Module 2: Quantum Linear Algebra, Part I

Module-2 solved problems

Module objectives

By *quantum linear algebra*, we mean the linear algebra needed for quantum computing.

This module's sole aim is to mathematically set us up for quantum computing:

- Complex numbers and vectors.
- Dirac notation, which takes getting used to.
- Unit length simplification
- Projector, Hermitian and unitary matrices.
- Change-of-basis.

Some of this material might be a bit dry, but is necessary before we get to the interesting stuff in later modules.

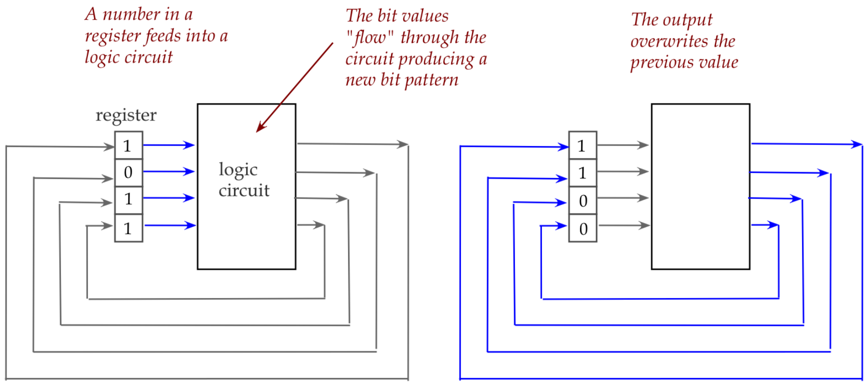
2.1 Why do we need the linear algebra of complex vectors?

We will over the course develop intuition for why we need complex-number vectors and certain types of matrices.

And we will see mathematical reasons for why complex vectors emerge from simple observations and basic assumptions.

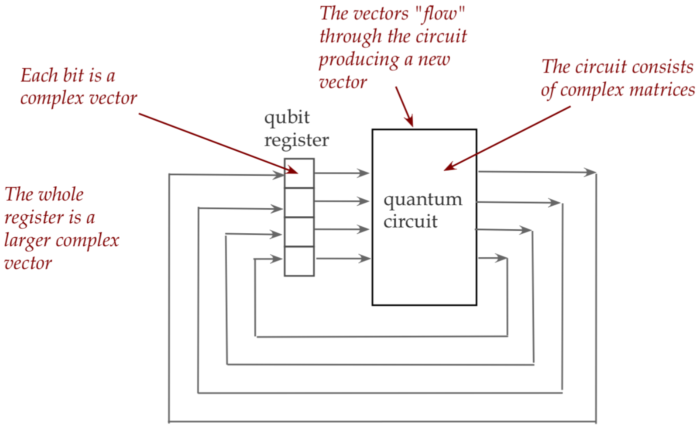
For now, we'll provide a high-level operational view as motivation.

Let's start by recalling how a standard electronic circuit (in a calculator or computer) performs an arithmetic operation:


A number in a register feeds into a logic circuit
The bit values "flow" through the circuit producing a new bit pattern
The output overwrites the previous value
register
1
0
1
1
logic circuit
1
1
0
0

- Here, a binary number sits in a device (called a register) that holds bits (the number's digits).
- These bits then flow into a logic circuit, which we've seen earlier consists of a collection of *Boolean gates* (like AND, OR and NOT gates).
- The gates together achieve the desired computation (increment, in the above example), and the resulting number is fed back into the register.
- This conceptual description is reasonably close to what happens physically inside a calculator or computer.

The quantum equivalent:


The vectors "flow" through the circuit producing a new vector
Each bit is a complex vector
The circuit consists of complex matrices
qubit register
The whole register is a larger complex vector
quantum circuit

- For the moment we will describe a *conceptual* view that is actually implemented in wildly different ways
  ⊳ The actual physical implementation can be quite different
- The starting point is not a binary number but a (complex) *vector*:
  - We will see that each qubit is actually a small vector.
  - A collection of qubits is also a (larger) vector.
- Just as logic gates "act" on regular bits, in quantum circuit it will be matrices that "act" on vectors.
- There are going to be two fundamentally different kinds of matrices, and they "act" quite differently.
  - One kind is called *unitary* and acts in the way we're already familiar with: the matrix multiplies into the input vector.
  - The other kind is called *Hermitian* and the way it "acts" on a vector is a little strange.
- In a physical realization, of course, there are physical devices (such as lasers) that achieve the job of these matrices.

Thus, there's no getting around the need to be really comfortable with the essential linear algebra needed for quantum computing: the linear algebra of complex vectors with two special types of matrices.

Shall we begin?

2.2 Complex numbers: a review

What are they and where did they come from?

- Consider an equation like $x^2 = 4$.
  ⊳ We know that $x = 2$ and $x = -2$ are solutions (written as $x = \pm 2$).

- What about $x^2 = 2$?
  ⊳ Doesn't change the concept: $x = \pm\sqrt{2}$.

- However, $x^2 = -2$ poses a problem.
  ⊳ No square of a real number is negative

- Suppose we invent the "number" $\sqrt{-1}$ and give it a symbol:

$$i \quad = \quad \sqrt{-1}$$

  Then using the rules of algebra

$$(i\sqrt{2})^2 \quad = \quad i^2(\sqrt{2})^2 \quad = \quad (\sqrt{-1})^2(\sqrt{2})^2 \quad = \quad (-1)(2) \quad = \quad -2$$

  gives a solution to the previous equation.

- In general, a *complex number* is written as $a + ib$ where a and b are *real* numbers.
  - The a in $a + ib$ is called the *real part*.
  - b is called the *imaginary part*.

- Why should this make sense?
  - We want to *include* all real numbers in the set of complex numbers
    ⊳ This works when $b = 0$ in $a + ib$
  - However, we need to define arithmetic operations carefully so that when $b = 0$ the same operations work for real numbers.
  - That is, operations defined on complex numbers should give the expected results when applied to complex numbers *with only real parts*

- Luckily, the straightforward algebra works: define
  - *Addition:* $(a + ib) + (c + id) \triangleq (a + c) + i(b + d)$
  - *Multiplication:* $(a + ib)(c + id) \triangleq (ac - bd) + i(ad + bc)$

  Notice: we simply treated the real numbers a, b, c, d and i as symbols and used standard algebra (factoring and distribution).

- Notice that the addition rule satisfies our original starting point with $a = 0, b = 1$:

$$(0 + i)(0 + i) \quad = \quad (0\;0 + 0i + 0i + i^2) \quad = \quad i^2 \quad = \quad -1$$

- Multiplication by a real-number scalar:

$$\beta(a + ib) \quad = \quad (\beta a + i(\beta b))$$

- Subtraction and division:
  - Subtraction simply negates the real numbers of the second complex number, as in:

$$(a + ib) - (c + id) \quad = \quad (a + ib) + ((-c) + i(-d))$$

  - Division needs a bit more thought:

$$\frac{a + ib}{c + id} \quad = \quad \frac{a + ib}{c + id}\,\frac{c - id}{c - id} \quad = \quad \frac{1}{c^2 + d^2}(a + ib)(c - id)$$

    Which now becomes a multiplication of two complex numbers, followed by a scalar multiplication.

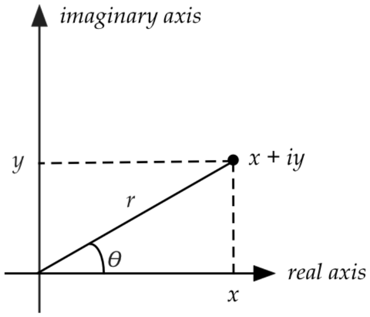
- Notation:
  - We will typically write a complex number symbolically as $a + ib$
  - But particular values might be written as $3 + 4i$ because that's more natural than $3 + i4$.
  - It takes some getting used to, but you should see $3 + 4i$ as *one number*.
  - When reading, your eyes should locate the i in $3 + 4i$ so that you separately see the imaginary part 4.

- About the *meaning* of a complex number:
  - A real number always has physical interpretations, like *length*.
  - A complex does not directly correspond to anything physical.
  - Instead, it's best to think of it as abstraction that leads to predictive power.
  - When we need to predict a physical quantity, we'll be sure to extract a real number.

- We should also point out a downside to complex numbers: there's no natural *ordering*
  - We can't say whether $3 + 4i$ is less than $4 + 3i$ or the other way around
  - Fortunately, this issue is not going to impact our needs.

Polar representation of complex numbers:

- One useful graphical representation is obvious when we write complex number z as $z = x + iy$.



- Given this, one can easily write

$$\begin{aligned} z &= x + iy \\ &= r\cos(\theta) + ir\sin(\theta) \\ &= r\left(\cos(\theta) + i\sin(\theta)\right) \end{aligned}$$

Once arithmetic has been defined, we can define *functions* on complex numbers:

- An important observation made by Euler:
  Suppose you *define*

$$e^z \triangleq 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \ldots$$

  analogous to the Taylor series for the real function e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

- Then substituting $z = i\theta$, and separating out alternate terms:

$$\begin{aligned} e^{i\theta} &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \ldots \\ &= (\text{series for } \cos(\theta)) + i\,(\text{series for } \sin(\theta)) \\ &= \cos(\theta) + i\sin(\theta) \end{aligned}$$

This is called Euler's relation.

- More generally, $re^{i\theta} = r(\cos(\theta) + i\sin(\theta))$
- Think of $z = re^{i\theta}$ as the polar representation of the complex number

$$z \;=\; x + iy \;=\; r(cos(\theta) + i\sin(\theta))$$

- Let's revisit some operations with the polar form $z_1 = r_1e^{i\theta_1}$ and $z_2 = r_2e^{i\theta_2}$

$$\begin{aligned} z_1z_2 &= r_1r_2e^{i(\theta_1+\theta_2)} \\ \frac{z_1}{z_2} &= \frac{r_1}{r_2}e^{i(\theta_1-\theta_2)} \\ z_1^{-1} &= \frac{1}{r_1}e^{-i\theta_1} \end{aligned}$$

- Sometimes an abundance of parentheses can be confusing, so we often simplify $r(cos(\theta) + i\sin(\theta))$ to

$$r(\cos\theta + i\sin\theta).$$

where it's understood that θ is the argument to cos and sin.

- Redundancy in θ:
  - While different values of a and b will result in different complex numbers, the same is not true for r, θ.
  - Because, for any integer k

$$\begin{aligned} \sin(\theta + 2\pi k) &= \sin\theta \\ \cos(\theta + 2\pi k) &= \cos\theta \end{aligned}$$

  - In polar form

$$e^{\theta+2\pi k} \;=\; e^{\theta}$$

  - Thus, for example, $3e^{\frac{\pi}{3}+6\pi}$ and $3e^{\frac{\pi}{3}}$ are the same number.
- Thus, two numbers in polar form $r_1e^{i\theta_1}$ and $r_2e^{i\theta_2}$ are equal if and only if $r_1 = r_2$ and $\theta_1 = \theta_2 + 2\pi k$ for some integer k.

Conjugates:

- It turns out to be really useful to define something called a *conjugate* of a complex number $z = a + ib$:

$$\begin{aligned} z^* &= (a+ib)^* \\ &\triangleq a - ib \end{aligned}$$

- Then,

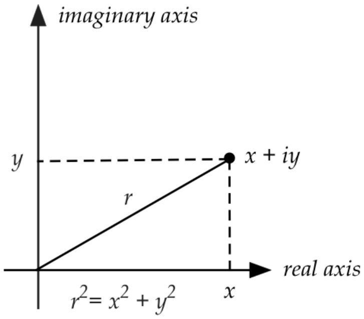
$$\begin{aligned} zz^* &= (a+ib)\,(a+ib)^* \\ &= (a+ib)\,(a-ib) \\ &= a^2 + b^2 \end{aligned}$$

Which is a real number.

- The *magnitude* of a complex number $z = a + ib$ is the real number

$$|z| \triangleq \sqrt{a^2 + b^2} = \sqrt{zz^*} = \sqrt{z^*z}$$

- The magnitude, we have seen, has a *geometrical interpretation* to complex numbers: the distance from the origin to the point representing z.



- Thus, in polar form

$$\left|re^{i\theta}\right| = |r||e^{i\theta}| = r$$

because $|e^{i\theta}|^2 = |\cos\theta + i\sin\theta|^2 = \cos^2\theta + \sin^2\theta = 1$

- Useful rules to remember about conjugation:

$$\begin{aligned}(z^*)^* &= z\\ (z_1 + z_2)^* &= z_1^* + z_2^*\\ (z_1 - z_2)^* &= z_1^* - z_2^*\\ (z_1 z_2)^* &= z_1^* z_2^*\\ (\frac{z_1}{z_2})^* &= \frac{z_1^*}{z_2^*}\end{aligned}$$

- And, most importantly, the *polar* form of conjugation:

$$\begin{aligned}(re^{i\theta})^* &= (r(\cos\theta + i\sin\theta))^*\\ &= r(\cos\theta - i\sin\theta)\\ &= r(\cos\theta + i\sin(-\theta))\\ &= re^{-i\theta}\end{aligned}$$

Note: we used parens only when needed above.

- We will commonly see the case when $r = 1$ (unit length), in which case it simplifies to $e^{-i\theta}$.
- Alternate notation for conjugation:
  - Math books often use a "bar" to denote conjugation:

$$\begin{aligned}\bar{z} &= \overline{a + ib}\\ &= a - ib\end{aligned}$$

  - We will generally prefer the * notation.

Additional notation:

- Notation for the real and imaginary parts of $z = a + ib$:

$$\begin{aligned} \text{Re}\, z &= a \\ \text{Im}\, z &= b \end{aligned}$$

- To extract the polar angle, one uses arg, but because many angles are equivalent, one gets a *set*:

$$\arg z \quad = \quad \{\theta : re^{\theta} = z\}$$

- When a particular angle is specified, as in $2e^{\frac{\pi}{3}}$, then the angle (in radians) is often called the *phase*.

- Notice that multiplication by e^{ϕ} changes the phase:

$$e^{\phi}\ 2e^{\frac{\pi}{3}} \quad = \quad 2e^{\frac{\pi}{3}+\phi}$$

  This is a notion we will return to frequently in the future.

In-Class Exercise 1: Review **these examples** and then solve:

a. Compute $|z|$ when $z = 4 - 3i$
b. Express $3 - 3i$ in polar form.
c. Show that $\text{Im}\,(2i(1+4i) - 3i(2-i)) = -4$
d. Write $\frac{1}{(1-i)(3+i)}$ in $a + ib$ form.
e. If $z = i^{\frac{1}{3}}$, what is the phase of z^* in degrees?

In-Class Exercise 2: Suppose $z = z^*$ for a complex number z. What can you infer about the imaginary part of z?

2.3 Complex vectors (in old notation)

Because the Dirac notation takes getting used to, we'll first look at complex vectors in the notation used in linear algebra courses.

A vector with possibly complex numbers as elements is a *complex vector*:

- Thus, if $2 + 3i$ and $5 - 4i$ are two complex numbers, the vector $(2 + 3i, 5 - 4i)$ is a 2D complex vector.

- But because the complex numbers include reals, $(1.5, 2)$ is also a complex vector.

- In general, a complex vector of n dimensions will have n complex numbers as elements:

$$(a_1 + ib_1, a_2 + ib_2, \ldots, a_n + ib_n)$$

- In column form:

$$\begin{bmatrix} a_1 + ib_1 \\ a_2 + ib_2 \\ \vdots \\ a_n + ib_n \end{bmatrix}$$

- What remains is to see how the operations on real vectors can be extended to complex vectors.

- Addition is a straightforward extension:
  - Let $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ be two complex vectors.
  - Here, each u_i and v_j are *complex numbers*, with real and imaginary parts.
  - Then,

$$\mathbf{u} + \mathbf{v} \quad = \quad (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$$

  - Each $u_i + v_i$ is *complex* addition.

- Example:

$$\begin{bmatrix} 1 + 2i \\ i \\ -3 + 4i \end{bmatrix} + \begin{bmatrix} -2 + i \\ 2 \\ 4 \end{bmatrix} \quad = \quad \begin{bmatrix} -1 + 3i \\ 2 + i \\ 1 + 4i \end{bmatrix}$$

- Scalar multiplication is a bit different in that the scalar can now be a complex number:

$$\begin{aligned} \alpha\mathbf{u} \quad &= \quad \alpha(u_1, u_2, \ldots, u_n) \\ &= \quad (\alpha u_1, \alpha u_2, \ldots, \alpha u_n) \end{aligned}$$

  - Here, both α and each u_i are complex numbers.
  - Thus, the rules of complex multiplication are needed for calculating each αu_i.

- Example:

$$(1 - 2i) \begin{bmatrix} 1 + 2i \\ 3 \end{bmatrix} \quad = \quad \begin{bmatrix} 5 \\ 3 - 6i \end{bmatrix}$$

- So far, operations for complex vectors look like their real counterparts.

- The dot product, however, is an exception.

- For complex vectors **u** and **v**, the dot (inner) product is defined as

$$\mathbf{u} \cdot \mathbf{v} \quad = \quad u_1^* v_1 + u_2^* v_2 + \ldots + u_n^* v_n$$

  Recall: for a complex number $z = a + bi$, $z^* = \text{conjugate}(z) = a - bi$.

- Example: $\mathbf{u} = (1, 2 - i, 3i)$, $\mathbf{v} = (2, 1, i)$

$$\mathbf{u} \cdot \mathbf{v} \quad = \quad (1, 2 - i, 3i)^* \ \cdot \ (2, 1, i) \quad = \quad 1 \times 2 + (2 + i) \times 1 + (-3i) \times i \quad = \quad 7 + i$$

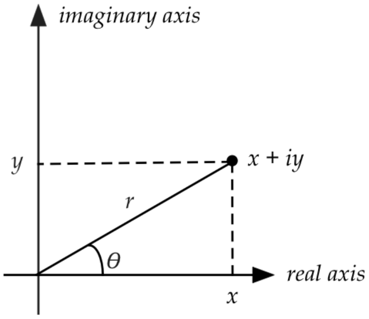
- **Important:** In the inner product, it's the ***left*** vector that gets conjugated. We'll say more about this below.

- The obvious question is, of course, why?
  ⊳ It has to do with the relationship between magnitude and dot-product.

- For a real vector $\mathbf{u} = (u_1, u_2, \ldots, u_n)$
  - $\text{magnitude}(\mathbf{u}) = |\mathbf{u}| = \sqrt{|u_1|^2 + |u_2|^2 + \ldots + |u_n|^2}$

  - In other words, $|\mathbf{u}| = \sqrt{\text{sum of squared magnitudes of elements of } \mathbf{u}}$
  - For a real number u_i, the squared-magnitude is simply $|u_i|^2 = u_i^2 = u_i \times u_i$
  - Not so for a complex number.

- The squared magnitude of the complex number $a + bi$ is $a^2 + b^2$, which is the distance from the origin:



Note: $a^2 + b^2 \neq (a + bi)(a + bi)$
But $a^2 + b^2 = (a + bi)(a - bi)$

- What does this have to do with the dot-product?
  - For real vectors,

$$\begin{aligned}
\mathbf{u} \cdot \mathbf{u} &= (u_1, u_2, \ldots, u_n) \cdot (u_1, u_2, \ldots, u_n) \\
&= u_1^2 + u_2^2 + \ldots + u_n^2 \\
&= |u_1|^2 + |u_2|^2 + \ldots + |u_n|^2 \\
&= |\mathbf{u}|^2
\end{aligned}$$

  - To make this work for complex numbers:

$$\begin{aligned}
\mathbf{u} \cdot \mathbf{u} &= (u_1^*, u_2^*, \ldots, u_n^*) \cdot (u_1, u_2, \ldots, u_n) \\
&= u_1^* u_1 + u_2^* u_2 + \ldots + u_n^* u_n \\
&= |u_1|^2 + |u_2|^2 + \ldots + |u_n|^2 \\
&= |\mathbf{u}|^2
\end{aligned}$$

Inner product convention:

- Most *math* books use a different inner-product definition, where the *right* vector is conjugated:

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1^* + u_2 v_2^* + \ldots + u_n v_n^*$$

- However, we will use the convention from physics, where the *left* vector is conjugated:

$$\mathbf{u} \cdot \mathbf{v} = u_1^* v_1 + u_2^* v_2 + \ldots + u_n^* v_n$$

- Consider the complex numbers $u_1 = (2 + 3i)$ and $v_1 = (3 + 4i)$. Then

$$\begin{aligned}
u_1 v_1^* &= (2 + 3i)(3 + 4i)^* &= (2 + 3i)(3 - 4i) &= (18 + i) \\
u_1^* v_1 &= (2 + 3i)^*(3 + 4i) &= (2 - 3i)(3 + 4i) &= (18 - i)
\end{aligned}$$

- So, the definitions will result in different inner-products but $\mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2$ still holds, and the real part is the same.

- The convention in physics and quantum computing is to *left-conjugate*, and that is what we will do.

- Incidentally, did you notice that in the above example $u_1 v_1^* = (u_1^* v_1)^*$?

2.4 Complex vectors (again) with Dirac notation

Let's first revisit two aspects of real vectors:

1. *Symbolic convention*:
    - Most linear algebra textbooks use boldface or arrow notation for vectors as in:

$$\mathbf{u} \;=\; \begin{bmatrix} 1 \\ 2+i \\ 3i \end{bmatrix} \qquad \vec{v} \;=\; \begin{bmatrix} 2 \\ 1 \\ i \end{bmatrix}$$

    - These are typically subscripted for multiple related vectors, as in:

$$\mathbf{w}_1 \;=\; \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \qquad \mathbf{w}_2 \;=\; \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

2. *Dot product*:
    - The dot product for real vectors is, not surprisingly, written with a dot:

$$\mathbf{w}_1 \cdot \mathbf{w}_2 \;=\; \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \;=\; 5$$

    The result is a *number*.

    - We can treat each real vector as a single-column matrix and instead write the dot-product as:

$$\mathbf{w}_1^T \mathbf{w}_2 \;=\; \begin{bmatrix} 1 & -2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \;=\; 5$$

    Here, with a slight abuse of notation, the 1×1 result can be interpreted as a number.

    - Note: for a complex vector, we'll need to *both* transpose and conjugate the left vector:

$$\mathbf{u} \cdot \mathbf{v} \;=\; (\mathbf{u}^*)^T \mathbf{v} \;=\; \begin{bmatrix} 1^* & (2-i)^* & (3i)^* \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ i \end{bmatrix} \;=\; 7+i$$

The first thing to do is to invent notation for *combined tranpose and conjugation*: the dagger notation

$$\mathbf{u}^\dagger \;=\; \begin{bmatrix} u_1^* & \dots & u_n^* \end{bmatrix}$$

Then, we can write

$$(\mathbf{u}^*)^T \mathbf{v} \;=\; \mathbf{u}^\dagger \mathbf{v}$$

Now on to Dirac notation for vectors:

- A column vector is written as:

$$|v\rangle = \begin{bmatrix}2\\1\\i\end{bmatrix}$$

- A *conjugated row* vector is written as

$$\langle u| = (|u\rangle)^\dagger = ((|u\rangle)^*)^T = \begin{bmatrix}1^* & (2-i)^* & (3i)^*\end{bmatrix} = \begin{bmatrix}1 & 2+i & -3i\end{bmatrix}$$

- And, most crucially, the dot product of $|u\rangle$ and $|v\rangle$ is written as

$$\underset{\textit{Inner product}}{\langle u|v\rangle} = \underset{\textit{Conjugated row vector of } \mathrm{u}}{\begin{bmatrix}1 & 2+i & -3i\end{bmatrix}} \underset{\textit{Column vector } \mathrm{v}}{\begin{bmatrix}2\\1\\i\end{bmatrix}} = \underset{\textit{Complex number}}{7+i}$$

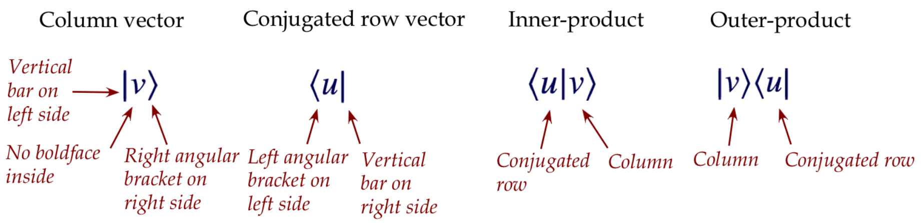
- Symbolically for general vectors $|u\rangle = (u_1, \dots, u_n)$ and $|v\rangle = (v_1, \dots, v_n)$,

$$\langle u|v\rangle = \sum_i u_i^* v_i$$

- This is also commonly called the *inner product*.
- Next, we'll use the same notation to write an *outer product*:

$$\underset{\textit{Outer product}}{|v\rangle\langle u|} = \underset{\textit{Column vector } \mathrm{v}}{\begin{bmatrix}2\\1\\i\end{bmatrix}} \underset{\textit{Conjugated row vector of } \mathrm{u}}{\begin{bmatrix}1 & 2+i & -3i\end{bmatrix}} = \underset{\textit{A matrix}}{\begin{bmatrix}2 & 4+2i & -6i\\ 1 & 2+i & -3i\\ i & 2i-1 & 3\end{bmatrix}}$$

- To summarize:



One needs to be careful with subscripts:

- In traditional notation, we wrote a collection of vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ with subscripts not boldfaced.

- In Dirac notation, a collection of n vectors is described as: $|u_1\rangle, |u_2\rangle, \ldots, |u_n\rangle$
  ▷ The subscripts are *inside* the asymmetric brackets.

- Unfortunately, this can lead to confusion when we want to describe the individual numbers in a vector, as in: $|v\rangle = (v_1, v_2, \ldots, v_n)$

- Thus, one needs to infer the correct meaning from the context.

Scalars and scalar conjugation:

- First consider an example with real scalars and vectors:

$$3 \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -6 \\ 9 \end{bmatrix}$$

  Note:
  - The result is the same in row form:

$$3 \begin{bmatrix} 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} 3 & -6 & 9 \end{bmatrix}$$

  - Symbolically, if $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ is a real vector and α is a real number, then:

$$(\alpha\mathbf{w})^T = \alpha\mathbf{w}^T$$

  - In our example:

$$\left(3 \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \right)^T = 3 \begin{bmatrix} 1 & -2 & 3 \end{bmatrix}$$

- But, because we conjugate complex vectors when transposing, the scalar can get conjugated.

- Let's first look at this symbolically:
  - Suppose $|w\rangle = (w_1, \ldots, w_n)$ is a complex vector and α a complex number.
  - Then we'll use the notation $|\alpha w\rangle$ to mean

$$|\alpha w\rangle = \alpha \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \alpha w_1 \\ \vdots \\ \alpha w_n \end{bmatrix}$$

    where α multiplies into each number in the vector.
  - Then observe that

$$\begin{aligned} \langle \alpha w| &= \begin{bmatrix} (\alpha w_1) & \cdots & (\alpha w_n) \end{bmatrix}^* \\ &= \begin{bmatrix} (\alpha w_1)^* & \cdots & (\alpha w_n)^* \end{bmatrix} \\ &= \begin{bmatrix} \alpha^* w_1^* & \cdots & \alpha^* w_n^* \end{bmatrix} \\ &= \alpha^* \begin{bmatrix} w_1^* & \cdots & w_n^* \end{bmatrix} \\ &= \alpha^* \langle w| \end{aligned}$$

    Thus, when a scalar is factored out of a conjugated-row, the scalar becomes conjugated.

- An example:

- Suppose

$$\alpha = (2-3i), \qquad |w\rangle = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$$

- Then, in this case

$$\langle w| = \begin{bmatrix} 1 & -2 & 3 \end{bmatrix}$$

(The conjugate of a real number is the same real number.)

- Next,

$$\alpha\,|w\rangle \;=\; (2-3i)\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \;=\; \begin{bmatrix} 2-3i \\ -4+6i \\ 6-9i \end{bmatrix}$$

That is, the scalar multiplies each element in the usual way.

- If we conjugate and transpose the result, we get

$$(\alpha\,|w\rangle)^{\dagger} \;=\; \left(\begin{bmatrix} 2-3i \\ -4+6i \\ 6-9i \end{bmatrix}\right)^{\dagger} \;=\; \begin{bmatrix} 2+3i & -4-6i & 6+9i \end{bmatrix}$$

which is NOT equal to $(2-3i)\begin{bmatrix} 1 & -2 & 3 \end{bmatrix}$, i.e., not equal to $\alpha\,\langle w|$

- But

$$\begin{bmatrix} 2+3i & -4-6i & 6+9i \end{bmatrix} \;=\; (2+3i)\begin{bmatrix} 1 & -2 & 3 \end{bmatrix} \;=\; \alpha^{*}\,\langle w|$$

- That is,

$$(\alpha\,|w\rangle)^{\dagger} \;=\; \alpha^{*}\,\langle w|$$

- Lastly, for the *magnitude* of a vector $|u\rangle$, note that

$$\left||u\rangle\right|^{2} \;=\; \langle u|u\rangle$$

Why? Recall: that's how we arrived at the definition of inner product!

- When the context makes it clear, we'll simplify the magnitude notation to $|u|$.

Vector operations:

- We've already seen scalar multiplication and inner product.

- The only other operation needed is plain old addition, which is the same as in real vectors, for example:

$$\begin{bmatrix} 2 \\ 1 \\ i \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \;=\; \begin{bmatrix} 3 \\ -1 \\ 3+i \end{bmatrix}$$

- Notationally, we write this in two equivalent ways:

$$|v+w\rangle \;\triangleq\; |v\rangle + |w\rangle$$

Note:

- The right side is easy: it's merely element-by-element addition of two vectors to give a third.
- The left side is a bit strange because we haven't said anything about what $v + w$ means inside the Dirac brackets.
  - ⊳ The above definition clarifies.

Scalar "movement":

- There is a type of algebraic simplification we often see in quantum computing that's worth highlighting.

- We'll do so with real vector examples, but the same idea applies to complex vectors.

- Consider $|v\rangle = (3, 1)$ and $\alpha = 5$:
  - We typically write the scalar multiplication as:

$$5 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

  - We could just as correctly write it as:

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} 5$$

  - In this sense, the scalar is "movable" when applied as a multiplier.

- This matters when the scalar itself comes as a result of an inner product.

- For example, suppose $|u\rangle = (1, 2)$:
  - Then, consider the outer-product (matrix) $|v\rangle\langle v|$ times the vector $|u\rangle$:

$$\begin{aligned} (|v\rangle\langle v|)\ \ |u\rangle &= \left(\begin{bmatrix} 3 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \left(\begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 15 \\ 5 \end{bmatrix} \end{aligned}$$

  - Instead, observe that we can do this differently by exploiting matrix-associativity:

$$\begin{aligned} (|v\rangle\langle v|)\ \ |u\rangle &= |v\rangle\, (\langle v|u\rangle) \\ &= \begin{bmatrix} 3 \\ 1 \end{bmatrix} \left(\begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right) \\ &= \begin{bmatrix} 3 \\ 1 \end{bmatrix} (5) \\ &= 5 \begin{bmatrix} 3 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 15 \\ 5 \end{bmatrix} \end{aligned}$$

- Symbolically, a movable scalar might represent an inner-product, which when moved, results in simplification:

$$(|v\rangle\langle v|)\ |u\rangle \quad = \quad |v\rangle\,(\langle v|u\rangle) \quad = \quad (\langle v|u\rangle)\,|v\rangle$$

Linear combinations:

- Combining scalar multiplication and addition gives us a linear combination, and some notation for it:

$$|\alpha u + \beta v\rangle \ \triangleq\ \alpha\,|u\rangle + \beta\,|v\rangle$$

- For conjugated rows, we need to conjugate the scalars:

$$\langle \alpha u + \beta v| \ \triangleq\ \alpha^*\,\langle u| + \beta^*\,\langle v|$$

- This leads to two forms of inner products with linear combinations:

$$\begin{aligned} \langle u \mid \alpha v + \beta w\rangle &= \alpha\,\langle u|v\rangle + \beta\,\langle u|w\rangle && \text{Linearity on the right} \\ \langle \alpha u + \beta v \mid w\rangle &= \alpha^*\,\langle u|w\rangle + \beta^*\,\langle v|w\rangle && \text{Conjugate linearity on the left} \end{aligned}$$

  Both are important to remember!

- How to read the notation $|\alpha u + \beta v\rangle$:
  - First start with u and v as regular column vectors, as in

$$u \quad = \quad \begin{bmatrix} 1 \\ 2-i \\ 3i \end{bmatrix} \qquad v \quad = \quad \begin{bmatrix} 2 \\ 1 \\ i \end{bmatrix}$$

  - Then compute the column vector $\alpha u + \beta v$, for example with $\alpha = (2-3i), \beta = -1$:

$$\alpha u + \beta v \quad = \quad (2-3i)\begin{bmatrix} 1 \\ 2-i \\ 3i \end{bmatrix} + (-1)\begin{bmatrix} 2 \\ 1 \\ i \end{bmatrix} \quad = \quad \begin{bmatrix} -3i \\ -8i \\ 9+5i \end{bmatrix}$$

  - This is already a column, so we can write it as

$$|\alpha u + \beta v\rangle \quad = \quad \begin{bmatrix} -3i \\ -8i \\ 9+5i \end{bmatrix}$$

- How to read the notation $\langle \alpha u + \beta v|$:
  - First think of $\alpha u + \beta v$ as the column

$$\alpha u + \beta v \quad = \quad \begin{bmatrix} -3i \\ -8i \\ 9+5i \end{bmatrix}$$

  - Now conjugate and transpose:

$$\langle \alpha u + \beta v| \quad = \quad \begin{bmatrix} 3i & 8i & 9-5i \end{bmatrix}$$

- *Important:* The above notation for linear combinations is worth re-reading several times: we will use this frequently throughout the course.

In-Class Exercise 3: Prove the two results stated above, and one more:

a. $\langle u \mid \alpha v + \beta w \rangle = \alpha \langle u|v \rangle + \beta \langle u|w \rangle$
b. $\langle \alpha u + \beta v \mid w \rangle = \alpha^* \langle u|w \rangle + \beta^* \langle v|w \rangle$.
c. $\langle u|v \rangle = \langle v|u \rangle^*$.

- **In-Class Exercise 4:** Review these examples and then solve the following. Given $|u\rangle = (\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}),\ |v\rangle = (\frac{1}{\sqrt{2}}, -\frac{i}{\sqrt{2}}),\ w = (1,0),\ \alpha = i,\ \beta = -\sqrt{2}i$, calculate (without converting $\sqrt{2}$ to decimal format):

a. $\alpha |v\rangle, |\alpha v\rangle, \langle \alpha u|,\ \alpha^* \langle u|$;
b. $\langle u|v \rangle$;
c. $|u\rangle\langle v|$, $|u\rangle\langle v|\ |w\rangle$, and $\langle v|w\rangle\ |u\rangle$, and compare the latter two results;
d. $||u\rangle|^2, \langle u|u\rangle$;
e. $\langle w \mid \alpha |u\rangle + \beta |v\rangle \rangle$ and $\alpha \langle w|u\rangle + \beta \langle w|v\rangle$;
f. $\langle \alpha |u\rangle + \beta |v\rangle \mid w\rangle$ and $\alpha * \langle u|w\rangle + \beta^* \langle v|w\rangle$.

2.5 Vector spaces, spans, bases, dimension, orthogonality

We're already familiar with these from prior linear algebra but let's do a quick review using Dirac notation.

In the definitions below, we assume that the vectors in any set have the same number of elements:

- It is never the case that we want to put, for example, $(1, -2, 3)$ and $(5, 6)$ in the same set.

Definitions:

- **Span**. Given a collection of vectors $|v_1\rangle, |v_2\rangle, \ldots, |v_k\rangle$, the *span* of these vectors is the *set* of all possible linear combinations of these (with complex scalars):

  $$\text{span}(|v_1\rangle, |v_2\rangle, \ldots, |v_k\rangle) \ \triangleq\ \{\alpha_1 |v_1\rangle + \alpha_2 |v_2\rangle + \ldots + \alpha_k |v_k\rangle : \alpha_i \in \mathbb{C}\}$$

  Note: just like $\mathbb{R}$ is the set of all real numbers, we use $\mathbb{C}$ for the set of all complex numbers.

- **Vector space**. A *vector space* V is a *set* of vectors such that, for any subset $|v_1\rangle, |v_2\rangle, \ldots, |v_k\rangle$ where $|v_i\rangle \in V$,

  $$\text{span}(|v_1\rangle, |v_2\rangle, \ldots, |v_k\rangle) \ \subseteq\ V$$

  Think of a vector space as: it contains all the linear combinations of anything inside it.

- The special vector space $\mathbb{C}^n$:

  $$\mathbb{C}^n \quad = \quad \{|v\rangle = (v_1, \ldots, v_n) : v_i \in \mathbb{C}\}$$

  That is, the set of *all* vectors with n elements, where each element is a complex number.
  - Thus, for example, $(1, 2+i, 3i) \in \mathbb{C}^3$.
  - And $(1, 0, 1, 1, 0) \in \mathbb{C}^5$.

    - Note: $\mathbb{C}^3$ is *not* a subset of $\mathbb{C}^5$.

- **Linear independence**. A collection of vectors $|v_1\rangle, |v_2\rangle, \ldots, |v_k\rangle$ is *linearly independent* if

$$\alpha_1 |v_1\rangle + \alpha_2 |v_2\rangle + \ldots + \alpha_k |v_k\rangle \quad = \quad 0$$

  implies $\forall i : \alpha_i = 0$.

- **Basis**. There are multiple equivalent definitions:
    1. A *basis* for a given vector space V, is a set of vectors $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$ from V such that:
        i. $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$ are linearly independent.
        ii. $V = \text{span}(|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle)$
    2. A *basis* for a given vector space V, is a set of vectors $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$ from V such that any vector $|u\rangle \in V$ is uniquely expressible as a linear combination of the $|v_i\rangle$'s.
       That is, there is only one linear combination $|u\rangle = \sum_i \alpha_i |v_i\rangle$.

  Note: all bases have the same number of vectors. That is, if one basis has n vectors, so does any other basis.

- **Dimension**. The dimension of a vector space V is the *number* of vectors in any basis, written as $\dim(V)$.

- As a consequence, if $\dim(V) = n$, any n linearly independent vectors from V forms a basis for V.

- **Orthogonal vectors**. Two vectors $|u\rangle$ and $|v\rangle$ from a vector space V are *orthogonal* if $\langle u|v\rangle = 0$.

- Note: if $\langle u|v\rangle = 0$, then $\langle v|u\rangle = 0$.

- **Orthonormal vectors**. Two vectors $|u\rangle$ and $|v\rangle$ from a vector space V are *orthonormal* if $\langle u|v\rangle = 0$ and $|u| = |v| = 1$ (that is, each is of unit length).

- **Orthonormal basis**. An orthonormal basis $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$ for a vector space V is a basis such that $\left\langle v_i|v_j\right\rangle = 0$ and $\langle v_i|v_i\rangle = 1$.

Expressing a vector in an orthonormal basis:

- If $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$ is an orthonormal basis for V, and $|u\rangle \in V$ is expressed as

$$|u\rangle \quad = \quad \alpha_1 |v_1\rangle + \ldots + \alpha_n |v_n\rangle$$

  then it's easy to calculate each coefficient α_i as

$$\alpha_i \quad = \quad \langle v_i|u\rangle$$

- It's worth examining why:

$$\begin{aligned}
\langle v_i|u\rangle \quad &= \quad \langle v_i \mid \alpha_1 v_1 + \alpha_2 v_2 \ldots + \ldots + \alpha_n v_n\rangle && \text{Expand } |u\rangle \\
&= \quad \alpha_1 \langle v_i|v_1\rangle + \alpha_2 \langle v_i|v_2\rangle + \ldots + \alpha_n \langle v_i|v_n\rangle && \text{Right-linearity of inner-product} \\
&= \quad \alpha_i \langle v_i|v_i\rangle && \text{All others are 0} \\
&= \quad \alpha_i |v_i|^2 && \\
&= \quad \alpha_i && \text{All } |v_i\rangle\text{'s are unit length}
\end{aligned}$$

  Note: all but one of the inner products are zero.

- Some terminology:
    - Writing a vector as a linear combination of basis vectors

$$|u\rangle \quad = \quad \alpha_1 |v_1\rangle + \ldots + \alpha_n |v_n\rangle$$

is called *expanding* a vector in a basis.
    - The complex scalars α_i are called *coefficients* or *amplitudes*.

What's important to know about vectors and bases in quantum computing:

- Nearly all vectors encountered will be unit vectors (magnitude = 1).
    - When an exception occurs, we typically *normalize* the vector to make it unit-length:

$$|v\rangle \quad = \quad \frac{1}{|u|}|u\rangle$$

- Nearly all bases will be orthonormal bases.

- These properties simplify expressions and calculations but at first can be a bit confusing.

In-Class Exercise 5: Suppose $|u\rangle = (1, 0, 0, 0)$, $|v\rangle = (0, 0, 1, 0)$ and $W = \text{span}(|u\rangle, |v\rangle)$.

a. Show that $|u\rangle, |v\rangle$ are linearly independent.
b. What is the dimension of W?
c. Express $|x\rangle = (-i, 0, i + 1, 0)$ in terms of $|u\rangle, |v\rangle)$.
d. Is $|u\rangle, |v\rangle)$ a basis for W? Explain.
e. Is $W = \mathbb{C}^k$ for any value of k?

The standard basis for $\mathbb{C}^n$:

- Define the n vectors $|e_1\rangle, \ldots, |e_n\rangle$ where $|e_i\rangle = (0, 0, \ldots, 1, \ldots, 0)$
  ▷ All 0's with a 1 as the i-th element.

- Example: for $\mathbb{C}^3$:

$$|e_1\rangle \quad = \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad |e_2\rangle \quad = \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad |e_3\rangle \quad = \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

- This basis is often called the *standard basis* and sometimes the *computational basis*.

- Note: the standard basis itself has no complex numbers.
  ▷ Any complex vector is nonetheless expressible via complex coefficients.

- We'll shortly see a somewhat unusual but commonly used notation for the standard basis.

2.6 More about Dirac notation

At this point you may be wondering why we bother with this "asymmetric" notation?

There are four reasons. The first two are:

- The notation nicely tracks conjugation and transpose.
  - For example, consider an outer-product matrix $|u\rangle\langle v|$ times a vector $|w\rangle$:

$$\begin{aligned}
\underbrace{|u\rangle\langle v|}_{\text{outer product matrix}}\,\overset{\text{vector}}{|w\rangle} &= |u\rangle\,\overset{\text{this is already a conjugated row}}{\overbrace{\langle v|w\rangle}^{\text{inner product (a number)}}} && \text{products of matrices are associative}\\
&= |u\rangle\,(\langle v|w\rangle) && \text{the inner-product number is movable}\\
&= (\langle v|w\rangle)\,|u\rangle\\
&= \alpha\,|u\rangle && \text{inner-product evaluates to a number}
\end{aligned}$$

  - The asymmetric brackets make the conjugation status obvious.

- Because the brackets delineate a single vector, one can write something more elaborate in between as in:
  - $|\text{3rd qubit}\rangle$
  - $\langle Au|$, where the operator A is applied to the vector u and then the result is turned into a conjugated row.

We'll now get to the third reason, which is that a convention has been developed with the notation that greatly eases descriptions for quantum computing:

- Let's start with the standard basis for $\mathbb{C}^2$: $(1,0)$ and $(0,1)$.
  - These are named in Dirac notation as

$$\begin{aligned}
|0\rangle &\triangleq \begin{bmatrix} 1 \\ 0 \end{bmatrix}\\
|1\rangle &\triangleq \begin{bmatrix} 0 \\ 1 \end{bmatrix}
\end{aligned}$$

  - We'll later see that these will correspond to two special states of a single qubit.

- **Note:** The vector $|0\rangle$ is *not* the zero vector.

- The next larger size is two qubits, with four states and four standard-basis vectors:

$$|00\rangle \triangleq \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad |01\rangle \triangleq \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad |10\rangle \triangleq \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \qquad |11\rangle \triangleq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

- Here are two 3-qubit example states:

$$|000\rangle \;\triangleq\; \begin{bmatrix}1\\0\\0\\0\\0\\0\\0\\0\end{bmatrix} \qquad |101\rangle \;\triangleq\; \begin{bmatrix}0\\0\\0\\0\\1\\0\\0\\0\end{bmatrix}$$

- At the moment, this won't make sense but pay attention to the compactness of notation:
  - If we go to 8 bits, the Dirac notation gives us vectors like $|10010001\rangle$.
  - The same vector in standard column form will have $2^8 = 256$ numbers!

- The problem gets worse once we have expressions that combine such vectors:
  - For example, the outer product (a matrix)

    $$|10010001\rangle \langle 10010001|$$

    is written just like that.
  - In conventional linear algebra, this is a 256×256 matrix.

- We will also have to get used to further shortcuts like these:
  - Recall the four two-qubit vectors above: $|00\rangle, |01\rangle, |10\rangle, |11\rangle$.
  - These are also often written as: $|0\rangle, |1\rangle, |2\rangle, |3\rangle$
  - Here, it's understood that the numbers 0, 1, 2, 3 are the decimal version of 00, 01, 10, 11.
  - In this way, the three-qubit vectors from above get compactly written as: $|0\rangle, |1\rangle, |2\rangle, |3\rangle, |4\rangle, |5\rangle, |6\rangle, |7\rangle$.

- We will also need to get used to placing unusual symbols in this notation, for example:
  - $|+\rangle$: Yes, that's the plus symbol in there, representing the vector $|+\rangle = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.
  - $|\uparrow\rangle$: The "up arrow" represents vertical polarization (in a polarized light setting), as $|\rightarrow\rangle$ represents horizontal polarization.
  - Other examples are: $|\nearrow\rangle$ and $|\nwarrow\rangle$
  - Think of these as alternatives to Greek-letter variable names.

  This can sometimes make for challenging reading as in:

$$\frac{1}{\sqrt{2}}|+\rangle \;+\; \frac{1}{\sqrt{2}}|-\rangle$$

This is **not** *addition. Instead, it's a* variable *symbol like x.*

This is the standard addition operator (for vectors)

The variable −

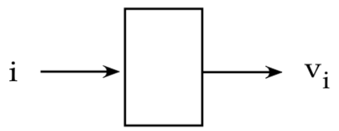
  (This is a linear combination of the two vectors $|+\rangle$ and $|-\rangle$

- Note: while conventional linear algebra likes to use letters like u, v, w for vectors, the quantum literature often uses ψ, pronounced "sigh".
  ▷ Thus, we will commonly see the notation $|\psi\rangle$.

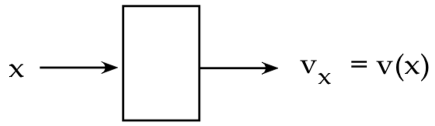
The fourth reason is valuable in quantum *mechanics*:

- It turns out that finite-sized vectors generalize to infinite-sized vectors quite easily:
  - A regular vector like $|v\rangle = (v_1, v_2, \ldots, v_n)$ has its elements *indexed* by the integers $1, 2, \ldots, n$.

- Then, when looking at v_i, the i is an integer between 1 and n.
- Think of the index i as input, and the element v_i as output:



- One could define vectors $|v\rangle = (v_1, v_2, \ldots)$ where the index set is all the natural numbers: $1, 2, 3, \ldots$
- And one can define vectors with a real-valued index, where the elements are v_x.
  - ▷ This is really nothing other than the *function* $v(x)$.



a function

- The Dirac notation is the same for functions: $|f(x)\rangle$
- This fits in with the other kind of generalization that's needed when going from "finite and discrete" to "infinite and continuous":
  - ▷ An *operator* is the equivalent generalization of a matrix.
- Again, Dirac notation treats both the same way, allowing for compact multi-use notation.

Without knowing much more than we've just seen, we can already work with this notation:

- For example, consider the two vectors

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

- We can easily compute inner products:

$$\begin{aligned}
\langle 0|0\rangle &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \\
\langle 0|1\rangle &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \\
\langle 1|0\rangle &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0 \\
\langle 1|1\rangle &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1
\end{aligned}$$

- And outer products:

$$\begin{aligned}
|0\rangle\langle 0| &= \begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}1 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\\
|0\rangle\langle 1| &= \begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}0 & 1\end{bmatrix} = \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\\
|1\rangle\langle 0| &= \begin{bmatrix}0\\1\end{bmatrix}\begin{bmatrix}1 & 0\end{bmatrix} = \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}\\
|1\rangle\langle 1| &= \begin{bmatrix}0\\1\end{bmatrix}\begin{bmatrix}0 & 1\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}
\end{aligned}$$

In-Class Exercise 6: For the vectors $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ defined above, compute the outer products

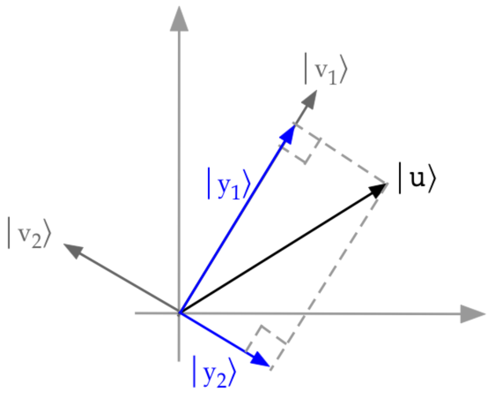
a. $|00\rangle\langle 00|$
b. $|01\rangle\langle 01|$
c. $|10\rangle\langle 10|$
d. $|11\rangle\langle 11|$

In-Class Exercise 7: Using the vectors $|0\rangle, |1\rangle$, compute the vectors

a. $|h_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$
b. $|h_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$
c. Show that $|h_1\rangle, |h_2\rangle$ are orthonormal (unit length and orthogonal).
d. Show that $|h_1\rangle, |h_2\rangle$ are a basis for $\mathbb{C}^2$. [Hint: start by expressing $|0\rangle, |1\rangle$ in terms of $|h_1\rangle, |h_2\rangle$.]

2.7 Projections and projectors

Because we don't have a convenient way to visualize a 2D *complex vector*, let's start with a description that uses *real* vectors:



- Here, $|v_1\rangle, |v_2\rangle$ are orthonormal basis vectors, and $|u\rangle$ is a vector.
- The picture shows the projection of $|u\rangle$ onto each of $|v_1\rangle$ and $|v_2\rangle$.

- Geometrically with real vectors, a projection is the shadow cast by one vector on another.
- For real vectors, we showed that

$$|y_1\rangle \quad = \quad \left(\frac{\langle v_1|u\rangle}{\langle v_1|v_1\rangle} \right) |v_1\rangle$$

- When $|v_1\rangle, |v_2\rangle$ are unit-length, $\langle v_i|v_i\rangle = |v_i|^2 = 1$. Thus

$$|y_1\rangle \quad = \quad \langle v_1|u\rangle \;\; |v_1\rangle$$

  Similarly

$$|y_2\rangle \quad = \quad \langle v_2|u\rangle \;\; |v_2\rangle$$

- Most importantly, the two projections geometrically add up to the original vector:

$$\begin{aligned} |u\rangle \quad &= \quad |y_1\rangle + |y_2\rangle \\ &= \quad \langle v_1|u\rangle \;\; |v_1\rangle \;\; + \;\; \langle v_2|u\rangle \;\; |v_2\rangle \end{aligned}$$

What does a *projection* mean for complex vectors?

- Instead of thinking geometrically, we'll instead focus on this question:
  What parts of $|v_1\rangle, |v_2\rangle$ add up to $|u\rangle$?
- So, let's imagine (complex) numbers α_1, α_2 where

$$\alpha_1 |v_1\rangle + \alpha_2 |v_2\rangle \quad = \quad |u\rangle$$

- We need to solve for the α's. Take the inner product with $|v_1\rangle$ on both sides:

$$\begin{array}{cccll} & \langle v_1 \mid \alpha_1 v_1 + \alpha_2 v_2\rangle & = & \langle v_1|u\rangle & \text{Use right-side linearity} \\ \Rightarrow & \alpha_1 \langle v_1|v_1\rangle + \alpha_2 \langle v_1|v_2\rangle & = & \langle v_1|u\rangle & \langle v_1|v_1\rangle = 1, \langle v_1|v_2\rangle = 0 \\ \Rightarrow & \alpha_1 & = & \langle v_1|u\rangle & \end{array}$$

  (Recall: the v_i's are orthonormal.)

- Thus, the coefficient-of-projection α_i is simply the inner product $\langle v_i|u\rangle$.
- In general, we'll have an orthonormal basis $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$, where for each for any vector $|u\rangle$:

Coefficient when u *is projected on* $\mathrm{v_i}$ $\longrightarrow$ $\langle v_i|u\rangle$ *Read as "on* $\mathrm{v_i}$*, project* u*"*

Complex number

$\mathrm{y_i}$ = *actual vector of projection along* $\mathrm{v_i}$

$$|y_i\rangle = \underbrace{\langle v_i|u\rangle}_{\text{Complex number that "scales" } \mathrm{v_i}} \; |v_i\rangle$$

- Example:
  - Suppose $|u\rangle = (\frac{\sqrt{3}}{2}, \frac{1}{2})$ and $|v_1\rangle = (\frac{1}{2}, \frac{\sqrt{3}}{2})$

    - Then,

$$\begin{aligned} |y_1\rangle &= \langle v_1|u\rangle \ |v_1\rangle \\ &= \left(\begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} \right) \begin{bmatrix} \frac{\sqrt{1}}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix} \ = \ \frac{\sqrt{3}}{2} \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix} \ = \ \begin{bmatrix} \frac{\sqrt{3}}{4} \\ \frac{1}{4} \end{bmatrix} \end{aligned}$$

Projectors or projection matrices:

- Now $|y_1\rangle$ is the vector that results from projecting the vector $|u\rangle$ on $|v_1\rangle$.
- We know that a matrix multiplying into a vector transforms it into another vector.
- Thus, one can ask: is there a matrix P_1 that would achieve this transformation, i.e.

$$|y_1\rangle \ = \ P_1 \, |u\rangle?$$

- Such a matrix is called a *projector* matrix.
- Observe that

$$\begin{aligned} |y_1\rangle &= (\langle v_1|u\rangle) \ |v_1\rangle && \text{From earlier} \\ &= |v_1\rangle \ (\langle v_1|u\rangle) && \text{The scalar in parens can be moved} \\ &= (|v_1\rangle\langle v_1|) \ |u\rangle && \text{Associativity of matrix multiplication} \end{aligned}$$

- But we've seen that the outerproduct $|v_1\rangle\langle v_1|$ is in fact a matrix. So, let's define

$$P_1 \triangleq |v_1\rangle\langle v_1|$$

- Just like we dropped boldface notation for vectors, we will do the same with matrices.
    - Generally, we will use unbolded capital letters for matrices, (the convention in quantum computing/mechanics).
    - It does take some getting used to.
- In our example:

$$|v_1\rangle\langle v_1| \ = \ \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \ = \ \begin{bmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix} \triangleq P_1$$

- Then,

$$P_1 \, |u\rangle \ = \ \begin{bmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} \ = \ \begin{bmatrix} \frac{\sqrt{3}}{4} \\ \frac{1}{4} \end{bmatrix} \ = \ |y_1\rangle$$

- For our two-vector basis, the other projector is:

$$P_2 \ = \ |v_2\rangle\langle v_2| \ = \ \begin{bmatrix} \frac{-\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} \begin{bmatrix} -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \ = \ \begin{bmatrix} \frac{3}{4} & -\frac{\sqrt{3}}{4} \\ -\frac{\sqrt{3}}{4} & \frac{1}{4} \end{bmatrix}$$

The *completeness relation for projectors*:

- We'll explain the idea with a two-vector orthonormal basis $|v_1\rangle, |v_1\rangle$.
- First, we'll write $|u\rangle$ as the addition of $|u\rangle$'s projections on $|v_1\rangle, |v_1\rangle$:

$$\begin{aligned} I\ |u\rangle &= |u\rangle \\ &= P_1\,|u\rangle + P_2\,|u\rangle \\ &= (P_1 + P_2)\ |u\rangle \end{aligned}$$

  Thus

$$(P_1 + P_2 - I)\,|u\rangle \quad = \quad \mathbf{0}$$

  for all u.

- And so, the sum of projectors is the identity matrix:

$$P_1 + P_2 \quad = \quad I$$

- Let's see this at work with the projectors written as outerproducts and for an n-vector orthonormal basis:

$$\begin{array}{rcll} |u\rangle & = & (\langle v_1|u\rangle)\ |v_1\rangle + \ldots + (\langle v_n|u\rangle)\ |v_n\rangle & \text{Each vector with each coefficient} \\ & = & |v_1\rangle\ (\langle v_1|u\rangle) + \ldots + |v_n\rangle\ (\langle v_n|u\rangle) & \text{Scalar movement} \\ & = & (|v_1\rangle\langle v_1|)\ |u\rangle + \ldots + (|v_n\rangle\langle v_n|)\ |u\rangle & \text{Associativity} \\ & = & (|v_1\rangle\langle v_1| + \ldots + |v_n\rangle\langle v_n|)\ |u\rangle & \text{Factoring} \end{array}$$

- That is,

$$|v_1\rangle\langle v_1| + \ldots + |v_n\rangle\langle v_n| \quad = \quad I$$

  Or

$$P_1 + \ldots + P_n \quad = \quad I$$

- Let's work out the completeness relation for the $|h_1\rangle, |h_2\rangle$ vectors seen earlier:
  - Recall: $|h_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|h_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$
  - Then,

$$\begin{aligned} |h_1\rangle\langle h_1| + |h_2\rangle\langle h_2| &= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} + \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \\ &= I \end{aligned}$$

- *Notation*: Sometimes, for a vector $|v\rangle$ we'll use the notation P_v to denote the projector $P_v = |v\rangle\langle v|$

In-Class Exercise 8: For any projector $|v\rangle\langle v|$ show that $(|v\rangle\langle v|)^\dagger = |v\rangle\langle v|$. [Hint: you can transpose and then conjugate.]
Note: a matrix A like $P_v = |v\rangle\langle v|$ that satisfies $A^\dagger = A$ is called *Hermitian*, as we'll see below.

The above result is worth codifying as a formal result:

- ***Proposition 2.1:***
  A projector $P_v = |v\rangle\langle v|$ equals its own adjoint: $P_v^\dagger = P_v$. That is, $(|v\rangle\langle v|)^\dagger = |v\rangle\langle v|$.

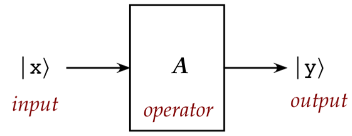
In-Class Exercise 9: Using the vectors $|0\rangle, |1\rangle$, compute the vectors

a. $|y_1\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{i}{\sqrt{2}}|1\rangle$
b. $|y_2\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{i}{\sqrt{2}}|1\rangle)$
c. Show that $|y_1\rangle, |y_2\rangle$ are a basis for $\mathbb{C}^2$.
d. Show the completeness relation for this basis.

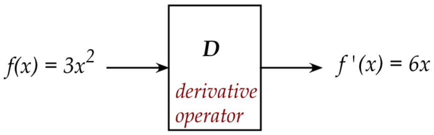
2.8 Two important types of operators (matrices): Hermitian and unitary

Recall the meaning of the term *operator*:

- Think of the matrix-vector multiplication $A|x\rangle$ as *producing* the vector $|y\rangle = A|x\rangle$:



- The idea of an operator can be extended to taking a *function* as input and producing a function as output:



  In this case D is the derivative operator so that $Df(x) = f'(x)$.

- In quantum computing, all operators are matrices.

- Why then do we use the term *operator*?
  - It's a good idea to use the more general term if you someday venture into quantum mechanics or more abstract linear algebra.
  - Many papers and books use the term operator.
  - Operators can be analyzed independent of whether they "act" on anything (vector or function).
  - Operators can be combined
    - ▷ Example: A product of two matrix operators is a matrix operator.

  For the rest of the course, when you see *operator*, think *matrix*.

The *adjoint* of a matrix operator:

- Let's define adjoint for matrix operators.

- For any matrix A, its *adjoint* is the matrix $A^\dagger$.

- That is: $A^\dagger$ is the transpose and conjugate of A.

- Example:
  Let

  $$A \;=\; \begin{bmatrix} 1+i & -i \\ i & 1+1 \end{bmatrix}$$

  Then the transpose is

  $$A^T \;=\; \begin{bmatrix} 1+i & i \\ -i & 1+i \end{bmatrix}$$

  And the conjugate of the transpose is:

  $$A^\dagger \;=\; (A^T)^* \;=\; \begin{bmatrix} 1-i & -i \\ i & 1-i \end{bmatrix}$$

In-Class Exercise 10: Consider these matrices. (Yes, the latter two have special names.)

$$A \;=\; \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \qquad B \;=\; \begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix} \qquad Y \;=\; \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad H \;=\; \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Compute

1. The adjoints $A^\dagger, B^\dagger, Y^\dagger, H^\dagger$
2. $A^\dagger A$, $Y^\dagger Y$ and $H^\dagger H$
3. $AA^\dagger$, $YY^\dagger$ and $HH^\dagger$

In computing the adjoint, do we get the same result if we first apply conjugation and then transpose?

The two kinds of operators we're going to need are: *Hermitian* and *unitary*.

Let's start with *Hermitian*:

- A *Hermitian* operator is an operator that satisfies $A = A^\dagger$

- Thus for example

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

  is Hermitian.

- While

  $$B \;=\; \begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix}$$

  is not.

- Think of *Hermitian* as the complex generalization of a *symmetric matrix* for real numbers.
  ▷ A real matrix that's symmetric is Hermitian.

A unitary operator:

- A *unitary* operator A is one that satisfies $A^\dagger A = AA^\dagger = I$
  ▷ That is, $A^{-1} = A^\dagger$

- Examples:
  - We saw that with

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

    $H^\dagger H = HH^\dagger = I$ and so H is unitary.
  - But

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

    is not.

We're going to include a third type of operator that's closely related to Hermitian operators: *projection*

- Recall that we used vectors to compute a projection *matrix*.

- In this spirit of "operator" terminology, we'll call this a *projection operator* or *projector* (as it's sometimes called).

- Recall that when $|u\rangle$ is projected on $|v\rangle$, the resulting vector (along $|v\rangle$) is:

  $$|v\rangle\langle v| \; |u\rangle$$

- We wrote the outerproduct $|v\rangle\langle v|$ earlier as the projector matrix

  $$P \quad = \quad |v\rangle\langle v|$$

  This is what we mean by the projector or projection operator.

So far all we have are definitions of these three types of matrices, all of which play a crucial role in anything quantum-related.

As a preview:

- We'll use unitary matrices to modify qubits.
  ▷ This will occur by the usual "matrix changes a vector by multiplication"

- Hermitian matrices are quite different:
  - They will be used for something called *measurement*, a feature unique to quantum systems.
  - And we won't be multiplying a vector by a Hermitian matrix
    ▷ Instead, the matrix will be "applied" to a vector in a rather unusual way

- Projector matrices play a role in the unusual application of Hermitian matrices, which is why the two kinds are intimately connected.

Before we get to using these matrices, we'll need to understand some useful general properties.

As a first step, let's point out something common to all three:

- All three matrices are *square*.

- Clearly, this is true for any projector:
  - A projector is constructed from an outerproduct of an $n \times 1$ vector $|v\rangle$ and a $1 \times n$ vector $\langle v|$:

$$(|v\rangle\langle v|)_{n\times n}$$

  - Example:

$$\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 3 & 1 & 0 \end{bmatrix} \quad = \quad \begin{bmatrix} 9 & 3 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- For a Hermitian operator, the squareness arises from requiring $A = A^{\dagger}$:
  - Suppose $A_{m\times n}$ has m rows, n columns.
  - Because $A^{\dagger}$ is the conjugate transpose, $A^{\dagger}$ has n rows, m columns.
  - The only way we can have $A = A^{\dagger}$ is if $m = n$.

- For unitary operator, the squareness comes from $A^{-1} = A^{\dagger}$:
  - We can of course compute $A^{\dagger}$ for a non-square matrix.
  - But for this to equal the inverse, we must have $AA^{\dagger} = A^{\dagger}A = I$.
  - Thus, for multiply compatibility on both sides, A must be square.

- Because these are the key operators that motivate the theory, the term *operator* itself is defined to be "square".

- This makes sense for matrix operators.

- What about something like the differential operator?

- For more general operators, the generalization of "square" is the following:
  - Consider a vector space V, a set of vectors closed under linear combinations
    - ▷ Any linear combination of vectors in V will be in V.
  - An operator is something that acts on a vector $v \in V$ to produce a vector in V.
  - Sometimes this is expressed as: an operator is a mapping from a vector space to itself.

2.9 Useful properties: adjoints and inner products

Note:

- For the most part, we will use "matrix proofs" because they are less abstract and amenable to examples.

- More abstract and generalized "slick" proofs do exist, which we'll use only occasionally.

- Where it makes sense, we'll accompany a proof with an example to illustrate the main ideas.

- More matrix notation:
  - Sometimes we'll use a_{ij} or $a_{i,j}$ to denote the element in row i and column j of a matrix A.
  - The above introduces a new symbol, so one often uses $(A)_{ij}$ or $(A)_{i,j}$ directly.
  - This has the advantage of conveniently describing the i-j-th element of a sum as in $(A+B)_{ij}$.

Let's start with a really useful property when adjoints occur in an inner-product:

- Let's unpack the multiplication $A\,|x\rangle$ and write:

$$A\,|x\rangle \;=\; \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- Then, taking the tranpose and conjugate of the result:

$$(A\,|x\rangle)^\dagger \;=\; \begin{bmatrix} x_1^* & \cdots & x_n^* \end{bmatrix} \begin{bmatrix} a_{11}^* & \cdots & a_{n1}^* \\ \vdots & \vdots & \vdots \\ a_{1n}^* & \cdots & a_{nn}^* \end{bmatrix} \;=\; \langle x|\,A^\dagger$$

  Thus, although we're used to applying an operator from the left, one can apply from the right when it makes sense.

- Let's treat $A\,|x\rangle$ as the single vector $|Ax\rangle$ and write

$$\langle Ax| \;=\; \left(\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right)^\dagger \;=\; \begin{bmatrix} x_1^* & \cdots & x_n^* \end{bmatrix} \begin{bmatrix} a_{11}^* & \cdots & a_{n1}^* \\ \vdots & \vdots & \vdots \\ a_{1n}^* & \cdots & a_{nn}^* \end{bmatrix} \;=\; \langle x|\,A^\dagger$$

- Thus, we get the all important relation

$$\langle Ax| \;=\; \langle x|\,A^\dagger \;=\; |Ax\rangle^\dagger$$

- Because of this,

$$\left\langle A^\dagger x\right| \;=\; \langle x|\,A$$

- Now let's see how the same idea works with an inner-product.
  - Suppose $|w\rangle$ is a vector and consider the inner-product $\langle w|Ax\rangle$.
  - Unpack this as

$$\langle w|Ax\rangle \;=\; \begin{bmatrix} w_1^* & \cdots & w_n^* \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \;=\; \left(\begin{bmatrix} a_{11}^* & \cdots & a_{n1}^* \\ \vdots & \vdots & \vdots \\ a_{1n}^* & \cdots & a_{nn}^* \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \right)^\dagger \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \;=\; \left\langle A^\dagger w \middle| x\right\rangle$$

- These properties are useful enough to memorialize as a formal result.

- ***Proposition 2.2:*** (Operator movement in inner-product)
  Let A be an operator and $|x\rangle$, $|w\rangle$ be vectors. Then,
  i. $\langle Ax| = \langle x|\,A^\dagger = |Ax\rangle^\dagger$
  ii. $\left\langle A^\dagger x\right| = \langle x|\,A$
  iii. $\langle w|Ax\rangle = \left\langle A^\dagger w \middle| x\right\rangle$

iv. $\langle Aw|x\rangle = \left\langle w\middle|A^\dagger x\right\rangle$

- The latter two are particularly useful, so let's put them in a diagram:

$$\langle w|Ax\rangle = \left\langle A^\dagger w\middle|x\right\rangle \qquad\qquad \langle Aw|x\rangle = \left\langle w\middle|A^\dagger x\right\rangle$$

An operator becomes its adjoint when moving across an inner product

Recall: A's adjoint is $A^\dagger$, and $A^\dagger$'s adjoint is A.

Next, some basic properties of adjoints:

- ***Proposition* 2.3:**
  If A and B are operators and α is a complex number, then
  i. $(A^\dagger)^\dagger = A$
  ii. $(\alpha A)^\dagger = \alpha^* A^\dagger$
  iii. $(A+B)^\dagger = A^\dagger + B^\dagger$
  iv. $(AB)^\dagger = B^\dagger A^\dagger$

  Proof:
  For any matrix A, let $(A)_{ij}$ represent the element in row i and column j.

  i. Observe that $(A)_{ij}$, the i-j-th element of A becomes $(A)^*_{ji}$ in $A^\dagger$. Thus, applying conjugate transpose again results in $(A)_{ij}$.
  ii. (αA) multiplies each element of A by α. Thus, the i-j-th element of $(\alpha A)^\dagger$

$$\begin{aligned}
\left((\alpha A)^\dagger\right)_{ij} &= \left((\alpha A)_{ji}\right)^* && \text{Transpose first}\\
&= \left(\alpha (A)_{ji}\right)^* && \text{Pull out } \alpha\\
&= \alpha^* (A)^*_{ji} && \alpha \text{ gets conjugated outside}\\
&= \alpha^* (A^\dagger)_{ij} && \text{By definition of conjugate transpose}
\end{aligned}$$

  iii. Again, we examine how the i-j-th element of $(A+B)^\dagger$ comes about:

$$\begin{aligned}
\left((A+B)^\dagger\right)_{ij} &= \left((A+B)^*\right)_{ji} && \text{Conjugate first, then transpose}\\
&= \left(A^* + B^*\right)_{ji} && \text{Conjugate distributes over } +\\
&= \left(A^*\right)_{ji} + \left(B^*\right)_{ji} && \text{Matrix addition}\\
&= \left(A^\dagger\right)_{ij} + \left(B^\dagger\right)_{ij} && \text{Conjugate transpose}
\end{aligned}$$

  iv. Now for $(AB)^\dagger = B^\dagger A^\dagger$. Unfortunately, getting at the i-j-th element of AB is messy because it is computed with row i of A and column j of B. Instead, we'll demonstrate a rather slick proof using inner-products:

$$\begin{aligned}
\left\langle w \,\middle|\, (AB)^\dagger x\right\rangle &= \langle (AB)w \mid x\rangle && \text{Operator movement in inner-product}\\
&= \langle A(Bw) \mid x\rangle && \text{Associativity}\\
&= \left\langle Bw \,\middle|\, A^\dagger x\right\rangle && \text{Operator movement in reverse}\\
&= \left\langle w \,\middle|\, B^\dagger (A^\dagger x)\right\rangle && \text{Operator movement again}\\
&= \left\langle w \,\middle|\, (B^\dagger A^\dagger) x\right\rangle && \text{Associativity}
\end{aligned}$$

Since this is true for all vectors $|w\rangle$, $|x\rangle$, $(AB)^\dagger = B^\dagger A^\dagger$

In-Class Exercise 11: Use

$$A = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \qquad B = \begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix} \qquad \alpha = 1 + i$$

as examples in confirming (i)-(iv) above.

2.10 Useful properties: Hermitian operators

Recall the definition of a Hermitian operator:

> An operator A is Hermitian if $A^\dagger = A$.

Now let's run through some properties:

- ***Proposition** 2.4:*
  Let A, B be Hermitian operators. Then
  i. $A + B$ is Hermitian
  ii. αA is Hermitian for real numbers α.

  Proof:
  For the first part

$$\begin{aligned}
(A+B)^\dagger_{ij} &= (A+B)^*_{ji} && \text{Conjugate, then transpose} \\
&= (A^* + B^*)_{ji} && \text{Conjugate distributes} \\
&= A^*_{ji} + B^*_{ji} && \text{Matrix addition} \\
&= A^\dagger_{ij} + B^\dagger_{ij} && \text{Definition of conjugate-transpose} \\
&= A_{ij} + B_{ij} && \text{A, B are Hermitian} \\
&= (A+B)_{ij}
\end{aligned}$$

  For the second

$$(\alpha A)^\dagger_{ij} = (\alpha A)^*_{ji} = (\alpha^* A^*)_{ji} = \alpha A^*_{ji} = \alpha A^\dagger_{ji} = \alpha A_{ij} = (\alpha A)_{ij}$$

In-Class Exercise 12: If α were not real above, show with an 2×2 example that αA is not Hermitian when A is Hermitian.

- ***Proposition** 2.5:*
  The diagonal elements of a Hermitian matrix are real numbers.

Proof:
Let a_{kk} be the k-th diagonal element of A. Since the transpose-conjugate of the diagonal remains on the diagonal, the k-th diagonal element of $A^\dagger$ is a^*_{kk}. Then, $A^\dagger = A$ implies $a^*_{kk} = a_{kk}$, which implies a_{kk} is real.

Example:
Consider

$$A = \begin{bmatrix} \ddots & \cdots & \\ \vdots & a+ib & \vdots \\ & \cdots & \ddots \end{bmatrix}$$

and let's focus on one of the diagonal elements shown. Then the transpose conjugate will end up conjugating this element

$$A^\dagger = \begin{bmatrix} \ddots & \cdots & \\ \vdots & a-ib & \vdots \\ & \cdots & \ddots \end{bmatrix}$$

Thus, $A^\dagger = A$ implies $a + ib = a - ib$ or $b = 0$.

- ***Proposition 2.6:***
  The eigenvalues of a Hermitian operator are real.

  Proof:

  - Suppose the Hermitian operator A has eigenvalue λ for eigenvector v, i.e. $A\,|v\rangle = \lambda\,|v\rangle$.
  - Recall from earlier, for any operator A and vectors $|w\rangle\,,|x\rangle$

  $$\langle w \mid Ax\rangle = \left\langle A^\dagger w \,\middle|\, x\right\rangle$$

  - For a Hermitian operator

  $$\langle w \mid Ax\rangle = \langle Aw \mid x\rangle$$

  - Next, substitute $w = v, x = v$ to get

  $$\langle Av \mid v\rangle = \langle v \mid Av\rangle$$

  - Applying the eigenvalue relation on both sides,

  $$\langle \lambda v \mid v\rangle = \langle v \mid \lambda v\rangle$$

    and factoring out λ,

  $$\lambda^*\,\langle v|v\rangle = \lambda\,\langle v|v\rangle$$

  - Thus, $\lambda^* = \lambda$ which makes it real.

  Application:

  - This turns out to matter when a Hermitian matrix is written in its eigenbasis, in which case the only non-zero elements (the eigenvalues) are on the diagonal.
  - These eigenvalues correspond to real-world quantities that are observed (such as energy, for example).

- ***Proposition 2.7:***
  Eigenvectors corresponding two distinct eigenvalues of a Hermitian operator are orthogonal.

Proof:

- Let λ_1, λ_2 be two distinct eigenvalues corresponding to eigenvectors $|v_1\rangle, |v_2\rangle$.
- Then

$$\langle Av_1|v_2\rangle \quad = \quad \langle \lambda_1 v_1|v_2\rangle \quad = \quad \lambda_1^* \langle v_1|v_2\rangle \quad = \quad \lambda_1 \langle v_1|v_2\rangle$$

  (The last step because eigenvalues are real.)
- Next,

$$\langle v_1 \mid Av_2\rangle \quad = \quad \langle v_1 \mid \lambda_2 v_2\rangle \quad = \quad \lambda_2 \langle v_1|v_2\rangle$$

- These two must be equal because

$$\langle Av_1 \mid v_2\rangle \quad = \quad \left\langle v_1 \,\middle|\, A^\dagger v_2\right\rangle \quad = \quad \langle v_1 \mid Av_2\rangle$$

  since $A = A^\dagger$.
- Thus, subtracting

$$(\lambda_1 - \lambda_2)\langle v_1|v_2\rangle \quad = \quad 0$$

  Which means $\langle v_1|v_2\rangle = 0$ since the eigenvalues are assumed to be distinct.

Application:

- This is critical in quantum mechanics because of the way measurements (as we'll see) work. When a measurement results one of two different eigenvalues, the resulting state (an eigenvector) will be exactly one or the other.
- In quantum computing, we will exploit this to reason about Hermitians.

- ***Proposition 2.8:*** (The spectral theorem for Hermitian operators)
  A Hermitian operator A on an n-dimensional vector space V has n orthonormal eigenvectors $|v_1\rangle, \ldots, |v_n\rangle$ such that:
  i. $|v_1\rangle, \ldots, |v_n\rangle$ form a basis for V.
  ii. A is a diagonal matrix

  when written in the basis $|v_1\rangle, \ldots, |v_n\rangle$.
  iii. $A = \sum_{i=1}^n \lambda_i |v_i\rangle\langle v_i|$.

  Proof:
  The proof is somewhat long and uses induction; we'll refer you to Axler's book.

2.11 Useful properties: orthonormality and projector operators

Squared *amplitudes* add up to 1:

- ***Proposition 2.9:***
  Let α_i's be the amplitudes (coefficients) when $|u\rangle = \sum_i \alpha_i |v_i\rangle$ is expressed in in the (orthonormal) basis $|v_1\rangle, \ldots, |v_n\rangle$. Then

$$\sum_i |\alpha_i|^2 \quad = \quad 1$$

Proof:

$$\begin{aligned}
\sum_i |\alpha_i|^2 &= \sum_i |\langle v_i|u\rangle|^2 \\
&= \sum_i \langle v_i|u\rangle^* \, \langle v_i|u\rangle && \text{Squared magnitude of a complex number} \\
&= \sum_i \langle u|v_i\rangle \, \langle v_i|u\rangle && \text{Inner-product property} \\
&= \sum_i \langle u| \left(|v_i\rangle\langle v_i|\right) |u\rangle && \text{Associativity} \\
&= \langle u| \left(\sum_i |v_i\rangle\langle v_i| \right) |u\rangle && \text{Algebra} \\
&= \langle u| \; I \; |u\rangle && \text{Completeness} \\
&= \langle u|u\rangle && \text{Identity times u} \\
&= 1
\end{aligned}$$

Application:

- As we'll see in the next module, this is fundamental to the theory when the $|\alpha_i|^2$'s are interpreted as probabilities.
- A full set of probabilities, as we know, must sum to 1.

Projectors and Hermitians:

- ***Proposition 2.10:***
  A projector is *idempotent*. That is, the projector

  Proof:

  $$P_v^2 \quad = \quad P_v P_v \quad = \quad |v\rangle\langle v| \; |v\rangle\langle v| \quad = \quad |v\rangle \left(\langle v|v\rangle\right) \langle v| \quad = \quad |v\rangle \times 1 \times \langle v| \quad = \quad |v\rangle\langle v| \quad = \quad P_v$$

  Application: this matches intuition

  - When we project a vector $|u\rangle$ onto $|v\rangle$, we get a vector along $|v\rangle$.
  - Suppose we call that vector $|y\rangle = P_v |u\rangle$.
  - Applying P_v twice is the same as applying P_v to $|y\rangle$, i.e., $P_v P_v |u\rangle = P_v |y\rangle$.
  - But the projection of a vector that's already along $|v\rangle$ leaves the vector unchanged.
  - Thus, $P_v |y\rangle = |y\rangle$.
  - Which means $P_v P_v |u\rangle = P_v |u\rangle$ or $P_v^2 = P_v$

- ***Proposition 2.11:***
  A projector is Hermitian. That is, the projector $P_v = |v\rangle\langle v|$ satisfies $P^\dagger = P$.

  Proof:
  See earlier exercise.

  Application: We will see a stronger result below, but we'll need this one when projectors need to be combined for multiple qubits.

- ***Proposition 2.12:***
  The projector $P_v = |v\rangle\langle v|$ has eigenvector $|v\rangle$ with eigenvalue 1.

  Proof:
  Clearly

  $$P_v |v\rangle \quad = \quad |v\rangle\langle v| \ |v\rangle \quad = \quad |v\rangle \ \langle v|v\rangle \quad = \quad |v\rangle$$

- ***Proposition 2.13:***
  For real number λ and projector $P_v = |v\rangle\langle v|$, the operator λP_v is Hermitian with eigenvector $|v\rangle$ and eigenvalue λ.

  Proof:
  The fact that λP_v is Hermitian follows from the proposition earlier that showed that for any real scalar λ and Hermitian A, the operator λA is Hermitian. Next,

  $$(\lambda P_v) |v\rangle \quad = \quad \lambda P_v |v\rangle \quad = \quad \lambda |v\rangle$$

  Which makes $|v\rangle$ an eigenvector of λP_v with eigenvalue λ.

- ***Proposition 2.14:***
  Let $P_{v_1}, \ldots, P_{v_n}$ be projectors for basis vectors $|v_1\rangle, \ldots, |v_n\rangle$. Next, let $\lambda_1, \ldots, \lambda_n$ be *real* numbers. Then the *real* linear combination

  $$A \quad = \quad \sum_i \lambda_i P_{v_i}$$

  is a Hermitian operator with eigenvectors $|v_1\rangle, \ldots, |v_n\rangle$ and corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$.

  Proof:
  The previous proposition shows that each $\lambda_i P_{v_i}$ is Hermitian. The sum of Hermitians is Hermitian. Next,

  $$A |v_i\rangle \quad = \quad (\sum_k \lambda_k P_{v_k}) |v_i\rangle \quad = \quad \lambda_i P_{v_i} |v_i\rangle \quad = \quad \lambda_i |v_i\rangle$$

  Application:

  - What we've shown is that *some* Hermitians have the particular linear-combination-of-projectors structure as abov.
  - These, in fact, will be the only type of Hermitians we'll encounter in quantum computing.
  - And this structure is the key to understanding one of the most counter-intuitive aspects of quantum computing: measurement.

2.12 Useful properties: unitary operators

Recall:

> A *unitary* operator A is one that satisfies $A^\dagger A = AA^\dagger = I$
> ▷ That is, $A^{-1} = A^\dagger$

Let's work through some useful properties of unitary operators

- ***Proposition 2.15:***
  A unitary operator preserves inner-products: $\langle Au|Av\rangle = \langle u|v\rangle$.

  Proof:
  We'll exploit the operator-across-inner-product property here:

  $$\langle Au|Av\rangle \quad = \quad \left\langle A^{\dagger}(Au) \,\middle|\, v \right\rangle \quad = \quad \left\langle (A^{\dagger}A)u) \,\middle|\, v \right\rangle \quad = \quad \langle Iu \mid v\rangle \quad = \quad \langle u|v\rangle$$

- ***Proposition 2.16:***
  A unitary operator preserves lengths: $|Au| = |u|$.

  Proof:
  From the above property

  $$|Au|^2 \quad = \quad \langle Au \mid Au\rangle \quad = \quad \langle u|u\rangle = |u|^2$$

- ***Proposition 2.17:***
  If A is unitary so are $A^{\dagger}$ and A^{-1}.

  Proof:
  The definition of unitary is symmetric: A is unitary if $AA^{\dagger} = A^{\dagger}A = I$. Thus, $A^{\dagger}$ is also unitary, and because $A^{-1} = A^{\dagger}$, it too is unitary.

- ***Proposition 2.18:***
  The columns of A are orthonormal, as are the rows.

  Proof:

  - Consider the product $A^{\dagger}A = I$ and the i-j-th element of I.
  - This is formed by multiplying the i-th row of A^dagger (conjugate of the i-th column) into the j-th column of A.
  - When $i \neq j$ we get 0, meaning the i and j columns are orthogonal.
  - Similarly, when $i = j$, we get a diagonal element of I, which is 1, meaning each column is of unit length.

  The argument for orthonormality of the rows uses the same arguments with $AA^{\dagger} = I$.

- ***Proposition 2.19:***
  If $|v_1\rangle, \ldots, |v_n\rangle$ are orthonormal, so are $|Av_1\rangle, \ldots, |Av_n\rangle$.

  Proof:
  What we need to show is that $\left\langle Av_i \,\middle|\, Av_j \right\rangle = 0$ if $i \neq j$ and $\left\langle Av_i \,\middle|\, Av_j \right\rangle = 1$ if $i = j$. This follows from the preservation of inner products: $\left\langle Av_i \,\middle|\, Av_j \right\rangle = \left\langle v_i|v_j \right\rangle$

- ***Proposition 2.20:***
  If A, B are unitary, then so are AB and BA.

  Proof:

  $$(AB)^{\dagger}\,(AB) \quad = \quad (B^{\dagger}A^{\dagger})\,(AB) \quad = \quad B^{\dagger}\,(A^{\dagger}A)\,B \quad = \quad B^{\dagger}(I)B \quad = \quad B^{\dagger}B \quad = \quad I$$

  The proof for BA is similar.

  Application:

  - The product rule above is probably the most frequently applied. Think of A and B as two *gates* that occur in sequence. The net result is the product, as we will see.

- The other properties above are useful in reasoning about unitary matrices and building the theory.

In-Class Exercise 13: Use a 2×2 example to show that the sum of two unitary matrices is not necessarily unitary.

2.13 The operator sandwich

Let's now return to applying an operator from the left, and from the right:

- We have seen that

  $$A\,|v\rangle \quad = \quad |Av\rangle$$

  and thus there is no ambiguity in writing either way.

- Similarly, when applying from the right:

  $$\langle Au| \quad = \quad \langle u|\,A^{\dagger}$$

  which means

  $$\left\langle A^{\dagger}u\right| \quad = \quad \langle u|\,A$$

- Next, consider the expression

  $$\langle u|A|v\rangle$$

  - If A were applied to the left of $|v\rangle$, this would become:

    $$\langle u|A|v\rangle \quad = \quad \langle u \mid Av\rangle$$

  - If A were applied to the right of $\langle u|$:

    $$\langle u|A|v\rangle \quad = \quad \left\langle A^{\dagger}u \,\middle|\, v\right\rangle$$

  - But both result in the same inner product:

    $$\langle u|Av\rangle \quad = \quad \left\langle A^{\dagger}u \,\middle|\, v\right\rangle$$

- Thus, the actual *calculation* could be done either way.

- We will use the so-called *operator sandwich* to write this as

  $$\langle u|A|v\rangle$$

  where it's implied that we have two ways of performing the calculation.

- **Note**:
  - $\langle u|A|v\rangle$ is still an inner-product, and will result in a *number*.
  - This result applies to any operator, not just Hermitian and unitary operators.

Let's see how this works through an example:

- Suppose

$$A \quad = \quad \begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix} \qquad |u\rangle \quad = \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad |v\rangle \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

- We'll first calculate

$$\begin{aligned} \langle u|A|v\rangle \quad &= \quad \langle u|Av\rangle \\ &= \quad \begin{bmatrix} 1 & 0 \end{bmatrix} \left(\begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right) \\ &= \quad \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 + \frac{i}{\sqrt{2}} \\ 1 - \frac{i}{\sqrt{2}} \end{bmatrix} \\ &= \quad 1 + \frac{i}{\sqrt{2}} \end{aligned}$$

- And now the other way:

$$\langle u|A|v\rangle \quad = \quad \left\langle A^{\dagger}u \,\middle|\, v \right\rangle$$

  First, note that

$$A^{\dagger}u \quad = \quad \begin{bmatrix} -i & 1 \\ 1 & i \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad = \quad \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

  The inner-product of the conjugated row $\left\langle A^{\dagger}u \right|$ with $|v\rangle$ then is

$$\left\langle A^{\dagger} \middle| v \right\rangle \quad = \quad \begin{bmatrix} i & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad = \quad 1 + \frac{i}{\sqrt{2}}$$

  Which is, as expected, the same result (number) as before.

How it's used:

- The operator sandwich is used frequently and is yet another notational and conceptual idea we need to get comfortable with.

- Here's one example with a projector:
  - Consider the projection of $|u\rangle$ on $|v\rangle$.
  - We know that the projector is written as the outer-product

$$P_v \quad = \quad |v\rangle\langle v|$$

  - Now consider the sandwich

$$\langle u|P_v|u\rangle \quad = \quad \langle u| \; |v\rangle\langle v| \; |u\rangle \quad = \quad \langle u|v\rangle \; \langle v|u\rangle \quad = \quad (\langle v|u\rangle)^* \; \langle v|u\rangle \quad = \quad |\langle v|u\rangle|^2$$

  - Thus, $\langle u|P_v|u\rangle$ is the squared magnitude of the coefficient in the projection.

- Here's another example with projectors:
    - The projected vector when projecting $|u\rangle$ along $|v\rangle$ is: $P_v\,|u\rangle$.
    - Suppose we want the length of this vector: $|P_v\,|u\rangle|$.
    - We can write the squared magnitude as:

$$\begin{aligned}
|P_v\,|u\rangle|^2 &= (P_v\,|u\rangle)^\dagger\; P_v\,|u\rangle && \text{Inner product of } P_v\,|u\rangle \text{ with itself}\\
&= \left(\langle u|\,P_v^\dagger\right)\; P_v\,|u\rangle && \text{Apply adjoint}\\
&= \langle u|\,P_v\,P_v\,|u\rangle && \text{A projector is Hermitian}\\
&= \langle u|P_v|u\rangle && \text{A projector is idempotent}
\end{aligned}$$

This is the same as $|\langle v|u\rangle|^2$ derived earlier because we're working with unit-length vectors.

Let's work through a problem to get some practice:

- The problem: compute $\langle u|P_{v_1}|u\rangle$ given $|u\rangle = \alpha\,|v_1\rangle + \beta\,|v_2\rangle$.

- Let's start with

$$\begin{aligned}
P_{v_1}\,|u\rangle &= |v_1\rangle\langle v_1|\;\left(\alpha\,|v_1\rangle + \beta\,|v_2\rangle\right) && \text{Projector as outer-product}\\
&= |v_1\rangle\;\left(\alpha\,\langle v_1|v_1\rangle + \beta\,\langle v_1|v_2\rangle\right) && \text{Right-linearity of inner-product}\\
&= |v_1\rangle\,\alpha && \text{Basis vector inner-products}\\
&= \alpha\,|v_1\rangle
\end{aligned}$$

- Next,

$$\begin{aligned}
\langle u|P_{v_1}|u\rangle &= \langle \alpha v_1 + \beta v_2|\;\;\alpha\,|v_1\rangle && \text{Sub for } |u\rangle \text{ on left, and above-computed } P_{v_1}\,|u\rangle\\
&= \alpha\,\langle \alpha v_1 + \beta v_2|v_1\rangle && \text{Move second } \alpha\\
&= \alpha\left(\alpha^*\,\langle v_1|v_1\rangle + \beta^*\,\langle v_2|v_1\rangle\right) && \text{Left-side linearity}\\
&= \alpha\alpha^* && \text{Basis vector inner-products}\\
&= |\alpha|^2
\end{aligned}$$

In-Class Exercise 14: Suppose $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ and $|u\rangle = \alpha\,|e_1\rangle + \beta\,|e_2\rangle$. Show that $\langle u|A|u\rangle = 2|\alpha|^2 + 3|\beta|^2$

2.14 A key question: what is the basis of the moment?

In the regular linear algebra we've studied:

- We see multiple bases but we typically only use *coordinates* from the standard basis.

- We rarely if ever switch coordinates to a non-standard basis.

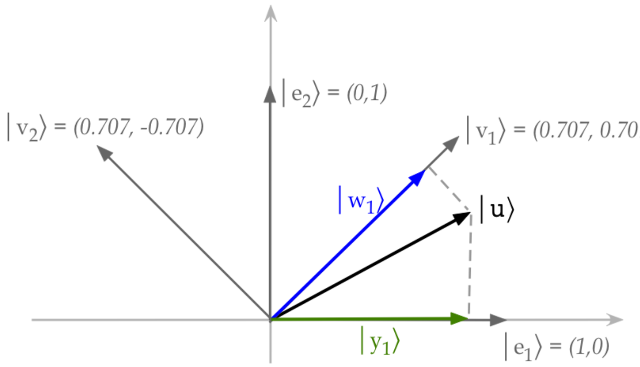
- And almost never express a matrix in a non-standard basis.

In quantum computing (and mechanics), on the other hand:

- We use different bases at different times.
- It is essential to know at any time *which basis* is being used, and for what.
- When using a non-standard basis we have two options:
  1. Use the non-standard basis but using coordinates from the standard basis for calculations.
  2. Convert to the coordinates of the non-standard basis and then calculate from there.

  We'll illustrate both below.

Let's work through an example to clarify:



- We'll use real numbers so that we can draw vectors (above).
- Here, there are two bases shown:
  - The standard basis:

$$|e_1\rangle \quad = \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad |e_2\rangle \quad = \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

  - Another basis (called the Hadamard basis):

$$|v_1\rangle \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \qquad |v_2\rangle \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

    Because $\frac{1}{\sqrt{2}} \approx 0.707$, the coordinates are labeled as such.

- Notice that the diagram shows projections of the vector

$$|u\rangle \quad = \quad \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$$

  along $|e_1\rangle$ (in green) and along $|v_1\rangle$ (in blue).

- Let

$$\begin{aligned} |y_1\rangle \ &\triangleq \ \text{projection along } |e_1\rangle \\ |w_1\rangle \ &\triangleq \ \text{projection along } |v_1\rangle \end{aligned}$$

Let's first use *standard-basis coordinates* for calculations:

- The projector for $|e_1\rangle$ is:

$$P_{e_1} \quad = \quad |e_1\rangle\langle e_1| \quad = \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \quad = \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

- Then, the projection of $|u\rangle$ along $|e_1\rangle$ is:

$$|y_1\rangle \quad = \quad P_{e_1} |u\rangle \quad = \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} \quad = \quad \begin{bmatrix} \frac{\sqrt{3}}{2} \\ 0 \end{bmatrix}$$

- The projector for $|v_1\rangle$ is:

$$P_{v_1} \quad = \quad |v_1\rangle\langle v_1| \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad = \quad \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

- And therefore the projection of $|u\rangle$ along $|v_1\rangle$ is:

$$|w_1\rangle \quad = \quad P_{v_1} |u\rangle \quad = \quad \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} \quad = \quad \begin{bmatrix} \frac{\sqrt{3}+1}{2} \\ \frac{\sqrt{3}+1}{2} \end{bmatrix}$$

  This should make sense because $|v_1\rangle$ is at a 45° angle and so both coordinates should be the same.

- Let's point out:
  - Although $|v_1\rangle, |v_2\rangle$ is a different basis than $|e_1\rangle, |e_2\rangle$, all calculations have been done using coordinates from $|e_1\rangle, |e_2\rangle$ (standard-basis).
  - We know this because $|v_1\rangle, |v_2\rangle$ themselves were written in standard-basis coordinates:

$$|v_1\rangle \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \qquad |v_2\rangle \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

  - Thus, *all the calculations have been done in the standard basis*.

Now let's see what it's like to perform calculations in the other basis, the $|v_1\rangle, |v_2\rangle$ basis:

- First, we'll ask: what are the coordinates of the basis vectors $|v_1\rangle, |v_2\rangle$ in their own basis?
  - To get the coordinates of $|v_1\rangle$ in the $|v_1\rangle, |v_2\rangle$ basis, we need to solve

$$\alpha_1 |v_1\rangle + \alpha_2 |v_2\rangle \quad = \quad |v_1\rangle$$

  - That is, what linear combination of the basis vectors gives us the target vector $|v_1\rangle$, in this case)?
  - Clearly, the solution is $\alpha_1 = 1, \alpha_2 = 0$.
  - Similarly, to get the coordinates of $|v_2\rangle$, we solve

$$\beta_1 |v_1\rangle + \beta_2 |v_2\rangle \quad = \quad |v_2\rangle$$

    to get $\beta_1 = 0, \beta_2 = 1$.
  - Thus the coordinates of $|v_1\rangle, |v_2\rangle$ in the new basis are:

$$|v_1\rangle \quad = \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad |v_2\rangle \quad = \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- Surprised by this (simple) result? We'll have more to say about this below.

- Next, let's write $|u\rangle$ in the $|v_1\rangle, |v_2\rangle$ basis:
    - We need to solve

$$\gamma_1 |v_1\rangle + \gamma_2 |v_2\rangle \quad = \quad |u\rangle$$

    - How do we solve for γ_1, γ_2 without knowing $|u\rangle$?
    - The key insight is that γ_1, γ_2 will be the same no matter which basis is used.
    - So, since we know $|u\rangle$ in the standard basis, we'll first use the standard basis: That is,

$$\gamma_1 \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} + \gamma_2 \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \quad = \quad \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$$

    - Which we can write in matrix form:

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} \quad = \quad \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$$

    - We could solve this the usual linear-equation way, but notice that the matrix is unitary, which means the inverse is the adjoint.
    - Thus

$$|u\rangle \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix} \quad = \quad \begin{bmatrix} \frac{\sqrt{3}+1}{2\sqrt{2}} \\ \frac{\sqrt{3}-1}{2\sqrt{2}} \end{bmatrix}$$

- Now that we know the coordinates of $|u\rangle$ in the new basis, let's compute the projection along $|v_1\rangle$:
    - The projector in new coordinates is:

$$P_{v_1} \quad = \quad |v_1\rangle\langle v_1| \quad = \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \quad = \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

    - Applying this to $|u\rangle$:

$$P_{v_1} |u\rangle \quad = \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}+1}{2\sqrt{2}} \\ \frac{\sqrt{3}-1}{2\sqrt{2}} \end{bmatrix} \quad = \quad \begin{bmatrix} \frac{\sqrt{3}+1}{2\sqrt{2}} \\ 0 \end{bmatrix}$$

In-Class Exercise 15: Using $|v_1\rangle, |v_2\rangle$ as the basis,

a. Find the coordinates of $|e_1\rangle, |e_2\rangle$;
b. Then, compute $|y_2\rangle = P_{e_2} |u\rangle$ in this basis.
c. Check your calculations by obtaining $|y_2\rangle$ in the standard basis and converting that to the $|v_1\rangle, |v_2\rangle$ basis.

Let's make a few comments about bases:

- Let's start by asking: when do we exercise a choice in basis?

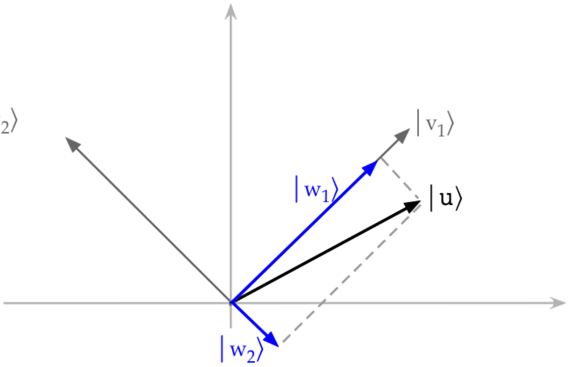
- There are only two reasons to use a particular basis:
    1. When the use of the basis is part of the design:

        - This occurs when we seek a particular outcome from a particular basis.
    2. When a different basis makes calculations easier, and insight possible.

- Most often, the first reason will drive our choices.
    - That is, we'll do calculations in the standard basis.
    - But we'll want projections and such onto vectors in a non-standard basis.
    - We'll calculate these projections with standard-basis coordinates.

- It's equally important to understand that much of the theory does not need a choice of basis to be made when proving a result.

- To see why, let's go back to our projection picture:



    - Here, the (real) vectors are abstract "arrows".
    - We "numerify" an arrow when an arrow is written in terms of basis vectors.
    - For example:

$$|u\rangle \quad = \quad \frac{\sqrt{3}}{2}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad = \quad \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$$

    This provides *coordinates* in the chosen basis (the standard basis in this case).

- Next, observe that
    - The notion of projection can be described *geometrically* (for real vectors) without any reference to which basis the vectors.
    - More valuably, the notion of project can be described *algebraically* without reference to a basis, that is, without "numerifying".
    - For example, the projection of $|u\rangle$ along $|v1\rangle$ is just

$$|v_1\rangle\langle v_1| \ |u\rangle$$

    in whichever basis is used.

- What is (mathematically) interesting is that:
    - Much of the theory can be laid out in a basis-free manner.
    - And, in fact, can be proved for more general and abstract forms of vectors.
    - For example: functions.
    - One defines a *vector space* as a set of mathematical objects that satisfy certain rules like addition, scalar multiplication and so on.
    - Certainly, the vectors we've seen satisfy this definition.
    - But so do functions, which allows the theory to apply more broadly.

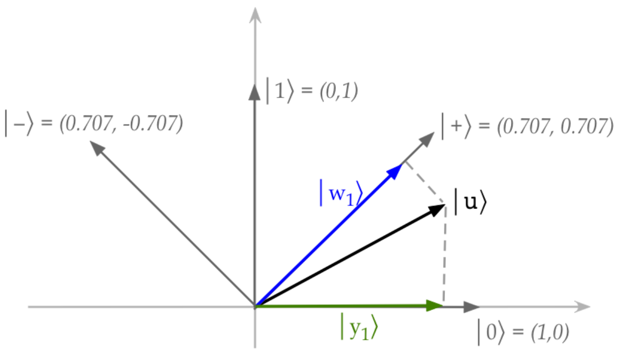
- To perform calculations (numbers!) we of course need to use some basis.

- The choice of basis is often up to us.

2.15 Getting used to Dirac notation

Because we will see certain Dirac symbols and notation repeatedly, let's revisit parts of the previous section in this notation:

- First, we'll relabel the picture



- Next, the standard basis vectors are written as

$$\begin{aligned} |0\rangle &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ |1\rangle &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

- The second basis is called the *Hadamard basis* and is written as:

$$\begin{aligned} |+\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \\ |-\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \end{aligned}$$

- Observe that

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

- The projector for $|0\rangle$ is:

$$P_0 = |0\rangle\langle 0| = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

And the projector for $|1\rangle$ is:

$$P_1 \;=\; |1\rangle\langle 1| \;=\; \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} \;=\; \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

- Similarly, the projector for $|+\rangle$ is:

$$P_+ \;=\; |+\rangle\langle +| \;=\; \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \;=\; \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

In-Class Exercise 16: This is just a writing exercise to get used to the new symbols. Rewrite the previous exercise with the new symbols. That is, using $|+\rangle, |-\rangle$ as the basis,

a. Find the coordinates of $|0\rangle, |1\rangle$;
b. Then, compute $|y_2\rangle = P_0 |u\rangle$ in this basis.
c. Check your calculations by obtaining $|y_2\rangle$ in the standard basis and converting that to the $|+\rangle, |-\rangle$ basis.

2.16 Operator-matrix and change-of-basis via sandwich and inner product

We'll now work through two useful results:

1. If an operator is not in matrix form, then how do we obtain its matrix?
2. How do we change coordinates from one basis to another?

Although we've seen the second one before (in the Review), we'll use our new found approach with the operator sandwich, and inner products.

And at the end, we'll explain why this is a powerful tool.

Let's start with the *matrix of an operator*:

- This may seem an odd objective because ... aren't all operators already matrices?
  - In fact, no. An operator can be specified by text description.
  - For example, let A be the operator that multiplies any 2D vector element by 3:

$$A\,|\psi\rangle \;=\; 3\,|\psi\rangle$$

  - It turns out, in this case, the matrix is:

$$A \;=\; \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

    This is called the *matrix representation* of the operator A.
  - What we seek: a systematic way to obtain the matrix representation.

- What we also seek: to directly obtain the matrix of an operator in a given basis.

- We'll do both with this powerful result:
  - Suppose $|v_1\rangle, \ldots, |v_n\rangle$ is the desired basis in which we want the matrix for operator A.
  - Define the numbers

$$A_{ij} \triangleq \langle v_i|A|v_j\rangle$$

  - Then the matrix for A is simply the matrix with

$$A_{ij} \quad = \quad \text{element in row i, column j}$$

- ***Proposition 2.21:***
  The matrix of an operator A in a basis $|v_1\rangle, \ldots, |v_n\rangle$ is formed by elements $A_{ij} \triangleq \langle v_i|A|v_j\rangle$ in row i and column j.

  Proof:
  We'll do this in two steps:

  1. First show that the matrix

$$\hat{A} \triangleq \sum_{i,j} A_{ij} |v_i\rangle\langle v_j|$$

     produces the same effect as the operator A.
  2. Then show that the i-j-th element of $\hat{A}$ is in fact $\langle v_i|A|v_j\rangle$.

  Let's start with the first step:

  - First, suppose

$$|\phi\rangle \quad = \quad A|\psi\rangle$$

    for two vectors $|\psi\rangle$ and $|\phi\rangle$.
  - Then, express both in the basis:

$$\begin{aligned} |\psi\rangle &= \alpha_1|v_1\rangle + \ldots + \alpha_n|v_n\rangle \\ |\phi\rangle &= \beta_1|v_1\rangle + \ldots + \beta_n|v_n\rangle \end{aligned}$$

    where $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_n$ are the *coordinates*, respectively, of each of these vectors.
  - Now, any outer-product like $|v_i\rangle\langle v_j|$ is a matrix and so

$$\hat{A} \quad = \quad \sum_{i,j} A_{ij} |v_i\rangle\langle v_j|$$

    which is a linear combination of matrices is a matrix.
  - Let's apply this matrix to $|\psi\rangle$:

$$\begin{aligned} \hat{A}|\psi\rangle &= \left(\sum_{i,j} A_{ij}|v_i\rangle\langle v_j|\right)\left(\sum_k \alpha_k|v_k\rangle\right) \\ &= \sum_i\sum_j\sum_k \alpha_k A_{ij}|v_i\rangle\langle v_j| \; |v_k\rangle \\ &= \sum_i\sum_j\sum_k \alpha_k A_{ij}|v_i\rangle \; \langle v_j|v_k\rangle \\ &= \sum_i\sum_j \alpha_j A_{ij}|v_i\rangle \end{aligned}$$

where the last simplification comes from the orthonormality simplification of the inner-products (the case $j = k$ and $j \neq k$).

- Next, write this last expression as

  $$\sum_i \left(\sum_j A_{ij} \alpha_j \right) |v_i\rangle$$

  Now observe: the sum inside the parentheses is the product of the i-th row of $\hat{A}$ and $|\psi\rangle$
- Now compare

  $$|\phi\rangle \;\; = \;\; A\,|\psi\rangle$$

  with what we just derived

  $$\hat{A}\,|\psi\rangle \;\; = \;\; \sum_i \left(\sum_j A_{ij} \alpha_j \right) |v_i\rangle$$

- Because coefficients in a linear combination of basis vectors are unique,

  $$\beta_i \;\; = \;\; \left(\sum_j A_{ij} \alpha_j \right)$$

- And thus, the matrix $\hat{A}$ produces the same transformation as the operator A.

Next, for step 2, let's ask: in using $\hat{A}$ for A, do we get the i-j-th element as $\left\langle v_i | A | v_j \right\rangle$?

- Consider $\left\langle v_i | A | v_j \right\rangle$.
- Substitute the matrix $\hat{A}$:

  $$\begin{aligned} \left\langle v_i \middle| \hat{A} \middle| v_j \right\rangle &= \left\langle v_i \middle| \sum_{m,n} A_{m,n} |v_m\rangle\langle v_m| \middle| v_j \right\rangle \\ &= \left\langle v_i \middle| \sum_m A_{m,j} \middle| v_m \right\rangle \\ &= A_{ij} \qquad\qquad\qquad\qquad = \left\langle v_i | A | v_j \right\rangle \end{aligned}$$

- So, the i-j-th element of $\hat{A}$ is in fact $\left\langle v_i | A | v_j \right\rangle$.

Next, we'll see that a change-of-basis matrix can be written entirely with inner-products:

- ***Proposition 2.22:***
  Let $|v_1\rangle, \ldots, |v_n\rangle$ (v-basis) and $|w_1\rangle, \ldots, |w_n\rangle$ (w-basis) be two bases and suppose that vector $|\psi\rangle$ has coordinates $\alpha_1, \ldots, \alpha_n$ in the v-basis and coordinates $\beta_1, \ldots, \beta_n$ in the w-basis. Then, the matrix with elements

  $$U_{ij} \;\; = \;\; \left\langle w_i | v_j \right\rangle$$

  converts coordinates from the v-basis to the w-basis. That is,

$$\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} \langle w_1|v_1\rangle & \cdots & \langle w_1|v_n\rangle \\ \vdots & \vdots & \vdots \\ \langle w_n|v_1\rangle & \cdots & \langle w_n|v_n\rangle \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Proof:
First note that the coordinates (coefficients) can be written as:

$$\begin{aligned} \alpha_i &= \langle v_i|\psi\rangle \\ \beta_i &= \langle w_i|\psi\rangle \end{aligned}$$

Next, in the matrix-vector product above, consider the k-th row times the vector:

$$\begin{aligned} &\langle w_k|v_1\rangle\,\alpha_1 + \ldots + \langle w_k|v_n\rangle\,\alpha_n \\ &= \ \langle w_k|v_1\rangle\ \langle v_1|\psi\rangle + \ldots + \langle w_k|v_n\rangle\,\langle v_n|\psi\rangle && \text{Sub for } \alpha_i \\ &= \ \left\langle w_k \,\middle|\, \sum_i |v_1\rangle\langle v_1| \,\middle|\, \psi \right\rangle && \text{Associativity} \\ &= \ \langle w_k\,|\,I\,|\,\psi\rangle && \text{Completeness} \\ &= \ \beta_k \end{aligned}$$

Which is indeed the k-th coordinate in the w-basis.

Whew! Both were detailed multi-step proofs that required careful attention to sums.

Why is this approach useful?

- First, it's practically useful in calculations, as we'll see below.
- Second, the entire derivation was done from an operator (not matrix) point of view and thus, it's useful when we only have operator descriptions (as opposed to matrices).

Let's apply what we've learned above:

- Example 1: Consider the operator that multiplies any 2D vector element by 3:

$$A\,|\psi\rangle \ = \ 3\,|\psi\rangle$$

    - The first thing to observe is that any matrix that multiplies basis vectors by 3 will also multiple *any* vector by three
    - This means, if we use the approach for deriving the matrix with basis vectors, it'll apply to all vectors (as desired).
    - Let's derive the matrix using $A_{ij} = \left\langle v_i|A|v_j \right\rangle$ using the standard basis $|0\rangle\,, |1\rangle$:

$$\begin{bmatrix} \langle 0|A|0\rangle & \langle 0|A|1\rangle \\ \langle 1|A|0\rangle & \langle 1|A|1\rangle \end{bmatrix} = \begin{bmatrix} 3\,\langle 0|0\rangle & 3\,\langle 0|1\rangle \\ 3\,\langle 0|1\rangle & 3\,\langle 1|1\rangle \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

    - Now check that it works for all vectors $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$:

$$A\,|\psi\rangle \ = \ \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \ = \ \begin{bmatrix} 3\alpha \\ 3\beta \end{bmatrix} \ = \ 3\,|\psi\rangle$$

    Which it does.

- Example 2: Let's write the coordinate conversion matrix going from the standard to the Hadamard basis:

- Recall the Hadamard basis:

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

  Note that, from this, we get

$$\begin{aligned} \langle +| &= \frac{1}{\sqrt{2}}\langle 0| + \frac{1}{\sqrt{2}}\langle 1| \\ \langle -| &= \frac{1}{\sqrt{2}}\langle 0| - \frac{1}{\sqrt{2}}\langle 1| \end{aligned}$$

- Next, the four inner products that will result in elements of the conversion matrix:

$$\begin{aligned} \langle +|0\rangle &= \left\langle \frac{1}{\sqrt{2}}\langle 0| + \frac{1}{\sqrt{2}}\langle 1| \,\middle|\, 0 \right\rangle &= \frac{1}{\sqrt{2}}\langle 0|0\rangle + \frac{1}{\sqrt{2}}\langle 1|0\rangle &= \frac{1}{\sqrt{2}} \\ \langle +|1\rangle &= \left\langle \frac{1}{\sqrt{2}}\langle 0| + \frac{1}{\sqrt{2}}\langle 1| \,\middle|\, 1 \right\rangle &= \frac{1}{\sqrt{2}}\langle 0|1\rangle + \frac{1}{\sqrt{2}}\langle 1|1\rangle &= \frac{1}{\sqrt{2}} \\ \langle -|0\rangle &= \left\langle \frac{1}{\sqrt{2}}\langle 0| - \frac{1}{\sqrt{2}}\langle 1| \,\middle|\, 0 \right\rangle &= \frac{1}{\sqrt{2}}\langle 0|0\rangle - \frac{1}{\sqrt{2}}\langle 1|0\rangle &= \frac{1}{\sqrt{2}} \\ \langle -|1\rangle &= \left\langle \frac{1}{\sqrt{2}}\langle 0| - \frac{1}{\sqrt{2}}\langle 1| \,\middle|\, 1 \right\rangle &= \frac{1}{\sqrt{2}}\langle 0|1\rangle - \frac{1}{\sqrt{2}}\langle 1|1\rangle &= -\frac{1}{\sqrt{2}} \end{aligned}$$

- Thus, the change-of-basis matrix is:

$$U = \begin{bmatrix} \langle +|0\rangle & \langle +|1\rangle \\ \langle -|0\rangle & \langle -|1\rangle \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

  Which matches what we saw earlier.

In-Class Exercise 17: Suppose A is an operator that "flips" the Hadamard basis vectors. That is, $A\,|+\rangle = |-\rangle$ and $A\,|-\rangle = |+\rangle$. Use the approach above to derive the matrix for A. Then, apply the matrix to $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$. Is the resulting vector orthogonal to $|\psi\rangle$?

Finally, let's reiterate the change-of-basis matrix for *matrices*:

- We saw this in the Review. Let's now write the same result in our new notation.
- Let $|v_1\rangle, \ldots, |v_n\rangle$ (v-basis) and $|w_1\rangle, \ldots, |w_n\rangle$ (w-basis) be two bases and suppose that we have an operator matrix A_v written in the v-basis. We want the same operator matrix in the w-basis, i.e., A_w.
- Let $U_{v\to w}$ be the change-of-basis that converts v-basis *vectors* to w-basis vectors.
- And $U_{w\to v}$ be the change-of-basis that converts w-basis *vectors* to v-basis vectors.
- Then

$$A_w = U_{v\to w}\, A_v\, U_{w\to v}$$

To highlight this, we'll write this up as a proposition.

- ***Proposition 2.23:***
  Let $|v_1\rangle, \ldots, |v_n\rangle$ (v-basis) and $|w_1\rangle, \ldots, |w_n\rangle$ (w-basis) be two bases and suppose that we have an operator matrix A_v written in the v-basis. Then the same operator in the w-basis is:

$$A_w \quad = \quad U_{v \to w} \, A_v \, U_{w \to v}$$

  where $U_{v \to w}$ and $U_{w \to v}$ are the change-of-basis matrices going from one to the other basis.

  Proof:
  See earlier Review.

- Let's apply this in an example:
  - Suppose we are in the $|+\rangle, |-\rangle$ basis (H-basis) and we have the operator

$$X \quad = \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

    in the same basis.
  - As an example of applying it:

$$X\,|+\rangle \quad = \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad = \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad = \quad |-\rangle$$

  - We'd like to have the operator in standard coordinates.
  - We've already derived the conversion matrix from standard to H-basis:

$$U_{S \to H} \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

    It's easy to show that $U_{H \to S}$ is the same: $U_{S \to H}$.
  - From this, we can convert X to the standard basis. Let's call the result X_S

$$X_S \quad = \quad U_{S \to H} \, X \, U_{H \to S} \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \quad = \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

  - Then, applying it in the standard basis, we get the same result as before $X\,|+\rangle = |-\rangle$ but now in standard coordinates:

$$X_S\,|+\rangle \quad = \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad = \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \quad = \quad |-\rangle$$

2.17 Orthogonal subspaces

Let's return to the idea of a vector space:

- Suppose $|v_1\rangle, \ldots, |v_n\rangle$ is a basis for n-component complex vectors.

- The set

$$\begin{aligned} V &= \text{span}(|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle) \\ &= \{\alpha_1 |v_1\rangle + \alpha_2 |v_2\rangle + \ldots + \alpha_k |v_n\rangle : \alpha_i \in \mathbb{C}\} \end{aligned}$$

  is a *vector space*, meaning that all linear combinations of any subset of vectors in V is in V.

- Another way of saying this: V is closed under linear combinations.

- Now consider just the span of first two of these vectors:

$$\begin{aligned} W &= \text{span}(|v_1\rangle, |v_2\rangle) \\ &= \{\alpha_1 |v_1\rangle + \alpha_2 |v_2\rangle : \alpha_i \in \mathbb{C}\} \end{aligned}$$

  Is this a vector space? Let's test:
    - Let's, for example, pick three vectors in W and ask if the linear combination is in W:

$$\begin{aligned} |x_1\rangle &= a_1 |v_1\rangle + a_2 |v_2\rangle \\ |x_2\rangle &= b_1 |v_1\rangle + b_2 |v_2\rangle \\ |x_3\rangle &= c_1 |v_1\rangle + c_2 |v_2\rangle \end{aligned}$$

    - A linear combination of these three is:

$$|z\rangle = \beta_1 |x_1\rangle + \beta_2 |x_2\rangle + \beta_3 |x_3\rangle$$

    - We want to ask: is $|z\rangle \in W$?
    - The exercise below shows that it does.

- Observe:
    - Every vector in W is in V.
    - Not every vector in V is in W.

  Thus, we say W is a *subspace* of V.

- Now consider the set W':

$$W' = \text{span}(|v_3\rangle, \ldots, |v_n\rangle)$$

  One can see that W' is also a subspace of V.

- Now, there's a *complementary* relationship between W and W':
    - Suppose $|x\rangle \in W$ and $|y\rangle \in W\prime$.
    - Then we can use the basis vectors to express them:

$$\begin{aligned} |x\rangle &= a_1 |v_1\rangle + a_2 |v_2\rangle \\ |y\rangle &= a_3 |v_3\rangle + \ldots + a_n |v_n\rangle \end{aligned}$$

    - Then

$$\begin{aligned} \langle x|y\rangle &= \left\langle a_1^* \langle v_1| + a_2^* \langle v_2| \,\middle|\, a_3 |v_3\rangle + \ldots + a_n |v_n\rangle \right\rangle \\ &= 0 \end{aligned}$$

      because the $|v_i\rangle$'s are orthogonal.
    - Thus, *every* vector in W is orthogonal to *every* vector in W'.

- Two subspaces that satisfy this property are said to be orthogonal complements:
    - We use the notation $W^\perp$ = orthogonal complement subspace to W.

- Here, $W^{\perp} = W'$

- Notice also that

$$W \cup W' \quad = \quad V$$

That is, if we pull into one set every vector in each of W, W' we'll get all the vectors in V.

- This notion can extend to multiple subspaces:
  - For example, suppose $W_3 = \text{span}(|v_3\rangle), W_4 = \text{span}(|v_4\rangle), \ldots, W_n = \text{span}(|v_n\rangle)$
  - Then any two of these are orthogonal, and all of them are orthogonal to $W = \text{span}(|v_1\rangle, |v_2\rangle)$.
  - And

$$W \cup W_3 \cup \ldots \cup W_n \quad = \quad V$$

In-Class Exercise 18: Show that $|z\rangle \in W$.

Projectors for subspaces:

- Recall the definitions of V and W from above:

$$\begin{aligned} V &= \text{span}(|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle) \\ W &= \text{span}(|v_1\rangle, |v_2\rangle) \end{aligned}$$

- Let's pick a generic vector in V from above:

$$|u\rangle \quad = \quad a_1 |v_1\rangle + a_2 |v_2\rangle + \ldots + a_n |v_n\rangle$$

- Define the matrix

$$\begin{aligned} P_{v_1,v_2} &= P_{v_1} + P_{v_2} && \text{Projectors for each of } |v_1\rangle, |v_2\rangle \\ &= |v_1\rangle\langle v_1| + |v_2\rangle\langle v_2| && \text{Each projector is an outer-product} \end{aligned}$$

- Let's apply this to the vector $|u\rangle$:

$$\begin{aligned} P_{v_1,v_2} |u\rangle &= (|v_1\rangle\langle v_1| + |v_2\rangle\langle v_2|) \; (a_1 |v_1\rangle + a_2 |v_2\rangle + \ldots + a_n |v_n\rangle) \\ &= (\, a_1 |v_1\rangle \; \langle \mathbf{v_1}|\mathbf{v_1}\rangle + a_2 |v_1\rangle \; \langle v_1|v_2\rangle + \ldots + a_n |v_1\rangle \; \langle v_1|v_n\rangle \,) \\ &\quad + \; (\, a_1 |v_1\rangle \; \langle v_2|v_1\rangle + a_2 |v_1\rangle \; \langle \mathbf{v_2}|\mathbf{v_2}\rangle + \ldots + a_n |v_1\rangle \; \langle v_2|v_n\rangle \,) \\ &= a_1 |v_1\rangle + a_2 |v_2\rangle \end{aligned}$$

(Only the two bolded inner products are non-zero.)
  - The result is a vector in W.
  - This is called the projection of $|u\rangle$ onto the subspace W.

- In this way, if $S \subset V$ is any subspace of V, with basis $|s_1\rangle, |s_2\rangle, \ldots, |s_k\rangle$, then the projector for this subspace is:

$$P_S \quad = \quad |s_1\rangle\langle s_2| + \ldots + |s_k\rangle\langle s_k|$$

This means that if $|u\rangle \in V$ is any vector in V, the vector $P_S |u\rangle \in S$.

- *Important:* the projected vectors are not necessarily unit-length.

- Example:
  - Suppose $V = \text{span}(|0\rangle, |1\rangle)$.
  - Let $S = \text{span}(|0\rangle)$ be a subspace.

- Then the projector is $P_S = |0\rangle\langle 0|$.
- Let

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

  be a vector in V.
- Then

$$\begin{aligned}
P_S\,|u\rangle &= |0\rangle\langle 0| \left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \right) && \text{Apply projector} \\
&= \frac{1}{\sqrt{2}}|0\rangle\ \langle 0|0\rangle + \frac{1}{\sqrt{2}}|0\rangle\ \langle 0|1\rangle && \text{Projector distributes} \\
&= \frac{1}{\sqrt{2}}|0\rangle && \text{Associativity groups inner-products}
\end{aligned}$$

  (Recall: $|0\rangle\langle 1| = 0$).
- Then the length of this result is:

$$|P_S\,|u\rangle| \;=\; \left|\frac{1}{\sqrt{2}}|0\rangle\right| \;=\; \frac{1}{\sqrt{2}}$$

- Why are projectors important?
  - Projectors are how the critical action of *measurement* occurs in quantum computing (and mechanics).
  - A stronger understanding of this part of the theory makes it less mysterious than it first appears to be.

- Next, let's work through an example:
  - Recall

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

  - Let's apply the projector for span($|+\rangle$) to the vector $|0\rangle$.
  - First, the projector:

$$P_+ \;=\; \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \;=\; \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

  - Now apply to $|0\rangle$:

$$P_+\,|0\rangle \;=\; \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \;=\; \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \;=\; \frac{1}{\sqrt{2}}|+\rangle$$

In-Class Exercise 19: Write down the projector P_- for the subspace span($|-\rangle$).

1. Show that $P_-\,|0\rangle = \frac{1}{\sqrt{2}}|-\rangle$
2. What is the result of applying P_- to $|+\rangle$?

The special case of 2D:

- Consider the 2D vector $|0\rangle$.

- Let's ask: which unit-length 2D vectors $\left|w^{\perp}\right\rangle$ are orthogonal to $|w\rangle = |0\rangle$?
    - Write

$$\left|w^{\perp}\right\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$$

    in terms of the standard basis.
    - For orthogonality,

$$\begin{array}{lrcl} & \left\langle 0 \,\middle|\, w^{\perp} \right\rangle & = & 0 \\ \Rightarrow & \left\langle 0 \,\middle|\, \alpha\,|0\rangle + \beta\,|1\rangle \right\rangle & = & 0 \\ \Rightarrow & \left\langle 0 \,\middle|\, \alpha\,|0\rangle + \beta\,|1\rangle \right\rangle & = & 0 \\ \Rightarrow & \alpha = 0 & & \\ \Rightarrow & \left|w^{\perp}\right\rangle = \beta\,|1\rangle & & \end{array}$$

    - For unit-length, we must have

$$\begin{array}{lrcl} & \left\langle w^{\perp} \middle| w^{\perp} \right\rangle & = & 1 \\ \Rightarrow & \left\langle \beta^{*}\,\langle 1| \,\middle|\, \beta\,|1\rangle \right\rangle & = & 1 \\ \Rightarrow & \beta^{*}\beta\,\langle 1|1\rangle & = & 1 \\ \Rightarrow & |\beta| \;\; = \;\; 1 & & \end{array}$$

    - Thus, β can be any complex number of unit magnitude.
    - Example: $\beta = 1$
        - ▷ $\left|w^{\perp}\right\rangle = |1\rangle$

- Notice that all the vectors orthogonal to $|w\rangle = |0\rangle$ are along $\left|w^{\perp}\right\rangle = |1\rangle$.

- In some sense, there's only one vector along which lie the orthogonal vectors to $|w\rangle = |0\rangle$.

- This is not true for 3D and higher:
    - For example, with real vectors, consider $|w\rangle = (1, 0, 0)$ in 3D.
    - Any vector in the y-z plane is perpendicular (orthogonal) to $|w\rangle$.
    - And there are an infinite number of them in an infinite number of directions.

- When working with single qubits, this observation will be useful, so let's generalize and observe the following:
    - Suppose $|v\rangle$ and $\left|v^{\perp}\right\rangle$ are two basis vectors for 2D.
    - Let $|w\rangle = a\,|v\rangle + b\,\left|v^{\perp}\right\rangle$
    - Then $\left|w^{\perp}\right\rangle = b^{*}\,|v\rangle - a^{*}\,\left|v^{\perp}\right\rangle$ is orthogonal because

$$\begin{array}{rcl} \left\langle w \middle| w^{\perp} \right\rangle & = & \left\langle a^{*}\,\langle v| + b^{*}\,\left\langle v^{\perp}\right| \,\middle|\, b^{*}\,|v\rangle - a^{*}\,\left|v^{\perp}\right\rangle \right\rangle \\ & = & a^{*}b^{*}\,\langle v|v\rangle - b^{*}a^{*}\,\left\langle v^{\perp}\middle|v^{\perp}\right\rangle \\ & = & 0 \end{array}$$

In-Class Exercise 20: Suppose $|w\rangle = \frac{\sqrt{3}}{2}|0\rangle - \frac{1}{2}|1\rangle$.

1. Show that $\left|w^{\perp}\right\rangle = \frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle$ is orthogonal and of unit-length.
2. Express $|+\rangle$ in this basis by computing the coefficients $\langle w|+\rangle$ and $\left\langle w^{\perp}|+\right\rangle$.
3. Check your calculations by showing $|\langle w|+\rangle|^2 + \left|\left\langle w^{\perp}|+\right\rangle\right|^2 = 1$.

In-Class Exercise 21:

1. Write down the two projectors P_0, P_1 for $|0\rangle, |1\rangle$, respectively. Then compute the matrix product $P_0 P_1$.
2. Write down the two projectors P_+, P_- for $|+\rangle, |-\rangle$, respectively. Then compute the matrix product $P_+ P_-$.

Explain the resulting product matrices.

2.18 A mathematical aside

For the theoretically inclined, we'll point out a few things, none of which are essential to the course but which might stimulate further curiosity:

- All the theory laid out in this module was done so with complex vectors and matrices.
- The theory of vector spaces and linear operators is much more general, and applies to any set of mathematical objects, with addition and scalar multiplication, that satisfies the required vector space properties: commutativity, associativity, additive identity and inverse, multiplicative identity, and distributive properties.
- Thus, while we defined an inner-product using complex-vector inner-product, a more abstract approach starts with an abstract definition of what an inner-product should satisfy.
- Similarly, a more general form of adjoint operator (not necessarily a matrix) is defined using the last relation: the adjoint of an operator A is an operator $A^\dagger$ that satisfies $\langle w|Ax\rangle = \left\langle A^\dagger w|x\right\rangle$ for every pair of vectors $|x\rangle, |w\rangle$.
- The advantage of abstraction is wider applicability, and the satisfaction of seeing the common features work for both quantum computing and mechanics.
- The advantage of the vector approach, instead, is that it's generally easier to work with and understand.
- What we recommend: start from the concrete to develop skill and concepts, and then move on to the abstract as necessary.

Some theoretical results of interest:

- There is a broader category of operator called *normal*: A is normal if $AA^\dagger = A^\dagger A$:
  - Any Hermitian or unitary operator is also normal.
  - Normal operators have an orthonormal basis of eigenvectors.
- Many results have an if-and-only-if character, for example:
  - If an operator A satisfies *any* of the properties of unitary operators described above, it must be unitary.
  - Any operator whose eigenvectors form a basis must be normal.

(See Axler's book for proofs)

2.19 Summary

This was one long module, packed with new concepts and notation all at once.

Understandably, this will take time to digest and make your own.

The most reliable way is: repetition. The third time you return to something and read/write/solve, you will be a lot more at ease.

The notion that one can compute using the elements of complex linear algebra seems alien at this time:

- After all, with regular linear algebra, vectors and matrices are things you "do math with".
- Writing code for matrices/vectors involves the usual loops, conditionals, arrays.
- But in quantum computing, there are no such things as loops, conditionals and arrays.
- Instead, one writes "programs" with vectors and unitary matrices, with the occasional use of Hermitians.

At this time, you might be curious about:

- What exactly do complex vectors represent?
- Why did we use such gnarly numbers like $\frac{1}{\sqrt{2}}$?
- How do complex vectors feature in algorithms that, for example, factor integers (Shor's algorithm that breaks conventional crypto).
- Why are complex vectors the fundamental building blocks? Couldn't a simpler non-complex theory work?

We'll answer these in due time.

The best thing to do right now is to re-read this module so as to get more comfortable with Dirac notation.