Module 3: The Single Qubit

Module-3 solved problems

Module objectives

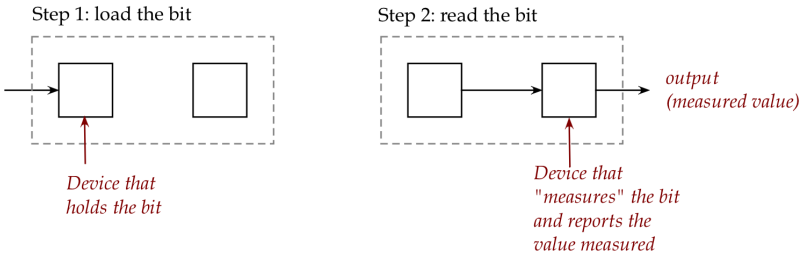
In this module, we'll look at a single qubit:

- What exactly is a qubit?
- How does one "work" with a qubit?
- How does all the linear algebra come in?
- Applications: polarization, hack-proof communication

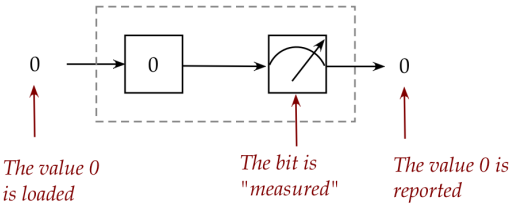
3.1 Classical bit vs. quantum qubit

Let's start by examining the behavior of a *classical bit*:

- By *classical bit* we mean a regular *binary* bit as seen in the circuitry of standard calculator or computer.
- A classical bit takes on only one of two values: 0 or 1.
- Now suppose we have a device that can hold a classical bit, and we wish to observe the value in it.
- We'll break down this process into two steps: *loading* the bit, and *reading* the bit:



- We'll now do an example and coalesce the two steps into a single diagram:

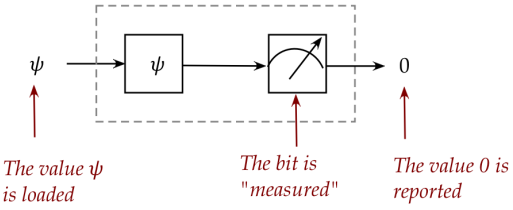


  Let's point out:

  - In practice the holding or storage device on the left is called a *flip-flop*.
  - A flip-flop is a two-state device that can be thought of as:
    - Low-voltage ⇒ 0
    - High-voltage ⇒ 1
  - The act of observing or measuring involves "spilling" of the electricity from the device to determine whether it's high or low (1 or 0).
  - We're using an old-fashioned dial to label a measuring device.

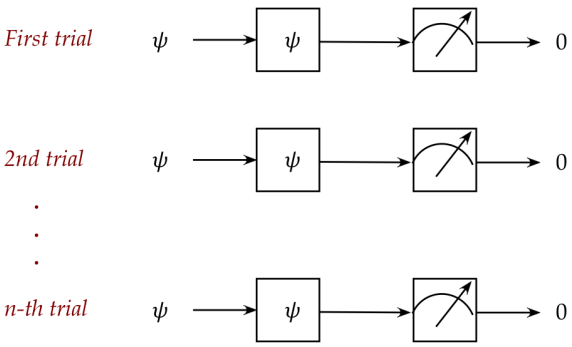
- Notice that we've used the term *measure* instead of *observe* or *read*. This is the terminology used in quantum systems.

- Next, we're going to repeat this load-and-measure for several times with the following set up:



  - Someone, without telling us, load the bit with a value unknown to us.
  - We'll call this unknown value ψ.
  - We then measure this value and report it.
  - Let's say we observe 0.

- We now repeat the whole experiment several times:



  We can assert *three* things with high confidence:

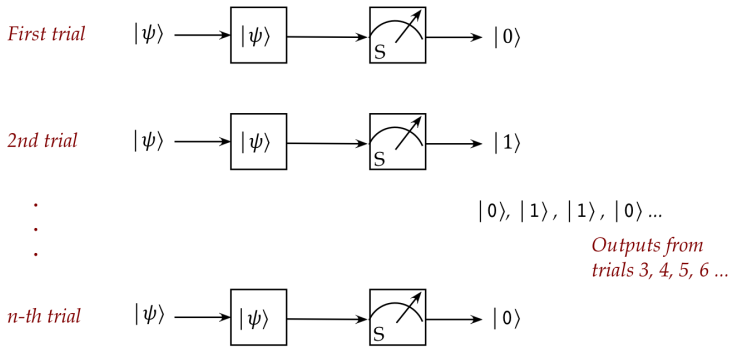
  - The unknown value ψ is in fact 0.
  - Any number of repetitions will yield exactly the same result.

- An alternative third-part measuring device will also yield exactly the same results (for this particular ψ).

- Surprisingly, none of these assertions can be made for a *qubit*.

Let's now do the same experiment with a *qubit*:

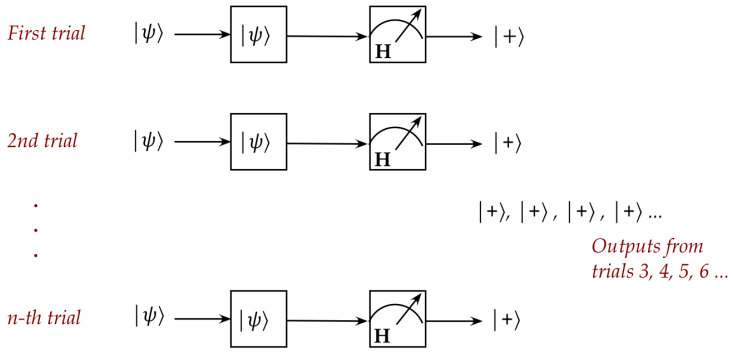
- Qubit = Quantum bit
- Here's an example of some trials for a quantum bit:



  Note:

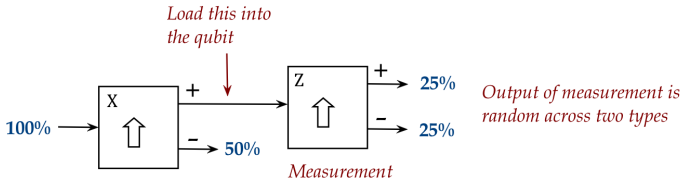
  - We now cannot assert a definitive value for the unknown value $|\psi\rangle$.
  - Multiple repetitions do *not* yield exactly the same value.
  - In fact, statistical analysis over many trials shows that the pattern of output is completely random between $|0\rangle$ and $|1\rangle$
  - The measuring device performs an "S" measurement.
    (Whatever that means. We'll clarify later.)

- What if we used a different measuring device?

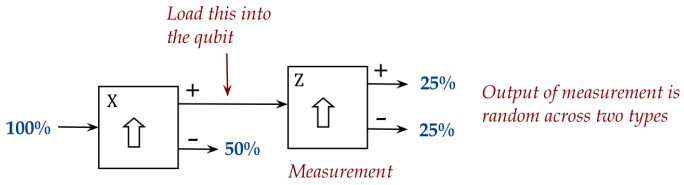


  - Here, we have replaced an "S" device with an "H" measurement device.
  - This time the output is entirely predictable!

- Does this remind you of anything from Module 1?
  - Here's one Stern-Gerlach experiment:



- The first state is X-aligned, the second is Z-aligned.
- Recall: the second-stage output is random betweeen two types of output.
- A related SG experiment:



Here, the output is consistently the same each time.

- In fact, a Stern-Gerlach apparatus *is* a qubit:
  - There are many such *two-level* quantum systems that can serve as qubit technology.
  - The SG apparatus is not at all practical.
    - ▷ Other technologies are much more practical and successful.
  -

What's useful: the same theory works for all two-level systems (devices).

So, what exactly is a qubit?

- Answer: *A qubit is a device that holds a unit-length 2D complex vector as its value.*
- Consider for example, the complex vector

$$|\psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Is $|\psi\rangle$ a qubit value?
  - Yes!
  - And it's a qubit value for any choice of the complex numbers α, β.

- Example qubit values:

$$|\psi\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |\psi\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad |\psi\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{bmatrix},$$

- Note that a qubit vector can be expressed in any desired basis:
  - Example: $|\psi\rangle = (\alpha, \beta)$ in the standard basis:

$$|\psi\rangle \;=\; \begin{bmatrix}\alpha\\ \beta\end{bmatrix} \;=\; \alpha\begin{bmatrix}1\\ 0\end{bmatrix} + \beta\begin{bmatrix}0\\ 1\end{bmatrix} \;=\; \alpha\,|0\rangle + \beta\,|1\rangle$$

  - Example: $|\psi\rangle = |0\rangle$ in the Hadamard basis:

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

  This is important, as we'll soon see.

- *Terminology*:
  - The vector $|\psi\rangle$ that we've called the qubit's value is more often called the *state* of the qubit.
  - $|\psi\rangle$ is also called the *quantum state vector*, and occasionally the *wavefunction*.

- We can already see one implication:
  - The set of possible value for a single qubit is (uncountably) infinite.

- Let's address the obvious question: do we impose the set of qubit values *and then* build hardware to support such qubits?
  - No! This is what qubits are, and the theory follows from their natural behavior.
  - All two-level quantum systems are modeled by 2D complex vectors
    - ▷ This might be unsettling, but this is how nature *is*.
  - There's more that's perplexing, as we'll see.

- We still need to address two issues:
  - Where does the randomness come in?
  - Why do we get different results with different types of measurements?

3.2 Nature's indeterminism and the puzzle of measurement

Let's start with the measuring device:

- Think of a qubit measuring device as a 2D *orthonormal basis*:
  - Example 1: Thus, for example, the basis

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

    is one measuring device.
  - Example 2: The basis

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

is another measuring device.

- In general, any orthonormal 2D basis

$$\begin{aligned} |v\rangle &= a|0\rangle + b|1\rangle \\ |v^{\perp}\rangle &= b^{*}|0\rangle - a^{*}|1\rangle \end{aligned}$$

is a single qubit measuring device.

- Now, any qubit state $|\psi\rangle$ can be expressed in terms of a measurement device's basis:
  - For example, consider the basis $|0\rangle, |1\rangle$, and

$$\psi = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Then,

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

  - Another example: consider the basis $|+\rangle, |-\rangle$, and the vector $\psi = |0\rangle$:

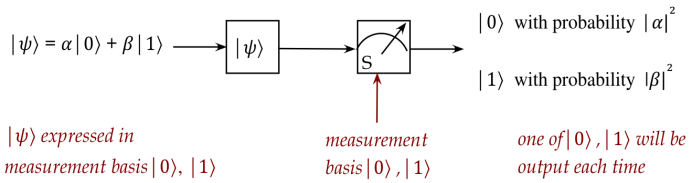
$$\psi = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle$$

Now let's explain randomness and measurement outcome:

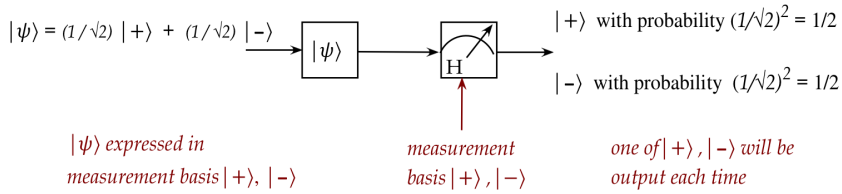
- The three aspects to a single measurement:
  1. Every measurement of a qubit value is associated with a *measurement basis*.
     ▷ Example: $|0\rangle, |1\rangle$
  2. A particular measurement outcome will be a random selection from *one* of the basis vectors.
     ▷ Thus, either $|0\rangle$ or $|1\rangle$
  3. Which basis vector becomes the random outcome is based on these probabilities:
     - Express the qubit's state in the measurement basis.
       ▷ Example: $\psi = \langle 0|\psi\rangle|0\rangle + \langle 1|\psi\rangle|1\rangle$
     - This means each basis vector will have a coefficient in the expression:
       ▷ Example: $\psi = \alpha|0\rangle + \beta|1\rangle$
     - The probability of seeing a particular basis vector is squared magnitude of the coefficient.

$$\begin{aligned} \text{Probability outcome is } |0\rangle &= |\alpha|^2 \\ \text{Probability outcome is } |1\rangle &= |\beta|^2 \end{aligned}$$

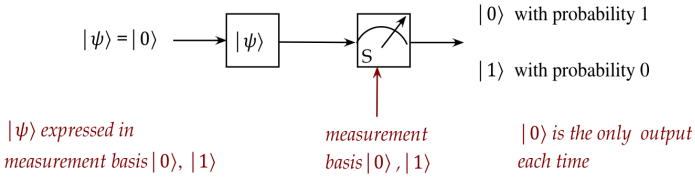
- Example: Suppose we use the measurement basis $|0\rangle, |1\rangle$ and $\psi = \alpha|0\rangle + \beta|1\rangle$:



- Example: suppose we use the measurement basis $|+\rangle, |-\rangle$ and $\psi = |0\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle$:



- Example: Suppose we use the measurement basis $|0\rangle, |1\rangle$ and $\psi = |0\rangle$



- There is a more general way of describing measurement that we'll see later.

In-Class Exercise 1: Suppose the qubit state is $|\psi\rangle = |+\rangle$ and the measurement basis is: $|w\rangle = \frac{\sqrt{3}}{2}|0\rangle - \frac{1}{2}|1\rangle$, and $|w^{\perp}\rangle = \frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle$. What are the possible output vectors, and with what probabilities do they occur? Start by expressing $|+\rangle$ in the $|w\rangle, |w^{\perp}\rangle$ basis by computing the coefficients $\langle w|+\rangle, \langle w^{\perp}|+\rangle$.

Let's say a bit more about measurement:

- First, why is the output random?
  - One might wonder: are measurement devices deliberately *constructed* to perform random selection?
  - And if so, why would anyone want to do that?

  The answer to the first: No!

  ▷ This is just how nature *is*.

- The *only* time the measurement is not random is when the outcome probability for *one* of the basis vectors is 1.

  ▷ The case when one basis vector occurs with certainty.

- In general, any measurement of any quantum state (in any quantum device) is going to involve random output.
- That is, one *cannot* construct a measurement device that avoids the randomness.
- Why do we bother with measurement when the outcome is uncertain? Why not just *observe* the qubit value?
  - This is another unsettling issue: *the only way to know a qubit's state is to measure it*.
- To summarize so far:
  - To examine a qubit's state, one must measure.
  - A measurement outcome is random, and is one of the basis vectors.
- There's yet another issue:
  - Measurement *changes* the qubit's state into one of the basis vectors.

  We'll say more about this in the next section.

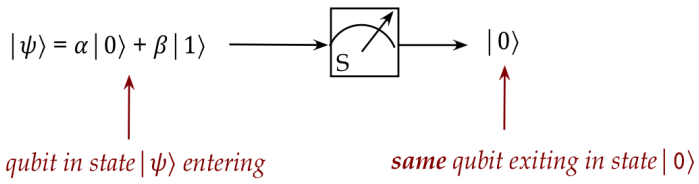
3.3 Stationary vs flying qubits

We will distinguish between are two types of hardware configurations: stationary and flying qubits.

Both are identical in terms of theory, and so we will (as do all books) use only the latter because it's more convenient.

Think of a flying qubit as an atom or photon in motion:

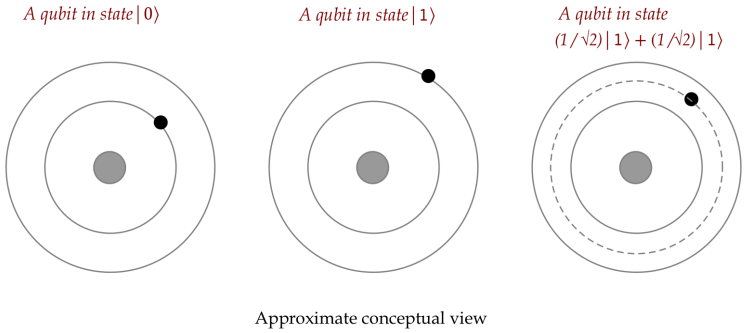
- Consider a flying qubit $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ entering a standard-basis measuring device.
- Suppose measurement resulted in $|0\rangle$.



- A qubit enters from the left (in this picture) in the $\alpha\,|0\rangle + \beta\,|1\rangle$ state.
- After measurement, it's new state is $|0\rangle$.
- That is, there's only one qubit in play here.
- The changed qubit can now be directed towards other devices, perhaps to a measurement in another basis.

Think of a stationary qubit as the state of an atom's outermost electron:

- For example, here are three qubits in three different states.



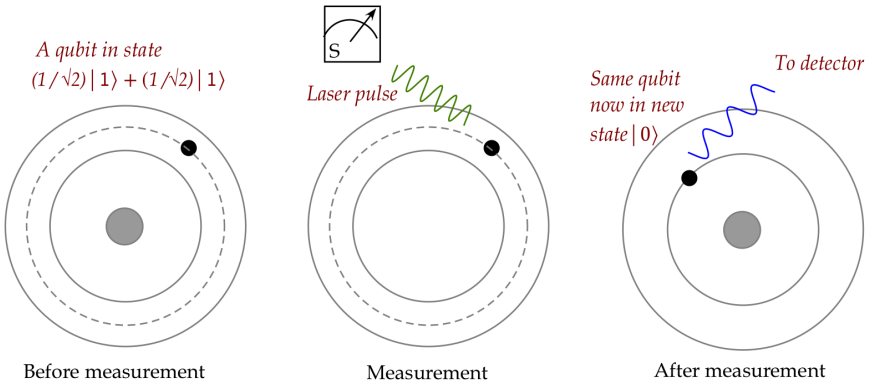
Approximate conceptual view

  Here:

  - Here, $|0\rangle$ is typically the lowest energy or *ground state* of the electron.
  - $|1\rangle$ is the excited state.
  - The state $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ is depicted above, on the right.

- Note:
  - The above depiction is only an approximation for a simplified high-level conceptual view.
  - The actual location of an electron, and this "planetary motion" view is actually not correct.
  - It's more accurate to say the electron is in a quantum state, to which various measurements can be made, including location, momentum, and other properties.

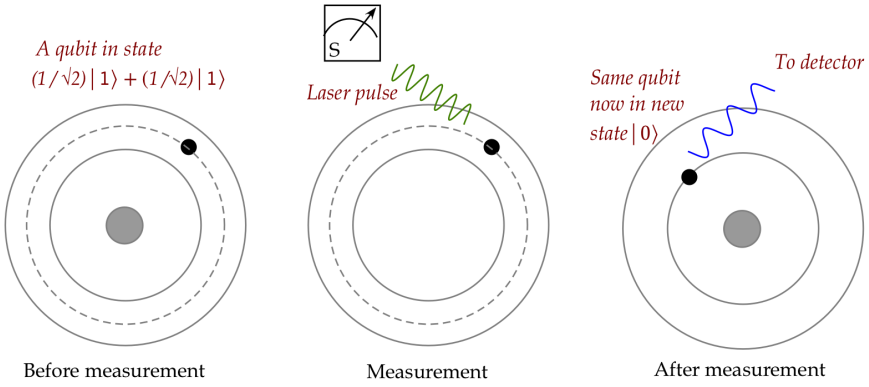
- How does a measurement work for a stationary qubit?



Before measurement | Measurement | After measurement

  - Here, the measurement device is brought to the qubit and activated.
  - This changes the state of the qubit.
  - There is some additional output that tells a detector whether the post-measurement state is $|0\rangle$ or $|1\rangle$
  - The measurement device is then deactivated.

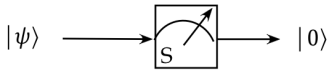
- Thus, using our conceptual boxes, we should really have drawn something like this:



- The stationary kind of qubit is the more common technology, and the one considered the most likely for quantum computing.

The convention in drawing:

- The theory is exactly the same for both types of qubit configurations.
- The convention in books is to use the flying qubit diagram,



Measurement and initialization:

- Suppose we want to initalize a qubit into state $|0\rangle$.
- At the start, suppose a qubit is in an unknown state $|\psi\rangle$.
- We perform a measurement in the $|0\rangle, |1\rangle$ basis.
  - If this results in state $|0\rangle$, we're done.
  - We know that the state is now $|0\rangle$.
- If not, we can change the state (next section) and apply measurement again, repeating this whole process until a measurement outcome is $|0\rangle$.

3.4 What can you do with a qubit?

If all we could do is measure, we would not get far with quantum computing.

Luckily, nature lets us *modify* a qubit's state without measurement.

However, there's one limitation:

▷ The only modification allowed is a *unitary operation*.

Let's look at an example:

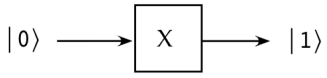
- Consider the unitary operator

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

- Let's ask: what would we get if we applied this to a qubit in state $|0\rangle$?

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

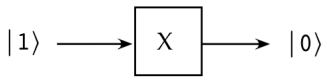
- We will diagram this as:



- Note that

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

  which we can depict as



- Observe that applying X twice results in

$$XX = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

  That is, the original state is returned.

In-Class Exercise 2: What does the X gate do to the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$?

Another example:

- Consider the Hadamard matrix

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

This is unitary as can be seen:

$$H^\dagger H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = I$$

Thus, it can be used to modify a qubit's state.

- An aside:
  - We've seen the H matrix in many roles.
  - For example, it was a change-of-basis matrix.
  - And because it's Hermitian, it served as an example then.
  - The Hadamard matrix is one of the most important tools in quantum computing, as we will see later.

- Now let's see what it does to $|0\rangle, |1\rangle$:

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

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

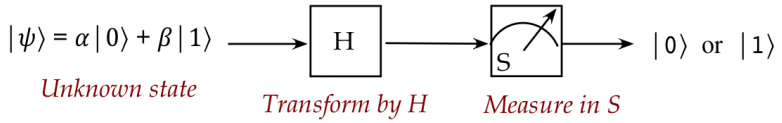
- Thus, the H operator
  - transforms the state $|0\rangle$ into the state $|+\rangle$;
  - transforms the state $|1\rangle$ into the state $|-\rangle$;

- What does it do to the general standard-basis vector $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$?

$$H|\psi\rangle = \frac{\alpha+\beta}{\sqrt{2}}|0\rangle + \frac{\alpha-\beta}{\sqrt{2}}|1\rangle$$

In-Class Exercise 3: Derive the above result that applies H to $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Also show that $H|\psi\rangle = \alpha|+\rangle + \beta|-\rangle$.

Now let's build our first circuit:



- We have an unknown state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ entering from the left.
- Applying the H transform results in

$$H|\psi\rangle = \frac{\alpha+\beta}{\sqrt{2}}|0\rangle + \frac{\alpha-\beta}{\sqrt{2}}|1\rangle$$

- Measuring in the standard basis yields

$$|0\rangle \text{with probability} \quad \left(\frac{\alpha+\beta}{\sqrt{2}}\right)^2$$
$$|1\rangle \text{with probability} \quad \left(\frac{\alpha-\beta}{\sqrt{2}}\right)^2$$

- Suppose we perform this experiment several times and obtain $|0\rangle$ *much more often* than $|0\rangle$.
  - This means the probability of obtaining $|0\rangle$ is much higher than the probability of getting $|0\rangle$.
  - That is,

$$\left(\frac{\alpha+\beta}{\sqrt{2}}\right)^2 \approx 1$$
$$\left(\frac{\alpha-\beta}{\sqrt{2}}\right)^2 \approx 0$$

  - From which we conclude $\alpha \approx \beta$.
- Thus, we've learned something about the unknown state through a simple transformation.
- Of course, we could measure the unknown state repeatedly and estimate in that manner too, but if we skew the probabilities as above, we can arrive at the same result sooner.

Another example:

- Consider the strangely-named matrix

$$\sqrt{X} = \begin{bmatrix} \frac{1+i}{2} & \frac{1-i}{2} \\ \frac{1-i}{2} & \frac{1+i}{2} \end{bmatrix}$$

- Is it unitary?

$$\left(\sqrt{X}\right)^\dagger = \begin{bmatrix} \frac{1-i}{2} & \frac{1+i}{2} \\ \frac{1+i}{2} & \frac{1-i}{2} \end{bmatrix}$$

  Then it's easy to calculate

$$\left(\sqrt{X}\right)^\dagger \sqrt{X} \;=\; I$$

- Observe that

$$\sqrt{X}\,\sqrt{X} \;=\; \begin{bmatrix} \frac{1+i}{2} & \frac{1-i}{2} \\ \frac{1-i}{2} & \frac{1+i}{2} \end{bmatrix} \begin{bmatrix} \frac{1+i}{2} & \frac{1-i}{2} \\ \frac{1-i}{2} & \frac{1+i}{2} \end{bmatrix} \;=\; \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \;=\; X$$

  Thus, the effect of applying $\sqrt{X}$ twice is the same as applying X.

- While X has a natural classical analog, the NOT Boolean gate, $\sqrt{X}$ has no classical analog.

Some terminology:

- A unitary operator is called a *quantum gate*, or just *gate*.
- *Superposition*:
  - A vector such as $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ or $\alpha|0\rangle + \beta|1\rangle$ that is a linear combination of basis vectors is sometimes called a *superposition*.
- About superpositions:
  - Any superposition requires specifying the basis used in the linear combination.
  - A superposition is *not* a simultaneous occurence of the basis vectors involved.
  - It's just a vector that happens to be expressible as a linear combination of basis vectors.
  - Clearly, a particular vector can be a superposition in one basis (e.g., $|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$) but not so in another basis (e.g, $|+\rangle$ in the $|+\rangle, |-\rangle$ basis).

In-Class Exercise 4: Consider the two circuits below, each given the same input.

1. Write down the possible states of the outputs.
2. Calculate the probabilities associated with each output state.

3.5 The promise and challenge of quantum computing viewed through a qubit

What we've seen so far gives us a hint of the promise and challenge in quantum computing.

First, the promise:

- Compared to a regular bit, a qubit can be in one of an infinite number of states.

- A unitary operator can be applied to any qubit, not just the two standard-basis states $|0\rangle, |1\rangle$.
  - It's natural to equate $|0\rangle, |1\rangle$ with regular bits 0 and 1.
  - We will later see how regular logic circuits can be converted to quantum equivalents, showing that a quantum circuit can do what a logic circuit can do.
  - But quantum states are many more!

- Consider the state

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

  One can think of this as combining the two states $|0\rangle, |1\rangle$ into a single state, with equal parts $|0\rangle$ and $|1\rangle$.

- We will later see that it's possible to create a 2-qubit state like

$$\frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$$

  That is, a combination of all four 2-qubit basis states.

- In fact, we'll see that with n qubits, one can create a state that combines all 2^n basis states!

- This suggests that we can compute simultaneously with all of them.
  ▷ In fact, this is often the starting point for many quantum algorithms.

- There is a *multi-qubit* property called *entanglement* that's both mystifying and powerful, which we will exploit.

- In combination, all these properties show that quantum computers can:
  - perform some computations that no classical computer can;
  - perform some computations provably faster than a classical computer (so-called quantum supremacy).

But ... there are many challenges in exploiting this power:

- Measurement has probabilistic outcomes:
  - This may mean repeating a computation until one has a result with high probability.

- Measurement destroys quantum state, replacing it with one of the measurement basis vectors:
  - This means one typically uses measurement only after a sequence of unitary operations.

- Although, as we will see later, unitary operations can simulate any classical logic circuit, there are added gates needed, which impacts efficiency.

- These are just the theoretical "in principle" challenges.

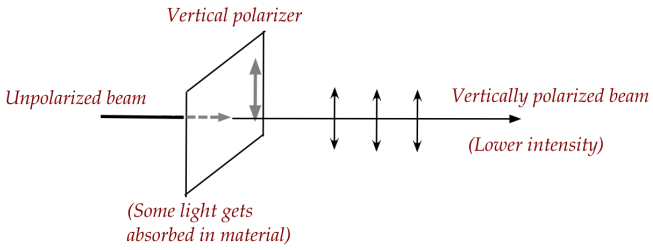
- The hardware challenges are even more significant:
  - Quantum states don't last long and thus unitary operations must be performed quickly.

    - The slightest bit of noise can derail any computation.
        - ▷ This is called *decoherence*.
    - It's difficult to work with multiple qubits.
    - Currently, hardware needed for cooling is expensive.

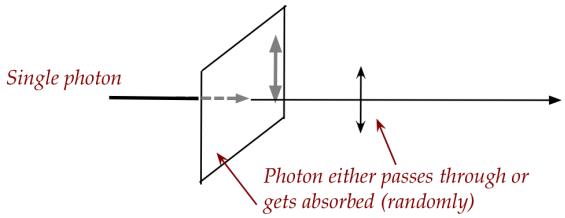
- Nonetheless, as of 2022, leading vendors are claim to have demonstrated computations on over a dozen error-corrected qubits.

3.6 Revisiting the light polarization experiment

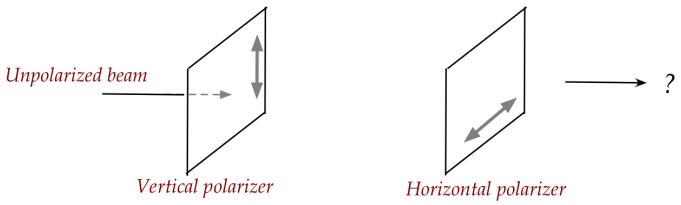
Recall the phenomenon of polarization:



The same experiment with single photons:



Now consider the first of two experiments:

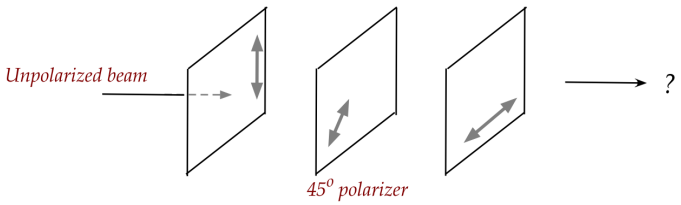


In this scenario:

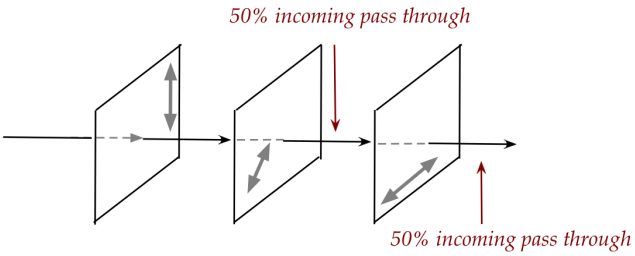
- Careful experimentation shows that: no light emerges from the second polarizer.
- The same result obtains when sending a single photon at a time.

- For single photons:
  - The photon either gets absorbed or passes through the first polarizer.
  - A photon that passes the first is always absorbed by the second polarizer.

Now consider the second scenario:



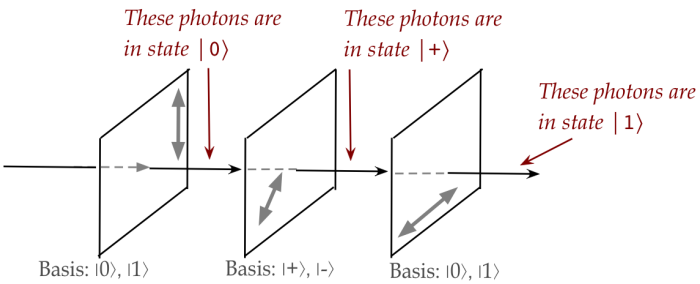
In this case:



- Photons are randomly either absorbed or pass through the first polarizer.
- Here's the key observation: 50% of the photons that reach the middle polarizer pass through.
- Of the photons that reach the third, 50% pass through.

While a classical explanation can be constructed for beams of light, no classical reasoning explains the single-photon experiments.

Now let's provide a quantum explanation:



- Each polarizer is a *measuring device*.
- The bases used by the first and third are:

- The first polarizer uses:

$$\begin{aligned} |0\rangle &= \text{pass through, vertically polarized} \\ |1\rangle &= \text{absorb} \end{aligned}$$

- The *third* polarizer uses:

$$\begin{aligned} |1\rangle &= \text{pass through, horizontally polarized} \\ |0\rangle &= \text{absorb} \end{aligned}$$

- The middle one is more interesting:

$$\begin{aligned} \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle &= \text{pass through, polarized at } 45^\circ \\ \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle &= \text{absorb} \end{aligned}$$

- Now consider a photon of unknown polarization arriving at the first polarizer:
  - The photon's polarization can be written in terms of the measuring device's $|0\rangle, |1\rangle$ basis.
  - Since it can be anything, we model the photon's polarization as $\alpha|0\rangle + \beta|1\rangle$.
  - The probability that it passes through, therefore, is

$$|\alpha|^2 = \text{probability of passing through}$$

  - After passing through, it will be in the $|0\rangle$ state.
    - ▷ Vertically polarized.

- Now let's look at a photon arriving at the middle polarizer:
  - The photon arrives in the state $|0\rangle$.
  - The middle polarizer is a measuring device with basis vectors

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle &&= \text{pass through, polarized at } 45^\circ \\ |-\rangle &= \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle &&= \text{absorb} \end{aligned}$$

  - We now express the incoming photon in the measuring device's basis:

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

  - The probability of passing through is therefore

$$\Pr[\text{pass through}] = \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$$

    This is what gives us the 50% pass through rate.
  - The photon that passes through now has the state $|+\rangle$.

- Finally, let's look at the last polarizer:

- The arriving photon is in the $|+\rangle$ state.
- The basis of the polarizer is:

$$\begin{aligned} |1\rangle &= \text{pass through, horizontally polarized} \\ |0\rangle &= \text{absorb} \end{aligned}$$

- Express the photon's state in the measurement basis:

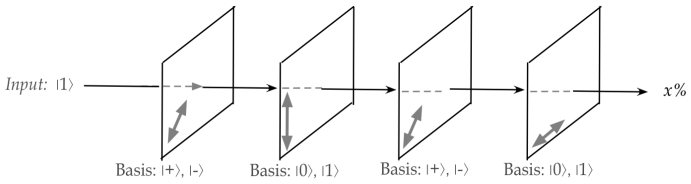
$$|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$

- Thus, the probability that it passes through is:

$$\Pr[\text{pass through}] = \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$$

- Note:
  - Photons can have their state changed by unitary operations.
  - Thus, photons can be made into flying qubits for computation.
  - This will result in our first real application - in the next section.

In-Class Exercise 5: Consider the following set up:



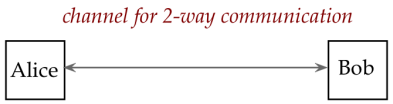
What percentage of $|1\rangle$ photons arriving on the left reach the output? Work through your probability calculations as shown above.

3.7 An application of single qubits: quantum key distribution

The goal of QKD (Quantum Key Distribution) is to enable unconditionally secure communication between two parties, typically called *Alice* and *Bob*.

Let's start by exploring some terms and the context:

- We use the abstract term *channel* to represent any medium of communication between Alice and Bob:



  - Example: mission control (Alice) to satellite (Bob) via radio waves (channel).
  - Example: Headquarters (Alice) to regional bank office (Bob) over the internet (channel).

- Alice and Bob can use a *one-time pad*, which is unconditionally secure:
  - They meet or securely-courier a *shared* book of random integers.
    ▷ This is not done using the channel but up front and securely.
  - Subsequent text messages are encoded using these numbers, with a fresh page used for each message.
  - Because the number patterns are never re-used, an adversary cannot gain any information statistically.
  - And there are (incorrect) number patterns that will decode to valid text, making it impossible to judge whether a lucky guess is correct.

- The big downside of course is the inconvenience of sharing the one-time pads, and keeping them secure.

- Alternatively, Alice and Bob can use a short-term *secret key*, which is a sequence of random integers, and re-use that key for near-term messages:
  - These keys can be created afresh periodically.
  - The problem of sharing such keys is called *key distribution*.
  - To share such keys without meeting, Alice and Bob use *public-key cryptography*.

- However, public-key cryptography relies on the hypothesis that the underlying mathematical problem cannot be easily solved:
  - Factoring (large) integers, as in the RSA public-key system.
  - The discrete log problem for elliptic-curve cryptography.

- As we will see in this course, a quantum computer can efficiently solve the former directly using Shor's algorithm, which can also used indirectly to solve the latter.

The adversary:

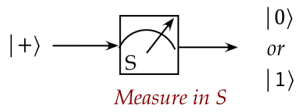
- In any security solution, one must model the adversary
  ▷ Often called the *threat model*.

- The primary attack we'll consider is "tapping into the line"
  - Here, the attacker can receive the photons Alice sends.
  - The attacker can either try to decipher, or both decipher and send photons along to Bob in order to evade detection.

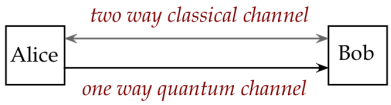
- There are other, more sophisticated attacks we'll consider in a later module.

Main ideas in quantum key distribution using the BB84 Protocol:

- BB84: Bennett and Brassard in 1984 (first such paper).
- BB84 aims to get a bit pattern securely from Alice to Bob.
- The protocol depends on the ability to send single qubits in particular states:
  - In today's technology, the only practical flying qubit is a photon.
  - Thus, the protocol relies on reliable single-photon transmission.
    - ▷ Which is hard to do, but has been demonstrated over distances beyond 100km.
- The protocol also relies on being able to produce truly random bits:
  - This is easy with quantum technology.
  - For example, one simply feeds $|+\rangle$ qubits to a standard-basis measurement:



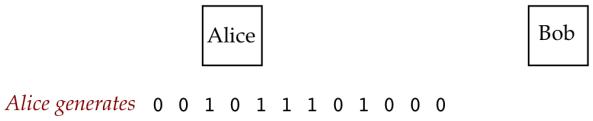
- The protocol uses two channels:



  - The quantum channel transports qubits.
  - The classical channel (which may have weaker security) is used for messages.

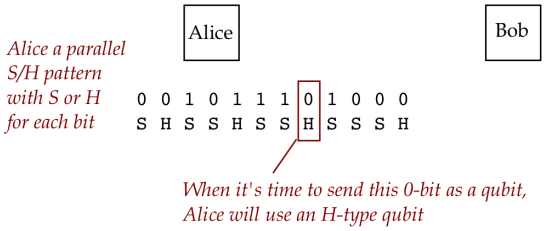
Let's examine the steps in BB84:

- *Step 1* (Alice):
  Alice creates a random bit string (the key) on her side, for example



  We'll call these the classical *potential* key bits.

- *Step 2* (Alice):
  Alice creates a random string using the letters "S" and "H", lining them up with the key bit pattern.

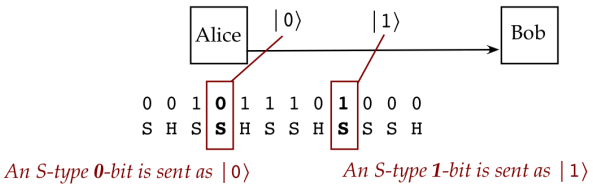


- If a key bit is aligned with "S", she uses an S-type qubit *for that bit*.
- Otherwise, an "H" type qubit.

- *Step 3* (Bob):

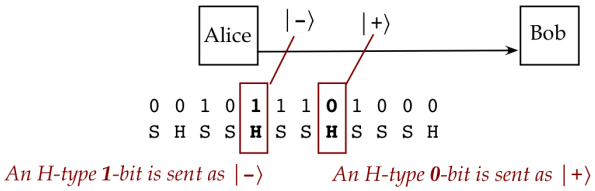
Bob's random string using the letters "S" and "H", will be used later in lining them up with the qubits that Alice will send.

- *Step 4* (Alice):
Alice then sends these as qubits, but using two different types of qubits:
  - If a particular qubit has been designated an *S-type* qubit, then



    - 0 key bit ⇒ send $|0\rangle$
    - 1 key bit ⇒ send $|1\rangle$

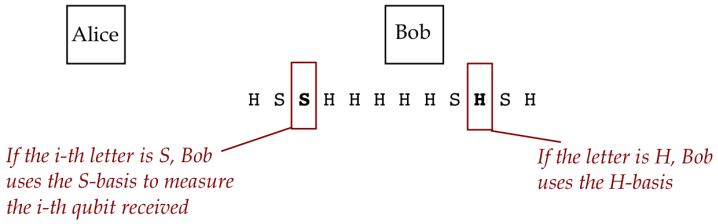
  - If a particular qubit has been designated an *H-type* qubit, then



    - 0 key bit ⇒ send $|+\rangle$
    - 1 key bit ⇒ send $|-\rangle$

  In this manner, all the key bits get sent as S or H qubits.

- *Step 5:* (Bob)

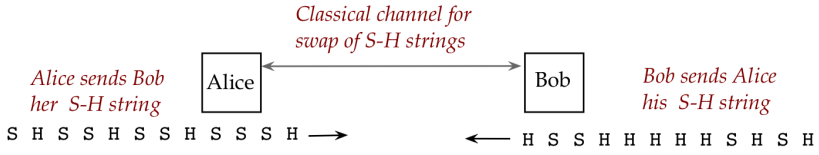


Bob measures these qubits using two bases, depending on his "S-H" string.

- If the i-th letter is S, he uses the Standard-basis, $|0\rangle, |1\rangle$, to measure the i-th qubit.
  - If he obtains $|0\rangle$, he infers 0 as the key bit.
  - Otherwise 1.
- Otherwise he uses the Hadamard basis, $|+\rangle, |-\rangle$,
  - If he obtains $|+\rangle$, he infers 0 as the key bit.
  - Otherwise 1.

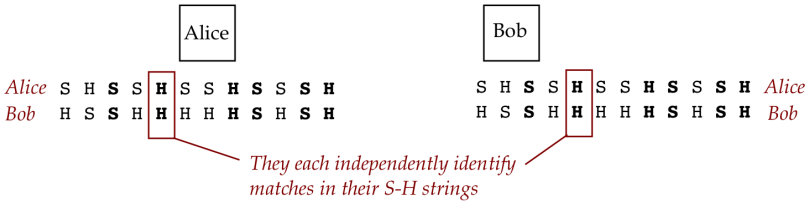
So, at the end of receiving all the qubits, Bob will have a key bit-string aligned with his S-H string:

- *Step 6:* (Alice and Bob)



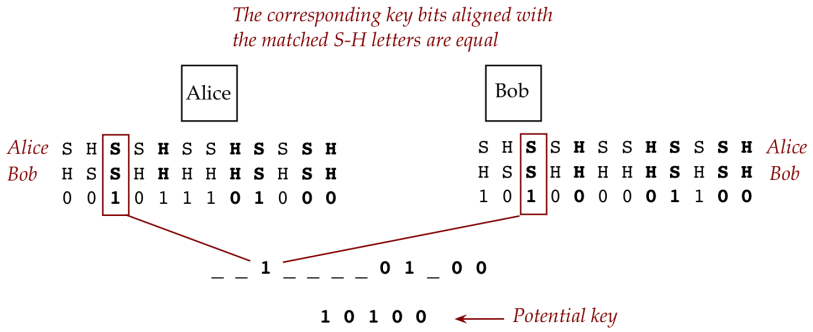
They use the classical channel to send each other their S-H strings.

- *Step 7:* (Alice and Bob)



They each identify the positions in their S-H strings where their letters match.

- *Step 8:* (Alice and Bob)



The bits in the positions where S-H aligned are identical

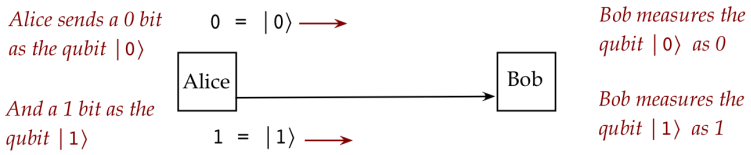
- These are the key bits.

Note: the other bits may differ.

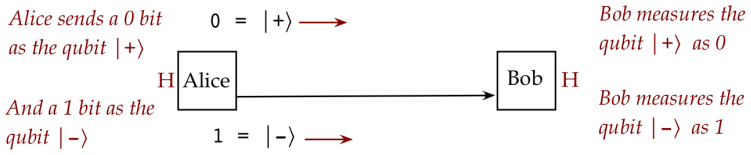
Let's first examine correctness, assuming no attacker (Eve) is present.

- We'll consider two cases:
  - When Alice and Bob use the same letter for a bit (S or H).
  - When their letters are different.

- *Case 1: same letter*
  - *Case 1(A)*: Alice uses S, Bob uses S.



    - If Alice has 0 as the classical bit, she sends $|0\rangle$
      - ▷ Bob uses S-basis and will see $|0\rangle$ with probability 1.
      - ▷ Bob's bit is 0, same as Alice's
    - If Alice has 1 as the classical bit, she sends $|1\rangle$
      - ▷ Bob will measure and see $|1\rangle$ with probability 1. Bob's bit is 1, same as Alice's
  - *Case 1(B)*: Alice uses H, Bob uses H

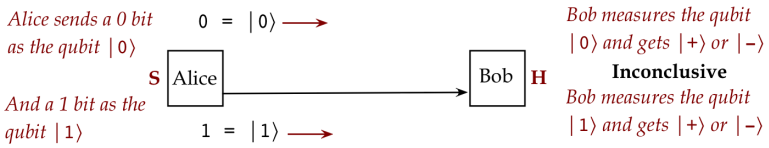


    - If Alice has 0 as the classical bit, she sends $|+\rangle$
      - ▷ Bob uses *H-basis* and will see $|+\rangle$ with probability 1.
      - ▷ Bob's bit is 0, same as Alice's
    - If Alice has 1 as the classical bit, she sends $|-\rangle$
      - ▷ Bob will measure and see $|-\rangle$ with probability 1.

▷ Bob's bit is 1, same as Alice's

Thus, in all cases when they use the same letter, the Bob correctly infers Alice's original classical bits with certainty.

- *Case 2: different letter*
  - *Case 2(A)*: Alice uses S, *Bob uses H to measure*.



    - If Alice has 0 as the classical bit, she sends $|0\rangle$
    - Bob uses H-basis to measure
      - Bob gets $|+\rangle$ with probability 0.5,
        - ▷ Infers 0 with probability 0.5
      - Bob gets $|-\rangle$ (infers 1) with probability 0.5
    - If Alice has 1 as the classical bit, she sends $|1\rangle$
    - With H-basis measurement, Bob sees $|+\rangle$ with probability 0.5, and infers 0.
    - And with probability 0.5, sees $|-\rangle$ and infers 1.

  To summarize: when their letter choices differ, Bob infers the correct bit with probability 0.5.

- Thus, by throwing out the bits where they differed in S-H, the *remaining bits agree exactly*.

- Note:
  - The actual classical bits were not transmitted.
  - The S-H choices were exchanged *after* Bob performs all measurements.

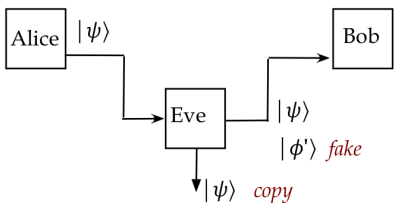
In-Class Exercise 6:

1. What would go wrong if the S-H strings were exchanged before Bob performs qubit measurements?
2. Write out the details of Case 2(B) above, explaining the details of measurement and probabilities.

Let's now focus on Eve:

- Eve the eavesdropper, that is.

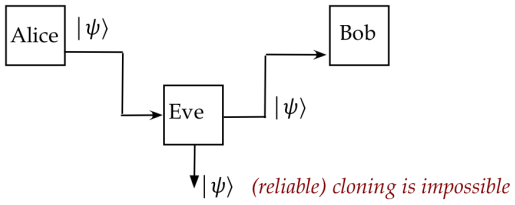
- What can Eve do?



- Eve can try to copy (some or all of) the qubits by tapping the optical fiber.
- Eve can capture Alice's qubits and send fake qubits to Bob.
- Eve can pose as Alice to Bob (impersonate).

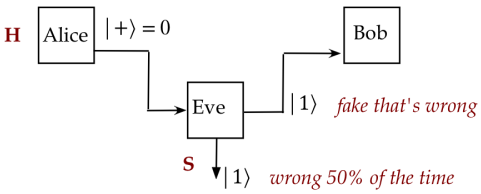
Let's look at these one by one.

- We will later prove a simple but dramatic result: *it is physically impossible, in general, to copy qubits*



  - This is called the *No-Cloning Theorem*.
  - There are some caveats, but in general, a clean accurate copy of a qubit can't be made without affecting the original.

- If Eve tries to intercept a few qubits, she'll have to guess which basis to use in measurement.



  - Which means she'll be wrong 50% of the time (with random guessing).
  - These qubits will change state because of measurement
  - If she sends the changed qubits to Bob, then Alice and Bob can detect this (as we'll show below).

- What happens if Eve impersonates Alice?
  - BB84 does not solve this problem.
  - The assumption is that the Alice and Bob can somehow authenticate each other through the classical channel.

An improvement: detecting Eve

- Think of the agreed bits as the "good bits".

- Alice and Bob, through the classical channel, share 50% of their "good bits".
  ▷ This means the openly shared bits can't be used for the secret key.

- If Eve intercepted qubits and sent fake ones, then 50% of these supposedly good bits will differ between Alice and Bob.
  ▷ Eve will be caught.

- The remaining unshared good bits becomes the actual secret key to be used for future encryption.

3.8 Phase equivalence in vectors

Consider these two vectors:

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

$$|\psi'\rangle = \left(\frac{2-\sqrt{2}i}{3}\right)|0\rangle + \left(\frac{\sqrt{2}-i}{3}\right)|1\rangle = \begin{bmatrix} \frac{2-\sqrt{2}i}{3} \\ \frac{\sqrt{2}-i}{3} \end{bmatrix}$$

- Clearly, $|\psi\rangle$ and $|\psi'\rangle$ are different vectors because there's no simplification of the components of $|\psi'\rangle$ that will result in $|\psi\rangle$.
- Now consider the probability of obtaining $|0\rangle$ when each vector above is measured by in the S-basis (standard basis):

  - When $|\psi\rangle$ is the input vector, the probability of obtaining $|0\rangle$ is:

$$\Pr[|0\rangle] = \left|\frac{\sqrt{2}}{\sqrt{3}}\right|^2 = \frac{2}{3}$$

  - With $|\psi'\rangle$ we get

$$\Pr[|0\rangle] = \left|\frac{2-\sqrt{2}i}{3}\right|^2 = \frac{2}{3}$$

  Both give the same results.

- Thus, experimentally, it is impossible to distinguish them with the S-basis measurement.
- Now suppose we define the number

$$a = \frac{2}{\sqrt{6}} - i\frac{\sqrt{2}}{\sqrt{6}}$$

- The exercise below shows that

$$|\psi'\rangle = a\,|\psi\rangle$$

and

$$|a|^2 = 1$$

- Thus, we can write

$$|\psi'\rangle = \frac{a\sqrt{2}}{\sqrt{3}}|0\rangle + \frac{a}{\sqrt{3}}|0\rangle$$

where the probability of obtaining $|0\rangle$ in S-basis is

$$\Pr[|0\rangle] = \left|\frac{a\sqrt{2}}{\sqrt{3}}\right|^2 = |a|^2\left|\frac{\sqrt{2}}{\sqrt{3}}\right|^2 = \left|\frac{\sqrt{2}}{\sqrt{3}}\right|^2 = \frac{2}{3}$$

- To clarify further, let's use

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

and

$$|\psi'\rangle = a|\psi\rangle = a\alpha|0\rangle + a\beta|1\rangle$$

Then, $|\psi'\rangle$

$$\Pr[|0\rangle] = |a\alpha|^2 = |a|^2|\alpha|^2 = |\alpha|^2$$

which is the same probability when using $|\psi\rangle$.

- Thus, for *any* complex number a such that $|a|^2 = 1$, the two vectors $|\psi\rangle$ and $|\psi'\rangle$ are indistinguisable with regard to S-basis measurements.
- This will be true for any measurement basis.
- Thus, no experiment can reveal a *measurable* difference between $|\psi\rangle$ and $|\psi'\rangle$.
- These two vectors are considered to represent the *same quantum state* and are called *globally phase-equivalent* vectors.

In-Class Exercise 7: Show that with a, $|\psi\rangle$ and $|\psi'\rangle$ defined above,

1. $|a|^2 = 1$
2. $|\psi'\rangle = a|\psi\rangle$

There is another, commonly used way of looking at phase-equivalence:

- We know that any complex number a can be written as

$$a = re^{i\theta}$$

When a has magnitude 1, then $r = 1$ and thus

$$a = e^{i\theta}$$

- Now consider any vector $|\psi\rangle$ expressed in any basis $|v\rangle, |v^{\perp}\rangle$:

$$|\psi\rangle = \alpha|v\rangle + \beta|v^{\perp}\rangle$$

and let

$$\psi' = a|\psi\rangle = e^{i\theta}|\psi\rangle$$

- Then

$$\psi' = e^{i\theta}\alpha|v\rangle + e^{i\theta}\beta|v^{\perp}\rangle$$

- Now when measured in the new $|v\rangle, |v^{\perp}\rangle$ basis, the probability of obtaining $|v\rangle$

$$\Pr[|v\rangle] = |e^{i\theta}\alpha|^2 = |e^{i\theta}|^2|\alpha|^2 = |\alpha|^2$$

Thus, no measurement can distinguish the two vectors: both represent the *same quantum state*.

Now let's look at the concept of *relative phase*:

- Consider two two vectors we're already familiar with:

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

- At first glance, it may seem that the probability of obtaining $|0\rangle$ is the same for both vectors and they should therefore be considered indistinguishable.
- However, if we switch measurement to the H-basis, that is the $|+\rangle, |-\rangle$ basis, they are distinguishable.
- Thus, even though

$$|-\rangle = \frac{1}{\sqrt{2}}|0\rangle + e^{i\pi}\frac{1}{\sqrt{2}}|1\rangle$$

has one coefficient modified by an $e^{i\theta}$ term, it is distinguishable if we change the basis.

- In general we say that the two vectors

$$\begin{aligned} |\psi\rangle &= \alpha |v\rangle + \beta \left|v^{\perp}\right\rangle \\ \psi' &= \alpha |v\rangle + e^{i\theta}\beta \left|v^{\perp}\right\rangle \end{aligned}$$

differ by *relative phase* and represent *different quantum states*.

- So, how do we know that a change of basis cannot distinguish two vectors in globally equivalent phases?
  - Suppose $|\psi\rangle$ and ψ' are two vectors such that

$$\psi' = a |\psi\rangle$$

  - Next, let's apply a change-of-basis matrix M to ψ'

$$M\psi' = M\, a\, |\psi\rangle = aM\, |\psi\rangle$$

    Thus, the constant a multiplies into the changed coordinates and does not change the probabilities in the new basis.

To summarize:

- *Global phase* does not matter and so, when there's a chance for simplification, we'll use this fact.
- *Relative phase* does matter and will, in fact, be exploited in algorithms, as we'll later see.

3.9 Projective measurement

Let's start with recalling what we've learned about measurement so far:

- Measuring starts with: "what are we measuring?":
  - The target of measuring is always a quantum state, a vector $|\psi\rangle$.
  - In this module, we are focusing on the quantum state of a qubit.
  - Later we will apply measurement to multiple qubits.
- A measurement device, we've seen, is a basis $|v\rangle, \left|v^{\perp}\right\rangle$.
- The act of measurement involves:
  - Express the state in terms of the measurement basis:

$$|\psi\rangle = \alpha |v\rangle + \beta \left|v^{\perp}\right\rangle$$

  - The outcomes of measurement will be *one* of the basis vectors $|v\rangle$ or $\left|v^{\perp}\right\rangle$.
  - Which one occurs is probabilistically determined by nature, based on the coefficients α, β:

$$\begin{aligned} \Pr[\text{observe } |v\rangle] &= |\alpha|^2 \\ \Pr[\text{observe } |v^\perp\rangle] &= |\beta|^2 \end{aligned}$$

We will now look at a broader theoretical framework for measurement based on *projections*:

- We'll still have a basis for measurement.
- But we'll use projection matrices on a state $|\psi\rangle$ to determine outcomes and probabilities.
- Why go through the extra trouble?
  - Once we start to use projection operators (matrices), we'll be able to combine them smoothly using matrix properties.
  - The theory so developed will naturally extend to including eigenvalues when projectors are combined
    - ▷ This is where Hermitian operators come in.
  - The broader theory unifies quantum computing and mechanics.

Recall projectors:

- In general, a projector for a vector $|v\rangle$ is an operator P_v such that $P_v |\psi\rangle$ is along $|v\rangle$ for any $|\psi\rangle$.
- The operator is easily constructed: it's the *outer-product* created from $|v\rangle$:

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

- Let's revisit some examples:
  - Example 1:

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

    are the projectors for the basis vectors $|0\rangle, |1\rangle$.
  - Example 2:

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

    are the projectors for the basis vectors $|+\rangle, |-\rangle$.

In-Class Exercise 8: Write down the two projector matrices $P_v, P_{v^\perp}$ for the general 2D basis $|v\rangle = a\,|0\rangle + b\,|1\rangle$, and $|v^\perp\rangle = b^*\,|0\rangle - a^*\,|1\rangle$, and check your calculations by showing $P_v P_{v^\perp} = 0$.

Now let's apply projectors to measurement:

- When we say "apply", we mean to a current qubit state $|\psi\rangle$.
- We will do *three* things when applying a projector P to a vector $|\psi\rangle$:
  1. Compute the resulting projection

$$|\psi_P\rangle \;=\; P\,|\psi\rangle$$

  2. Compute the squared magnitude of this resulting vector

$$\big|\,|\psi_P\rangle\big|^2 \;=\; \big|P\,|\psi\rangle\big|^2$$

     This will necessarily be less than or equal to 1, and will be the probability of obtaining the normalized-projection (next).
  3. Compute the normalized projection:

$$|\psi_N\rangle \;=\; \frac{|\psi_P\rangle}{||\psi_P\rangle|} \;=\; \frac{P\,|\psi\rangle}{|P\,|\psi\rangle|}$$

     This will be one of the possible *outcomes* of measurement.

Projective measurement for a single qubit via an example:

- Let's use the S-basis $|0\rangle, |1\rangle$ for measuring

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

- The projectors for each basis vector are:

$$P_0 \;=\; \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$
$$P_1 \;=\; \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

- *Step 1:* compute the projected vectors

$$|\psi_{P_0}\rangle \;=\; P_0\,|\psi\rangle \;=\; \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \;=\; \begin{bmatrix} \alpha \\ 0 \end{bmatrix} \;=\; \alpha\,|0\rangle$$
$$|\psi_{P_1}\rangle \;=\; P_1\,|\psi\rangle \;=\; \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \;=\; \begin{bmatrix} 0 \\ \beta \end{bmatrix} \;=\; \beta\,|1\rangle$$

- *Step 2:* compute the squared magnitude of each projection:

$$\left| |\psi_{P_0}\rangle \right|^2 = \left|\alpha |0\rangle\right|^2 = \left\langle \alpha \langle 0| \,\middle|\, \alpha |0\rangle \right\rangle = \alpha^* \alpha = |\alpha|^2$$
$$\left| |\psi_{P_1}\rangle \right|^2 = \left|\beta |1\rangle\right|^2 = \left\langle \beta \langle 1| \,\middle|\, \beta |1\rangle \right\rangle = \beta^* \beta = |\beta|^2$$

- *Step 3:* compute the normalized projections:

$$|\psi_{N_0}\rangle = \frac{P_0 |\psi\rangle}{|P_0 |\psi\rangle|} = \frac{\alpha |0\rangle}{|\alpha|}$$
$$|\psi_{N_1}\rangle = \frac{P_1 |\psi\rangle}{|P_1 |\psi\rangle|} = \frac{\beta |1\rangle}{|\beta|}$$

- Now apply these steps to measurement:
  - The two possible *outcomes* of measurement are the normalized projections:

$$|\psi_{N_0}\rangle = \frac{\alpha |0\rangle}{|\alpha|}$$
$$|\psi_{N_1}\rangle = \frac{\beta |1\rangle}{|\beta|}$$

  - They each occur with probabilities

$$\Pr[\text{observe } \psi_{N_0}] = |\alpha|^2$$
$$\Pr[\text{observe } \psi_{N_1}] = |\beta|^2$$

- There is one matter to clear up:
  - In the earlier approach to measurement in S-basis, the outcomes were

$$|0\rangle \qquad \text{or } |1\rangle$$

  - But the projective approach has outcomes

$$\frac{\alpha |0\rangle}{|\alpha|} \qquad \text{or } \frac{\beta |1\rangle}{|\beta|}$$

  - Do the two approaches differ in outcomes?

- They are in fact the same because of global-phase equivalence:

$$\frac{\alpha}{|\alpha|} |0\rangle = |0\rangle$$

because

$$\left| \frac{\alpha}{|\alpha|} \right| = \frac{|\alpha|}{|\alpha|} = 1$$

- Knowing this, we'll of course opt for the simpler $|0\rangle, |1\rangle$ as outcomes.

In-Class Exercise 9: Suppose the qubit state is $|\psi\rangle = |+\rangle$ and the measurement basis is: $|w\rangle = \frac{\sqrt{3}}{2}|0\rangle - \frac{1}{2}|1\rangle$, and $|w^{\perp}\rangle = \frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle$. Apply the projective measurement approach:

1. Compute the projectors $P_w, P_{w^{\perp}}$.
2. Apply the three steps, simplifying where possible.

Confirm that you get the same results as in Exercise 1.

Let's summarize by going to general n-dimensional space (since we'll be getting there soon):

- Suppose $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$ represents a measurement basis.
- Next, suppose $|\psi\rangle$ is an n-dimensional complex vector representing a quantum state of some quantum object.
- The three steps (with simplified subscripts) are
    1. Compute each projected vector

$$|\psi_{P_i}\rangle \quad = \quad P_i\,|\psi\rangle$$

    2. Compute the squared magnitude of each projection:

$$||\psi_{P_i}\rangle|^2$$

    as the probability of seeing the i-th normalized projection.
    3. Compute each normalized projection (each potential outcome):

$$|\psi_{N_i}\rangle \quad = \quad \frac{|\psi_{N_i}\rangle}{||\psi_{P_i}\rangle|}$$

    Then we have the outcomes and their probabilities.

3.10 Hermitians and measurement

Hermitian operators play a central role in measurement, and purpose of this section is to connect the dots between the projector approach and Hermitians.

Let's start by recall some facts about Hermitians:

- A Hermitian operator A satisfies $A^{\dagger} = A$.
- Hermitian operators have *real eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n$.
- The eigenvectors of a Hermitian operator, $|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_n\rangle$, form an orthonormal basis.

- By the spectral theorem, we can write

$$A = \sum_i \lambda_i |\phi\rangle\langle\phi|$$

  which is a real-linear combination of the projectors $P_i = |\phi\rangle\langle\phi|$

- We can think of this result as "Did you know that a Hermitian can be decomposed into a linear combination of its eigenvector projectors, where the coefficients are the eigenvalues?"

Now let's go the other way: from projectors to Hermitians

- Suppose we have an orthonormal basis $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$.
- And suppose $\gamma_1, \gamma_2, \ldots, \gamma_n$ are any *distinct* real numbers.
- Then, the linear combination of projectors

$$\sum_i \gamma_i P_{v_i} = \sum_i \gamma_i |v_i\rangle\langle v_i|$$

  is a Hermitian operator.
  (We proved this in Module 2.)

- Moreover, this operator has *eigenvalues* $\gamma_1, \gamma_2, \ldots, \gamma_n$ and associated *eigenvectors* $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$.
- How do we know this?

$$\left(\sum_i \gamma_i |v_i\rangle\langle v_i|\right)|v_i\rangle = \gamma_i |v_i\rangle\langle v_i|v_i\rangle = \gamma_i |v_i\rangle$$

- Let's go a step further and see this in an example:
  - Here, we pick the numbers $\gamma_1, \gamma_2, \ldots, \gamma_n$.
  - In 2D, we'll pick, say $\gamma_1 = 2, \gamma_2 = 3$.
  - Consider the S-basis and let

$$A = \gamma_1 |0\rangle\langle 0| + \gamma_2 |1\rangle\langle 1| = 2\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 3\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

  - Then, A is Hermitian with eigenvalues 2 and 3.

- Think of this artificially constructed Hermitian as "packaging" all the projectors together in a way that lets the spectral decomposition recover the individual projectors.
- The Hermitian also includes eigenvalues, one per projector.

In-Class Exercise 10: Use the projector matrices for the basis and the numbers $\gamma_1 = 3, \gamma_2 = 5$ to construct the Hermitian. Then show that these are the eigenvalues with $kt+, |-\rangle$ as eigenvectors.

How does any of this matter?

- Recall: we've said that the outcome of a measurement is one of the basis vectors involved.
- There is in fact another outcome: the associated *eigenvalue*:
  - For a Hermitian that's derived from an actual physical quantum object, there are two outcomes:
    - The eigenvector that will be the resulting state.
    - The associated eigenvalue.
- In real physical devices, the associated eigenvalue (a real number) is a physical quantity that's observable, like energy or frequency.
- Thus, the eigenvalue matters in physical models because it is the quantifiable output that a device reports.
- But in quantum *computing*, such eigenvalues play no role. We only care about the resulting *eigenvectors*.
- Nonetheless, we will use the Hermitian formulation because:
  - When you later encounter theory it will be in this form, as we'll see in the next section.
  - Because Hermitian matrices package projectors, the theory develops powerful ways to combine Hermitians, as we'll see.

3.11 The Hermitian sandwich

Recall that an eigenvalue is one of the measurement outcomes in a physical experiment.

Each repetition of an experiment will result in one of the eigenvalues appearing, according to the probabilities involved:

- Suppose $|\psi\rangle$ is a state vector.
- Next, let the Hermitian operator A represent a measurement.
- Suppose that A has eigenvectors $|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_n\rangle$, with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.
- Then, we know that:
  - The resulting state after measurement will be one of: $|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_n\rangle$.
  - If the physical state $|\phi_i\rangle$ results, the physical quantity λ_i will be observed.

Alternatively: if λ_i is observed, one concludes that the resulting state is $|\phi_i\rangle$.

- Let's focus on which λ_i's we see in several repetitions of measuring the original $|\psi\rangle$. For example, we might see
  - Trial #1: λ_3
  - Trial #2: λ_7
  - Trial #3: λ_1
  - Trial #4: λ_3
  - ...
- We are interested now in the *mean* (average) eigenvalue obtained:

$$\frac{\lambda_3 + \lambda_7 + \lambda_1 + \lambda_3 + \ldots}{\text{number of trials}}$$

  This is from an experiment.
- We'd like to *calculate* this exactly using the probabilities from the model.
  ▷ For this, we use the expected-value calculation
- If p_i = probability of obtaining λ_i then the mean is

$$\sum_i \lambda_i p_i$$

- Such an *expected value* and is often written as:

$$\langle \lambda \rangle \quad = \quad \sum_i \lambda_i p_i$$

  where λ is a random variable representing the possible values $\lambda_1, \ldots, \lambda_n$.
- Now, the probabilities depend on the particular state $|\psi\rangle$.
- Let's write $|\psi\rangle$ in terms of the measurement basis

$$|\psi\rangle \quad = \quad \alpha_1 \phi_1 + \ldots + \alpha_n \phi_n$$

  Then

$$p_i \quad = \quad |\alpha_i|^2$$

Using the sandwich:

- Let's compute $\langle \psi | A | \psi \rangle$:

$$\begin{aligned}
\langle\psi|A|\psi\rangle &= \left\langle \sum_i \alpha_i^* \langle\phi| \,\middle|\, \sum_i \lambda_i |\phi\rangle\langle\phi| \,\middle|\, \sum_i \alpha_i |\phi\rangle \right\rangle \\
&= \left\langle \sum_i \alpha_i^* \langle\phi| \,\middle|\, \lambda_i \alpha_i |\phi\rangle \right\rangle \\
&= \sum_i \alpha_i^* \alpha_i \lambda_i \langle\phi_i|\phi_i\rangle \\
&= \sum_i \lambda_i |\alpha_i|^2 \\
&= \langle\lambda\rangle
\end{aligned}$$

- Thus

$$\langle\lambda\rangle = \langle\psi|A|\psi\rangle$$

- Since each expectation depends on the state, we could make that explicit in the notation:

$$\langle\lambda_\psi\rangle = \langle\psi|A|\psi\rangle$$

- The sandwich has other uses:
  - For example, one can compute the expected value of λ_ψ^2 (called the 2nd moment).

$$\langle\lambda_\psi^2\rangle = \left\langle\psi\middle|A^2\middle|\psi\right\rangle$$

    where A^2 is the operator applied twice (matrix multiplication).
  - And from there, the variance:

$$\text{var}(\lambda) = \left\langle\psi\middle|A^2\middle|\psi\right\rangle - \langle\psi|A|\psi\rangle^2$$

- One can then reason about such statistics with multiple operators and their covariance, and analyze them over all possible $|\psi\rangle$.
  ▷ One such famous result: Heisenberg's uncertainty principle

---

3.12 Matrix calculations with Dirac notation

Let's revisit the X operator that "flipped" a qubit:

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

Notice that:

- In applying the operator, we wrote the operator in matrix form, then wrote the $|0\rangle$ vector as a column vector.

There is an alternate way that uses Dirac notation:

- First express the operator using outer products:

$$X = |0\rangle\langle 1| + |1\rangle\langle 0|$$

- Then apply:

$$X|0\rangle = (|0\rangle\langle 1| + |1\rangle\langle 0|)|0\rangle = |0\rangle\langle 1|\,|0\rangle + |1\rangle\langle 0|\,|0\rangle = \langle 1|0\rangle\,|0\rangle + \langle 0|0\rangle\,|1\rangle = |1\rangle$$

  This exploits the simplifications possible with orthonormal vectors when most inner products evaluate to 0.

- Let's go back to the outer-product representation and clarify:

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

- It is not always the case that the Dirac approach will simplify, but we will see cases where it turns out to be simpler.

- This is the case when we work with muliple qubits where the matrix sizes are large.

- Let's apply X to a general vector $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$

$$X|\psi\rangle = (|0\rangle\langle 1| + |1\rangle\langle 0|)(\alpha|0\rangle + \beta|1\rangle) = \beta\langle 1|1\rangle|0\rangle + \alpha\langle 0|0\rangle|1\rangle = \beta|0\rangle + \alpha|1\rangle$$

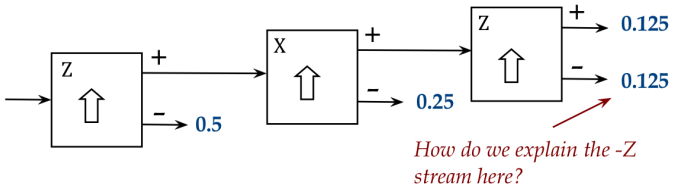
  as expected.

In-Class Exercise 11: Show that the Hadamard matrix can be written as $H = \frac{1}{\sqrt{2}}(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 1|)$ and then apply it to $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$.

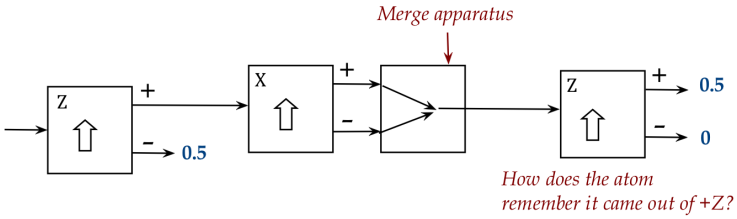
3.13 Stern-Gerlach explained

Let's revisit two Stern-Gerlach experiments from Module 1 that will be sufficient:

- First



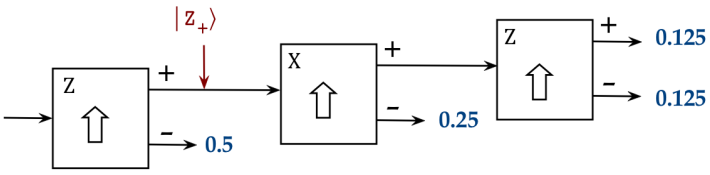
- And then



The quantum framework:

- We model the state of an atom as a 2D vector $|\psi\rangle$.
- Each apparatus behaves like a measuring device.
- The basis for the Z-aligned apparatus has two vectors
  - $|z_+\rangle$: atoms in this state come out of the '+' port
  - $|z_-\rangle$: atoms in this state come out of the '-' port
- Likewise, for an X-aligned apparatus, the basis vectors are $|x_+\rangle$, $|x_-\rangle$.
- And for the Y-direction: $|y_+\rangle$, $|y_-\rangle$.
- Each pair being a basis, have orthonormal vectors.

Consider the first experiment:



- The stream of atoms coming out of the + port of the first (Z-aligned) SG are all in the state $|z_+\rangle$.
  - ▷ This is our measurement assumption.
- Now single-atom experiments show that when the $|z_+\rangle$ atoms enter the second stage, they come out of + and - with equal probability.
- Since the second apparatus is described by the basis $|x_+\rangle$, $|x_-\rangle$, we express the input $|z_+\rangle$ in terms of this basis

$$|z_+\rangle \quad = \quad \alpha\,|x_+\rangle + \beta\,|x_-\rangle$$

from which we must have equal probabilities.

- Thus, $\alpha = \frac{1}{\sqrt{2}} = \beta$:

$$|z_+\rangle \quad = \quad \frac{1}{\sqrt{2}}|x_+\rangle + \frac{1}{\sqrt{2}}|x_-\rangle$$

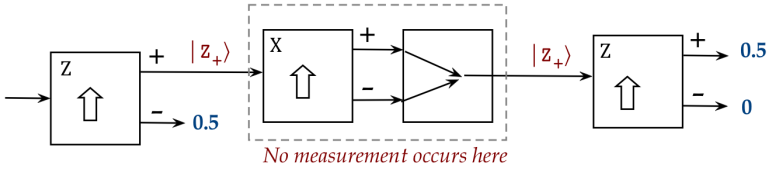
- This explains the second-stage split.
- When picking off the + stream in the second stage, we know that the state of those atoms is $|x_+\rangle$.
- By symmetry

$$|x_+\rangle \quad = \quad \frac{1}{\sqrt{2}}|z_+\rangle + \frac{1}{\sqrt{2}}|z_-\rangle$$

where $|z_+\rangle\,, |z_-\rangle$ is the third stage basis.

- This is why we get equal probabilities in the third stage.

Now for the "merged" second stage:



- If the merging is done correctly *no measurement* occurs in the middle.
- This means that the exiting state is the same as the entering state: $|z_+\rangle$.
- When $|z_+\rangle$ is measured in the $|z_+\rangle\,, |z_-\rangle$ basis, the outcome is $|z_+\rangle$ with probability 1!

3.14* Mathematical aside: why complex numbers are needed

So far most of our examples have involved real numbers like $\frac{1}{\sqrt{2}}$, even when using the Hadamard basis.

One could ask: why do we need complex numbers at all? Why not just work with real vectors?

We will use the Stern-Gerlach set up to show that complex numbers arise naturally from the calculations needed to explain what we saw.

Let's start with an observation:

- Suppose

$$|w\rangle \quad = \quad a_1 |v_1\rangle + a_2 |v_2\rangle$$

  expresses some vector $|w\rangle$ in some orthonormal 2D basis $|v_1\rangle, |v_2\rangle$.

- Write the coefficients in polar form:

$$|w\rangle \quad = \quad r_1 e^{i\theta_1} |v_1\rangle + r_2 e^{i\theta_2} |v_2\rangle$$

- These is global-phase equivalent to:

$$e^{-i\theta_1} \left(r_1 e^{i\theta_1} |v_1\rangle + r_2 e^{i\theta_2} |v_2\rangle \right) \quad = \quad r_1 |v_1\rangle + r_2 e^{i\gamma} |v_2\rangle$$

  where $\gamma = \theta_2 - \theta_1$.

- Thus, every qubit vector can be written equivalently with a vector whose first coefficient is real.

With this insight, let's look at the Z-to-X measurement:

- Write

$$\begin{aligned} |x_+\rangle &= \alpha_1 |z_+\rangle + \beta_1 |z_-\rangle \\ |x_-\rangle &= \alpha_2 |z_+\rangle + \beta_2 |z_-\rangle \end{aligned}$$

- From experiments we know that

$$|\alpha_1|^2 \quad = \quad \frac{1}{2} \quad = \quad |\alpha_2|^2$$

  and with the insight that we can choose a phase-equivalence where these are real, it must be that

$$\alpha_1 \quad = \quad \alpha_2 \quad = \quad \frac{1}{\sqrt{2}}$$

- So, now write

$$\begin{aligned} |x_+\rangle &= \frac{1}{\sqrt{2}} |z_+\rangle + r_1 e^{i\gamma_1} |z_-\rangle \\ |x_-\rangle &= \frac{2}{\sqrt{2}} |z_+\rangle + r_2 e^{i\gamma_2} |z_-\rangle \end{aligned}$$

- The second coefficient also results in a probability of $\frac{1}{2}$, and so

$$r_1 \;=\; r_2 \;=\; \frac{1}{\sqrt{2}}$$

- Next, experiments have shown that no $|x_+\rangle$-state atom comes out of a $|x_-\rangle$ port, which means these are orthogonal vectors:

$$\langle x_+|x_-\rangle \;=\; 0$$

- Expanding from the expressions for each:

$$\left\langle \frac{1}{\sqrt{2}}\langle z_+| + \frac{1}{\sqrt{2}}e^{-i\gamma_1}\langle z_-| \;\middle|\; \frac{2}{\sqrt{2}}|z_+\rangle + \frac{1}{\sqrt{2}}e^{i\gamma_2}|z_-\rangle \right\rangle \;=\; 0$$

  Or

$$e^{i(\gamma_2-\gamma_1)} + 1 \;=\; 0$$

- We can choose γ_1 because:
  - The actual Z-X directions are only relative to each other.
  - We have not specified the orientation of the X-aligned SG in the plane perpendicular to Z.
- The most convenient choice is $\gamma_1 = 0$.
- This implies

$$e^{i\gamma_2} \;=\; -1$$

  and therefore

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

- So far, no complex numbers.

Next, we need to set up the equations implied by the third basis: $|y_+\rangle, |y_-\rangle$

- Write

$$\begin{aligned} |y_+\rangle &= \alpha_3|z_+\rangle + \beta_3|z_-\rangle \\ |y_-\rangle &= \alpha_4|z_+\rangle + \beta_4|z_-\rangle \end{aligned}$$

- Orthogonality amongst $|y_+\rangle, |y_-\rangle$ means

$$\langle y_+|y_-\rangle = \alpha_3^*\alpha_4 + \beta_3^*\beta_4 \;=\; 0$$

- Recall that

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

- Thus, we can calculate

$$
\begin{aligned}
\langle y_+|x_+\rangle &= \frac{1}{\sqrt{2}}(\alpha_3^* + \beta_3^*) \\
\langle y_-|x_+\rangle &= \frac{1}{\sqrt{2}}(\alpha_4^* + \beta_4^*) \\
\langle y_+|x_-\rangle &= \frac{1}{\sqrt{2}}(\alpha_3^* - \beta_3^*) \\
\langle y_-|x_-\rangle &= \frac{1}{\sqrt{2}}(\alpha_4^* - \beta_4^*)
\end{aligned}
$$

- Next, use these four in the experimental results that split X-to-Y streams:

$$
\begin{aligned}
\langle y_+|x_+\rangle\ \langle x_+|y_+\rangle &= \frac{1}{2} \\
\langle y_-|x_+\rangle\ \langle x_+|y_-\rangle &= \frac{1}{2} \\
\langle y_+|x_-\rangle\ \langle x_-|y_+\rangle &= \frac{1}{2} \\
\langle y_-|x_-\rangle\ \langle x_-|y_-\rangle &= \frac{1}{2}
\end{aligned}
$$

- Let's do one of these:

$$
\begin{aligned}
\langle y_+|x_+\rangle\ \langle x_+|y_+\rangle &= \frac{1}{\sqrt{2}}(\alpha_3^* + \beta_3^*)\frac{1}{\sqrt{2}}(\alpha_3 + \beta_3) \\
&= \frac{1}{2}(\alpha_3^*\alpha_3 + \beta_3^*\beta_3 + \alpha_3^*\beta_3 + \alpha_3\beta_3^*) \\
&= \frac{1}{2}(1 + \alpha_3^*\beta_3 + \alpha_3\beta_3^*)
\end{aligned}
$$

- Equating this to $\frac{1}{2}$ results in

$$
\alpha_3^*\beta_3 + \alpha_3\beta_3^* = 0
$$

  or

$$
\alpha_3^*\beta_3 + (\alpha_3^*\beta_3)^* = 0
$$

- This is in the form $z + z^* = 0$, which implies z's real part is 0.
  (Try writing $z = a + ib$ to see why.)

- Recall that we can choose the first coefficient α_3 to be real.

- This implies β_3 cannot be purely real.
  ▷ β_3 must have a non-zero imaginary part.

- Thus, with

$$|\alpha_3|^2 \;=\; \frac{1}{2} \;=\; |\beta_3|^2$$

  we can choose

$$\alpha_3 \;=\; \frac{1}{\sqrt{2}} \qquad \beta_3 \;=\; \frac{i}{\sqrt{2}}$$

- Continuing the analysis, we finally get

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

Thus, we have not only explained what necessitated complex numbers, but we also completed the full analysis of all three SG bases.

3.15* Mathematical aside: is the set of qubit vectors a vector space?

Recall that a vector space is closed under addition: add any two vectors in the vector space and you should get a vector in that space.

Yet

$$|\psi\rangle \;=\; |0\rangle + |1\rangle$$

is not a valid qubit state since it's not unit-length.

What went wrong?

It turns out that a further tweak is needed for mathematical consistency:

- An analogy will help:
  - If we want to perform arithmetic (with regular integers) so that all the integers fall in the range $0, 1, \ldots m-1$, then we use mod-m with *every operation*.
  - Thus, for example, with $m = 5$

$$3 + 4 \text{ (mod m)} \;=\; 2$$

- We can think of this as:
  - Perform the operation with regular arithmetic.
  - Take the result and compute mod-m.

- Once we do this, all kinds of arithmetic operations will result in keeping all the results with our desired set of $0, \ldots, m-1$.

- A similar thing can be done with vectors:
  - Perform addition or scalar multiplication in the regular way.
  - Take the result and *normalize* it.

  This way, the result is of unit-length.

- This action of normalizing is sometimes also called a projection, and the type of vector space is called a projective complex vector space.

- We won't worry about this issue any further.
