor of two states, $|0\rangle$ with $|2\rangle$. Consider the associated density matrix

$$\rho = |\psi\rangle\langle\psi| = |02\rangle\langle 02| = E_{33} \in M_6(\mathbb{C}).$$

By definition, ρ is a pure state. Note that $E_{33} \in M_6(\mathbb{C})$ can be written as $e_{11} \otimes e_{33} \in M_2(\mathbb{C}) \otimes M_3(\mathbb{C})$. Since $e_{11} \in M_2(\mathbb{C})$ is a density matrix and since $e_{33} \in M_3(\mathbb{C})$ is a density matrix, we can conclude that ρ is uncorrelated. ♦

Exercise 3.15 Prove or disprove: If $|\psi\rangle \in \mathbb{C}^n \otimes \mathbb{C}^m$ is a separable state, then $\rho := |\psi\rangle\langle\psi| \in M_n \otimes M_m$ is uncorrelated. ♦

Lemma 3.16 *If ρ is a separable density matrix, then the* partial trace ρ^{pt} *of ρ is a density matrix.*

Example 3.17 Let $|\psi\rangle := \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle) = \frac{1}{\sqrt{2}}\begin{bmatrix}0 & 1 & 1 & 0\end{bmatrix}^T$ be a bipartite state. Let's first determine if $|\psi\rangle$ is separable or entangled. Suppose $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$ for some states $|\psi_1\rangle = \begin{bmatrix}a\\b\end{bmatrix}, |\psi_2\rangle = \begin{bmatrix}c\\d\end{bmatrix}$, for some $a, b, c, d \in \mathbb{C}$. Then $ad \neq 0$ and $bc \neq 0$ imply $a \neq 0$ and $c \neq 0$, which violates $ac = 0$. Therefore, $|\psi\rangle$ is entangled.

The associated density matrix is

$$\begin{aligned}\rho &= |\psi\rangle\langle\psi| \\ &= \frac{1}{2}(|10\rangle\langle 10| + |10\rangle\langle 01| + |01\rangle\langle 10| + |01\rangle\langle 01|) \\ &= \frac{1}{2}(E_{22} + E_{23} + E_{32} + E_{33}) \\ &= \frac{1}{2}\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}\end{aligned}$$

We want to determine if ρ is uncorrelated, separable, or inseparable. One can show that ρ is definitely not correlated, similar to the argument showing $|\psi\rangle$ is entangled. Note that by the previous lemma, if ρ^{pt} is *not* a density matrix, then ρ is *not* separable. Observe

$$\rho^{\text{pt}} = \frac{1}{2}(|10\rangle\langle 10| + |11\rangle\langle 00| + |00\rangle\langle 11| + |01\rangle\langle 01|) = \frac{1}{2}\begin{bmatrix}0 & 0 & 0 & 1\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 1 & 0 & 0 & 0\end{bmatrix}$$

It is straight forward to check that ρ^{pt} is Hermitian and has trace 1. However, its eigenvalues are $-1, 1, 1, 1$, so it is **not** positive semi-definite, and therefore is **not** a density matrix. Therefore, ρ is inseparable. ⧫

Definition 3.18 Let $\Phi \in H_A \otimes H_B$. The *partial trace of* Φ with respect to its B component, denoted by Φ_A, is defined to be

$$\text{tr}_B(\Phi) := \sum_k (I \otimes \langle k|)\Phi(I \otimes |k\rangle).$$

Remark 3.19 Consider $A = \sum_i c_i A_i \otimes B_i \in M_n \otimes M_m$, where each A_i acts on $\mathcal{H}_1 = \mathbb{C}^n$ and each B_i acts on $\mathcal{H}_2 = \mathbb{C}^m$. Then the partial trace of A over $\mathcal{H}_2$ is a matrix in M_n, denoted $\text{tr}_2(A)$, given by

$$\text{tr}_2(A) = \sum_k (I \otimes \langle k|)A(I \otimes |k\rangle) = \sum_i c_i A_i \otimes \left(\sum_k \langle k|B_i|k\rangle\right) = \sum_i c_i \text{tr}(B_i)A_i.$$

Similarly, the partial trace of A over $\mathcal{H}_1$ is a matrix in M_m, denoted $\text{tr}_1(A)$, given by

$$\text{tr}_1(A) = \sum_i c_i \text{tr}(A_i)B_i.$$

⧫

Example 3.20 Let $|\psi\rangle := \frac{1}{\sqrt{2}} = (|00\rangle + |11\rangle) \in \mathbb{C}^2 \otimes \mathbb{C}^2$ (so $H_A = \mathbb{C}^2$ and $H_B = \mathbb{C}^2$). The corresponding (pure state) density matrix with respect to the $\{|0\rangle, |1\rangle\}$ basis is

$$\rho = \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|) = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \frac{1}{2}(E_{11} + E_{14} + E_{41} + E_{44})$$

Let's compute both partial traces of ρ:

$$\rho_A = \mathrm{tr}_B(\rho) = \sum_{i,j=1}^{2} (I \otimes \langle ij|)\rho(I \otimes |ij\rangle) =$$

♦

$$\rho_B = \mathrm{tr}_A(\rho) =$$

Proposition 3.21 *Let $\rho = |\psi\rangle\langle\psi|$ be a density matrix with respect to some bipartite state $|\psi\rangle \in C^n \otimes \mathbb{C}^m$. Then ρ is separable if and only if its partial traces are pure states.*

Example 3.22 Let $|\psi\rangle := \frac{1}{2}\begin{bmatrix} 1 & 0 & i & i & 0 & 1 \end{bmatrix}^T \in \mathbb{C}^6$. You found a Schmidt decomposition for $|\psi\rangle$ in $\mathbb{C}^2 \otimes \mathbb{C}^3$ in Exercise 2.57:

$$|\psi\rangle = \frac{1}{2}\left(\begin{bmatrix} 1 \\ i \end{bmatrix} \otimes |0\rangle + \begin{bmatrix} i \\ 1 \end{bmatrix} \otimes |2\rangle\right).$$

In particular, this Schmidt composition tells you that $|\psi\rangle$ is entangled.

Let $|u_1\rangle := \begin{bmatrix} 1 & i \end{bmatrix}^T$ and $|u_2\rangle := \begin{bmatrix} i & 1 \end{bmatrix}^T$, and note that

$$|u_1\rangle\langle u_1| = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \quad |u_2\rangle\langle u_2| = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$$

$$|u_1\rangle\langle u_2| = \begin{bmatrix} -i & 1 \\ 1 & i \end{bmatrix} \quad |u_2\rangle\langle u_1| = \begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix}.$$

The associated (pure state) density matrix can be written both as

$$\rho = \frac{1}{4}\begin{bmatrix} 1 \\ 0 \\ i \\ i \\ 0 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -i & -i & 0 & 1 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 & 0 & -i & -i & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ i & 0 & 1 & 1 & 0 & i \\ i & 0 & 1 & 1 & 0 & i \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -i & -i & 0 & 1 \end{bmatrix}$$

and as

$$\begin{aligned} \rho &= \frac{1}{4}\left(|u_1\rangle\langle u_1| \otimes |0\rangle\langle 0| + |u_1\rangle\langle u_2| \otimes |0\rangle\langle 2| + |u_2\rangle\langle u_1| \otimes |2\rangle\langle 0| + |u_2\rangle\langle u_2| \otimes |2\rangle\langle 2|\right) \\ &= \frac{1}{4}\left(|u_1\rangle\langle u_1| \otimes E_{11} + |u_1\rangle\langle u_2| \otimes E_{13} + |u_2\rangle\langle u_1| \otimes E_{31} + |u_2\rangle\langle u_2| \otimes E_{33}\right) \end{aligned}$$

Although ρ is written as a convex combination of elementary tensors, notice that two of the four elementary tensors (the mixed terms) are NOT comprised of density matrices. In particular, none of $|u_1\rangle\langle u_2|$, $|u_2\rangle\langle u_1|$, E_{13}, nor E_{31} are Hermitian, so they are automatically disqualified from being density matrices.

So we want to know: is ρ separable? It possible to find $\rho_1, i \in M_2(\mathbb{C})$, $\rho_2, i \in M_3(\mathbb{C})$, and probabilities $0 < p_i \leq 1$ satisfying $\sum_i p_i = 1$ such that $\rho = \sum_i pi\rho_{1,i} \otimes \rho_{2,i}$? ♦

3.3 Qubits and Information Theory

A classical (Boolean) bit is an element $\mathtt{x} \in \{0, 1\}$. (A bit can be described as "the *information entropy* of a binary random variable that is 0 or 1 with equal probability.)

Information can be transferred in an ordered string of 0's and 1's, called a bit string. A bit string of length 8 is called a *byte*. Each 0 or 1 in the string is called a *bit*, which is short for "binary digit." We'd call 100110 a bit string of length 6, while we'd call 1022 a trit string of length 4, which each entry being called a *trit*, short for "ternary digit." In general, if we want to have k options $\{0, 1, 2, ..., k-1\}$ Error correction and detection is an important concept in classical information. We want to know how reliable received data is as it is transferred over a noisy communication channel. One scheme to detect error is to introduce redundancy.

Example 3.23 The byte `00100100` represents \$ in UTF-8 encoding. ♦

Definition 3.24 A *qubit* is a (unit) vector in $\mathbb{C}^2$ written in terms of the standard basis $\{|1\rangle, |2\rangle\}$, which we now write as $\{|0\rangle, |1\rangle\}$. In other words, a *qubit* is a vector $|\psi\rangle \in \mathbb{C}^2$ such that

$$|\psi\rangle = a|0\rangle + b|1\rangle = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \text{where } |a|^2 + |b|^2 = 1.$$

♦

Even though a qubit can assume infinitely many states, our method of extracting information from a qubit – namely, measurement – will still only give $|0\rangle$ or $|1\rangle$ with some probability due to collapse. For example, suppose a qubit is in the state $|\psi\rangle = a|0\rangle + b|1\rangle$. Even by making measurements on many identical copies of the system, we are not able to exactly determine the coefficients a and b, i.e., we cannot know the exact state of the system. At best, we know that the probability that $|\psi\rangle$ is state $|0\rangle$ is given by the expectation value $\langle\psi|P_1|\psi\rangle = |a|^2$ and the probability that $|\psi\rangle$ is state $|1\rangle$ is given by the expectation value $\langle\psi|P_2|\psi\rangle = |b|^2$, where $P_1 = |0\rangle\langle 0|$ and $P_2 = |1\rangle\langle 1|$.

Remark 3.25 The states $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ represent the same quantum state – we cannot distinguish them. So often we will work with the rank-one orthogonal projection $\rho = |\psi\rangle\langle\psi|$, called a *pure state*, to represent a quantum state. Note that ρ is a **density matrix** with $|a|^2$ and $|b|^2$ diagonal entries of ρ when ρ is expressed in the same basis as the state vector $|\psi\rangle$.♦

Definition 3.26 A group/system of n qubits is called a *(quantum) register*. ♦

Example 3.27 In a classical system, the state of the system is determined by the specifying the state of each component, e.g., `xyzw` is completely determined by the values of `x,y,z,w` independently – 4 values or 1 choice out of 2^4 choices.

Suppose we describe each qubit in a register of n qubits, $|\psi_1\rangle, \ldots |\psi_n\rangle$, in analogy with the classical case. Then we can describe each qubit as $a_i|0\rangle + b_i|1\rangle$ where $|a_i|^2 + |b_i|^2 = 1$. So we would only need $2n$ complex numbers (complex amplitudes) to describe the state of the system, e.g.,

$$(a_1|0\rangle + b_1|1\rangle) \otimes \cdots \otimes (a_n|0\rangle + b_n|1\rangle).$$

But we know that not every state is decomposable this way, i.e., superposition requires us to consider **entangled** states. Thus, a general state in the register has the form

$$|\psi\rangle = \sum_{i_k \in \{0,1\}} a_{i_1 \cdots i_n} |i_1\rangle \otimes \cdots \otimes |i_n\rangle \in \mathbb{C}^{2^n}$$

♦

Example 3.28 The set

$$\left\{ |\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), |\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), |\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), |\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \right\}$$

is an orthonormal basis of a two-qubit system and is called the *Bell basis*. Each vector is called a *Bell state* and each Bell state is **entangled**!

Show that the Bell basis is obtained from the binary basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ by a unitary transformation. ♦

Exercise 3.29 Find the expectation value of $\sigma_x \otimes \sigma_z$ measured in each of the Bell states. ♦

3.3.1 More Evidence of Classical vs. Quantum Information Theory

Example 3.30 (Classical Repetition Code) Suppose the user wants to transmit the information `1011`. One can encode this information using a repetition code of length 4, i.e., we map `1011` to `1111 0000 1111 1111` and transmit those four blocks of four bits. The received blocks are decoded by a majority decision. For example, `1111 0100 0111 1000` has 5 corrupted bits and is decoded as `1010`. Even though the decoded message wasn't "correct," the corrupted bit was much more likely to be "correct" than "incorrect" with our redundancy.

This error correction scheme is based on the idea that a bit `x` can be encoded as `xx` by having "blank bits" available for the copying, i.e., `x_` is sent to `xx`. ♦

This scheme works because classical bits are easy to measure and repeat. The following distinguishes classical and quantum information processes. Recall that there are two quantum operations we can perform on a quantum system:

1. An observation via applying a spectral projection which collapses the system into some (eigen)state of the (Hermitian matrix associated to the) observable – clearly corrupting the information (such as a particle being passed through a polarized filter).

2. A unitary operation that sends states to states – unit vectors to unit vectors (a change of orthonormal basis, aka coordinates). This operation is reversible and is called a *quantum gate*. These operations are called gates as a shoutout to classical computing operations called *logic gates* – `AND`, `OR`, `XOR`, `NOT`, `NAND`, `NOR`, and `XNOR`, which are implemented by transistors, diodes, etc. in a classical computer.

Start with this on 3/29/22

Definition 3.31 A *quantum gate* in an n-qubit system is a unitary transformation of $\mathbb{C}^{2^n}$, which we represent by a unitary matrix U acting on $\mathbb{C}^{2^n}$, i.e., U is a $2^n \times 2^n$ unitary matrix.⧫

Example 3.32 For a single bit `x`, there are two logic gates (operations) – the identity gate and `NOT`. ⧫

Example 3.33 We can define gates on a single qubit system by sending orthonormal bases to orthonormal bases. For example,

$$\begin{aligned} I &: |0\rangle \to |0\rangle, |1\rangle \to |1\rangle \Rightarrow I = I_2 \\ X &: |0\rangle \to |1\rangle, |1\rangle \to |0\rangle \Rightarrow X = \sigma_x \\ Y &: |0\rangle \to -|1\rangle, |1\rangle \to |0\rangle \Rightarrow Y = -i\sigma_y = XZ \\ Z &: |0\rangle \to |0\rangle, |1\rangle \to -|1\rangle \Rightarrow Z = \sigma_z \end{aligned}$$

⧫

The idea of repeating *qu*bits, by essentially saying "slot A has bit 1 and slot B has bit 0" would be represented in quantum as "slot A is in state $|1\rangle$ and slot B is in state $|0\rangle$," which we encode as "the system involving slots A and B (in that order) is in state $|1\rangle \otimes |0\rangle = |10\rangle$."

Theorem 3.34 (No-Cloning Theorem, Wootters and Zurek, Dieks) *An unknown quantum system cannot be cloned by unitary transformations. In other words, there is no unitary operator U on $\mathcal{H} \otimes \mathcal{H}$ such that for all normalized states $|\psi\rangle, |e\rangle \in \mathcal{H}$ one has $U(|\psi e\rangle) = |\psi\psi\rangle$, i.e., it's not possible to have "the blank state" $|e\rangle$ evolve into $|\psi\rangle$, regardless of the state $|\psi\rangle$.*

Proof. Suppose there exists a unitary U such that $U|\psi e\rangle = U|\psi\psi\rangle$ for any $|\psi\rangle \in \mathcal{H}$. Let $\{|\varphi\rangle, |\phi\rangle\}$ be a set of orthogonal states in $\mathcal{H}$. By our assumption, we must have $U|\varphi e\rangle = |\varphi\varphi\rangle$ and $U|\phi e\rangle = |\phi\phi\rangle$. Since $|\varphi\rangle$ and $|\phi\rangle$ are orthogonal, $|\psi\rangle = \frac{1}{\sqrt{2}}(|\varphi\rangle + |\phi\rangle)$ is a state. So our assumption yields

$$U|\psi e\rangle = |\psi\psi\rangle = \frac{1}{2}(|\varphi\varphi\rangle + |\varphi\phi\rangle + |\phi\varphi\rangle + |\phi\phi\rangle).$$

But we must also have

$$U|\psi e\rangle = \frac{1}{\sqrt{2}}(U|\varphi e\rangle + U|\phi e\rangle) = \frac{1}{\sqrt{2}}(|\varphi\varphi\rangle + |\phi\phi\rangle),$$

which is a contradiction by linear independence and matching coefficients. ■

Moral. We're going to need a bigger boat.

End of 3/22/22.

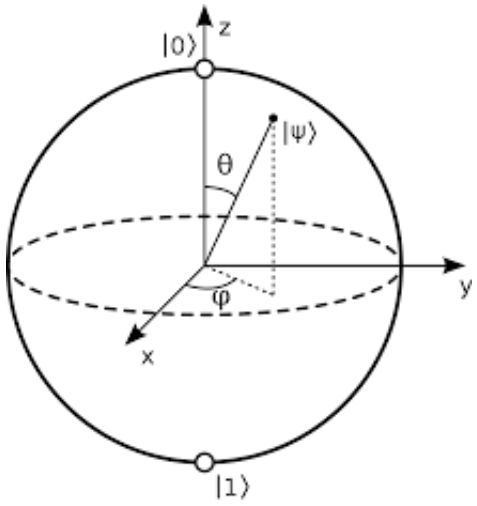


Figure 1: Bloch Sphere

3.3.2 Visualizing Single Qubits and Quantum Gates on the Bloch Sphere

Each point on the Bloch sphere corresponds to a *pure state*–and remember, a single pure state could represent more than one state so long as those states are equal up to a factor of $e^{i\gamma}$, a global phase shift. Indeed, if $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is a state (so $\sqrt{|\alpha|^2 + |\beta|^2} = 1$), then $\alpha = |\alpha| e^{i\gamma_1}$ and $\beta = |\beta| e^{i\gamma_2}$ for some $\gamma_1, \gamma_2 \in [0, 2\pi)$. Observe

$$\begin{aligned} |\psi\rangle &= \alpha|0\rangle + \beta|1\rangle \\ &= |\alpha| e^{i\gamma_1}|0\rangle + |\beta| e^{i\gamma_2}|1\rangle \\ &= e^{\gamma_1}\left(|\alpha||0\rangle + |\beta| e^{i(\gamma_2-\gamma_1)}|1\rangle\right) \end{aligned}$$

Define $\theta := \gamma_2 - \gamma_1$ and choose ϕ so that $\cos(\phi/2) = |\alpha|$ and $\sin(\phi/2) = |\beta|$. Note that φ can take values in $[0, 2\pi)$ while θ will be limited to $[0, \pi]$. We can graph this spherical coordinate $(1, \varphi, \theta)$ to represent $|\psi\rangle$. Note that the global phase shift is ignored, as is mentioned above.

So,

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

Exercise 3.35 Show that $|0\rangle$ is represented on the Bloch sphere by the spherical coordinate $(1, 0, 0)$ and $|1\rangle$ is represented on the Bloch sphere by the spherical coordinate $(1, 0, \pi)$. ♦

Exercise 3.36 Explain why $\beta|1\rangle$ is represented by the *same* spherical coordinate on the Bloch sphere as $|1\rangle$ for all complex numbers β such that $|\beta|^2 = 1$. THEN, explain (mathematically) how/why a state $|\psi\rangle \in \mathbb{C}^2$ is represented by the same point on the Bloch sphere as $c|\psi\rangle$ for any $c \in \mathbb{C}$. ♦

Exercise 3.37 For each of the following quantum gates, graph the images $|\psi_1\rangle$ and $|\psi_2\rangle$ of $|0\rangle$ and $|1\rangle$, respectively, by finding θ_j and ϕ_j such that

$$|\psi_j\rangle = \cos\left(\frac{\phi_j}{2}\right)|0\rangle + e^{i\theta_j}\sin\left(\frac{\phi_j}{2}\right)|1\rangle$$

for each $j = 1, 2$.

$$\begin{aligned} I &: |0\rangle \to |0\rangle, |1\rangle \to |1\rangle \Rightarrow I = I_2 \\ X &: |0\rangle \to |1\rangle, |1\rangle \to |0\rangle \Rightarrow X = \sigma_x \\ Y &: |0\rangle \to -|1\rangle, |1\rangle \to |0\rangle \Rightarrow Y = -i\sigma_y = XZ \\ Z &: |0\rangle \to |0\rangle, |1\rangle \to -|1\rangle \Rightarrow Z = \sigma_z \end{aligned}$$

♦

Remark 3.38 Orthogonal vectors are represented as antipodal points on the Bloch sphere.♦

If we wish to prepare our system to be in an unusual orthornomal basis $\{|\psi_1\rangle, |\psi_2\rangle\}$, how do we build a quantum gate that will take our usual orthonormal basis $\{|0\rangle, |1\rangle\}$ to this new basis? We'll need to use change of basis techniques from linear algebra!

Exercise 3.39 Find the quantum gate U which carries the standard orthonormal basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ to the Bell basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$. ♦

3.4 Measuring Qubits

One can measure single qubits for a state vector in an n qubit system.

Example 3.40 Let $|\psi\rangle$ be a state in a two-qubit system, i.e.,

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle, \quad |a|^2 + |b|^2 + |c|^2 + |d|^2 = 1.$$

Suppose we want to find the probability that the first qubit will be in state $|0\rangle$ (outcome +1) or $|1\rangle$ (outcome -1) upon measurement of σ_z – measuring the polarization of the first qubit.

Recall that if we want to find the probability for a qubit $|x\rangle = \alpha|0\rangle + \beta|1\rangle$ to be in state $|0\rangle$ (outcome +1) or $|1\rangle$ (outcome -1) upon measurement of σ_z, we use the spectral projections $M_0 = |0\rangle\langle 0|$ and $M_1 = |1\rangle\langle 1|$. In particular, we obtain $|0\rangle$ with probability

$$p(0) = |\alpha|^2 = \|M_0|x\rangle\|^2 = \langle x|M_0^\dagger M_0|x\rangle = \langle x|M_0|x\rangle$$

and obtain $|1\rangle$ with probability

$$p(1) = |\beta|^2 = \|M_1|x\rangle\|^2 = \langle x|M_1^\dagger M_1|x\rangle = \langle x|M_1|x\rangle.$$

To measure the first qubit, we use $A = \sigma_z \otimes I_2$ and observe that $M_0 = |0\rangle\langle 0| \otimes I_2$ and $M_1 = |1\rangle\langle 1| \otimes I_2$ for this experiment. Write $|\psi\rangle$ as

$$\begin{aligned} |\psi\rangle &= |0\rangle \otimes (a|0\rangle + b|1\rangle) + |1\rangle \otimes (c|0\rangle + d|1\rangle) \\ &= u|0\rangle \otimes \left(\frac{a}{u}|0\rangle + \frac{b}{u}|1\rangle\right) + v|1\rangle \otimes \left(\frac{c}{v}|0\rangle + \frac{d}{v}|1\rangle\right), \end{aligned}$$

where $u = \sqrt{|a|^2 + |b|^2} \neq 0$ and $v = \sqrt{|c|^2 + |d|^2} \neq 0$. ♦

Applying M_0 and M_1, we obtain $|0\rangle$ with probability $\langle\psi|M_0|\psi\rangle = u^2$ and $|1\rangle$ with probability $\langle\psi|M_1|\psi\rangle = v^2$, and the state $|\psi\rangle$ collapses to

$$|0\rangle \otimes \left(\frac{a}{u}|0\rangle + \frac{b}{u}|1\rangle\right) \quad \text{and} \quad |1\rangle \otimes \left(\frac{c}{v}|0\rangle + \frac{d}{v}|1\rangle\right),$$

respectively upon measurement.

Measuring the second qubit is similar.

Remark 3.41 "Measurement gives an alternative viewpoint on entanglement. Suppose that the first qubit of the state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is measure to be 0 (or 1). Then the outcome of the second qubit is **definitely** 0 (or 1)." (Why?)

Indeed, rewriting

$$|\Phi^+\rangle = \left(\frac{1}{\sqrt{2}}|0\rangle\right) \otimes (|0\rangle + 0\,|1\rangle) + \left(\frac{1}{\sqrt{2}}|1\rangle\right) \otimes (0\,|0\rangle + |1\rangle),$$

then applying M_0 and M_1 from the previous example gives $|\Phi^+\rangle$ collapses onto $|00\rangle$ and $|11\rangle$, respectively upon measurement. **Hence, measurement on the first qubit affects the outcome of the measurement on the second qubit – the initial state is entangled, i.e., there is a strong correlation between the two qubits.**

Conversely, the vector $\frac{1}{\sqrt{2}}(|00\rangle + |01\rangle) = |0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ is separable, the second qubit is measured to be 0 (or 1) with probability 1/2, independent of whether or not the first qubit was measured. ♦

3.5 Einstein-Podolsky-Rosen (EPR) Paradox – Bell Formulation

Suppose a particle source produces the EPR state $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ and sends one particle to Alice and the other to Bob, who may be separated far away. Note that though the particles may be separated, the particles combined are described by the state $|\Psi^-\rangle$.

Now, if Alice measures her particle to read $|0\rangle$ (or $|1\rangle$) with 1/2 probability, then the state instantly collapses onto $|1\rangle$ (or $|0\rangle$). This means that Bob will **definitely** observe $|1\rangle$ (or $|0\rangle$) upon his measurement.

This is called *nonlocal behavior* and it pissed Einstein off. (He was not a quantum fan. He called nonlocal behavior "spooky action at a distance.") On first glance, it seems like Alice's measurement (and the resulting information) propogated to Bob's qubit instantaneously, thus violating special relativity – nothing can travel faster than the speed of light.

The issue (or non-issue) with this argument is that no energy, no information is actually being transmitted. Alice cannot control her measurement, and thus, she cannot control the reading of Bob's particle. Her measurement is **random** – she has a 50-50 chance of seeing $|0\rangle$ or $|1\rangle$. Alice and Bob can only independently measure a large number of EPR pairs – observing some random sequence of 0's and 1's, then notice the correlation between their sequences after the fact (by some means of classical communication).

3.6 An Application of Single Qubit Measurements: BB84 Protocol for Quantum Key Distribution (QKD)

Example 3.42 Alice and Bob are using a *one-time pad* encryption technique to communicate about selling stocks. Alice wants to send the byte 00100100 to Bob as she just received insider info that they should sell their stocks. They use the BB84 protocol to distribute the encryption key for their communication.

1. **Alice and Bob use the bases $\mathcal{B}_1$ and $\mathcal{B}_2$ to prepare and measure $4N = 32$ photons**:

$$\mathcal{B}_1 = \left\{ \begin{matrix} 0 & \mapsto & |0\rangle \\ 1 & \mapsto & |1\rangle \end{matrix} \right\} \quad \text{and} \quad \mathcal{B}_2 = \left\{ \begin{matrix} 0 & \mapsto & |\nearrow\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ 1 & \mapsto & |\nwarrow\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \end{matrix} \right\} \quad \blacklozenge$$

2. **Alice sends Bob $4N = 32$ photons, each prepared with one of the two bases at random.**

Alice sends	0	1	0	0	1	1	0	1	0	0	1	0	0	0	1	0	$\cdots$
Alice's basis	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\cdots$

3. **Bob measures each received photon, each with one of the two bases at random.**

Fill-in the bits that Bob should expect to read: 0, 1, or ? (for 0 or 1)

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

4. **Alice and Bob talk and identify the photons that were prepared and measured using the same basis – there should be roughly $2N = 16$ photons:**

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

The remaining roughly $2N = 16$ photons should have been prepared and measured with respect to the same basis.

Alice sends	0	1	1	0	0	1	0	0	1	0	0	0	1	1	1	0
Alice's basis	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_2$
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														

5. **Alice and Bob select $N = 8$ photons among the $2N = 16$ photons at random to detect tampering:**

Alice sends	0	1	1	0	0	1	0	0	1	0	0	0	1	1	1	0
Alice's basis	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_2$
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														

6. **If Eve was listening, then Eve had to have intercepted and measured the photons using $\mathcal{B}_1$ or $\mathcal{B}_2$ at random before passing the collapsed state to Bob.**

Fill-in the bits that Eve reads/send and what Bob reads as a consequence: 0, 1, or ? (for 0 or 1)

Alice sends	0	1	1	0	0	1	0	0	1	0	0	0	1	1	1	0
Alice's basis	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_1$	$\mathcal{B}_2$	$\mathcal{B}_2$	$\mathcal{B}_2$
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					
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														

7. Can Alice and Bob detect if Eve is droppin' the eves? Explain.

4 Quantum Computation

4.1 Quantum Gates

Definition 4.1 Recall that a *quantum gate* for an n-qubit register is a unitary $U \in M_n$. ♦

We've already seen several examples. Let's look at some examples which mimic classical information and transfer of classical bits.

Example 4.2 Recall some classical bit operations `NOT`, `AND`, `OR`, and `EXOR`. All of these are called (classical) **logic gates**–`NOT` takes single bit inputs to single bit outputs, while the other three take 2-bit inputs to single bit outputs.

- `NOT`: $0 \mapsto 1$, $1 \mapsto 0$
- `AND`: $00 \mapsto 0$, $01 \mapsto 0$, $10 \mapsto 0$, $11 \mapsto 1$.
- `OR`: $00 \mapsto 0$, $01 \mapsto 1$, $10 \mapsto 1$, $11 \mapsto 1$.
- `EXOR`: $00 \mapsto 0$, $01 \mapsto 1$, $10 \mapsto 1$, $11 \mapsto 0$.

There are others. But quantum gates must be reversible, which in particular means they must send n-qubit inputs to n-qubit outputs. The `CNOT` gate, for example, takes a 2-qubit, `xy`, and checks the first bit `x`, to see if it's a 1. If it is, it will flip the `y` bit (applying `NOT`). This is the intuition for its name, `CNOT`, which stands for controlled-`NOT`.

- `X`: $0 \mapsto 1$, $1 \mapsto 0$ appropriately implements the classical `NOT` gate.
- `CCNOT`: $|11x\rangle = |11\neg x\rangle$, while $|00x\rangle \mapsto |00x\rangle$, $|10x\rangle \mapsto |10x\rangle$, $|01x\rangle \mapsto |01x\rangle$. This is the quantum analogue of the classical `AND` gate.
- `OR`: Recall that, in classical logic, $P \vee Q \equiv \neg P \wedge \neg Q$. Combining the X gate with the `AND` gate helps us to define a quantum analogue of `OR`:

$$(I \otimes I \otimes X) \circ \texttt{AND} \circ (X \otimes X \otimes I).$$

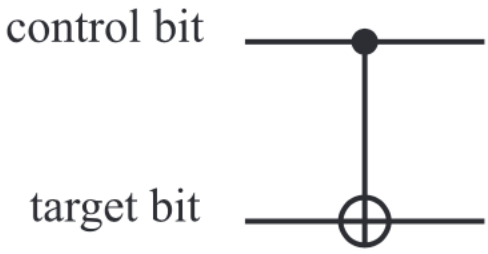
- `CNOT`: $00 \mapsto 00$, $01 \mapsto 01$, $10 \mapsto 11$, $11 \mapsto 10$. This gate corresponds to the classical `EXOR`. ♦

Definition 4.3 The quantum `CNOT` gate is defined by:

$$|00\rangle \mapsto |00\rangle, \quad |01\rangle \mapsto |01\rangle, \quad |10\rangle \mapsto |11\rangle, \quad |11\rangle \mapsto |10\rangle.$$

and is denoted in a circuit by:

Example 4.4 Let's examine the `CNOT` gate.



1. Find a matrix representation for `CNOT` $\in M_4$ with respect to the standard basis for $\mathbb{C}^4$.
2. Show that `CNOT` cannot be written as $A \otimes B \in M_2 \otimes M_2$.
3. Explain how `CNOT` is the quantum analogue of the classical `XOR` gate. ♦

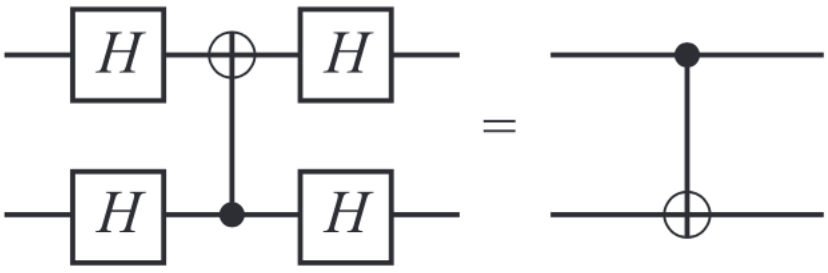
If we think about quantum gates as being analogous to classical logic gates, which logic gate does `CNOT` correspond to? Surprisingly, it's not the `NOT` gate. It's the `EXOR` gate! How do we do this? Keep the first bit, leaving it alone, and then apply `EXOR` to the second bit. By the way, `CNOT` $= \langle 0| \, |0\rangle \otimes I + \langle 1| \, |1\rangle \otimes X$.

Let's try to transfer the `CNOT` gate to actually depend on the second bit, and flip the first bit only when the second bit is 1. Recall the Hadamard gate, $H : |0\rangle \mapsto \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, $|1\rangle \mapsto \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Show that the Hadamard gate is unitary and idempotent.

For a very long list and implementations of quantum gates, visit
https://qiskit.org/documentation/apidoc/circuit_library.html.

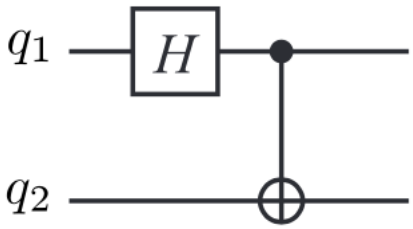
Definition 4.5 A *quantum circuit* is a computational routine consisting of coherent quantum operations on quantum data, such as qubits. It is an ordered sequence of quantum gates, measurements and resets, which may be conditioned on real-time classical computation. ♦

Exercise 4.6 Show that the two circuits below are equivalent:



Explain, using tensor products, what this says about the relationship between the `CNOT` gate with the Hadamard gate H. ♦

Exercise 4.7 Find the outputs of the qubits $|00\rangle$, $|01\rangle$, $|10\rangle$,$|11\rangle$ when passed through the circuit



4.2 Density Matrices in Quantum Information

Recall that, given a quantum state $|\psi\rangle \in \mathbb{C}^n$, we can associate a matrix $\rho \in M_n(\mathbb{C})$ by defining ρ to be the projection given by the outer product of $|\psi\rangle$ with itself: $\rho := |\psi\rangle\langle\psi|$. Suppose $|\psi\rangle \in \mathbb{C}^2$ is a unit vector given by $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Then

$$\rho = |\psi\rangle\langle\psi| = \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{bmatrix}.$$

Notice that the diagonal entries of ρ clearly illustrate the probabilities that $|\psi\rangle$, upon measurement in the $\{|0\rangle, |1\rangle\}$ basis, will collapse to $|0\rangle$ or $|1\rangle$. This ρ matrix has other properties, like $\mathrm{tr}(\rho) = 1$, $\rho^\dagger = \rho$, and $\langle x|\rho|x\rangle \geq 0$. In particular, ρ is positive semi-definite and trace 1.

Recall that, in general, there is an entire class of matrices which have the above properties but which do *not* come directly from a single outer product (those that do are called pure states), like $|\psi\rangle\langle\psi|$. We call all such matrices *density matrices*, and they are defined to be the set of all *convex combinations* of pure states: $\{\sum_{i=1}^k p_i |\psi_i\rangle\langle\psi_i| \ : \ 0 \leq p_i \leq 1, \sum_{i=1}^k p_i = 1\}$, where $|\psi_i\rangle$ are unit vectors in $\mathbb{C}^n$.

The set of all density matrices **does not form a subspace** of $M_n(\mathbb{C})$:

4.2.1 Density matrices through quantum gates

Given a pure state $\rho = |\psi\rangle\langle\psi|$, we can see ρ evolve under the action of a unitary quantum gate U by considering the effect of U on $|\psi\rangle$:

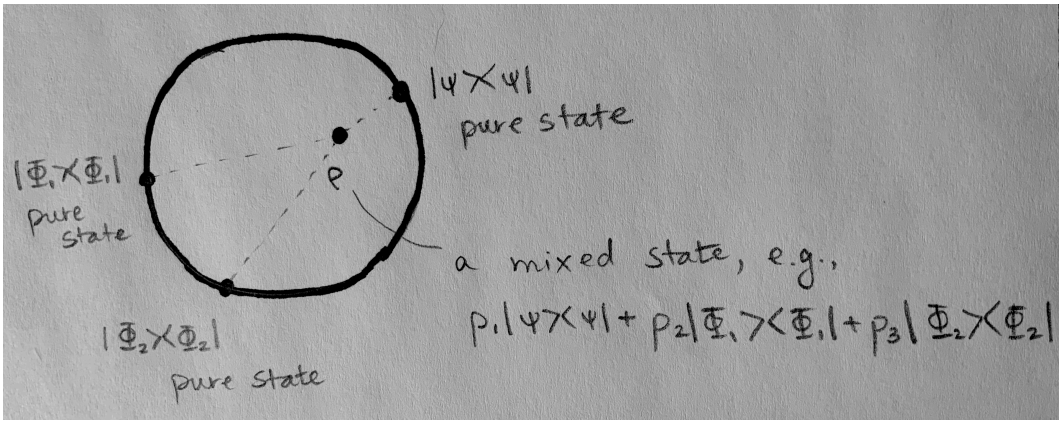
$$\rho \mapsto (U|\psi\rangle)(U|\psi\rangle)^\dagger = U|\psi\rangle\langle\psi|U^\dagger = U\rho U^\dagger.$$

If $\rho = \sum_{i=1}^k p_i |\psi_i\rangle\langle\psi_i|$ is a mixed state, it evolves similarly when passed through gate U:

$$\rho \mapsto U\left(\sum_{i=1}^k p_i |\psi_i\rangle\langle\psi_i|\right)U^\dagger = \sum_{i=1}^k p_i U|\psi_i\rangle\langle\psi_i|U^\dagger.$$

Exercise 4.8 Prove that, given any density matrix ρ (possibly a mixed state), the evolution of ρ through a unitary gate U is still a density matrix. ♦

What we have to consider, realistically, is that we have an imperfect channel–that our quantum gates might not *always* act on our states as we would hope. Consider the following example.



Example 4.9 Suppose we want to apply the `CNOT` gate to the state $|10\rangle$, but the gate only actually works as X 80% of the time. Unfortunately, 12% of the time, the gate actually behaves as U_2 and the other 8% it behaves as U_3. We could observe this phenomenon by measuring $U|\psi\rangle$ a lot of times. Then, we'd want to describe the state on the **other** side of the gate not as $X|\psi\rangle$, but as:

$$|10\rangle \mapsto .8\texttt{CNOT}|10\rangle + .12U_1|10\rangle + .08U_2|10\rangle$$

What if we want to push this through another gate, unitary or projection? It's potentially no longer an elementary tensor of the form $|\psi_1\rangle \otimes |\psi_2\rangle$, and we know that we need to send the entire state through each gate simultaneously so that collapse of a part of the state doesn't affect a later gate's output. The solution is to encode $|10\rangle$ as a density matrix, so that we can always send density matrices through the gates and produce **density matrices**. *Compute*

References

[NO08] Mikio Nakahara and Tetsuo Ohmi, *Quantum computing*, CRC Press, Boca Raton, FL, 2008, From linear algebra to physical realizations. MR 2387891