Quantum Mechanics

Qubits and Entanglement

Quantum States

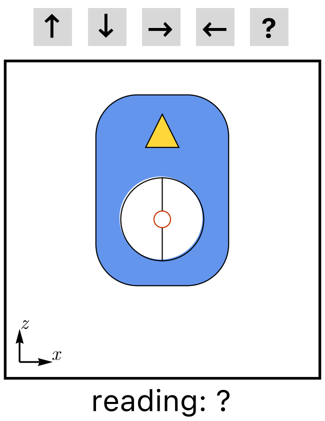
▪ System and Measurement

▪ Spins and Qubits

- The concept of **spin** is derived from particle physics. It is an *internal degree of freedom* attached to a particle (say electron).
    - Naively, spin can be pictured as a little arrow pointing in some direction.
    - But that classical picture is not precise and sometimes misleading.
- We can *isolate* the quantum spin from the particle that carries it ⇒ we can abstract the concept of **qubit**, or **quantum bit**: a *two*-state *quantum* system.
    - A qubit is the *simplest* quantum system, yet it exhibits all the most *essential* properties of quantum mechanics.
    - It is also used as a **unit** of **quantum information**, like classical bit for classical information in our computer.
    - Some believe that qubits are the **building blocks** of (maybe all) quantum systems. There is an on-going research collaboration called "**it from qubit**" (Simons foundation): to unify **matter**, **spacetime** (**gravity**) and **information**.

▪ A Toy Experiment

Let us try to understand qubit by *probing* it. Here is a *toy experiment* (simulated by a classical computer based on rules of quantum mechanics).



- A **qubit** (spin) is contained in an apparatus.
- The **apparatus** is a black box with a window that displays the result of the measurement.
- The apparatus has an **orientation** in the space (indicated by the direction of △)
- The apparatus has two modes:
  - △: detached from the qubit (no readings in this case),
  - ▲: interacting with the qubit (to make measurement), result displayed.

  We found the following behaviors:
- The apparatus only has *two* possible outcomes $\sigma = +1$ and $\sigma = -1 \Rightarrow$ A qubit is a **two-state system**.
- After a measurement, without disturbing the qubit, if we make the measurement again, **same** result will be obtained $\Rightarrow$ an isolated qubit has no dynamics, it acts as a **quantum memory**.
  - This is good, we can *confirm* the result of an experiment (otherwise we could learn nothing).
  - Initial *measurement* **prepares** the qubit in one of the two states.
  - Subsequent *measurement* **confirm** that state.
- Flip the apparatus upside down $\Rightarrow$ get **opposite** reading $\sigma \to -\sigma \Rightarrow$ we might conclude σ is a degree of freedom associated with a sense of *direction* in the space $\Rightarrow$ conjecture: the spin observable

  $$\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z) \tag{1}$$

  should be an **oriented vector** of some sort, we have measured one **component** of the vector along the axis set by the apparatus.

So far, no difference between classical and quantum physics.

- We should be able to measure σ^x by rotating the apparatus to the x-direction.
  - Classical: would get $\sigma^x = 0$,
  - Quantum: actually get $\sigma^x = \pm 1$ still! Moreover, the two out comes $\sigma^x = +1$ and $\sigma^x = -1$ appears **randomly**!

- We can repeat the procedure: *prepare* the qubit in $\sigma^z = +1$ state → rotate the apparatus along x-axis → *measure* σ^x.
  - Collect the results and analyze the statistics, we found

    $$p(\sigma^x = +1) = 1/2, \ p(\sigma^x = -1) = 1/2. \tag{2}$$

  - The **average of repeated measurements** is zero (we use $\langle * \rangle$ to denote the expectation value of an observable)

    $$\langle \sigma^x \rangle = (+1)(1/2) + (-1)(1/2) = 0. \tag{3}$$

    This matches with the *classical* expectation.
- The *measurement* of σ^x has *prepared* the qubit in either one of the $\sigma^x = \pm 1$ state. Now if we go back to measure σ^z, we get **random** results of $\sigma^z = \pm 1$, the *initial* $\sigma^z = +1$ state has been **destroyed by the measurement** of σ^x.
- If we *prepare* the qubit in $\sigma^z = +1$ state → measure σ along the direction of the unit vector $\boldsymbol{n} = (\sin\theta, 0, \cos\theta)$,
  - Classical: would get $\sigma = \cos\theta$,
  - Quantum: still get $\sigma = \pm 1$ randomly, but the statistics is biased, such that the average $\langle \sigma \rangle = \cos\theta$ matches the classical expectation.
- Even more general, if we *prepare* the $\sigma = +1$ state along unit vector $\boldsymbol{m}$ and *measure* σ along the unit vector $\boldsymbol{n}$, the result is still *randomly* $\sigma = \pm 1$, however the *average* is classical

  $$\langle \sigma \rangle = \boldsymbol{n} \cdot \boldsymbol{m}. \tag{4}$$

  Conclusion:
- *Quantum* systems are **not deterministic**, result of experiments can be statistically *random*.
- But if the *same* experiment is repeated *many times*, the **expectation value** can match the *classical* physics.
- Experiments are not *gentle*. **Measurement** can *change* the **quantum state**.

Question: Can we build a mathematical model to consistently describe the experimental properties of a qubit?

State and Representation

Qubit State

- We denote a **quantum state** by a **ket-vector** (or **ket**) $|\psi\rangle$. It could be considered as a mathematical object containing the data which is sufficient to describe all measurable properties of the state.
- Take a **qubit** for example, suppose we place the *apparatus* along the z-axis and make *measurement*,

- If the outcome is $\sigma^z = +1$, we say that the qubit has been *prepared* to the **up-spin** state, denoted as $|\uparrow\rangle$.
- If the outcome is $\sigma^z = -1$, we say that the qubit has been *prepared* to the **down-spin** state, denoted as $|\downarrow\rangle$.

- By calling a **ket** $|\psi\rangle$ as a vector, it can indeed be **represented** as a **column vector**.
  - For example, we can *choose* a **basis** (like a *coordinate system*) and write

$$|\uparrow\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; |\downarrow\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{5}$$

  - $\doteq$ implies the representation is **basis dependent** and may change if we view the same state in a different basis.
  - The *vector representation* of a *quantum state* is also called a **state vector**.
  - By saying that a qubit is a **two-state system**, its *state vector* has *two components*. Each component is a *complex number*.

- The *state vector* $|\psi\rangle$ of a *qubit* is **different** from the *spin vector* $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ that describes the spin orientation.
  - For example,

$$\begin{array}{lll} |\psi\rangle & \text{rep.} & \langle\boldsymbol{\sigma}\rangle \\ \hline |\uparrow\rangle & \begin{pmatrix} 1 \\ 0 \end{pmatrix} & (0, 0, +1) \\ |\downarrow\rangle & \begin{pmatrix} 0 \\ 1 \end{pmatrix} & (0, 0, -1) \end{array} \tag{6}$$

  - The components of the state vector are **complex** (in general), while the components of $\langle\boldsymbol{\sigma}\rangle$ are **real**.
  - But the information about $\langle\boldsymbol{\sigma}\rangle$ (*3 real* numbers) is fully **encoded** in the state vector $|\psi\rangle$(*2 complex = 4 real* numbers) in an implicit way (which we will analyze later).

- Similar to a vector, a **ket** $|\psi\rangle$ admits the following two basic mathematical operations
  - **Scalar multiplication**: $|\psi\rangle \mapsto z\,|\psi\rangle$ ($z \in \mathbb{C}$). For example

$$\begin{aligned} |A\rangle &= z_1\,|\uparrow\rangle \doteq z_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} z_1 \\ 0 \end{pmatrix}, \\ |B\rangle &= z_2\,|\downarrow\rangle \doteq z_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ z_2 \end{pmatrix}. \end{aligned} \tag{7}$$

  - **Addition**: $|A\rangle, |B\rangle \mapsto |A\rangle + |B\rangle$. For example

$$|A\rangle + |B\rangle \doteq \begin{pmatrix} z_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ z_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}. \tag{8}$$

- Put together, *multiplying* states by *complex scalars* and then *adding* them together, the combined operation is called a **linear superposition** of the states.

- *Linear superposition* of quantum states of a system is still a quantum state of the same system.
- For example, a generic **qubit state**,

$$\boxed{|\psi\rangle = \psi_\uparrow |\uparrow\rangle + \psi_\downarrow |\downarrow\rangle \doteq \begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix}.} \tag{9}$$

- The *complex vector space* where the *state vector* lives in is called the **Hilbert space**. It is the *space* of **quantum states**.
  - The *qubit* has a **two-dimensional Hilbert space** ⇔ all possible qubit (spin) state can be **represented** as a **two-component** complex vector.
  - The **dimension** of the *Hilbert space* is the **number** of *basis states* that span the Hilbert space.

Statistical Interpretation

So the quantum state of a qubit is fully described by *two complex numbers* $\psi_\uparrow$ and $\psi_\downarrow$. What are their physical interpretations?

Given a spin that has been *prepared* in the state $|\psi\rangle = \psi_\uparrow |\uparrow\rangle + \psi_\downarrow |\downarrow\rangle$, and that the *apparatus* is *oriented* along z-axis,

- The quantity $\psi_\uparrow^* \psi_\uparrow \equiv |\psi_\uparrow|^2$ is the **probability** that the spin would be *measured* to be $\sigma^z = +1$. It is the probability of the spin being *up* if measured along z-axis.
- Likewise, $\psi_\downarrow^* \psi_\downarrow \equiv |\psi_\downarrow|^2$ is the **probability** the spin being *down* ($\sigma^z = -1$) if measured along z-axis.

Because the apparatus has only two outcomes $\sigma^z = \pm 1$, it is a convention to have the *probabilities adding up to* 1.

$$\boxed{|\psi_\uparrow|^2 + |\psi_\downarrow|^2 = 1.} \tag{10}$$

This is the **normalization condition** of the state vector. A state vector satisfying this condition is said to be **normalized**, otherwise we say it is **unnormalized**. In most cases, we deal with normalized states, but unnormalized states are also useful in quantum information.

Now we had a better understanding of why the representation in Eq. (5) was chosen. If the qubit is prepared to the $|\uparrow\rangle$ state, in the *subsequent measurement* of σ^z, we will get $\sigma^z = +1$ with probability 1, and $\sigma^z = -1$ with probability 0, so $|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is a valid choice. Similar argument for $|\downarrow\rangle$.

What about $\psi_\uparrow^* \psi_\downarrow$ or $\psi_\downarrow^* \psi_\uparrow$?

- First we identify that there is only *one* remaining piece of information in there, which is a **relative phase factor** $e^{i\varphi}$ between $\psi_\uparrow$ and $\psi_\downarrow$,

$$\psi_\uparrow^* \psi_\downarrow = |\psi_\uparrow| |\psi_\downarrow| e^{i\varphi}, \;\; \psi_\downarrow^* \psi_\uparrow = |\psi_\uparrow| |\psi_\downarrow| e^{-i\varphi}. \tag{11}$$

- The *amplitude* $|\psi_\uparrow| |\psi_\downarrow|$ becomes *large* when the spin is *not predominantly* in either $|\uparrow\rangle$ or $|\downarrow\rangle$ (along z-axis) ⇒ then it is likely to lie in the xy-plane if measured.
- The *phase angle* φ parameterize the **polar angle** in the xy-plane along which the spin is likely to orient.
- The information about $\langle\sigma^x\rangle$ and $\langle\sigma^y\rangle$ is stored in $\psi_\uparrow^* \psi_\downarrow$ (a kind of *interrelation* between $\psi_\uparrow$ and $\psi_\downarrow$).

We have discussed about the meaning of $|\psi_\uparrow|$, $|\psi_\downarrow|$ and φ. Those are just three real parameters, but the state vector $|\psi\rangle$ has *two complex* = *four real* components.

What is the fourth real parameter?

It turns out to be an **overall phase factor**, which can be changed by

$$|\psi\rangle \mapsto e^{i\theta} |\psi\rangle. \tag{12}$$

- The *overall phase* is an **redundancy** in the description.
- There should be *no physical meaning* associated with the *overall phase* of the state (jargon: the overall phase is a **gauge freedom**).

▪ Inner Product

- For each **ket-vector** $|\psi\rangle$, there is a **dual vector**, called the **bra-vector** $\langle\psi|$, living in the **dual Hilbert space**.
  - The *bra-vector* can be **represented** as a **row vector**, *conjugate transpose* to the *ket-vector*.

$$|\psi\rangle = \psi_\uparrow |\uparrow\rangle + \psi_\downarrow |\downarrow\rangle \doteq \begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix} \Rightarrow \langle\psi| = \psi_\uparrow^* \langle\uparrow| + \psi_\downarrow^* \langle\downarrow| \doteq (\psi_\uparrow^* \;\; \psi_\downarrow^*). \tag{13}$$

  - The names *bra* and *ket* come from *bra-ket* (or bracket) $\langle\psi|\phi\rangle$, which represents the **inner product** of two states $|\psi\rangle$ and $|\phi\rangle$.

$$\boxed{\langle\psi \mid \phi\rangle \doteq (\psi_1^* \;\; \psi_2^* \;\; \ldots) \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \end{pmatrix} = \psi_1^* \phi_1 + \psi_2^* \phi_2 + \ldots = \sum_i \psi_i^* \phi_i.} \tag{14}$$

  - *Interchange* bras and kets corresponds to **complex conjugation**,

$$\boxed{\langle\psi \mid \phi\rangle = \langle\phi \mid \psi\rangle^*.} \tag{15}$$

- **Normalized state**: a state $|\psi\rangle$ is **normalized** ⇔ Its *inner product* with *itself* is *one*, $\langle\psi \mid \psi\rangle = 1$.
  - For example, the *normalization condition* Eq. (10) can be written as

$$\langle\psi \mid \psi\rangle \doteq (\psi_\uparrow^* \;\; \psi_\downarrow^*) \begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix} = \psi_\uparrow^* \psi_\uparrow + \psi_\downarrow^* \psi_\downarrow = 1. \tag{16}$$

- $|\uparrow\rangle$ and $|\downarrow\rangle$ are normalized, because $\langle\uparrow\mid\uparrow\rangle = \langle\downarrow\mid\downarrow\rangle = 1$.

- **Orthogonal states**: two states $|\psi\rangle$ and $|\phi\rangle$ are **orthogonal** to each other $\Leftrightarrow$ their *inner product* is *zero*, $\langle\psi|\,\phi\rangle = 0$.

  - For example, $|\uparrow\rangle$ and $|\downarrow\rangle$ are *orthogonal*,

$$\langle\uparrow\mid\downarrow\rangle \doteq (\,1\;\;0\,)\begin{pmatrix}0\\1\end{pmatrix} = 0. \tag{17}$$

  By Eq. (15), also vanishes.

  - $|\uparrow\rangle$ and $|\downarrow\rangle$ are *orthogonal* for a good reason: they are **distinct** states of a qubit, i.e. if the spin is *up*, it is definitely *not down*, vice versa.

Inner product allows us to do calculation on the abstract level (without involving vectors explicitly).

$$\begin{aligned}\langle\psi\mid\psi\rangle &= (\psi_\uparrow^*\,\langle\uparrow| + \psi_\downarrow^*\,\langle\downarrow|)\,(\psi_\uparrow\,|\uparrow\rangle + \psi_\downarrow\,|\downarrow\rangle)\\ &= \psi_\uparrow^*\,\psi_\uparrow\,\langle\uparrow\mid\uparrow\rangle + \psi_\uparrow^*\,\psi_\downarrow\,\langle\uparrow\mid\downarrow\rangle + \psi_\downarrow^*\,\psi_\uparrow\,\langle\downarrow\mid\uparrow\rangle + \psi_\downarrow^*\,\psi_\downarrow\,\langle\downarrow\mid\downarrow\rangle\\ &= \psi_\uparrow^*\,\psi_\uparrow + \psi_\downarrow^*\,\psi_\downarrow = 1.\end{aligned} \tag{18}$$

- **Orthonormal basis**: a *complete* set of **normalized** states $|i\rangle$ which are also **orthogonal** to each other and span the Hilbert space (meaning that there will be no more candidate state in the Hilbert space that is orthogonal to all of the current basis states).

$$\boxed{\langle i\mid j\rangle = \delta_{ij} = \begin{cases}1 & i = j,\\ 0 & i \neq j.\end{cases}} \tag{19}$$

  - Example: $|\uparrow\rangle$ and $|\downarrow\rangle$ form an *orthonormal basis* of the *qubit* Hilbert space.
  - The **dimension** of the *Hilbert space* = the **number** of *basis states*.
  - Every state $|\psi\rangle$ in the Hilbert space can be written as a *linear superposition* of *orthonormal basis states*,

$$\boxed{|\psi\rangle = \psi_1\,|1\rangle + \psi_2\,|2\rangle + \ldots = \sum_i \psi_i\,|i\rangle.} \tag{20}$$

  - The *superposition coefficient* ψ_i are the **components** of the state vector, which can be *extracted* by the *inner product* with the basis state,

$$\boxed{\psi_i = \langle i\mid\psi\rangle.} \tag{21}$$

  - Eq. (20) and Eq. (21) can be written in a more elegant form in terms of bras and kets only

$$\boxed{|\psi\rangle = \sum_i |i\rangle\,\langle i\mid\psi\rangle.} \tag{22}$$

It could be helpful to check these statement explicitly by choosing an explicit *vector* *representations*

$$|1\rangle \doteq \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix},\ |2\rangle \doteq \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix},\ |3\rangle \doteq \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix},\ \ldots. \tag{23}$$

But such approach is not necessary. The bra-ket notation is powerful in that we will not need to work with vector representations explicitly.

HW 1

> Let us choose a different representation for the qubit, say,
> $|0\rangle \doteq \begin{pmatrix} e^{i\varphi/2}\cos\theta/2 \\ e^{-i\varphi/2}\sin\theta/2 \end{pmatrix},\ |1\rangle \doteq \begin{pmatrix} -e^{i\varphi/2}\sin\theta/2 \\ e^{-i\varphi/2}\cos\theta/2 \end{pmatrix},$
> where θ and φ are arbitrary real angles. Show that $|0\rangle$ and $|1\rangle$ form an orthonormal basis (for any choices of θ and φ).

Solution (HW 1)

▪ States Along Other Axes

Define the following qubit states

- Set the apparatus along z-axis, measure σ^z,

$$\sigma^z = \begin{cases} +1 & |\uparrow\rangle, \\ -1 & |\downarrow\rangle. \end{cases} \tag{25}$$

- Set the apparatus along x-axis, measure σ^x,

$$\sigma^x = \begin{cases} +1 & |\rightarrow\rangle, \\ -1 & |\leftarrow\rangle. \end{cases} \tag{26}$$

- Set the apparatus along y-axis, measure σ^y,

$$\sigma^y = \begin{cases} +1 & |\otimes\rangle, \\ -1 & |\odot\rangle. \end{cases} \tag{27}$$

They are three sets of *orthonormal basis*, each can be represented in the other two basis.

Let us represent the states in the σ^z basis

$$\begin{aligned} |\rightarrow\rangle &= \frac{1}{\sqrt{2}}|\uparrow\rangle + \frac{1}{\sqrt{2}}|\downarrow\rangle \doteq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\ |\leftarrow\rangle &= \frac{1}{\sqrt{2}}|\uparrow\rangle - \frac{1}{\sqrt{2}}|\downarrow\rangle \doteq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned} \tag{28}$$

$$\begin{aligned} |\otimes\rangle &= \frac{1}{\sqrt{2}}|\uparrow\rangle + \frac{i}{\sqrt{2}}|\downarrow\rangle \doteq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}, \\ |\odot\rangle &= \frac{1}{\sqrt{2}}|\uparrow\rangle - \frac{i}{\sqrt{2}}|\downarrow\rangle \doteq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}. \end{aligned} \tag{29}$$

The vector representation is *not unique*, but nevertheless, an explicit representation is always useful in helping us to gain some intuition.

Summary

Much of the toy experiment of the qubit can be understood in the framework

- As we measure σ^z and get $\sigma^z = +1$, we have prepare the qubit in the $|\uparrow\rangle$ state.
- Subsequent measurement will confirm $\sigma^z = +1$ with probability 1.
- When the apparatus is flipped upside down, *relative to the apparatus*, the qubit state rotates by

$$|\uparrow\rangle \to |\downarrow\rangle,\ |\downarrow\rangle \to -|\uparrow\rangle. \tag{30}$$

So the measurement outcome is $\sigma = -1$ with probability 1.

- When the apparatus is set along the x-axis, we can use

$$|\uparrow\rangle = \frac{1}{\sqrt{2}}|\rightarrow\rangle + \frac{1}{\sqrt{2}}|\leftarrow\rangle \tag{31}$$

to explain that we will measure either $\sigma^x = +1$ or $\sigma^x = -1$ with equal probability (both probability $= 1/2$).

- After the measurement of σ^x, suppose we get $\sigma^x = -1$, the quantum state **collapses** to $|\leftarrow\rangle$, then in the subsequent measurement of σ^z, we use

$$|\leftarrow\rangle = \frac{1}{\sqrt{2}}|\uparrow\rangle - \frac{1}{\sqrt{2}}|\downarrow\rangle \tag{32}$$

to explain that we will get either $\sigma^z = +1$ or $\sigma^z = -1$ with equal probability.

What is a quantum state **collapse**? How does it happen?

This is still an open question at the frontier of research. What we currently know

- **Measurement** is a kind of **interaction** between the qubit and the apparatus.
  - The interaction **entangles** (we will discuss this later) the *qubit* and the *apparatus* together, and the **quantum information** about the original qubit spreads to the *apparatus* and maybe further spreads to its *embedding environment*.
  - Depending on the interaction details,
    - Some information (such as the **quantum coherence**) gets **globally scrambled** into the environment, and is no longer retrievable (by local observers).
    - Some information (such as the **measurement outcome**) gets **locally duplicated** in the environment, and emerges as a **classical reality**.
- The randomness in the **quantum state collapse** originates from **quantum information scrambling**.

- The quantum information that scrambles into the environment is *effectively lost*, since we can not afford the huge computational effort the *decode* it. The **loss** of **quantum information** creates **entropy**, a.k.a. **ignorance**, a.k.a. **randomness**.
- The randomness in quantum mechanics may be a "illusion" of limited **quantum computational resources**. "Our resources limit our understanding". Given such limitation, we have to adopt a probabilistic description in quantum mechanics (Similar philosophy applies to statistical mechanics)

Quantum Operators

Hermitian Operators

How Operator Works?

Axioms of Quantum Mechanics (two of five)

Axiom 1 (States): **States** of a quantum system are described as (rays of) **vectors** in the associated Hilbert space.

Axiom 2 (Observables): **Physical observables** of a quantum system are described by **Hermitian operators** (represented by **Hermitian matrices**) acting on the associated Hilbert space.

Observables are things that we can *measure*. *Operators* are what we apply to a state to "*modify*" the state. How can these two seemly different concepts be related?

Well, let us first understand how operator works?

- An operator M (like a "machine") takes a state $|\psi\rangle$ and returns another state $|\phi\rangle$:

$$M\,|\psi\rangle = |\phi\rangle. \tag{33}$$

  - An operator is said to be **linear**, if it *preserves* the *linearity* of the state, i.e. $M(z_1\,|\psi\rangle + z_2\,|\phi\rangle) = z_1\,M\,|\psi\rangle + z_2\,M\,|\phi\rangle$.
  - In general, an **linear operator** can be written as a *linear superposition* of **basis operators** $|i\rangle\langle j|$ and can be represented as a **matrix**,

$$\boxed{M = \sum_{ij} |i\rangle\, M_{ij}\,\langle j| \doteq \begin{pmatrix} M_{11} & M_{12} & \cdots \\ M_{21} & M_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}} \tag{34}$$

  - Each matrix element M_{ij} is a *complex* number (in general).

Take a qubit for example, there are four *basis operators* $|i\rangle\langle j|$

$$|\uparrow\rangle\langle\uparrow| \doteq \begin{pmatrix}1\\0\end{pmatrix}(1\ \ 0) = \begin{pmatrix}1&0\\0&0\end{pmatrix},$$
$$|\uparrow\rangle\langle\downarrow| \doteq \begin{pmatrix}1\\0\end{pmatrix}(0\ \ 1) = \begin{pmatrix}0&1\\0&0\end{pmatrix},$$
$$|\downarrow\rangle\langle\uparrow| \doteq \begin{pmatrix}0\\1\end{pmatrix}(1\ \ 0) = \begin{pmatrix}0&0\\1&0\end{pmatrix},$$
$$|\downarrow\rangle\langle\downarrow| \doteq \begin{pmatrix}0\\1\end{pmatrix}(0\ \ 1) = \begin{pmatrix}0&0\\0&1\end{pmatrix}.$$

Each basis operator implements a "basic operation", e.g. $|\downarrow\rangle\langle\uparrow|$ takes the up-spin state $|\uparrow\rangle$ and returns the down-spin state $|\downarrow\rangle$. Any linear operator of a qubit will be a superposition of these four *basis operators*.

$$\begin{aligned}M &= M_{\uparrow\uparrow}|\uparrow\rangle\langle\uparrow| + M_{\uparrow\downarrow}|\uparrow\rangle\langle\downarrow| + M_{\downarrow\uparrow}|\downarrow\rangle\langle\uparrow| + M_{\downarrow\downarrow}|\downarrow\rangle\langle\downarrow|\\ &\doteq \begin{pmatrix}M_{\uparrow\uparrow}&M_{\uparrow\downarrow}\\M_{\downarrow\uparrow}&M_{\downarrow\downarrow}\end{pmatrix}.\end{aligned}\tag{36}$$

- Applying an *operator* to a *state* $\doteq$ multiplying a *matrix* to a *vector*. Consider the *vector* representations of *states*

$$\begin{aligned}|\psi\rangle &= \sum_i \psi_i |i\rangle \doteq \begin{pmatrix}\psi_1\\\psi_2\\\vdots\end{pmatrix},\\ |\phi\rangle &= \sum_i \phi_i |i\rangle \doteq \begin{pmatrix}\phi_1\\\phi_2\\\vdots\end{pmatrix},\end{aligned}\tag{37}$$

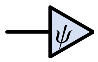
the two sides of Eq. (33) are

$$\begin{aligned}M|\psi\rangle &= \sum_{ij}|i\rangle M_{ij}\langle j|\sum_k \psi_k|k\rangle\\ &= \sum_{i\,j} M_{ij}\,\psi_j\,|i\rangle,\\ |\phi\rangle &= \sum_i \phi_i\,|i\rangle,\end{aligned}\tag{38}$$

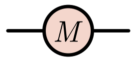
which will match iff

$$\begin{aligned}\phi_i &= \sum_j M_{ij}\,\psi_j,\\ \begin{pmatrix}\phi_1\\\phi_2\\\vdots\end{pmatrix} &= \begin{pmatrix}M_{11}&M_{12}&\cdots\\M_{21}&M_{22}&\cdots\\\vdots&\vdots&\ddots\end{pmatrix}\begin{pmatrix}\psi_1\\\psi_2\\\vdots\end{pmatrix}.\end{aligned}\tag{39}$$

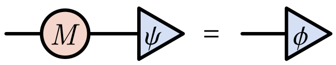
- **Tensor network**: a diagrammatic representation of tensor contractions
  - Each object is a **tensor** (multi-dimensions array).
  - **Vectors** are *rank-1 tensors*, represented by an object with *one* leg



- **Matrices** are *rank-2 tensors*, represented by an object with *two* legs



- **Tensor contraction**: indices on *internal legs* are automatically summed over. For example, matrix-vector multiplication can be expressed as a tensor contraction.



- On an orthonormal basis, the *matrix elements* of an operator M can be extracted by

$$\boxed{M_{ij} = \langle i|\, M\, |j\rangle,} \tag{40}$$

because the following identity holds

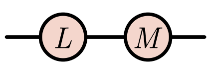
$$M = \sum_{ij} |i\rangle\, \langle i|\, M\, |j\rangle\, \langle j|, \tag{41}$$

given that $\sum_i |i\rangle\, \langle i| = \mathbb{1}$ is an **identity operator**. This trick is commonly used to find representations of states and operators, and is called the *resolution of identity*. See also Eq. (21).

- **Composition** of *operators*: one operation following by another (from *right* to *left*)

$$\begin{aligned} L\, M &= \sum_{ij} |i\rangle\, L_{ij}\, \langle j| \sum_{kl} |k\rangle\, M_{kl}\, \langle l| \\ &= \sum_{ij} |i\rangle \left(\sum_k L_{ik}\, M_{kj}\right) \langle j|. \end{aligned} \tag{42}$$

- Composing two *operators* ≐ multiplying two *matrices*.

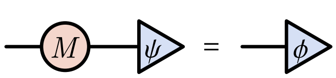


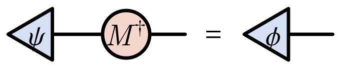
▪ Hermitian Conjugate

We have talked about how an operator acts on a *ket-vector* $|\psi\rangle$, what about its action on the *bra-vector* $\langle\psi|$?

Hilbert space	⇒	**dual Hilbert space**
ket–state $|\psi\rangle$	⇒	**bra–state** $\langle\psi|$
operator M	⇒	**Hermitian conjutate** operator $M^\dagger$

- If $M\, |\psi\rangle = |\phi\rangle$ then $\langle\psi|\, M^\dagger = \langle\phi|$ (which defines $M^\dagger$ as a dual/conjugate of M).
- In terms of *tensor networks*, this corresponds to *flipping* tensors around.





Recall from Eq. (13):

$$
\begin{aligned}
&|\psi\rangle = \sum_i \psi_i\,|i\rangle \doteq \begin{pmatrix}\psi_1\\ \psi_2\\ \vdots\end{pmatrix}\\
&\Rightarrow \langle\psi| = \sum_i \langle i|\,\psi_i^* \doteq (\,\psi_1^*\;\;\psi_2^*\;\;\cdots\,),
\end{aligned}
\tag{43}
$$

the way to get $\langle\psi|\,M^\dagger = \langle\phi|$ is to define

$$
\begin{aligned}
&M = \sum_{ij} |i\rangle\,M_{ij}\,\langle j| \doteq \begin{pmatrix} M_{11} & M_{12} & \cdots\\ M_{21} & M_{22} & \cdots\\ \vdots & \vdots & \ddots\end{pmatrix}\\
&\Rightarrow M^\dagger = \sum_{ij} |i\rangle\,M_{ji}^*\,\langle j| \doteq \begin{pmatrix} M_{11}^* & M_{21}^* & \cdots\\ M_{12}^* & M_{22}^* & \cdots\\ \vdots & \vdots & \ddots\end{pmatrix},
\end{aligned}
\tag{44}
$$

such that

$$
\begin{aligned}
&\langle\psi|\,M^\dagger = \sum_i \langle i|\,\psi_i^* \sum_{jk} |j\rangle\,M_{kj}^*\,\langle k|\\
&= \sum_k \phi_k^*\,\langle k| = \langle\phi|.
\end{aligned}
\tag{45}
$$

where Eq. (39) was used in the form of

$$
\begin{aligned}
&\phi_k^* = \sum_j M_{kj}^*\,\psi_j^*,\\
&(\,\phi_1^*\;\;\phi_2^*\;\;\cdots\,) = (\,\psi_1^*\;\;\psi_2^*\;\;\cdots\,)\begin{pmatrix} M_{11}^* & M_{21}^* & \cdots\\ M_{12}^* & M_{22}^* & \cdots\\ \vdots & \vdots & \ddots\end{pmatrix}.
\end{aligned}
\tag{46}
$$

In terms of *matrix* representation, **Hermitian conjugate** acts as

- **matrix transpose** (interchanges the rows and columns),
- followed by **complex conjugation** of each matrix element.

How to think of it: *Hermitian conjugate* ~ a generalization of *complex conjugate* from complex *numbers* to *matrices*.

Hermitian conjugate has the following properties:

- **Duality**: suppose A is an operator

$$\left(A^\dagger\right)^\dagger = A. \tag{47}$$

- **Linearity**: suppose A and B are operators, a and b are complex numbers,

$$(a\,A + b\,B)^\dagger = a^*\,A^\dagger + b^*\,B^\dagger. \tag{48}$$

- **Factor reversal**: suppose A and B are operators

$$(A\,B)^\dagger = B^\dagger\,A^\dagger. \tag{49}$$

Hermitian Operator

Real numbers play a special role in physics. The results of any measurements are real. If in quantum mechanics, physical observables are represented by *operators*, how do we impose the “reality” condition on operators?

- A **real number** is a number whose *complex conjugation* is itself.
- A ~~real operator~~ **Hermitian operator** is an linear operator whose *Hermitian conjugate* is itself.

For example, if $L = \sum_{ij} |i\rangle\, L_{ij}\, \langle j|$ is *Hermitian*, then

$$L = L^\dagger, \tag{50}$$

or in terms of matrix elements,

$$L_{ij} = L_{ji}^*. \tag{51}$$

- Given a complex number z, real part: $\mathrm{Re}\, z = (z + z^*)/2$, imaginary part: $\mathrm{Im}\, z = (z - z^*)/(2\,i)$. Similarity, given a linear operator M

$$\mathrm{Re}\, M = \frac{1}{2}\left(M + M^\dagger\right),\ \mathrm{Im}\, M = \frac{1}{2\,i}\left(M - M^\dagger\right). \tag{52}$$

  - Both Re M and Im M are Hermitian operators.

Eigenvalues and Eigenvectors

In general, a linear operator acting on a state will change the state. But for a fixed linear operator M, there can be special states $|\mu\rangle$ that remain the same under the operation. The only effect of M on these states is to rescale them by an overall factor μ (can be complex).

$$\boxed{M\,|\mu\rangle = \mu\,|\mu\rangle.} \tag{53}$$

- the μ (outside the ket) is a *number*, indicating how much the vector is rescaled under the action of M. This number is an **eigenvalue** of the operator.
- $|\mu\rangle$ is an **eigenvector** that is associated with its eigenvalue μ.

Given the matrix representation of an operator, its eigenvalues and eigenvectors can be found ~~by solving the eigen equation~~ by *Mathematica*.

```
Eigensystem[{{0, -1}, {1, 0}}]
```

```
{{i, -i}, {{i, 1}, {-i, 1}}}
```

- For bra vectors,

$$M\,|\mu\rangle = \mu\,|\mu\rangle \Rightarrow \langle\mu|\,M^\dagger = \langle\mu|\,\mu^*. \tag{54}$$

What is special about Hermitian operators?

- *Eigenvalues* of a Hermitian operator are **real**.
- *Eigenvectors* of a Hermitian operator for a **complete** set of basis. (Any vector can be expanded as a sum of these eigenvectors.)
  - If $\lambda_1 \neq \lambda_2$ are two *unequal* eigenvalues of a Hermitian operator, then their corresponding eigenvectors $|\lambda_1\rangle$ and $|\lambda_2\rangle$ are orthogonal (automatically).
  - Eigenvectors of the *same* eigenvalue can be made orthogonal (by orthogonalization, e.g. Gram-Schmidt procedure).

```
Orthogonalize[{{1, 2}, {3, 4}}]
```

$$\left\{\left\{\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right\}, \left\{\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right\}\right\}$$

  - For *bounded* Hermitian operators (e.g. finite matrices in finite dimensional Hilbert space), eigenvectors can be *normalized*.
- In conclusion, each **Hermitian operator** generates a set of **complete** and **orthonormal** basis for Hilbert space. The set of basis is also called the **eigenbasis** of a Hermitian operator.

Suppose L is Hermitian ($L = L^\dagger$) and

$$\begin{aligned} L\,|\lambda_1\rangle &= \lambda_1\,|\lambda_1\rangle, \\ L\,|\lambda_2\rangle &= \lambda_2\,|\lambda_2\rangle. \end{aligned} \tag{55}$$

We can flip the first equation $\langle\lambda_1|\,L^\dagger = \langle\lambda_1|\,L = \langle\lambda_1|\,\lambda_1^*$,

$$\begin{aligned} \langle\lambda_1|\,L\,|\lambda_2\rangle &= \lambda_1^*\,\langle\lambda_1 \mid \lambda_2\rangle, \\ \langle\lambda_1|\,L\,|\lambda_2\rangle &= \lambda_2\,\langle\lambda_1 \mid \lambda_2\rangle. \end{aligned} \tag{56}$$

- If $|\lambda_1\rangle = |\lambda_2\rangle$ (automatically implying $\lambda_1 = \lambda_2$), Eq. (56) implies $\langle\lambda|\,L\,|\lambda\rangle = \lambda^*\,\langle\lambda \mid \lambda\rangle = \lambda\,\langle\lambda \mid \lambda\rangle$, so λ is real.
- If $|\lambda_1\rangle$ and $|\lambda_2\rangle$ are two different (non-colinear) states,
  - with unequal eigenvalues $\lambda_1 \neq \lambda_2$, Eq. (56) implies $(\lambda_1 - \lambda_2)\,\langle\lambda_1 \mid \lambda_2\rangle = 0$, so $\langle\lambda_1 \mid \lambda_2\rangle = 0$.
  - but their eigenvalues $\lambda_1 = \lambda_2 = \lambda$ happen to be the same. In this case, $|\lambda_1\rangle$ and $|\lambda_2\rangle$ are **degenerate**. Degenerated states span a subspace, called the **degenerate subspace**. Any state in the degenerate subspace

$$|\lambda\rangle = z_1\,|\lambda_1\rangle + z_2\,|\lambda_2\rangle, \tag{57}$$

is an eigenvector of the Hermitian operator with the same eigenvalue λ, because

$$\begin{aligned} L\,|\lambda\rangle &= z_1\,L\,|\lambda_1\rangle + z_2\,L\,|\lambda_2\rangle \\ &= z_1\,\lambda\,|\lambda_1\rangle + z_2\,\lambda\,|\lambda_2\rangle \\ &= \lambda(z_1\,|\lambda_1\rangle + z_2\,|\lambda_2\rangle) \\ &= \lambda\,|\lambda\rangle. \end{aligned} \tag{58}$$

- *Hermitian operator* admits the following **spectral decomposition** in its own *eigenbasis*,

$$\boxed{L = \sum_i |\lambda_i\rangle\, \lambda_i\, \langle\lambda_i|.} \tag{59}$$

  - Note: unlike a generic matrix representation $L = \sum_{ij} |i\rangle\, l_{ij}\, \langle j|$, in the eigenbasis, the summation only run through the dimension of the Hilbert space once.
  - In the eigenbasis, the Hermitian operator is represented as a **diagonal matrix**. So the procedure of bring the *matrix* representation to its *diagonal* form by transforming to its *eigenbasis* is called **diagonalization**. (We will discuss more about it later.)

Measurement Postulate

Now we are well prepared to come back to Axiom 2.

Axiom 2 (Observables): **Physical observables** of a quantum system are described by **Hermitian operators** (represented by **Hermitian matrices**) acting on the associated Hilbert space.

Suppose we have a physical observable described the Hermitian operator L. It has a set of eigenvalues and eigenvectors

$$L = \sum_i |\lambda_i\rangle\, \lambda_i\, \langle\lambda_i|. \tag{60}$$

- The possible **outcomes** of a **measurement** are the **eigenvalues** λ_i. (Assuming they are *not degenerate* for now.)
- The **measurement** projects (collapses) the quantum state to the **eigenstate** $|\lambda_i\rangle$ that corresponds to the measurement outcome λ_i.

Now comes another **axiom of quantum mechanics**

Axiom 3 (Measurement): Given a quantum system in the **state** $|\psi\rangle$ and the **observable** L to be measured, the **probability** to observe the measurement outcome λ_i is $p(L = \lambda_i) = |\langle\lambda_i \mid \psi\rangle|^2$.

  - No way to tell for certain which outcome will be observed. There is only a probability $p(\lambda_i)$.
  - Probability is given by the *square* of the overlap. Why the square? Probability must be (i) real and positive, (ii) "gauge invariant" (i.e. independent of the overall phase of either states).
  - *Subsequent measurement* must *confirm* the result. ⇒ After the initial measurement, the state must have been collapsed to the eigenstate $|\lambda_i\rangle$ (but how?).

What if there is a *degenerate subspace* corresponding to the eigen value λ?

- **Projection operator** that projects to the eigenspace of L associated with the eigenvalue λ

$$P(L=\lambda) = \sum_{\lambda_i} |\lambda_i\rangle\, \delta(\lambda_i - \lambda)\, \langle\lambda_i|,$$
$$\delta(\lambda_i - \lambda) = \begin{cases} 1 & \lambda_i = \lambda, \\ 0 & \lambda_i \neq \lambda. \end{cases} \tag{61}$$

- The probability to observe the measurement outcome $L = \lambda$ will be

$$\boxed{p(L=\lambda) = \langle\psi|\, P(L=\lambda)\, |\psi\rangle.} \tag{62}$$

- If the outcome λ is observed, the state must have collapsed to

$$\boxed{|\psi\rangle \xrightarrow{\text{measure } L,\text{ get } \lambda} \frac{P(L=\lambda)\,|\psi\rangle}{\langle\psi|\, P(L=\lambda)\, |\psi\rangle^{1/2}}.} \tag{63}$$

- **Expectation value** of the observable. The averaged measurement outcome over many repeated experiments (initial state must be prepared each time). By definition and use $p(L=\lambda_i) = |\langle\lambda_i \mid \psi\rangle|^2$

$$\langle L\rangle = \sum_i \lambda_i\, p(L=\lambda_i) = \sum_i \langle\psi \mid \lambda_i\rangle\, \lambda_i\, \langle\lambda_i \mid \psi\rangle, \tag{64}$$

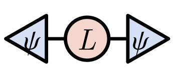
given $L = \sum_i |\lambda_i\rangle\, \lambda_i\, \langle\lambda_i|$ we have

$$\boxed{\langle L\rangle = \langle\psi|\, L\, |\psi\rangle.} \tag{65}$$

- The answer is a *real* scalar (as L is Hermitian).
- Represented as *vectors* and *matrices*,

$$\begin{pmatrix} \psi_1^* & \psi_2^* & \cdots \end{pmatrix} \begin{pmatrix} L_{11} & L_{12} & \cdots \\ L_{21} & L_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix}, \tag{66}$$

- or in terms of *tensor network*,



HW 2

(i) Construct the projection operators $P(\sigma^x = \pm 1)$ as 2×2 matrices in the $\{\quad,\quad\}$ basis.
(ii) Use the projection operator to calculate the probability $p(\sigma^x = \pm 1)$ of obtaining ± 1 outcome when σ^x is measured on the state.
(iii) Measuring σ^x on the state, if the measurement outcome turns out to be -1, compute the post-measurement state that the qubit collapses to.

Solution (HW 2)

▪ Example: Single-Qubit Operators

For a single qubit (spin), the physical observables are $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$.

- Each **observable** corresponds to a **Hermitian operator** acting in the 2-dimensional Hilbert space.
- In the $|\uparrow\rangle$ and $|\downarrow\rangle$ basis, their **matrix** representations are

$$\begin{aligned}\sigma^x &\doteq \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\\ \sigma^y &\doteq \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},\\ \sigma^z &\doteq \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.\end{aligned} \tag{74}$$

These matrices are called **Pauli matrices**.

- They are all *Hermitian* matrices.
- Their *eigenvectors* are given by Eq. (5), Eq. (28), and Eq. (29)

$$\begin{aligned}&|\sigma^x = +1\rangle \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix},\ |\sigma^x = -1\rangle \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -1\end{pmatrix};\\ &|\sigma^y = +1\rangle \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix},\ |\sigma^y = -1\rangle \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix};\\ &|\sigma^z = +1\rangle \doteq \begin{pmatrix}1\\ 0\end{pmatrix},\ |\sigma^z = -1\rangle \doteq \begin{pmatrix}0\\ 1\end{pmatrix}.\end{aligned} \tag{75}$$

Each set of eigenvectors form a set of *complete* and *orthonormal* basis of the qubit Hilbert space.

- Their corresponding *eigenvalues* are all ± 1: no matter we measure the qubit along x, y, z directions, we only get to possible outcomes ± 1.

▫ Math: Pauli Algebra

- Multiplication table, based on Eq. (74),

$$\begin{array}{c|cccc} & \mathbb{1} & \sigma^x & \sigma^y & \sigma^z \\ \hline \mathbb{1} & \mathbb{1} & \sigma^x & \sigma^y & \sigma^z \\ \sigma^x & \sigma^x & \mathbb{1} & i\,\sigma^z & -i\,\sigma^y \\ \sigma^y & \sigma^y & -i\,\sigma^z & \mathbb{1} & i\,\sigma^x \\ \sigma^z & \sigma^z & i\,\sigma^y & -i\,\sigma^x & \mathbb{1}\end{array} \tag{76}$$

Product of two Pauli matrices (treated as definition)

$$\boxed{\sigma^i\,\sigma^j = \delta^{ij}\,\mathbb{1} + i\,\epsilon^{ijk}\,\sigma^k,} \tag{77}$$

where i, j, $k = 1, 2, 3$ (stands for x, y, z).

- Another version of Eq. (77) using vector notation

$$\boldsymbol{a}\cdot\boldsymbol{\sigma}\,\boldsymbol{b}\cdot\boldsymbol{\sigma} = \boldsymbol{a}\cdot\boldsymbol{b}\,\mathbb{1} + i\,(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{\sigma}, \tag{78}$$

where $\boldsymbol{a}$, $\boldsymbol{b}$ are three-component vectors (each component is a scalar). Here $\boldsymbol{a}\cdot\boldsymbol{\sigma}$ means

$$\boldsymbol{a}\cdot\boldsymbol{\sigma} = a_1\,\sigma^1 + a_2\,\sigma^2 + a_3\,\sigma^3. \tag{79}$$

- Trace of Pauli matrices

$$\mathrm{Tr}\,\mathbb{1} = 2,\ \ \mathrm{Tr}\,\sigma^x = \mathrm{Tr}\,\sigma^y = \mathrm{Tr}\,\sigma^z = 0. \tag{80}$$

- Combining with Eq. (77) or Eq. (78), we have

$$\boxed{\mathrm{Tr}\left(\sigma^i\,\sigma^j\right) = 2\,\delta^{ij},} \tag{81}$$

$$\begin{aligned}&\mathrm{Tr}(\boldsymbol{a}\cdot\boldsymbol{\sigma}) = 0,\\ &\mathrm{Tr}(\boldsymbol{a}\cdot\boldsymbol{\sigma}\,\boldsymbol{b}\cdot\boldsymbol{\sigma}) = 2\,\boldsymbol{a}\cdot\boldsymbol{b}.\end{aligned} \tag{82}$$

HW 3

Let $\boldsymbol{m}$ and $\boldsymbol{n}$ be three-component real unit vectors. Define the operator $\boldsymbol{m}\cdot\boldsymbol{\sigma} = m_x\,\sigma^x + m_y\,\sigma^y + m_z\,\sigma^z$ for the vector $\boldsymbol{m} = (m_x,\,m_y,\,m_z)$, similarly for $\boldsymbol{n}$.
(i) Write down the matrix representation of $\boldsymbol{m}\cdot\boldsymbol{\sigma}$ in the $\{|\uparrow\rangle, |\downarrow\rangle\}$ basis.
(ii) If we measure the observable $\boldsymbol{m}\cdot\boldsymbol{\sigma}$, what could be the possible measurement outcomes?
(iii) Show that the probability of observing $\boldsymbol{n}\cdot\boldsymbol{\sigma} = +1$ when measuring the observable $\boldsymbol{n}\cdot\boldsymbol{\sigma}$ on the state $|\boldsymbol{m}\cdot\boldsymbol{\sigma} = +1\rangle$ is $\frac{1}{2}\,(1 + \boldsymbol{m}\cdot\boldsymbol{n})$.
(iv) What is the expectation value of the operator $\boldsymbol{n}\cdot\boldsymbol{\sigma}$ on the state $|\boldsymbol{m}\cdot\boldsymbol{\sigma} = +1\rangle$? (in terms of $\boldsymbol{m}$ and $\boldsymbol{n}$)

Solution (HW 3)

▪ Review: Measurement and Operator

We have learnt about:

- **observables** are described by *Hermitian* **operators**,
- **measuring** an *observable* on a quantum state could *change* the state (up on obtaining the outcome).

Is the *change* of state under the measurement implemented by the Hermitian *operator*? - **No!**

Example: prepare the qubit in $|\uparrow\rangle$, measure $\sigma^x \Rightarrow$ get $|\rightarrow\rangle$ or $|\leftarrow\rangle$ with probability 1/2 to 1/2. But σ^x operator does not take $|\uparrow\rangle$ to either $|\rightarrow\rangle$ or $|\leftarrow\rangle$. In fact $\sigma^x\,|\uparrow\rangle = |\downarrow\rangle$.

So what does the Hermitian operator really implement?

- **Hermitian operator** *attaches* measurement outcomes (*eigenvalues*) to its *eigenstates* (as prefactors).

$$L = \sum_i |\lambda_i\rangle\,\lambda_i\,\langle\lambda_i|. \tag{88}$$

Example: suppose we measure σ^x and obtain:

$$\sigma^x = \begin{cases} +1 & |\rightarrow\rangle, \\ -1 & |\leftarrow\rangle, \end{cases} \tag{89}$$

then the operator σ^x attaches the measurement outcome to the state

$$\begin{aligned} \sigma^x |\rightarrow\rangle &= (+1)\,|\rightarrow\rangle = |\rightarrow\rangle, \\ \sigma^x |\leftarrow\rangle &= (-1)\,|\leftarrow\rangle = -|\leftarrow\rangle. \end{aligned} \tag{90}$$

What if we apply σ^x to $|\uparrow\rangle$?

$$\begin{aligned} \sigma^x |\uparrow\rangle &= \sigma^x \frac{1}{\sqrt{2}} (|\rightarrow\rangle + |\leftarrow\rangle) \\ &= \frac{1}{\sqrt{2}} (\sigma^x |\rightarrow\rangle + \sigma^x |\leftarrow\rangle) \\ &= \frac{1}{\sqrt{2}} (|\rightarrow\rangle - |\leftarrow\rangle) \\ &= |\downarrow\rangle. \end{aligned} \tag{91}$$

- As an **operator**, σ^x *flips* the spin (exchanges $|\sigma^z = \pm 1\rangle$ states).
- As an **observable**, $\langle\psi|\,\sigma^x\,|\psi\rangle$ provides the *expectation value* of σ^x on any given state $|\psi\rangle$ (by the mechanism of attaching measurement outcomes).
- Although measuring σ^x on $|\uparrow\rangle \Rightarrow$ *collapse* $|\uparrow\rangle$ to either $|\rightarrow\rangle$ or $|\leftarrow\rangle$, this "*collapse*" operation is not implemented by the operator σ^x but by the **projection operators** (following a *normalization* procedure)

$$P(\sigma^x = \pm 1) = \frac{\mathbb{1} \pm \sigma^x}{2}. \tag{92}$$

In general, a *Hermitian operator* L can be used to define a family of *projection operators* (parameterized by λ)

$$L = \sum_i |\lambda_i\rangle\,\lambda_i\,\langle\lambda_i| \Rightarrow P(L = \lambda) = \sum_i |\lambda_i\rangle\,\delta(\lambda_i - \lambda)\,\langle\lambda_i|. \tag{93}$$

Quantum state collapse is implemented as

$$|\psi\rangle \xrightarrow{\text{measure } L, \text{ get } \lambda} \frac{P(L = \lambda)\,|\psi\rangle}{\langle\psi|\,P(L = \lambda)\,|\psi\rangle^{1/2}}. \tag{94}$$

- This is a *non-linear* operation on $|\psi\rangle \Rightarrow$ beyond the current framework of quantum mechanics (which only involves *linear* operators).

Unitary Operators

Basis Transformation

What operator should we apply to rotate one basis to another?

- Example:

$$U = |\uparrow\rangle\langle\rightarrow| + |\downarrow\rangle\langle\leftarrow|. \tag{95}$$

- U maps $|\rightarrow\rangle$ to $|\uparrow\rangle$ and maps $|\leftarrow\rangle$ to $|\downarrow\rangle$.
- Using explicit *vector* representations

$$\begin{aligned} &|\uparrow\rangle \doteq \begin{pmatrix}1\\0\end{pmatrix},\ |\downarrow\rangle \doteq \begin{pmatrix}0\\1\end{pmatrix};\\ &|\rightarrow\rangle \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, |\leftarrow\rangle \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}. \end{aligned} \tag{96}$$

we find

$$U \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}. \tag{97}$$

U is an example of **unitary operator**.

- It implements a **basis rotation**, as if we have redefined σ^x to σ^z. Every state in the Hilbert space will rotate correspondingly.

$$U(z_1|\rightarrow\rangle + z_2|\leftarrow\rangle) = z_1|\uparrow\rangle + z_2|\downarrow\rangle. \tag{98}$$

- A **unitary operator** is a linear operator whose **Hermitian conjugation** is its **inverse**, i.e.

$$\boxed{U^\dagger U = U U^\dagger = \mathbb{1}.} \tag{99}$$

- Two operators are *inverse* to each other ⇔ sequential application of them is equivalent to applying the *identity* (do-nothing) operator $\mathbb{1}$.
- The operation implemented by U is countered by that of $U^\dagger$, and vice versa.

- *Unitary operators* implements **basis rotation** (mapping $|\lambda_i\rangle$ to $|\mu_i\rangle$).

$$\boxed{U = \sum_i |\lambda_i\rangle\langle\mu_i|,} \tag{100}$$

- If $|\lambda_i\rangle$ and $|\mu_i\rangle$ are identical, $U = \mathbb{1}$ becomes the identity operator (which is also unitary).

One can verify that

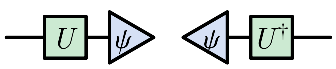
$$U^\dagger U = \sum_i |\mu_i\rangle\langle\lambda_i| \sum_j |\lambda_j\rangle\langle\mu_j|$$

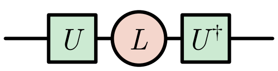
$$= \sum_{ij} |\mu_i\rangle \langle\lambda_i \mid \lambda_j\rangle \langle\mu_j|$$
$$= \sum_{ij} |\mu_i\rangle \, \delta_{ij} \, \langle\mu_j|$$
$$= \sum_{i} |\mu_i\rangle \langle\mu_i| = \mathbb{1},$$

and similar for $U\,U^\dagger = \mathbb{1}$. This means actually any basis transformation can be considered as a unitary operator.

- Applying basis transformations to

$$\begin{aligned} &\text{ket states}: \; |\psi\rangle \to U\,|\psi\rangle, \\ &\text{bra states}: \; \langle\psi| \to \langle\psi|\, U^\dagger, \\ &\text{operators}: \; L \to U\,L\,U^\dagger. \end{aligned} \tag{102}$$





  - Operator is made of ket and bra states, so the unitary operator must be applied from both sides.
  - The *expectation value* of an observable is *invariant* under *basis transformation*. (Measurement outcome should be *basis-independent*.)

$$\langle L\rangle = \langle\psi|\, L\, |\psi\rangle \to \langle\psi|\, U^\dagger \;\; U \;\; L \;\; U^\dagger \; U\, |\psi\rangle = \langle\psi|\, \mathbb{1} \;\; L \;\; \mathbb{1}\, |\psi\rangle = \langle L\rangle. \tag{103}$$

- **Diagonalization** of a *Hermitian operator*: find a unitary operator to bring the Hermitian operator to *diagonal form* by transforming to its *eigenbasis*.

$$\begin{aligned} U &= \sum_i |i\rangle \langle\lambda_i|, \\ L &= \sum_i |\lambda_i\rangle \, \lambda_i \, \langle\lambda_i|, \end{aligned} \tag{104}$$

such that under $L \to U\,L\,U^\dagger$,

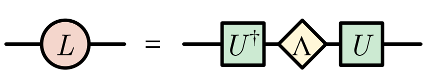
$$\Lambda = U\,L\,U^\dagger = \sum_i |i\rangle \, \lambda_i \, \langle i| \doteq \begin{pmatrix} \lambda_1 & 0 & \cdots \\ 0 & \lambda_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{105}$$

is diagonal.

  - Every *Hermitian* matrix can be written as

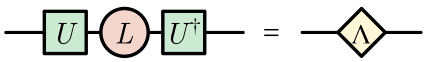
$$L = U^\dagger \, \Lambda \, U, \tag{106}$$

where Λ is *diagonal* and U is *unitary*.



- Or equivalently, the *unitary* transformation U brings the *Hermitian* matrix to its *diagonal* form,

$$U\,L\,U^\dagger = \Lambda. \tag{107}$$



■ Operator Functions

An **operator function** is a *function* which maps a *operator* (*matrix*) to a *operator* (*matrix*). There are two ways to promote a *scalar function* $f(x)$ to an *operator function* $f(M)$:

- For a diagonalizable operator $M = \sum_i |\mu_i\rangle\,\mu_i\,\langle\mu_i|$ (as long as $M\,M^\dagger = M^\dagger\,M$), we define

$$f(M) = \sum_i |\mu_i\rangle\,f(\mu_i)\,\langle\mu_i|. \tag{108}$$

- If $f(x)$ admits Taylor expansion

$$f(x) = f(0) + f'(0)\,x + \frac{f''(0)}{2!}\,x^2 + \ldots, \tag{109}$$

the corresponding operator function is

$$f(M) = f(0) + f'(0)\,M + \frac{f''(0)}{2!}\,M^2 + \ldots. \tag{110}$$

One special application of the operator function is to define the **operator (matrix) exponential**. For example,

$$e^{i\,\theta\,\sigma^y} = ?,\ \text{given}\ \sigma^y \doteq \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{111}$$

- Method I: *Mathematica*

$$e^{i\,\theta\,\sigma^y} \doteq e^{\begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{112}$$

```
MatrixExp[{{0, θ}, {-θ, 0}}]
{{Cos[θ], Sin[θ]}, {-Sin[θ], Cos[θ]}}
```

- Method II: Diagonalization. Switch to the eigenbasis of σ^y, which was given in Eq. (119)

$$e^{i\,\theta\,\sigma_y} = \left|\otimes\right\rangle e^{+i\,\theta} \left\langle\otimes\right| + \left|\odot\right\rangle e^{-i\,\theta} \left\langle\odot\right|, \tag{113}$$

then using Eq. (29) to show

$$\begin{aligned} e^{i\,\theta\,\sigma^y} &\doteq \frac{1}{2}\begin{pmatrix} 1 \\ i \end{pmatrix} e^{+i\,\theta} \begin{pmatrix} 1 & -i \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ -i \end{pmatrix} e^{-i\,\theta} \begin{pmatrix} 1 & i \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \end{aligned} \tag{114}$$

- Method III: Taylor expansion.

$$
\begin{aligned}
e^{i\theta\sigma^y} &= \mathbb{1} + i\,\theta\,\sigma^y + \frac{1}{2!}(i\,\theta\,\sigma^y)^2 + \frac{1}{3!}(i\,\theta\,\sigma^y)^3 + \frac{1}{4!}(i\,\theta\,\sigma^y)^4 + \ldots \\
&= \left(1 - \frac{1}{2!}\theta^2 + \frac{1}{4!}\theta^4 + \ldots\right)\mathbb{1} + i\left(\theta - \frac{1}{3!}\theta^3 + \ldots\right)\sigma^y \\
&= \cos\theta\,\mathbb{1} + i\sin\theta\,\sigma^y \\
&\doteq \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.
\end{aligned}
$$

Note that $(\sigma^y)^2 = \mathbb{1}$, then terms of even and odd powers of θ can be collected together respectively.

More generally,

$$\boxed{e^{i\,\theta\,\boldsymbol{n}\cdot\boldsymbol{\sigma}} = (\cos\theta)\,\mathbb{1} + i\,(\sin\theta)\,\boldsymbol{n}\cdot\boldsymbol{\sigma},} \tag{116}$$

where $\theta \in \mathbb{R}$ is real and $\boldsymbol{n}$ is a three-component unit vector. It can be shown using Taylor expansion technique as Eq. (115), by noting that $(\boldsymbol{n}\cdot\boldsymbol{\sigma})^2 = (\boldsymbol{n}\cdot\boldsymbol{n})\,\mathbb{1} = \mathbb{1}$, according to Eq. (78).

Hermitian Generators

If **Hermitian operators** are generalization of **real numbers**, then **unitary operators** are generalization of **phase factors**. ($u \in \mathbb{C}$ and $|u| = 1$)

$$u^*\,u = u\,u^* = |u|^2 = 1. \tag{117}$$

- For complex numbers, a phase factor can be written as $u = e^{i\,\theta}$, where $\theta \in \mathbb{R}$ is a *real* phase angle.
- Similar ideas apply to unitary operators: every **unitary operator** can be **generated** by a **Hermitian operator** in the form of

$$U = e^{i\,L}. \tag{118}$$

Given a Hermitian operator L, by $e^{i\,L}$ we mean

- in the eigen basis

$$e^{i\,L} = \sum_i |\lambda_i\rangle\, e^{i\,\lambda_i}\, \langle\lambda_i|. \tag{119}$$

- by operator Taylor expansion

$$e^{i\,L} = \mathbb{1} + i\,L + \frac{(i\,L)^2}{2!} + \frac{(i\,L)^3}{3!} + \ldots. \tag{120}$$

By definition, $e^{i\,L}$ is *unitary* if L is *Hermitian*, since

$$
\begin{aligned}
U^\dagger\,U &= \left(e^{i\,L}\right)^\dagger e^{i\,L} \\
&= \sum_i |\lambda_i\rangle\, e^{-i\,\lambda_i}\,\langle\lambda_i| \sum_j |\lambda_j\rangle\, e^{i\,\lambda_j}\,\langle\lambda_j|
\end{aligned}
$$

$$= \sum_i |\lambda_i\rangle\, e^{-i\lambda_i}\, e^{i\lambda_i}\, \langle\lambda_i|$$
$$= \sum_i |\lambda_i\rangle\, \langle\lambda_i|$$
$$= \mathbb{1},$$

and similar for $U\, U^\dagger = \mathbb{1}$.

Example: Consider $U(\theta) = e^{i\,\theta\,\sigma^y}$ in Eq. (112). It implements a **basis rotation** with θ being the **rotation angle**:

$$U(\theta)\,|\uparrow\rangle = \cos\,\theta\;|\uparrow\rangle - \sin\,\theta\;|\downarrow\rangle \doteq \begin{pmatrix} \cos\theta \\ -\sin\theta \end{pmatrix}. \tag{122}$$

Special case: when $\theta = 0$, $U(0) = \mathbb{1} \Rightarrow$ no rotation is performed.

More generally, let $U(\theta)$ be the **unitary operator** that implements certain *basis rotation* by an **angle** θ. When $\theta = \Delta\theta$ is **small**, we can Taylor expand

$$U(\Delta\theta) = U(0) + U'(0)\,\Delta\theta + \ldots = \mathbb{1} + U'(0)\,\Delta\theta + \ldots, \tag{123}$$

where $U'(0)$ is $\partial_\theta\, U(\theta)$ evaluated at $\theta = 0$.

$U'(0)$ is also an operator (matrix), usually denoted as $U'(0) = i\,L$. We call L the **generator** of the rotation/unitary operator, because it *generates* an **infinitesimal rotation**

$$U(\Delta\theta) = \mathbb{1} + i\,\Delta\theta\, L + \ldots. \tag{124}$$

$U(\Delta\theta)$ is **unitary** $\Rightarrow$ L is **Hermitian**.

$$\begin{aligned} &U(\Delta\theta)^\dagger\; U(\Delta\theta) \\ &= \left(\mathbb{1} - i\,\Delta\theta\, L^\dagger + \ldots\right)(\mathbb{1} + i\,\Delta\theta\, L + \ldots) \\ &= \mathbb{1} + i\,\Delta\theta\left(L - L^\dagger\right) + \ldots = \mathbb{1}. \end{aligned} \tag{125}$$

Large rotations can be *accumulated* from *small* rotations.

$$U(N\,\Delta\theta) = U(\Delta\theta)^N = (\mathbb{1} + i\,\Delta\theta\, L)^N. \tag{126}$$

As $\Delta\theta$ is small (but N can be large, s.t. $\theta = N\,\Delta\theta$ is finite),

$$\ln\, U(N\,\Delta\theta) = N\,\ln(\mathbb{1} + i\,\Delta\theta\, L) = i\,N\,\Delta\theta\, L, \tag{127}$$

So $U(N\,\Delta\theta) = e^{i\,N\,\Delta\theta\, L}$, we obtain the *exponential* form

$$\boxed{U(\theta) = e^{i\,\theta\, L}.} \tag{128}$$

Conclusion: every *Hermitian* operator generates a *unitary* operator.

▪ Time-Evolution is Unitary

Unitarity: *information* is never lost.

- Two *identical* and *isolated* systems, start out in **different** states $\Rightarrow$ **remains** in **different** states (in both future and past).

- Two *identical* and *isolated* systems, start out in the **same** state ⇒ follow **identical evolution** (in both future and past).

Although **measurement** seems to be **non-deterministic**, evolution of quantum **state** is **deterministic**: suppose you know the *state* at one time, then the quantum *equation of motion* tell you what it will be later.

$$\boxed{|\psi(t)\rangle = U(t)\,|\psi(0)\rangle,} \tag{129}$$

$|\psi(0)\rangle$ is the initial state, and $|\psi(t)\rangle$ is the state at time t. $U(t)$ is the **time-evolution operator** that takes $|\psi(0)\rangle$ to $|\psi(t)\rangle$. We will show that $U(t)$ should be *unitary*.

- *Different* states remain *different* (here, *different* states are states that can be told apart definitely by a measurement, due to their different outcomes, so they are actually *orthogonal*):

$$\langle\phi(0) \mid \psi(0)\rangle = 0 \Rightarrow \langle\phi(t) \mid \psi(t)\rangle = \langle\phi(0)|\, U(t)^\dagger\, U(t)\,|\psi(0)\rangle = 0. \tag{130}$$

- *Same* states remain the *same*

$$\langle\psi(0) \mid \psi(0)\rangle = 1 \Rightarrow \langle\psi(t) \mid \psi(t)\rangle = \langle\psi(0)|\, U(t)^\dagger\, U(t)\,|\psi(0)\rangle = 1. \tag{131}$$

  Or, the fact that the probability adds up to 1 is preserved.

Treat $|\psi(0)\rangle$ and $|\phi(0)\rangle$ as members of any orthonormal basis, then Eq. (130) and Eq. (131) implies

$$\langle i|\, U(t)^\dagger\, U(t)\,|j\rangle = \delta_{ij} \Rightarrow U(t)^\dagger\, U(t) = \mathbb{1}. \tag{132}$$

Therefore, the **time-evolution** operator $U(t)$ is **unitary**.

Hamiltonian and Schrödinger Equation

Hamiltonian *generates* time-evolution!

As a *unitary* operator, the *time-evolution* operator is also *generated* by a *Hermitian* operator, called the **Hamiltonian**,

$$H = i\, U'(0) = i\,\partial_t\, U(t)\,|_{t=0}\,. \tag{133}$$

For small Δt, *infinitesimal* evolution is given by

$$U(\Delta t) = \mathbb{1} - i\, H\, \Delta t + \ldots, \tag{134}$$

therefore the state evolves as

$$|\psi(\Delta t)\rangle = U(\Delta t)\,|\psi(0)\rangle = |\psi(0)\rangle - i\ \Delta\ t\ H\,|\psi(0)\rangle, \tag{135}$$

meaning that

$$i\ \partial_t|\psi(0)\rangle = i\ \frac{|\psi(\Delta t)\rangle - |\psi(0)\rangle}{\Delta t} = H\,|\psi(0)\rangle. \tag{136}$$

There is nothing special about $t = 0$. Eq. (136) should hold at any time.

$$i\,\partial_t|\psi(t)\rangle = H\,|\psi(t)\rangle. \tag{137}$$

This is the **Schrödinger equation**, the *equation of motion* for the quantum state.

- The Hamiltonian $H(t) = i\, U'(t)$ can be **time-dependent** in general.
- But in many cases, we consider H to be **time-independent**, by assuming the **time-translation symmetry**.

  What happens to Planck's constant?

$$\hbar = \frac{h}{2\pi} = 1.0545718\,(13) \times 10^{-34}\ \text{J s}. \tag{138}$$

In quantum mechanics, the *observable* associated with the **Hamiltonian** is the **energy**. To balance the *dimensionality* across the Schrödinger equation, *Planck's constant* is inserted for Eq. (137):

$$i\,\hbar\,\partial_t |\psi(t)\rangle = H\,|\psi(t)\rangle. \tag{139}$$

Why is $\hbar$ so small? Well, the answer has more to do with biology than with physics ⇒ Why we are so big, heavy and slow? A natural choice for quantum mechanics is to set the units such that $\hbar = 1$. It is a common practice in theoretical physics (we will also use this convention sometimes).

We conclude with another **axiom of quantum mechanics**

Axiom 4 (Dynamics): The **time-evolution** of the *state* of a quantum system is governed by the **Hamiltonian** of the system, according to the time-dependent **Schrödinger equation**.

$$\boxed{i\,\hbar\,\partial_t |\psi(t)\rangle = H\,|\psi(t)\rangle.} \tag{140}$$

If the Hamiltonian H is **time-independent**, we can first find its eigenvalues (**eigenenergies**) and eigenvectors (**energy eigenstates**).

$$\boxed{H\,|E_i\rangle = E_i\,|E_i\rangle.} \tag{141}$$

This is also called the *time-independent* **Schrödinger equation**. Without solving a *differential equation*, we just need to *diagonalize* a *Hermitian matrix* in this case.

Each *energy eigenstate* will evolve in time simply by a *rotating overall phase*,

$$|E_i(t)\rangle = e^{-\frac{i}{\hbar} E_i\, t}\,|E_i\rangle. \tag{142}$$

- $|E_i\rangle$ form a complete set of orthonormal basis, called **energy eigenbasis**.
- Verify that Eq. (142) is a solution of Eq. (140):

$$\begin{aligned} i\,\hbar\,\partial_t |E_i(t)\rangle &= i\,\hbar\,\partial_t\left(e^{-\frac{i}{\hbar} E_i\, t}\,|E_i\rangle\right) = E_i\,|E_i(t)\rangle, \\ H\,|E_i(t)\rangle &= e^{-\frac{i}{\hbar} E_i\, t}\,H\,|E_i\rangle = E_i\,|E_i(t)\rangle. \end{aligned} \tag{143}$$

  So the two sides matches.

Any initial state $|\psi(0)\rangle$ will evolve in time by first *representing* the initial state in the *energy eigenbasis*, and attaching to each energy eigenstate by its rotating overall phase,

$$
\begin{aligned}
|\psi(t)\rangle &= \sum_i e^{-\frac{i}{\hbar} E_i t} |E_i\rangle \langle E_i \mid \psi(0)\rangle \\
&= e^{-\frac{i}{\hbar} H t} |\psi(0)\rangle.
\end{aligned} \tag{144}
$$

A *time-independent* Hamiltonian generates the time-evolution via *matrix exponentiation*

$$
\boxed{U(t) = e^{-\frac{i}{\hbar} H t}.} \tag{145}
$$

However, for *time-dependent* Hamiltonian, there no such a clean formula. Evolution must be carried out step by step, denoted as a *time-ordered* exponential

$$
U(t) = \mathcal{T} \exp\left(-\frac{i}{\hbar} \int_0^t H(t')\, d t'\right). \tag{146}
$$

Example: Spin in a Magnetic Field

How to write down a Hamiltonian?

- derive it from experiment,
- borrow it from some theory we like,
- pick one and see what happens.

Hamiltonian must be *Hermitian* anyway. For a single qubit, the most general Hamiltonian takes the form of

$$
\begin{aligned}
H &= h_0 \mathbb{1} + h_x \sigma^x + h_y \sigma^y + h_z \sigma^z \\
&= h_0 \mathbb{1} + \boldsymbol{h} \cdot \boldsymbol{\sigma},
\end{aligned} \tag{147}
$$

where $h_0, h_x, h_y, h_z \in \mathbb{R}$ are all *real* coefficients. $\boldsymbol{h} = (h_x, h_y, h_z)$ is a vector of *numbers* and $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ is a vector of *operators*.

- The time-evolution operator (set $\hbar = 1$ in the following)

$$
\begin{aligned}
U(t) &= e^{-i H t} \\
&= e^{-i h_0 t}\left(\cos(|\boldsymbol{h}|\, t)\, \mathbb{1} - i \sin(|\boldsymbol{h}|\, t)\, \hat{\boldsymbol{h}} \cdot \boldsymbol{\sigma}\right),
\end{aligned} \tag{148}
$$

where $|\boldsymbol{h}| = \sqrt{\boldsymbol{h} \cdot \boldsymbol{h}}$ and $\hat{\boldsymbol{h}} = \boldsymbol{h} / |\boldsymbol{h}|$.

- A state $|\psi(0)\rangle$ will evolve with time following

$$
\begin{aligned}
|\psi(t)\rangle &= U(t)\, |\psi(0)\rangle \\
&= e^{-i h_0 t}\left(\cos(|\boldsymbol{h}|\, t)\, \mathbb{1} - i \sin(|\boldsymbol{h}|\, t)\, \hat{\boldsymbol{h}} \cdot \boldsymbol{\sigma}\right) |\psi(0)\rangle.
\end{aligned} \tag{149}
$$

- If we measure $\boldsymbol{\sigma}$ on the state $|\psi(t)\rangle$, the expectation value will be given by

$$
\begin{aligned}
\langle \boldsymbol{\sigma} \rangle_t &= \langle \psi(t) |\, \boldsymbol{\sigma}\, | \psi(t) \rangle \\
&= \cos(2\, |\boldsymbol{h}|\, t)\, \langle \boldsymbol{\sigma} \rangle_0 + \sin(2\, |\boldsymbol{h}|\, t)\, \hat{\boldsymbol{h}} \times \langle \boldsymbol{\sigma} \rangle_0 + (1 - \cos(2\, |\boldsymbol{h}|\, t))\, \hat{\boldsymbol{h}}\left(\hat{\boldsymbol{h}} \cdot \langle \boldsymbol{\sigma} \rangle_0\right).
\end{aligned} \tag{150}
$$

which also evolves with time.

HW 4

(i) Derive Eq. (148) from Eq. (147).
(ii) Derive Eq. (150) from Eq. (149).
Hint: Eq. (78) can make life much more easier.

Special case: assume $\boldsymbol{h} = (0, 0, h_z)$ along the z-direction, and parameterize the expectation of the spin vector by $\langle \boldsymbol{\sigma} \rangle = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$.

$$\langle \boldsymbol{\sigma} \rangle_t = (\sin\theta_0 \cos(\varphi_0 + 2\, h_z\, t),\ \sin\theta_0 \sin(\varphi_0 + 2\, h_z\, t), \cos\theta_0), \tag{151}$$

where θ_0 and φ_0 are the initial azimuthal and polar angles.

- The *spin* should *precess* around the axis of the *magnetic field* $\Rightarrow$ $\boldsymbol{h}$ has the physical meaning of the external *magnetic field*.
- *Energy* of a spin in the magnetic field is $\langle H \rangle = -\boldsymbol{h} \cdot \langle \boldsymbol{\sigma} \rangle$ (up to some constant energy shift h_0).

Solution (HW 4)

Operator Algebra

Commutator

- **Commutator** of two operators A and B

$$\boxed{[A, B] = A\, B - B\, A.} \tag{158}$$

  - Commutator is *antisymmetric*, $[A, B] = -[B, A]$. As a result, *commutator* of an operator with *itself* always *vanishes* $[A, A] = 0$.
  - If the commutator vanishes $[A, B] = 0$, we say that the two operators A and B **commute**.

Example of commutators:

$$\begin{aligned} &[\sigma^x, \sigma^y] = 2\, i\, \sigma^z, \\ &[\sigma^y, \sigma^z] = 2\, i\, \sigma^x, \\ &[\sigma^z, \sigma^x] = 2\, i\, \sigma^y. \end{aligned} \tag{159}$$

Or more compactly as

$$\boxed{\left[\sigma^a, \sigma^b\right] = 2\, i\, \epsilon^{abc}\, \sigma^c,} \tag{160}$$

for $a, b, c = 1, 2, 3$ (stand for x, y, z). This can be considered as the defining algebraic properties of *single-qubit operators* (Pauli matrices). Or even more compactly using the cross product of vectors

$$\boxed{\boldsymbol{\sigma} \times \boldsymbol{\sigma} = 2\, i\, \boldsymbol{\sigma}.} \tag{161}$$

Useful *rules* to evaluate commutators

- Bilinearity

$$[A,\ B+C]=[A,\ B]+[A,\ C],$$
$$[A+B,\ C]=[A,\ C]+[B,\ C]. \tag{162}$$

- Product rules

$$[A,\ B\,C]=[A,\ B]\,C+B[A,\ C],$$
$$[A\,B,\ C]=[A,\ C]\,B+A[B,\ C]. \tag{163}$$

- Jacobi identity (as a replacement of associative law)

$$[A,\ [B,\ C]]+[B,\ [C,\ A]]+[C,\ [A,\ B]]=0,$$
$$[[A,\ B],\ C]+[[B,\ C],\ A]+[[C,\ A],\ B]=0. \tag{164}$$

Commutation Relation

- A and B **commute**: $A\,B=B\,A$ (operators can *pass through* each other as if they were *numbers*) $\Rightarrow$ it does not matter which operator is applied first, the consequence will be the same.

  Examples:

  - A: put on the socks,
  - B: put on the shoes,
  - C: put on the hat,

A and B do *not commute* (changing the order leads to different result). But A and C *commute*, B and C also *commute* (changing the order does not affect the result).

- An operator always *commutes* with *itself*.
- *Identity* operator *commutes* with any operator.

Commutation Relation (Single-Qubit)

For a generic qubit state $|\psi\rangle \doteq \begin{pmatrix}\psi_\uparrow\\ \psi_\downarrow\end{pmatrix}$,

$$\sigma^z\,\sigma^x\,|\psi\rangle : \begin{pmatrix}\psi_\uparrow\\ \psi_\downarrow\end{pmatrix} \xrightarrow{\sigma^x} \begin{pmatrix}\psi_\downarrow\\ \psi_\uparrow\end{pmatrix} \xrightarrow{\sigma^z} \begin{pmatrix}\psi_\downarrow\\ -\psi_\uparrow\end{pmatrix},$$
$$\sigma^x\,\sigma^z\,|\psi\rangle : \begin{pmatrix}\psi_\uparrow\\ \psi_\downarrow\end{pmatrix} \xrightarrow{\sigma^z} \begin{pmatrix}\psi_\uparrow\\ -\psi_\downarrow\end{pmatrix} \xrightarrow{\sigma^x} \begin{pmatrix}-\psi_\downarrow\\ \psi_\uparrow\end{pmatrix}. \tag{165}$$

Conclusion: σ^x and σ^z do not commute. In fact, $[\sigma^z,\ \sigma^x]=2\,i\,\sigma^y\neq 0$, which can be readily verified from their matrix representations

$$\sigma^z \doteq \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},\ \sigma^x \doteq \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}. \tag{166}$$

$|\uparrow\rangle$ and $|\downarrow\rangle$ are eigenstates of σ^z with different eigenvalues. σ^z marks the states differently, and σ^x mixes the states. In general, "*markers*" and "*mixers*" do *not* commute.

Commutation Relation (Two-Qubit)

Define $\sigma^{a\,b}=\sigma^a\otimes\sigma^b$, e.g.

$$\sigma^{12} = \sigma^1 \otimes \sigma^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ \hline 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{array}\right),$$

$$\sigma^{23} = \sigma^2 \otimes \sigma^3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ \hline i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{array}\right).$$

Note: in general, tensor product of two matrices is given by

$$\begin{aligned} &\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \otimes \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \\ &= \begin{pmatrix} A_{11}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} & A_{12}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \\ A_{21}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} & A_{22}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \end{pmatrix} \\ &= \left(\begin{array}{cc|cc} A_{11} B_{11} & A_{11} B_{12} & A_{12} B_{11} & A_{12} B_{12} \\ A_{11} B_{21} & A_{11} B_{22} & A_{12} B_{21} & A_{12} B_{22} \\ \hline A_{21} B_{11} & A_{21} B_{12} & A_{22} B_{11} & A_{22} B_{12} \\ A_{21} B_{21} & A_{21} B_{22} & A_{22} B_{21} & A_{22} B_{22} \end{array}\right). \end{aligned} \tag{168}$$

You can use *Mathematica* to calculate tensor product like this:

```
MatrixForm@KroneckerProduct[PauliMatrix[1], PauliMatrix[2]]
```

$$\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$

Consider two Hermitian operators A and B in this four dimensional Hilbert space:

$$A \simeq \sigma^{12}, \; B \simeq \sigma^{23}. \tag{169}$$

Do A and B commute?

- Yes, because we can explicitly verify $\left[\sigma^{12}, \sigma^{23}\right] = 0$ using the matrix representation.

```
A = KroneckerProduct[PauliMatrix[1], PauliMatrix[2]];
B = KroneckerProduct[PauliMatrix[2], PauliMatrix[3]];
A.B - B.A // MatrixForm
```

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- But is there a better way to see this?

Switch to the *diagonal basis* of A: find a unitary operator (choice is not unique) to diagonalize A

$$U_1 \doteq e^{\frac{i\pi}{4}\sigma^{22}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & -i \\ 0 & 1 & i & 0 \\ 0 & i & 1 & 0 \\ -i & 0 & 0 & 1 \end{pmatrix}. \tag{170}$$

U_1 takes A and B to the **block diagonal** form

$$\begin{aligned} A \to A' &= U_1\, A\, U_1^\dagger \doteq \sigma^{30} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array}\right), \\ B \to B' &= U_1\, B\, U_1^\dagger \doteq -\sigma^{01} = \left(\begin{array}{cc|cc} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{array}\right). \end{aligned} \tag{171}$$

```
A = KroneckerProduct[PauliMatrix[1], PauliMatrix[2]];
B = KroneckerProduct[PauliMatrix[2], PauliMatrix[3]];
U1 = MatrixExp[ⅈ π / 4 KroneckerProduct[PauliMatrix[2], PauliMatrix[2]]];
MatrixForm[U1.#.ConjugateTranspose[U1]] & /@ {A, B}
```

$$\left\{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}\right\}$$

B' does not mix different eigenspaces of A' ⇒ A' and B' commute ⇒ A and B also commute.

HW 5

Show that the fact that two operator commute (or not commute) is independent of the choice of basis, i.e. suppose $A' = U\,A\,U^\dagger$ and $B' = U\,B\,U^\dagger$, then $[A, B] = 0 \Leftrightarrow [A', B'] = 0$.

Mixing within the block (by B') does not cause a problem, why? Because A' look like an identity matrix within each block, which commutes with any matrix within the same block.

Diagonal blocks can be further *diagonalized independently* (within each block). For example, we can take

$$U_2 \doteq e^{\frac{i\pi}{4}\sigma^{02}} = \frac{1}{\sqrt{2}}\left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{array}\right), \tag{172}$$

under which

$$A' \to A'' = U_2\, A'\, U_2^\dagger \simeq \sigma^{30} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array}\right),$$

$$B' \to B'' = U_2\, B'\, U_2^\dagger \simeq \sigma^{03} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{array}\right).$$

The combined unitary transformation $U = U_2\, U_1$ **simultaneously diagonalize** A and B, such that $A'' = U\, A\, U^\dagger$ and $B'' = U\, B\, U^\dagger$ are both *diagonal*.

Solution (HW 5)

▫ Commutation Relation (General Discussions)

In fact, **commuting operators** can always be *simultaneously* diagonalized.

- Suppose $\{A_1, A_2, \ldots\}$ is a set of commuting (Hermitian) operators, i.e. $\forall\, i, j : [A_i, A_j] = 0$, the general algorithm to simultaneous diagonalize them is to first form a *random* Hamiltonian

$$H = \sum_i r_i\, A_i, \tag{175}$$

with r_i being random *real* numbers. Find a unitary operator U to diagonalize the Hamiltonian H, the same unitary U would simultaneously diagonalize all A_i *with probability 1.*

```
As = {KroneckerProduct[PauliMatrix[1], PauliMatrix[2]],
   KroneckerProduct[PauliMatrix[2], PauliMatrix[3]]};
MatrixForm /@ As
```

$$\left\{\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}\right\}$$

```
H = RandomReal[{-1, 1}, Length@As].As;
MatrixForm@H
```

$$\begin{pmatrix} 0.+0.\,i & 0.+0.\,i & 0.+0.168926\,i & 0.-0.931765\,i \\ 0.+0.\,i & 0.+0.\,i & 0.+0.931765\,i & 0.-0.168926\,i \\ 0.-0.168926\,i & 0.-0.931765\,i & 0.+0.\,i & 0.+0.\,i \\ 0.+0.931765\,i & 0.+0.168926\,i & 0.+0.\,i & 0.+0.\,i \end{pmatrix}$$

```
U = Conjugate@Eigenvectors@H;
MatrixForm@Chop[U.#.ConjugateTranspose[U]] & /@ As
```

$$\left\{\begin{pmatrix} 1. & 0 & 0 & 0 \\ 0 & -1. & 0 & 0 \\ 0 & 0 & 1. & 0 \\ 0 & 0 & 0 & -1. \end{pmatrix}, \begin{pmatrix} -1. & 0 & 0 & 0 \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 1. & 0 \\ 0 & 0 & 0 & -1. \end{pmatrix}\right\}$$

Commuting operators can share a set of **common eigenvectors**, which can always be constructed by *simultaneous diagonalization*. For example, if $[A, B] = 0$, there exist a set of vectors $|\alpha, \beta\rangle$,

$$\boxed{\begin{aligned} A\,|\alpha, \beta\rangle &= \alpha\,|\alpha, \beta\rangle, \\ B\,|\alpha, \beta\rangle &= \beta\,|\alpha, \beta\rangle. \end{aligned}} \tag{176}$$

Each eigenvector is labeled jointly by the eigenvalues α and β.

- **Commuting physical observables** can be *simultaneously* measured.
  - The possible outcomes of a *joint measurement* of (A, B) are given by the **pairs of eigenvalues** (α, β).
  - On a given state $|\psi\rangle$, the *probability* to obtain the measurement outcome (α, β) is given by

    $$p(\alpha, \beta) = |\langle\alpha, \beta \mid \psi\rangle|^2. \tag{177}$$

  - After the measurement, the state is *projected* to the **common eigenstate** $|\alpha, \beta\rangle$ that corresponds to the measurement outcome (α, β).
- **Non-commuting physical observables** do not share common eigenstates, therefore do not support a consistent joint measurement. The amount of *inconsistency* (**uncertainty**) of the joint measurement is characterized by the **commutator**. This statement is more precisely formulated as the **uncertainty relation**.

▪ Uncertainty Relation

Statistics of measurement. Consider an *observable* L, whose eigenvalues are λ (i.e. $L\,|\lambda\rangle = \lambda\,|\lambda\rangle$), measured on a *state* $|\psi\rangle$ in *repeated* experiments (prepare $|\psi\rangle \to$ measure $L \to$ repeat). Possible outcomes λ appear with *probability* $p(\lambda) = |\langle\lambda \mid \psi\rangle|^2$.

- **Mean** (expectation value):

  $$\langle L\rangle = \sum_\lambda \lambda\, p(\lambda) = \langle\psi|\, L\,|\psi\rangle. \tag{178}$$

- **Variance** (2nd moment):

  $$\operatorname{var} L = \sum_\lambda (\lambda - \langle L\rangle)^2\, p(\lambda) = \langle\psi|\,(L - \langle L\rangle\, \mathbb{1})^2\,|\psi\rangle. \tag{179}$$

  Introduce the observable (the fluctuation of L around its expectation value)

  $$\Delta L = L - \langle L\rangle\, \mathbb{1}, \tag{180}$$

  The variance can be written as $\operatorname{var} L = \left\langle(\Delta L)^2\right\rangle$.

- **Standard deviation**: characterizes the **uncertainty** of the measurement of L

  $$\operatorname{std} L = (\operatorname{var} L)^{1/2} = \left\langle(\Delta L)^2\right\rangle^{1/2}. \tag{181}$$

Uncertainty Relation: for any pair of *observables* A and B measured on any given *state* (repeatedly),

$$\boxed{(\text{std } A)\,(\text{std } B) \geq \frac{1}{2}\,|\langle[A,\,B]\rangle|.} \tag{182}$$

- In words, the product of the *uncertainties* cannot be smaller than half of the magnitude of the expectation value of the *commutator*.
- For *commuting* observables ($[L,\,M] = 0$), $(\text{std } L)\,(\text{std } M) \geq 0$, it is possible to have $\text{std } L = \text{std } M = 0$ simultaneously, i.e. L and M can be jointly measured with perfect certainty.
- For *non-commuting* observables, if $|\langle[L,\,M]\rangle| \neq 0$, it is impossible to have $\text{std } L = \text{std } M = 0$ simultaneously, i.e. L and M can not be jointly measured with certainty.

Proof of the uncertainty relation:

Suppose A and B are Hermitian operators. Let $|\phi\rangle = (A + i\,x\,B)\,|\psi\rangle$. For any choice of $x \in \mathbb{R}$,

$$\langle\psi|\,(A - i\,x\,B)\,(A + i\,x\,B)\,|\psi\rangle = \langle\phi\mid\phi\rangle \geq 0. \tag{183}$$

On the other hand,

$$\begin{aligned}&\langle\psi|\,(A - i\,x\,B)\,(A + i\,x\,B)\,|\psi\rangle\\ &= \langle\psi|\,A^2 + i\,x\,[A,\,B] + x^2\,B^2\,|\psi\rangle\\ &= \left\langle B^2\right\rangle x^2 + i\,\langle[A,\,B]\rangle\,x + \left\langle A^2\right\rangle \geq 0,\end{aligned} \tag{184}$$

where $\langle * \rangle$ is a shorthand notation of $\langle\psi| * |\psi\rangle$. The quadratic equation $\left\langle B^2\right\rangle x^2 + i\,\langle[A,\,B]\rangle\,x + \left\langle A^2\right\rangle = 0$ has no (or only one) real root, implying that its discriminant Δ must be negative (or zero), i.e.

$$\Delta = (i\,\langle[A,\,B]\rangle)^2 - 4\left\langle B^2\right\rangle\left\langle A^2\right\rangle \leq 0. \tag{185}$$

Therefore for any A, B on any state $|\psi\rangle$,

$$\left\langle A^2\right\rangle^{1/2}\left\langle B^2\right\rangle^{1/2} \geq \frac{1}{2}\,|\langle[A,\,B]\rangle|. \tag{186}$$

The uncertainty relation Eq. (182) can be shown by replacing $A \to \Delta A$ and $B \to \Delta B$.

HW 6

> Suppose A and B are Hermitian operators.
> (i) Show that $\left\langle A^2\right\rangle$, $\left\langle B^2\right\rangle$ and $i\,\langle[A,\,B]\rangle$ are real.
> (ii) Show that $[\Delta A,\,\Delta B] = [A,\,B]$.

Solution (HW 6)

Operator Dynamics

Two *pictures* of the **quantum dynamics**:

- **Schrödinger picture**: state evolves in time, operator remains fixed,

$$\langle L(t)\rangle = \langle\psi(t)|\,L\,|\psi(t)\rangle. \tag{190}$$

- **Heisenberg picture**: operator evolves in time, state remains fixed,

$$\langle L(t)\rangle = \langle\psi|\, L(t)\,|\psi\rangle. \tag{191}$$

The two pictures are consistent, if

$$|\psi(t)\rangle = U(t)\,|\psi\rangle \Rightarrow L(t) = U(t)^{\dagger}\, L\, U(t), \tag{192}$$

such that Eq. (190) and Eq. (191) are consistent, as they both implies

$$\langle L(t)\rangle = \langle\psi|\, U(t)^{\dagger}\, L\, U(t)\,|\psi\rangle. \tag{193}$$

Note: one should only apply one picture at a time, i.e. either the state or the operator is time-dependent, **but not both**.

In the *Heisenberg picture*, the time-evolution of an operator

$$\boxed{L(t) = U(t)^{\dagger}\, L\, U(t),} \tag{194}$$

described by the **Heisenberg equation**

$$\boxed{i\,\hbar\,\partial_t L(t) = [L(t), H].} \tag{195}$$

A sketch of the derivation: for small Δt (with $\hbar = 1$)

$$\begin{aligned} L(\Delta t) &= U(\Delta t)^{\dagger}\, L\, U(\Delta t) \\ &= e^{i H \Delta t}\, L\, e^{-i H \Delta t} \\ &= (\mathbb{1} + i\, H\,\Delta t + \ldots)\, L\, (\mathbb{1} - i\, H\,\Delta t + \ldots) \\ &= L + i\,(H\, L - L\, H)\,\Delta t + \ldots \\ &= L - i\,[L, H]\,\Delta t + \ldots \end{aligned} \tag{196}$$

therefore

$$i\,\partial_t L = i\,\frac{L(\Delta t) - L}{\Delta t} = [L, H]. \tag{197}$$

Correspondingly, its expectation value evolves as

$$i\,\hbar\,\partial_t\langle L(t)\rangle = \langle[L(t), H]\rangle. \tag{198}$$

If $[L, H] = 0$, the *Heisenberg equation* Eq. (195) implies that $\partial_t L = 0$, i.e. L will be invariant in time. The observable L is a **conserved quantity** (or an **integral of motion**) if L *commutes* with the Hamiltonian H.

HW 7

Consider a single-qubit Hamiltonian $H = \boldsymbol{h}\cdot\boldsymbol{S}$, where $\boldsymbol{S} = \frac{\hbar}{2}\,\boldsymbol{\sigma}$ is the spin operator.
(i) Show that the expectation values of the spin operator evolves as $\partial_t\langle\boldsymbol{S}\rangle = \boldsymbol{h}\times\langle\boldsymbol{S}\rangle$.
(ii) Show that
$\langle\boldsymbol{S}(t)\rangle = \cos(|\boldsymbol{h}|\,t)\,\langle\boldsymbol{S}(0)\rangle + \sin(|\boldsymbol{h}|\,t)\,\hat{\boldsymbol{h}}\times\langle\boldsymbol{S}(0)\rangle + (1-\cos(|\boldsymbol{h}|\,t))\,\hat{\boldsymbol{h}}\left(\hat{\boldsymbol{h}}\cdot\langle\boldsymbol{S}(0)\rangle\right)$
is a solution of $\partial_t\langle\boldsymbol{S}\rangle = \boldsymbol{h}\times\langle\boldsymbol{S}\rangle$, where $\hat{\boldsymbol{h}} = \boldsymbol{h}/|\boldsymbol{h}|$.
This describes the dynamics of a spin in a magnetic field $\boldsymbol{h}$.
(iii) Show that the spin component along the magnetic field $\hat{\boldsymbol{h}}\cdot\boldsymbol{S}$ is a conserved quantity, that generates the SO(2) symmetry of the Hamiltonian.

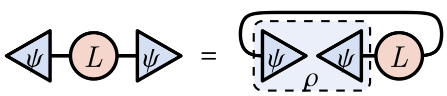
Solution (HW 7)

Density Matrix

Idea of Density Matrix

Motivation: an alternative way to think about the expectation value of an observable L

$$\langle L\rangle = \langle\psi|\, L\, |\psi\rangle = \mathrm{Tr}\, |\psi\rangle\, \langle\psi|\, L. \tag{203}$$



Introduce the **density matrix** (**density operator**) of a quantum state $|\psi\rangle$

$$\boxed{\rho = |\psi\rangle\, \langle\psi|,} \tag{204}$$

as an equivalent description of the state.

- The normalization of the *state* $\langle\psi \mid \psi\rangle = 1$ implies the **normalization** of the *density matrix*

$$\boxed{\mathrm{Tr}\, \rho = 1.} \tag{205}$$

- The **expectation value** of an *physical observable* L measured with respect to the state ρ is given by

$$\boxed{\langle L\rangle = \mathrm{Tr}\, \rho\, L.} \tag{206}$$

Example: density matrix of a qubit. Assume a qubit describe by the following state

$$|\psi\rangle = \psi_\uparrow\, |\uparrow\rangle + \psi_\downarrow\, |\downarrow\rangle \simeq \begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix}. \tag{207}$$

Density matrix can be constructed as

$$\rho = |\psi\rangle\, \langle\psi| \simeq \begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix} (\, \psi_\uparrow^* \;\; \psi_\downarrow^* \,) = \begin{pmatrix} |\psi_\uparrow|^2 & \psi_\uparrow\, \psi_\downarrow^* \\ \psi_\downarrow\, \psi_\uparrow^* & |\psi_\downarrow|^2 \end{pmatrix}. \tag{208}$$

Evaluate *expectation values* of *qubit operators* using density matrix

$$\begin{aligned} \langle\sigma^x\rangle &= \mathrm{Tr}\, \rho\, \sigma^x \simeq \mathrm{Tr}\begin{pmatrix} |\psi_\uparrow|^2 & \psi_\uparrow\, \psi_\downarrow^* \\ \psi_\downarrow\, \psi_\uparrow^* & |\psi_\downarrow|^2 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \psi_\uparrow^*\, \psi_\downarrow + \psi_\downarrow^*\, \psi_\uparrow, \\ \langle\sigma^y\rangle &= \mathrm{Tr}\, \rho\, \sigma^y \simeq \mathrm{Tr}\begin{pmatrix} |\psi_\uparrow|^2 & \psi_\uparrow\, \psi_\downarrow^* \\ \psi_\downarrow\, \psi_\uparrow^* & |\psi_\downarrow|^2 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = -i\, \psi_\uparrow^*\, \psi_\downarrow + i\, \psi_\downarrow^*\, \psi_\uparrow, \\ \langle\sigma^z\rangle &= \mathrm{Tr}\, \rho\, \sigma^z \simeq \mathrm{Tr}\begin{pmatrix} |\psi_\uparrow|^2 & \psi_\uparrow\, \psi_\downarrow^* \\ \psi_\downarrow\, \psi_\uparrow^* & |\psi_\downarrow|^2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = |\psi_\uparrow|^2 - |\psi_\downarrow|^2. \end{aligned} \tag{209}$$

What if there is a 50 × % possibility that the system is prepared in $|\psi\rangle$ and 50 × % probability in $|\phi\rangle$? The expectation value of an observable L should be

$$\begin{aligned}\langle L\rangle &= \frac{1}{2}\langle\psi| L|\psi\rangle + \frac{1}{2}\langle\phi| L|\phi\rangle \\ &= \frac{1}{2}\operatorname{Tr}|\psi\rangle\langle\psi| L + \frac{1}{2}\operatorname{Tr}|\phi\rangle\langle\phi| L \\ &= \operatorname{Tr}\left(\frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\phi\rangle\langle\phi|\right)L.\end{aligned} \tag{210}$$

We are just averaging over our *ignorance* of the **state preparation**. Now we can define a *density matrix* to describe our knowledge about the system

$$\rho = \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\phi\rangle\langle\phi|, \tag{211}$$

such that the rule to compute expectation value is still $\langle L\rangle = \operatorname{Tr}\rho L$ as in Eq. (206).

In general, the **density matrix** is defined for an **ensemble** of **quantum systems**, other than a single quantum system.

- Suppose the *system* is randomly prepared in the *state* $|\phi_i\rangle$ with *probability* p_i, the density matrix of the *ensemble* is given by

$$\boxed{\rho = \sum_i |\phi_i\rangle\, p_i\, \langle\phi_i|.} \tag{212}$$

- A density matrix should satisfy the following properties
  - **Hermitian**: $\rho^\dagger = \rho$.
  - **Normalization** (trace one): $\operatorname{Tr}\rho = 1$.
  - **Positive (semi)definite**: $\forall\, |\psi\rangle : \langle\psi|\rho|\psi\rangle \geq 0$.
- *Not* every density matrix can be expressed in the form of $|\psi\rangle\langle\psi| \Rightarrow$ A *density matrix* is richer and more general than a *state vector*.

Quantum Tomography: reconstruction of the *density matrix* from (repeated) *measurements* on the systems taken from the *ensemble*. For a single qubit, by measuring $\langle\boldsymbol{\sigma}\rangle$, the density matrix can be reconstructed as

$$\boxed{\rho = \frac{1}{2}(\mathbb{1} + \langle\boldsymbol{\sigma}\rangle\cdot\boldsymbol{\sigma}).} \tag{213}$$

As ρ is the only solution of the density matrix that is *normalized* and *reproduces* the expectation values of all *measurements* on the qubit.

HW 8 Check that the density matrix ρ in Eq. (213) is normalized $\operatorname{Tr}\rho = 1$ and reproduces all measurement expectation values $\operatorname{Tr}\rho\,\boldsymbol{\sigma} = \langle\boldsymbol{\sigma}\rangle$.

Solution (HW 8)

▪ Dynamics of Density Matrix

The *time-evolution* of the *density matrix* follows the **von Neumann equation** (also known as the Liouville-von Neumann equation)

$$\boxed{i\hbar\,\partial_t \rho(t) = [H, \rho(t)].} \tag{216}$$

- Here the density matrix is taken to be in the **Schrödinger picture**.
- Even though the *von Neumann equation* looks like the *Heisenberg equation* $i\hbar\,\partial_t L(t) = -[H, L(t)]$ (which governs the operator evolution in the Heisenberg picture), but there is a crucial sign difference.
- However in the **Heisenberg picture**, the density matrix is *time-independent*, because the *state* does not evolve in the Heisenberg picture and the density matrix *follows the state*.

HW 9 In the case of , derive the von Neumann equation Eq. (216) from the Schrödinger equation Eq. (140).

If the time-evolution of the state is described by the *unitary operator* $U(t)$, the density matrix evolves as

$$\rho(t) = U(t)\,\rho(0)\,U(t)^{\dagger}. \tag{217}$$

Example: Consider a single-qubit Hamiltonian $H = \frac{\omega}{2}\sigma^z$. Starting from the initial density matrix (in the diagonal basis of H)

$$\rho(0) \doteq \begin{pmatrix} |\psi_\uparrow|^2 & \psi_\uparrow \psi_\downarrow^* \\ \psi_\downarrow \psi_\uparrow^* & |\psi_\downarrow|^2 \end{pmatrix}. \tag{218}$$

Under time evolution (set $\hbar = 1$),

$$\rho(t) \doteq \begin{pmatrix} |\psi_\uparrow|^2 & \psi_\uparrow \psi_\downarrow^*\, e^{-i\omega t} \\ \psi_\downarrow \psi_\uparrow^*\, e^{i\omega t} & |\psi_\downarrow|^2 \end{pmatrix}. \tag{219}$$

The **diagonal** elements are *invariant*, the **off-diagonal** elements *rotates* in time following $e^{\pm i\omega t}$ (with an angular frequency of ω).

Solution (HW 9)

▪ Measurement and Decoherence

Measurement Postulate in terms of *density matrix*

- An **ensemble** of **quantum states** is described by a **density matrix** ρ.
- A **physical observable** is described by a **Hermitian operator** $L = \sum_i |\lambda_i\rangle\,\lambda_i\,\langle\lambda_i|$.

Define the **projection operator** $P(L = \lambda)$, which projects to the eigenspace of L of the eigenvalue λ (it is also fine if λ is not an eigenvalue of L, $P(L = \lambda)$ will then project out all states),

$$P(L = \lambda) = \sum_i |\lambda_i\rangle\, \delta(\lambda - \lambda_i)\, \langle\lambda_i|. \tag{223}$$

- The **probability** to observe the *measurement outcome* λ by measuring L on ρ is given by

$$\boxed{p(L = \lambda) = \operatorname{Tr} \rho\, P(L = \lambda).} \tag{224}$$

- The **expectation value** of the observable L is given by

$$\boxed{\langle L\rangle = \operatorname{Tr} \rho\, L.} \tag{225}$$

- The ensemble **post-selected** upon the *observation of outcome* λ is described by

$$\boxed{\rho \xrightarrow{\text{measure } L, \text{ get } \lambda} \frac{P(L = \lambda)\, \rho\, P(L = \lambda)}{p(L = \lambda)}.} \tag{226}$$

Measurement couples the quantum **system** to the **apparatus** (and eventually the entire **environment**). In the view of the system, suppose the coupling is resembled a *relative energy shift* between $|\uparrow\rangle$ and $|\downarrow\rangle$ states, i.e. $H = \frac{\omega}{2}\,\sigma^z$. The density matrix evolves as Eq. (219),

$$\rho(t) \doteq \begin{pmatrix} |\psi_\uparrow|^2 & \psi_\uparrow\, \psi_\downarrow^*\, e^{-i\,\omega\,t} \\ \psi_\downarrow\, \psi_\uparrow^*\, e^{i\,\omega\,t} & |\psi_\downarrow|^2 \end{pmatrix}. \tag{227}$$

If ω is *large* (the coupling is strong) and *noisy* (the environment is chaotic), $e^{\pm i\,\omega\,t}$ looks like a fast fluctuating **random phase**, which *averages* to zero over a short period of time.

$$\begin{aligned} \overline{\rho} &= \frac{1}{T}\int_0^\infty \rho(t)\, e^{-t/T}\, dt \\ &\doteq \begin{pmatrix} |\psi_\uparrow|^2 & \frac{\psi_\uparrow\,\psi_\downarrow^*}{1+i\,\omega\,T} \\ \frac{\psi_\downarrow\,\psi_\uparrow^*}{1-i\,\omega\,T} & |\psi_\downarrow|^2 \end{pmatrix} \xrightarrow{\omega\,T \gg 1} \begin{pmatrix} |\psi_\uparrow|^2 & 0 \\ 0 & |\psi_\downarrow|^2 \end{pmatrix}. \end{aligned} \tag{228}$$

The **off-diagonal** elements of the density matrix **decays** much more *quickly* than the **diagonal** elements, due to its fast oscillating phase (in this model). (We will come back later with a better model.)

Quantum Decoherence (brief idea): the **loss** of *off-diagonal* density matrix elements (**quantum coherence**) over time in the *measurement basis* determined by how the system is coupled to the apparatus.

After quantum decoherence, the *time-averaged density matrix*

$$\overline{\rho} = |\uparrow\rangle\, |\psi_\uparrow|^2\, \langle\uparrow| + |\downarrow\rangle\, |\psi_\downarrow|^2\, \langle\downarrow| \tag{229}$$

describes a qubit ensemble with

$$\begin{array}{ll} \text{probability} & \text{to be in the sate} \\ \quad |\psi_\uparrow|^2 & \quad |\uparrow\rangle, \\ \quad |\psi_\downarrow|^2 & \quad |\downarrow\rangle. \end{array} \tag{230}$$

Note: quantum *decoherence* dose *not* generate *actual* quantum state *collapse*. It only provides an *ensemble* of quantum states that matches the *measurement postulate*. The **measurement**

problem "How the *measurement* actually leads to the *realization* of precisely one state in the ensemble?" remains an issue of interpretation.

■ Quantum Channel*

Transmitting a particle through a **quantum channel**, its density matrix ρ may undergo

- a **unitary evolution** (time evolution)

$$\rho \to U\,\rho\,U^{\dagger}, \tag{231}$$

- a **projective measurement** (measure and obtain a definite outcome)

$$\rho \to P\,\rho\,P, \quad \text{(without normalization)} \tag{232}$$

$$\rho \to \frac{P\,\rho\,P}{\mathrm{Tr}(P\,\rho\,P)}, \quad \text{(with normalization)} \tag{233}$$

the normalization factor $\mathrm{Tr}\,P\,\rho\,P = \mathrm{Tr}\,\rho\,P$ is the probability to obtain the out come.

The *evolution* and *measurement* can be unified as **quantum operations**, described by a Kraus operator K

$$\rho \to K\,\rho\,K^{\dagger}, \quad \text{(without normalization)} \tag{234}$$

$$\rho \to \frac{K\,\rho\,K^{\dagger}}{\mathrm{Tr}\big(K\,\rho\,K^{\dagger}\big)}. \quad \text{(with normalization)} \tag{235}$$

A sequence of *quantum operations* put together forms a *quantum channel.*

Example: Quantum optics

- *Polarization* of photon
  - σ^z basis states: *horizontal* and *vertical* polarizations

$$|\leftrightarrow\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ |\updownarrow\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{236}$$

  - σ^x basis states: 45° polarizations

$$|\nearrow\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ |\searrow\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{237}$$

  - σ^y basis states: *circular* polarizations

$$|\circlearrowleft\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix},\ |\circlearrowright\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \tag{238}$$

  - Linear polarization along θ angle (with respect to x-axis)

$$|\theta\rangle \doteq \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \Rightarrow \rho_\theta = |\theta\rangle\,\langle\theta| \doteq \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}. \tag{239}$$

  - *Natural light*: an ensemble of all possible polarizations with equal probability $\Rightarrow$ *maximally mixed* state

$$\rho = \int \frac{d\theta}{2\pi}\, \rho_\theta \doteq \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{240}$$

The density matrix description of light polarization is also known as the **Jones matrix**.

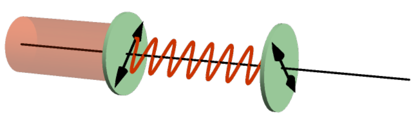
- Unitary evolution: *phase retarders* creates ϕ relative phase shift between horizontal and vertical polarizations

$$U_\phi = e^{i \frac{\phi}{2} \sigma^z} \doteq \begin{pmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{pmatrix}. \tag{241}$$

- Projective measurement: *polarizers* oriented along θ angle axis

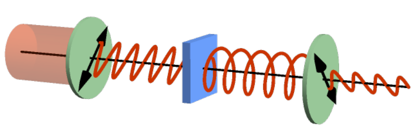
$$P_\theta = |\theta\rangle \langle\theta| \doteq \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}. \tag{242}$$

Natural light going through two perpendicular polarizers ⇒ no transmission.



$$\begin{aligned} \rho' &= P_{-\pi/4}\, P_{\pi/4}\, \rho\, P_{\pi/4}\, P_{-\pi/4} \\ &\doteq \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ &\Rightarrow \mathrm{Tr}\,\rho' = 0. \end{aligned} \tag{243}$$

Insert a phase retarder between the polarizers ⇒ $1/4$ transmission!



$$\begin{aligned} \rho' &= P_{-\pi/4}\, U_{\pi/2}\, P_{\pi/4}\, \rho\, P_{\pi/4}\, U^\dagger_{\pi/2}\, P_{-\pi/4} \\ &\doteq \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} e^{i\pi/4} & 0 \\ 0 & e^{-i\pi/4} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{1}{8} \end{pmatrix} \\ &\Rightarrow \mathrm{Tr}\,\rho' = \frac{1}{4}. \end{aligned} \tag{244}$$

Pure State and Mixed State

- **Pure state**: a coherent quantum state, described by a **state vector** $|\psi\rangle$, or a **pure state density matrix** *of the form* $\rho = |\psi\rangle \langle\psi|$.
- **Mixed state**: a statistical mixture of pure states, can *not* be described by any single state vector, described by a **mixed state density matrix** as a *superposition* of *pure state density matrices.*
- **Superposition** at different levels:
    - **Quantum superposition** (pure state superposition): superposition of **state vectors**

      $$|\psi\rangle = z_1 |\phi_1\rangle + z_2 |\phi_2\rangle + \ldots. \tag{245}$$

      The result is still a *pure state.*
    - **Statistical superposition** (mixed state superposition): superposition of **density matrices**

      $$\rho = p_1 |\phi_1\rangle \langle\phi_1| + p_2 |\phi_2\rangle \langle\phi_2| + \ldots, \tag{246}$$

      or more generally, $\rho = p_1 \rho_1 + p_2 \rho_2 + \ldots$. The result is generally a *mixed state.*

In terms of the *density matrix*, a *quantum superposition* of Eq. (245) is expressed as

$$\begin{aligned} |\psi\rangle \langle\psi| = {} & |z_1|^2 |\phi_1\rangle \langle\phi_1| + |z_2|^2 |\phi_2\rangle \langle\phi_2| + \\ & + z_1 z_2^* |\phi_1\rangle \langle\phi_2| + z_2 z_1^* |\phi_2\rangle \langle\phi_1| + \ldots, \end{aligned} \tag{247}$$

also involves *cross terms* that represents **quantum coherence**.

Spectral decomposition of the density matrix

$$\boxed{\rho = \sum_i p_i |\phi_i\rangle \langle\phi_i|.} \tag{248}$$

- As ρ is *Hermitian*, its *eigenvectors* $|\phi_i\rangle$ form an *orthonormal basis.*
- The *eigenvalues* p_i has the physical meaning of *probability*, with the following properties:
    - **Hermitian**: $\rho^\dagger = \rho \Leftrightarrow p_i \in \mathbb{R}$.
    - **Normalization** (trace one): $\operatorname{Tr} \rho = 1 \Leftrightarrow \sum_i p_i = 1$.
    - **Positive (semi)definite**: $\forall\, |\psi\rangle : \langle\psi| \rho |\psi\rangle \geq 0 \Leftrightarrow p_i \geq 0$.

The density matrix ρ describes an **ensemble** of quantum systems, where each *pure state* $|\phi_i\rangle$ is *prepared* with *probability* p_i.

- If p_i have only a single one followed by all zeros, e.g. $p_1 = 1$, $p_2 = p_3 = \ldots = 0$, the *density matrix* ρ is **pure**, since it can be written as $\rho = |\phi_1\rangle \langle\phi_1|$.
- Otherwise, for generic distribution of p_i, the *density matrix* ρ is **mixed**.

Purity: to quantify to which degree the density matrix is pure/mixed,

$$\operatorname{Tr} \rho^2 = \sum_i p_i^2 \tag{249}$$

By construction, $\mathrm{Tr}\,\rho^2 \in [0, 1]$. The criteria to determine if a density matrix ρ is pure or mixed is

$$\rho \text{ is } \begin{cases} \text{pure} & \text{if } \mathrm{Tr}\,\rho^2 = 1, \\ \text{mixed} & \text{if } \mathrm{Tr}\,\rho^2 < 1. \end{cases} \tag{250}$$

HW 10

(i) Show that for a single qubit, the purity is related to the spin expectation value $\langle \boldsymbol{\sigma} \rangle = \mathrm{Tr}\,\rho\,\boldsymbol{\sigma}$ by $\mathrm{Tr}\,\rho^2 = \left(1 + \langle \boldsymbol{\sigma} \rangle^2\right)/2$.
(ii) For pure state, what is the norm of the spin expectation value $|\langle \boldsymbol{\sigma} \rangle|$?
(iii) What is the minimal possible purity of a qubit? When the minimal purity is achieved (the qubit is maximally mixed) what is the spin expectation value $\langle \boldsymbol{\sigma} \rangle$?

Solution (HW 10)

▪ von Neumann and Rényi Entropy

von Neumann entropy of a density matrix

$$\boxed{S^{(1)} = -\mathrm{Tr}\,\rho \ln \rho.} \tag{252}$$

In terms of the *eigenvalues* p_i, $S^{(1)} = -\sum_i p_i \ln p_i$ matches the **Shannon entropy** of a *probability distribution* in the information theory. [Note: $0 \ln 0$ should be treated as 0 in this calculation]

HW 11

Consider a generic single-qubit density matrix of the following form
$\rho = \frac{1}{2}\,(\mathbb{1} + \boldsymbol{m} \cdot \boldsymbol{\sigma})$,
where $\boldsymbol{m}$ is a three-component real vector. Calculate its von Neumann entropy $S^{(1)}$. Show that $S^{(1)} = 0$ when $|\boldsymbol{m}| = 1$, and $S^{(1)} = \ln 2$ when $|\boldsymbol{m}| = 0$.

Rényi entropy of a density matrix

$$\boxed{S^{(n)} = \frac{1}{1-n} \ln \mathrm{Tr}\,\rho^n.} \tag{253}$$

In terms of the *eigenvalues* p_i, $S^{(n)} = (1-n)^{-1} \ln \sum_i p_i^n$.

- n is the **Rényi index**.
  - $n = 0$: **max-entropy**, simply counts the log of the Hilbert space dimension $S^{(0)} = \ln \dim \mathcal{H}$.
  - $n \to 1$ limit: equivalent to the **von Neumann entropy**, i.e. $S^{(1)} = \lim_{n \to 1} S^{(n)}$.

HW 12

Show that in the $n \to 1$ limit, the Rényi entropy reduces to the von Neumann entropy.

  - $n = 2$: the **2nd Rényi entropy** is directly related to **purity** by $S^{(2)} = -\ln \mathrm{Tr}\,\rho^2$.
  - $n = \infty$: **min-entropy**, lower bound of all Rényi entropies, $S^{(\infty)} = -\ln \max_i p_i$.
- The **spectrum** of the **density matrix**, i.e. all *eigenvalues* p_i, can be *reconstructed* from the family of *Rényi entropies* (by solving the following equations, in principle).

$$\sum_i p_i^n = e^{(1-n)\,S^{(n)}} \quad (\text{for } n = 1, 2, \ldots, \dim \mathcal{H}). \tag{254}$$

Solution (HW 10)

Solution (HW 11)

■ Entropy and Knowledge

The *Rényi entropy* (including the *von Neumann entropy* as a special case) can characterize how much the *ensemble* is *mixed*.

$$\rho \text{ is } \begin{cases} \text{pure} & \text{if } S^{(n)} = 0, \\ \text{mixed} & \text{if } S^{(n)} > 0, \end{cases} \text{for } n = 1, 2, \ldots. \tag{258}$$

Pure state has **no entropy**. A *pure* state represents the *maximal knowledge* we can have of a system.

Entropy *measures* our **ignorance** about the quantum system. If the ensemble is *pure*, the system is in a *definite* quantum state, hence no entropy. If the ensemble is *mixed*, there are several possible states that the system can take, our *ignorance* is quantified by the *entropy*.

- **Jensen's inequality**: Rényi entropy is generally *decreasing* with the Rényi index,

$$\ln \dim \mathcal{H} = S^{(0)} \geq S^{(1)} \geq S^{(2)} \geq \ldots \geq S^{(\infty)} \geq 0. \tag{259}$$

The *equality* is achieved (simultaneously) if all p_i are *equal*.

$$\forall\, i : p_i = \frac{1}{\dim \mathcal{H}} \Rightarrow \forall\, n \geq 0 : S^{(n)} = \ln \dim \mathcal{H}. \tag{260}$$

In this case, all *Rényi entropies* reach the *maximum*, and the *ensemble* is **maximally mixed**. The *density matrix* is proportional to *identity matrix* for *maximally mixed* ensemble.

$$\rho = \frac{1}{\dim \mathcal{H}}\, \mathbb{1}. \tag{261}$$

Any quantum state can be realized with *equal possibility* in a *maximally mixed* ensemble ⇒ we are *completely ignorant* about the system ⇒ *entropy* is therefore *maximized*.

Maximally mixed qubit: SU(2) symmetric, no preferred spin direction, i.e. $\langle \boldsymbol{\sigma} \rangle = 0$. Then according to Eq. (213),

$$\rho = \mathbb{1}/2 \doteq \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{262}$$

- Application: if the qubit basis corresponds to the *left-circular* and *right-circular* **photon polarization**, then the density matrix in Eq. (262) describes the **natural light** ensemble of photons.
- All Rényi entropies are identically $\ln 2$ for a maximally mixed qubit,

$$S^{(n)} = \frac{1}{1-n} \ln\left(\frac{1}{2^n} + \frac{1}{2^n}\right) = \ln 2 = 1 \text{ bit}. \tag{263}$$

- This is the *maximal entropy* that a qubit could have: our ignorance about a qubit is at most 1 bit. This is why a *qubit* is called a **quantum bit**.

Let us conclude our discussion in the following table:

ensemble	pure	mixed	maximally mixed
entropy	0	$\longleftrightarrow$	$\ln \dim \mathcal{H}$
knowledge	max	$\longleftrightarrow$	none

Quantum Entanglement

▪ Two-Qubit Systems

▪ Two-Qubit States

Each qubit has *two* basis states $|\uparrow\rangle$ and $|\downarrow\rangle$ (forming a 2-dim Hilbert space) $\Rightarrow$ two qubits together have *four* basis states

$$\begin{array}{cc|cc} & & \multicolumn{2}{c}{\text{qubit } B} \\ & & |\uparrow\rangle & |\downarrow\rangle \\ \hline \text{qubit } A & |\uparrow\rangle & |\uparrow\uparrow\rangle & |\uparrow\downarrow\rangle \\ & |\downarrow\rangle & |\downarrow\uparrow\rangle & |\downarrow\downarrow\rangle \end{array} \tag{264}$$

The precise meaning of $|\uparrow\uparrow\rangle$ is a **tensor product** of $|\uparrow\rangle_A$ and $|\uparrow\rangle_B$ states. In the *vector representation*,

$$|\uparrow\uparrow\rangle = |\uparrow\rangle_A \otimes |\uparrow\rangle_B \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \left(\begin{array}{c} 1 \\ 0 \\ \hline 0 \\ 0 \end{array}\right). \tag{265}$$

Similarly,

$$\begin{aligned} |\uparrow\downarrow\rangle &= |\uparrow\rangle_A \otimes |\downarrow\rangle_B \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \left(\begin{array}{c} 0 \\ 1 \\ \hline 0 \\ 0 \end{array}\right), \\ |\downarrow\uparrow\rangle &= |\downarrow\rangle_A \otimes |\uparrow\rangle_B \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \left(\begin{array}{c} 0 \\ 0 \\ \hline 1 \\ 0 \end{array}\right), \\ |\downarrow\downarrow\rangle &= |\downarrow\rangle_A \otimes |\downarrow\rangle_B \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \left(\begin{array}{c} 0 \\ 0 \\ \hline 0 \\ 1 \end{array}\right). \end{aligned} \tag{266}$$

These four *basis states* span the **two-qubit Hilbert space**.

Note: in general, the tensor product of vectors follows

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \otimes \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} z_1 \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \\ z_2 \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} z_1\, w_1 \\ z_1\, w_2 \\ \hline z_2\, w_1 \\ z_2\, w_2 \end{pmatrix}. \tag{267}$$

This is consistent with the tensor product of matrices in Eq. (168).

A *generic state* in the *two-qubit Hilbert space* is a superposition of these four basis states,

$$|\psi\rangle = \psi_1\, |\uparrow\uparrow\rangle + \psi_2\, |\uparrow\downarrow\rangle + \psi_3\, |\downarrow\uparrow\rangle + \psi_4\, |\downarrow\downarrow\rangle \doteq \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}. \tag{268}$$

Normalization is still expected: $\langle\psi \mid \psi\rangle = \sum_i |\psi_i|^2 = 1$.

- **Product state**: a state that can be *factorized* as a *tensor product* of *single-qubit states*.

Suppose $|z\rangle = z_1\, |\uparrow\rangle + z_2\, |\downarrow\rangle$ is a state of the *first* qubit and $|w\rangle = w_1\, |\uparrow\rangle + w_2\, |\downarrow\rangle$ is a state of the *second* qubit. A *two-qubit* **product state** takes the general form of

$$\begin{aligned} &|z\rangle \otimes |w\rangle = (z_1\, |\uparrow\rangle + z_2\, |\downarrow\rangle) \otimes (w_1\, |\uparrow\rangle + w_2\, |\downarrow\rangle) \\ &= z_1\, w_1\, |\uparrow\uparrow\rangle + z_1\, w_2\, |\uparrow\downarrow\rangle + z_2\, w_1\, |\downarrow\uparrow\rangle + z_2\, w_2\, |\downarrow\downarrow\rangle. \end{aligned} \tag{269}$$

The main feature of a *product state* is that each qubit behaves **independently** of the other: *measurement* or *unitary operation* of one qubit will *not affect* the other.

Not every state in the *two-qubit Hilbert space* can be written as *product state*. Why? Let us count the degrees of freedom:

- A generic state as $|\psi\rangle$ in Eq. (268) has *six* real parameters. $4 \times 2 - 1 - 1 = 6$.
- A generic *product* state as $|z\rangle \otimes |w\rangle$ in Eq. (269) has only *four* real parameters. $(2 \times 2 - 1 - 1) \times 2 = 4$.

A generic state has more freedom than a product state, the additional freedom has to do with **quantum entanglement**.

- **Entangled state**: any state that can *not* be factorized to *product states* are *entangled*.

  Example: the state $\frac{1}{\sqrt{2}}\,(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ is entangled.

HW 13 Prove that $\frac{1}{\sqrt{2}}\,(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ can not be written as a product state.

Question: Is the state $\frac{1}{2}\,(|\uparrow\uparrow\rangle + |\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle + |\downarrow\downarrow\rangle)$ entangled?

It is not obvious to see if a state is entangled or not ⇒ we need to develop *measures of entanglement*, such that by measuring these quantities, we can decide how much the state is entangled… (to be discussed later)

Solution (HW 13)

■ Two-Qubit Operators

Any *physical observable* of a two-qubit system is represented as a *Hermitian operator* acting on the two-qubit Hilbert space.

- **Single-qubit observables**:

$$\begin{aligned}\boldsymbol{\sigma}_A &= (\sigma_A^x, \sigma_A^y, \sigma_A^z),\\ \boldsymbol{\sigma}_B &= (\sigma_B^x, \sigma_B^y, \sigma_B^z).\end{aligned} \tag{271}$$

- **Two-qubit observables** (joint measurements):

$$\boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B = \begin{pmatrix} \sigma_A^x \sigma_B^x & \sigma_A^y \sigma_B^x & \sigma_A^z \sigma_B^x \\ \sigma_A^x \sigma_B^y & \sigma_A^y \sigma_B^y & \sigma_A^z \sigma_B^y \\ \sigma_A^x \sigma_B^z & \sigma_A^y \sigma_B^z & \sigma_A^z \sigma_B^z \end{pmatrix}. \tag{272}$$

The precise meaning of σ_A^x:

$$\sigma_A^x \otimes \mathbb{1}_B \doteq \sigma^{10} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right). \tag{273}$$

The precise meaning of $\sigma_A^z \sigma_B^y$:

$$\sigma_A^z \otimes \sigma_B^y \doteq \sigma^{32} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{array}\right). \tag{274}$$

Note: the *tensor product* of *matrices* should be *consistent* with that of *vectors*.

The *single-qubit observables* $\boldsymbol{\sigma}_A$, $\boldsymbol{\sigma}_B$, *two-qubit observables* $\boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B$ together with the *identity observable* $\mathbb{1}$ (altogether $3+3+3\times 3+1 = 16$ observables) form the **complete set of observables** for a *two-qubit system*, i.e. any *physical observables* of a two-qubit system must be a *linear superposition* of these 16 *basis observables*.

■ A Two-Qubit Model

Two-qubit Heisenberg model. Consider *two qubits* governed by the *Hamiltonian*

$$H = \frac{J}{4} \boldsymbol{\sigma}_A \cdot \boldsymbol{\sigma}_B = \frac{J}{4} \left(\sigma_A^x \sigma_B^x + \sigma_A^y \sigma_B^y + \sigma_A^z \sigma_B^z\right). \tag{275}$$

First write down the matrix representation,

$$H \doteq \frac{J}{4} \left(\sigma^{11} + \sigma^{22} + \sigma^{33}\right) = \frac{J}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{276}$$

Then diagonalize the Hamiltonian.

- Eigenvalue $E_s = -3\,J/4$: a unique eigenstate ⇒ **spin-singlet** state

$$|s\rangle = \frac{1}{\sqrt{2}}\,(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \tag{277}$$

- Eigenvalue $E_t = J/4$: three degenerated eigenstates ⇒ **spin-triplet** states (there is a basis freedom here, we make the following choice)

$$\begin{aligned} |t_1\rangle &= \frac{1}{\sqrt{2}}\,(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle), \\ |t_2\rangle &= \frac{1}{\sqrt{2}}\,(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle), \\ |t_3\rangle &= \frac{1}{\sqrt{2}}\,(|\uparrow\uparrow\rangle - |\downarrow\downarrow\rangle). \end{aligned} \tag{278}$$

The *lowest energy eigenstate* is called the **ground state**, the rest of the eigenstates are **excited states**. In this model, assuming $J > 0$, the *ground state* is the *spin-singlet* state.

- Classical picture: $H = (J/4)\,\boldsymbol{\sigma}_A \cdot \boldsymbol{\sigma}_B$ with $J > 0$ ⇒ energy is lowered if $\boldsymbol{\sigma}_A \cdot \boldsymbol{\sigma}_B < 0$, i.e. $\boldsymbol{\sigma}_A$ and $\boldsymbol{\sigma}_B$ are *anti-aligned*, or in an **antiferromagnetic correlation**.
- The *singlet state* is a superposition of $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$, consistent with the classical picture, but there is more to explore.

■ The Spin-Singlet State

Use the *vector representation* of the *spin-single state*

$$|s\rangle = \frac{1}{\sqrt{2}}\,(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \doteq \frac{1}{\sqrt{2}}\,(\,0\;\;1\;\;-1\;\;0\,)^{\mathrm{T}}. \tag{279}$$

- Expectation value of **single-qubit** observables

$$\begin{aligned} \langle s|\,\boldsymbol{\sigma}_A\,|s\rangle &= (0,\,0,\,0), \\ \langle s|\,\boldsymbol{\sigma}_B\,|s\rangle &= (0,\,0,\,0). \end{aligned} \tag{280}$$

- Expectation value of **two-qubit** observables

$$\langle s|\,\boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B\,|s\rangle = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{281}$$

There is something unusual!

- $|s\rangle$ is a **pure state** of the *two-qubit* system ⇒ the system is in a *definite* quantum state, *entropy* of the *entire system* $= 0$ ⇒ we have the *full knowledge* about the system.
- However $\langle s|\,\boldsymbol{\sigma}_A\,|s\rangle = 0$ implies nothing is know about qubit A, because qubit A is in a **maximally mixed state** with maximal *entropy* of the *subsystem* (1bit) ⇒ we are *completely ignorant* about the subsystems. (Same argument applies for qubit B)

The phenomenon that we may know *everything* about a *quantum system* yet *nothing* about its *subsystems* is a demonstration of **quantum entanglement**.

- **Classical information** is stored *locally* (bit-by-bit) in every single classical bit. Knowing the entire system = knowing the state of every classical bit.
- **Quantum information** can be stored *jointly* in the *interrelations* among qubits, but *not locally* in single qubits. Knowing the entire system does not imply the knowledge of its subsystem.

Entanglement Entropy

The **entanglement entropy** of the qubit A in a two-qubit state $|\psi\rangle$ is given by

$$\boxed{S(A) = -\operatorname{Tr} \rho_A \ln \rho_A.} \tag{282}$$

where ρ_A is the **reduced density matrix** of qubit A obtained by *tracing out* qubit B in the full **density matrix** $|\psi\rangle\langle\psi|$

$$\boxed{\rho_A = \operatorname{Tr}_B |\psi\rangle\langle\psi|.} \tag{283}$$

One may also define a more general *Rényi version* as

$$\boxed{S^{(n)}(A) = \frac{1}{1-n} \ln \operatorname{Tr} \rho_A^n.} \tag{284}$$

Example I: take the **spin-singlet state**

$$|\psi\rangle = |s\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \tag{285}$$

- Full density matrix

$$|s\rangle\langle s| \doteq \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix} = \frac{1}{2}\left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ \hline 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right). \tag{286}$$

- *Partial trace* over qubit $B \Rightarrow$ *reduced density matrix* of qubit A

$$\begin{aligned} \rho_A &= \operatorname{Tr}_B |s\rangle\langle s| \\ &\doteq \frac{1}{2}\begin{pmatrix} \operatorname{tr}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} & \operatorname{tr}\begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \\ \operatorname{tr}\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} & \operatorname{tr}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{287}$$

Note that ρ_A indeed describes a *maximally mixed* qubit.

- Compute the *entropy* of the *reduced density matrix*,

$$S(A) = -\operatorname{Tr} \rho_A \ln \rho_A = \ln 2 = 1\,\text{bit}. \tag{288}$$

Example II: take the **product state**

$$|\psi\rangle = \frac{1}{2} (|\uparrow\uparrow\rangle + |\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle + |\downarrow\downarrow\rangle). \tag{289}$$

- Full density matrix

$$\rho = |\psi\rangle\langle\psi| \doteq \frac{1}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}(1\ 1\ 1\ 1) = \frac{1}{4}\left(\begin{array}{cc|cc}1&1&1&1\\1&1&1&1\\\hline 1&1&1&1\\1&1&1&1\end{array}\right). \tag{290}$$

- *Partial trace* over qubit $B \Rightarrow$ *reduced density matrix* of qubit A

$$\begin{aligned}\rho_A &= \operatorname{Tr}_B \rho \\ &\doteq \frac{1}{4}\begin{pmatrix}\operatorname{tr}\begin{pmatrix}1&1\\1&1\end{pmatrix} & \operatorname{tr}\begin{pmatrix}1&1\\1&1\end{pmatrix}\\ \operatorname{tr}\begin{pmatrix}1&1\\1&1\end{pmatrix} & \operatorname{tr}\begin{pmatrix}1&1\\1&1\end{pmatrix}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1&1\\1&1\end{pmatrix}.\end{aligned} \tag{291}$$

- Compute the *entropy* of the *reduced density matrix*,

$$S(A) = -\operatorname{Tr} \rho_A \ln \rho_A = -(0 \ln 0 + 1 \ln 1) = 0\,\text{bit}. \tag{292}$$

Conclusion: The **entanglement entropy** characterizes the amount of **quantum entanglement** between subsystem A and its complement $\overline{A}$ (which is B here), given that the full system $A \cup \overline{A}$ is *pure*.

$\lvert\psi\rangle$ (pure)	product	entangled	maximally entangled
ρ_A	pure	mixed	maximally mixed
$S^{(n)}(A)$	0	$\longleftrightarrow$	$\ln \dim \mathcal{H}$
entanlement	none	$\longleftrightarrow$	max

(293)

For diagnostic purpose (to distinguish product state from entangled state), any *Rényi index* $n = 1, 2, \ldots$ will work.

Why entropy provides a measure of entanglement? Quantum entanglement: the *nonlocal* nature of *quantum information* in an *entangled* state (i.e. information shared jointly among subsystems) $\Rightarrow$ separating out a subsystem would lead to *lost* of *information* $\Rightarrow$ hence the *production* of (entanglement) *entropy*.

Open questions: The *system* must be *pure*, otherwise there are other source of entropy productions. What about entanglement in a *mixed* state? Good to describe *bipartite* entanglement. What about *multipartite* entanglement?

■ Mutual Information

The **mutual information** between qubit A and qubit B is

$$\boxed{I(A:B) = S(A) + S(B) - S(A \cup B).} \tag{294}$$

Or more generally, one may define the *Rényi version*,

$$\boxed{I^{(n)}(A:B) = S^{(n)}(A) + S^{(n)}(B) - S^{(n)}(A \cup B).} \tag{295}$$

- $I^{(n)}(A:B) =$ the amount of *information* shared by A and B.
- **Subadditivity of entropy** $S^{(n)}(A) + S^{(n)}(B) \geq S^{(n)}(A \cup B) \Leftrightarrow$ **positivity of mutual information** $I^{(n)}(A:B) \geq 0$.

Example: take the **spin-singlet state**, we have

$$\begin{aligned} &S^{(n)}(A) = S^{(n)}(B) = 1\,\text{bit}, \\ &S^{(n)}(A \cup B) = 0\,\text{bit}, \end{aligned} \tag{296}$$

hence 2 bit mutual information (regardless of the Rényi index n)

$$I^{(n)}(A:B) = S^{(n)}(A) + S^{(n)}(B) - S^{(n)}(A \cup B) = 2\,\text{bit}. \tag{297}$$

This is a surprising result!

- For classical systems, the *mutual information* between two *classical bits* will never exceed 1 bit. How can we tell more than 1 bit of information about B by measuring A?
- The maximal mutual information between two classical bits is achieved when they are perfectly correlated, e.g.

$$p(\uparrow\downarrow) = p(\downarrow\uparrow) = 1/2,\ \ p(\uparrow\uparrow) = p(\downarrow\downarrow) = 0. \tag{298}$$

- **Entanglement** is *more* than **correlation**: the *extra bit* of quantum information *shared* between qubits A and B is their *quantum entanglement*, that goes beyond the classical correlation.

For a two-qubit system, the *2nd Rényi* ($n = 2$) mutual information $I^{(2)}(A:B)$ between the two qubits is related to the *spin observables* in a relatively simple way

$$I^{(2)}(A:B) = \ln\left(1 + \frac{\|\langle \boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B \rangle\|^2 - \|\langle \boldsymbol{\sigma}_A \rangle\|^2\ \|\langle \boldsymbol{\sigma}_B \rangle\|^2}{\left(1 + \|\langle \boldsymbol{\sigma}_A \rangle\|^2\right)\left(1 + \|\langle \boldsymbol{\sigma}_B \rangle\|^2\right)}\right). \tag{299}$$

Note: $\|\langle \boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B \rangle\|^2 = \sum_{i,j=x,y,z} \left\langle \sigma_A^i \otimes \sigma_B^j \right\rangle^2$ and $\|\langle \boldsymbol{\sigma}_A \rangle\|^2 = \sum_{i=x,y,z} \left\langle \sigma_A^i \right\rangle^2$.

HW 14

> Prove Eq. (299). Hint: by quantum tomography, the two-qubit density matrix reads
> $\rho = \frac{1}{4}\left(\mathbb{1} + \langle \boldsymbol{\sigma}_A \rangle \cdot \boldsymbol{\sigma}_A + \langle \boldsymbol{\sigma}_B \rangle \cdot \boldsymbol{\sigma}_B + \boldsymbol{\sigma}_A \cdot \langle \boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B \rangle \cdot \boldsymbol{\sigma}_B\right)$.
> Hint: the following identities will be useful
> $\mathrm{Tr}(A \otimes B) = (\mathrm{Tr}\,A)(\mathrm{Tr}\,B)$, $\mathrm{Tr}\left(\sigma^i \sigma^j\right) = 2\,\delta^{ij}$.

- **Classical state**: *statistical* superposition

$$\rho = \frac{1}{2}\,|\uparrow\downarrow\rangle\,\langle\uparrow\downarrow| + \frac{1}{2}\,|\downarrow\uparrow\rangle\,\langle\downarrow\uparrow|, \tag{300}$$

- Observables

$$\begin{aligned}&\langle\boldsymbol{\sigma}_A\rangle = \langle\boldsymbol{\sigma}_B\rangle = (0, 0, 0),\\ &\langle\boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B\rangle = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.\end{aligned} \tag{301}$$

- Mutual information

$$I^{(2)}(A:B) = \ln\!\left(1 + \|\,\langle\boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B\rangle\,\|^2\right) = \ln(1+1) = \ln 2 = 1\text{ bit}. \tag{302}$$

- **Quantum state**: *quantum* superposition

$$\begin{aligned}&\rho = |s\rangle\,\langle s|,\\ &|s\rangle = \frac{1}{\sqrt{2}}\,(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle).\end{aligned} \tag{303}$$

- Observables

$$\begin{aligned}&\langle\boldsymbol{\sigma}_A\rangle = \langle\boldsymbol{\sigma}_B\rangle = (0, 0, 0),\\ &\langle\boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B\rangle = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.\end{aligned} \tag{304}$$

- Mutual information

$$I^{(2)}(A:B) = \ln\!\left(1 + \|\,\langle\boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B\rangle\,\|^2\right) = \ln(1+3) = \ln 4 = 2\text{ bit}. \tag{305}$$

In a spin-singlet state, not only $\sigma_A^z\,\sigma_B^z$ is perfectly correlated, but $\sigma_A^x\,\sigma_B^x$ and $\sigma_A^y\,\sigma_B^y$ are *also* perfectly correlated. Such additional correlations (by changing *measurement basis*) can not be realized by classical bits. The additional information channel enables the *two-qubit* system to store all its *two bits* of *quantum information* purely in the "cloud", as *shared information* between qubits, without using any "local storage".

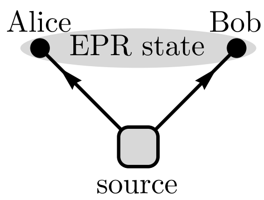
Solution (HW 14)

▪ EPR Pair and Bell Inequality

Bell states: *maximally entangled pure* states of two qubits. Also known as Einstein-Podolsky-Rosen **(EPR) pair** states. The *spin-singlet* state in Eq. (277) is one example. Here is another example:

$$|\text{EPR}\rangle = \frac{1}{\sqrt{2}}\,(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle). \tag{309}$$

Suppose a machine can repeatedly *prepare* such EPR pairs and *distribute* the qubits separately to Alice and Bob,



Alice and Bob can measure their own qubit and record the measurement outcome. After the measurement, the pair of qubits are discarded. New EPR pairs will be acquired from the source.

- Alice defines her set of observables:

$$\boldsymbol{\sigma}_A = \left(\sigma_A^x, \sigma_A^y, \sigma_A^z\right) \doteq \left(\sigma^{10}, \sigma^{20}, \sigma^{30}\right). \tag{310}$$

- Bob defines his set of observables:

$$\boldsymbol{\sigma}_B = \left(\sigma_B^x, \sigma_B^y, \sigma_B^z\right) \doteq \left(\sigma^{01}, -\sigma^{02}, \sigma^{03}\right). \tag{311}$$

  Note that σ_B^y is defined unusually with a *minus sign* (Bob has the freedom to define his σ^y).

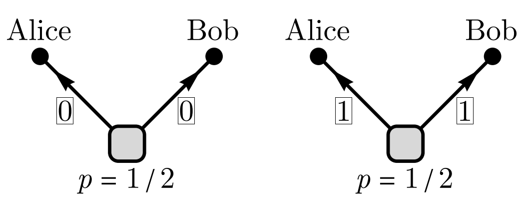
- Such *choice* of observables provides some convenience: the observables are *perfectly correlated* between Alice and Bob

$$\begin{aligned} &\langle \text{EPR} | \boldsymbol{\sigma}_A | \text{EPR} \rangle = \langle \text{EPR} | \boldsymbol{\sigma}_B | \text{EPR} \rangle = (0, 0, 0), \\ &\langle \text{EPR} | \boldsymbol{\sigma}_A \otimes \boldsymbol{\sigma}_B | \text{EPR} \rangle = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{312}$$

  If Alice and Bob both measure σ^z, they will find

$$\sigma_A^z = \sigma_B^z = \begin{cases} +1 & p = 1/2 \\ -1 & p = 1/2 \end{cases}. \tag{313}$$

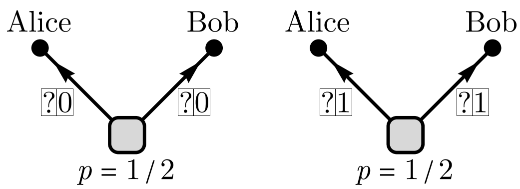
- *Quantum* explanation: can be inferred from $\langle \sigma_A^z \rangle = \langle \sigma_B^z \rangle = 0$ and $\langle \sigma_A^z \sigma_B^z \rangle = 1$.
- This is not too surprising: just a perfect correlation between two random variables. *Classically*, one may model the perfect correlation by a **hidden variable**:



  If Alice and Both both measure σ^x, they will find

$$\sigma_A^x = \sigma_B^x = \begin{cases} +1 & p = 1/2 \\ -1 & p = 1/2 \end{cases}. \tag{314}$$

- *Quantum* explanation: can be inferred from $\langle \sigma_A^x \rangle = \langle \sigma_B^x \rangle = 0$ and $\langle \sigma_A^x \sigma_B^x \rangle = 1$.
- To model this *classically*: we will need to introduce *another* hidden variable to encode the perfect correlation in σ^x channel.



As Alice and Bob can choose to measure either σ^z or σ^x at their *free will* $\Rightarrow$ *Classically*, both hidden variables about σ^z and σ^x must be sent with the qubit. (Although a single $|\mathrm{EPR}\rangle$ state is sufficient to explain all situations in the quantum way).

If Alice measures σ_A^z and Bob measures σ_B^x, they will find independently that

$$\sigma_A^z = \begin{cases} +1 & p = 1/2 \\ -1 & p = 1/2 \end{cases}, \ \sigma_B^x = \begin{cases} +1 & p = 1/2 \\ -1 & p = 1/2 \end{cases}. \tag{315}$$

- *Quantum* explanation: can be inferred from $\langle \sigma_A^z \rangle = \langle \sigma_B^x \rangle = 0$ and $\langle \sigma_A^z \, \sigma_B^x \rangle = 0$.
- The *classical* hidden variables can reproduce this behavior only if they follow the joint distribution

Alice	Bob	p
00	00	1/4
01	01	1/4
10	10	1/4
11	11	1/4

(316)

So far so good. But Alice and Bob can also decide to measure σ^y, or more generally, *any* linear combination of their observables ... What if Alice measures $\boldsymbol{n}_A \cdot \boldsymbol{\sigma}_A$ and Bob measures $\boldsymbol{n}_B \cdot \boldsymbol{\sigma}_B$? (where $\boldsymbol{n}_A$ and $\boldsymbol{n}_B$ are *unit* vectors) Their outcomes will follow the joint distribution

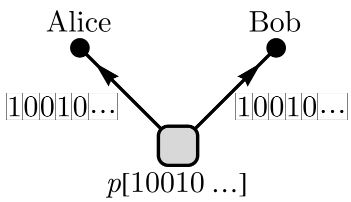
$\boldsymbol{n}_A \cdot \boldsymbol{\sigma}_A$	$\boldsymbol{n}_B \cdot \boldsymbol{\sigma}_B$	p
+1	+1	$(1 + \boldsymbol{n}_A \cdot \boldsymbol{n}_B)/4$
+1	−1	$(1 - \boldsymbol{n}_A \cdot \boldsymbol{n}_B)/4$
−1	+1	$(1 - \boldsymbol{n}_A \cdot \boldsymbol{n}_B)/4$
−1	−1	$(1 + \boldsymbol{n}_A \cdot \boldsymbol{n}_B)/4$

(317)

The probability that Alice and Bob obtain the same outcome is

$$p(\boldsymbol{n}_A \cdot \boldsymbol{\sigma}_A = \boldsymbol{n}_B \cdot \boldsymbol{\sigma}_B) = \frac{1 + \boldsymbol{n}_A \cdot \boldsymbol{n}_B}{2}. \tag{318}$$

- *Quantum* explanation: can be inferred from $\langle \boldsymbol{n}_A \cdot \boldsymbol{\sigma}_A \rangle = \langle \boldsymbol{n}_B \cdot \boldsymbol{\sigma}_B \rangle = 0$ and $\langle \boldsymbol{n}_A \cdot \boldsymbol{\sigma}_A \, \boldsymbol{n}_B \cdot \boldsymbol{\sigma}_B \rangle = \boldsymbol{n}_A \cdot \boldsymbol{n}_B$.
- *Classically*, to reproduce all these, we will need *many* (could be infinitely many) hidden variables. (This is ugly but not fatal yet.)



There should be complicated *correlation* among *hidden variables* in an *attempt* to match quantum predictions (but the attempt may fail). Suppose two of the hidden variables happen to determine the outcome of $\boldsymbol{n}_1 \cdot \boldsymbol{\sigma}$ and $\boldsymbol{n}_2 \cdot \boldsymbol{\sigma}$. After *marginalizing* (summing) over all the other hidden variables, the marginal distribution should be

$$\begin{array}{lll} \text{Alice} & \text{Bob} & p \\ \hline \ldots 00 \ldots & \ldots 00 \ldots & (1 + \boldsymbol{n}_1 \cdot \boldsymbol{n}_2)/4 \\ \ldots 01 \ldots & \ldots 01 \ldots & (1 - \boldsymbol{n}_1 \cdot \boldsymbol{n}_2)/4 \\ \ldots 10 \ldots & \ldots 10 \ldots & (1 - \boldsymbol{n}_1 \cdot \boldsymbol{n}_2)/4 \\ \ldots 11 \ldots & \ldots 11 \ldots & (1 + \boldsymbol{n}_1 \cdot \boldsymbol{n}_2)/4 \end{array}. \tag{319}$$

Now consider Alice and Bob can choose to measure any one of the *three* observables $\boldsymbol{n}_1 \cdot \boldsymbol{\sigma}$, $\boldsymbol{n}_2 \cdot \boldsymbol{\sigma}$ and $\boldsymbol{n}_3 \cdot \boldsymbol{\sigma}$ (on their own qubits respectively, where $\boldsymbol{n}_{1,2,3}$ are *unit* vectors).

- *Classically*, there must be *three* hidden variables associated with the *three* observables, following some marginal distribution

$$\begin{array}{lll} \text{Alice} & \text{Bob} & p \\ \hline \ldots 000 \ldots & \ldots 000 \ldots & p_1 \\ \ldots 001 \ldots & \ldots 001 \ldots & p_2 \\ \ldots 010 \ldots & \ldots 010 \ldots & p_3 \\ \ldots 011 \ldots & \ldots 011 \ldots & p_4 \\ \ldots 100 \ldots & \ldots 100 \ldots & p_5 \\ \ldots 101 \ldots & \ldots 101 \ldots & p_6 \\ \ldots 110 \ldots & \ldots 110 \ldots & p_7 \\ \ldots 111 \ldots & \ldots 111 \ldots & p_8 \end{array}. \tag{320}$$

  The probability must sum up to 1, i.e.

$$p_1 + p_2 + \ldots + p_8 = 1. \tag{321}$$

  - If Alice measures $\boldsymbol{n}_1 \cdot \boldsymbol{\sigma}_A$ and Bob measures $\boldsymbol{n}_2 \cdot \boldsymbol{\sigma}_B$, the probability that they obtain the same outcome is

$$p(\boldsymbol{n}_1 \cdot \boldsymbol{\sigma}_A = \boldsymbol{n}_2 \cdot \boldsymbol{\sigma}_B) = p_1 + p_2 + p_7 + p_8. \tag{322}$$

  - If Alice measures $\boldsymbol{n}_2 \cdot \boldsymbol{\sigma}_A$ and Bob measures $\boldsymbol{n}_3 \cdot \boldsymbol{\sigma}_B$, the probability that they obtain the same outcome is

$$p(\boldsymbol{n}_2 \cdot \boldsymbol{\sigma}_A = \boldsymbol{n}_3 \cdot \boldsymbol{\sigma}_B) = p_1 + p_4 + p_5 + p_8. \tag{323}$$

  - If Alice measures $\boldsymbol{n}_3 \cdot \boldsymbol{\sigma}_A$ and Bob measures $\boldsymbol{n}_1 \cdot \boldsymbol{\sigma}_B$, the probability that they obtain the same outcome is

$$p(\boldsymbol{n}_3 \cdot \boldsymbol{\sigma}_A = \boldsymbol{n}_1 \cdot \boldsymbol{\sigma}_B) = p_1 + p_3 + p_6 + p_8. \tag{324}$$

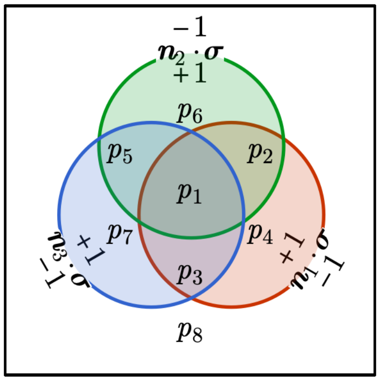
Put together,

$$\begin{aligned}&p(\boldsymbol{n}_1\cdot\boldsymbol{\sigma}_A=\boldsymbol{n}_2\cdot\boldsymbol{\sigma}_B)+p(\boldsymbol{n}_2\cdot\boldsymbol{\sigma}_A=\boldsymbol{n}_3\cdot\boldsymbol{\sigma}_B)+p(\boldsymbol{n}_3\cdot\boldsymbol{\sigma}_A=\boldsymbol{n}_1\cdot\boldsymbol{\sigma}_B)\\&=3\,p_1+p_2+p_3+p_4+p_5+p_6+p_7+3\,p_8\\&=1+2\,p_1+2\,p_8\end{aligned}\tag{325}$$

This leads to a (version of) **Bell inequality**.

$$\boxed{p(\boldsymbol{n}_1\cdot\boldsymbol{\sigma}_A=\boldsymbol{n}_2\cdot\boldsymbol{\sigma}_B)+p(\boldsymbol{n}_2\cdot\boldsymbol{\sigma}_A=\boldsymbol{n}_3\cdot\boldsymbol{\sigma}_B)+p(\boldsymbol{n}_3\cdot\boldsymbol{\sigma}_A=\boldsymbol{n}_1\cdot\boldsymbol{\sigma}_B)\geq 1.}\tag{326}$$

A diagrammatic illustration:

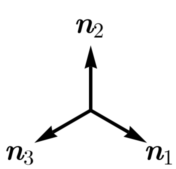


- Now what is the **quantum mechanical prediction**? Recall the *quantum* result in Eq. (318), the Bell inequality would require

$$\frac{1+\boldsymbol{n}_1\cdot\boldsymbol{n}_2}{2}+\frac{1+\boldsymbol{n}_2\cdot\boldsymbol{n}_3}{2}+\frac{1+\boldsymbol{n}_3\cdot\boldsymbol{n}_1}{2}\geq 1,\tag{327}$$

for three unit vectors $\boldsymbol{n}_1$, $\boldsymbol{n}_2$ and $\boldsymbol{n}_3$.

Consider a special case, where the three vectors are 120° to each other in a plane.



$$\boldsymbol{n}_1\cdot\boldsymbol{n}_2=\boldsymbol{n}_2\cdot\boldsymbol{n}_3=\boldsymbol{n}_3\cdot\boldsymbol{n}_1=-1/2.\tag{328}$$

Then Eq. (327) would require

$$\frac{1}{4}+\frac{1}{4}+\frac{1}{4}=\frac{3}{4}\geq 1,\tag{329}$$

which is not true.

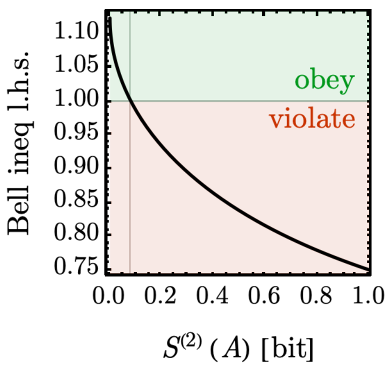
The *violation* of Bell inequality indicates that no classical model of *local hidden variables* can ever reproduce all the predictions of quantum mechanics. This is the **Bell's theorem**.

How does Bell inequality tell us about entanglement?

> Consider a two qubit state $|\psi\rangle=\cos\,\alpha\,|\uparrow\uparrow\rangle+\sin\,\alpha\,|\downarrow\downarrow\rangle$, where α is a phase angle.
> (i) Calculate the 2nd Rényi entanglement entropy $S^{(2)}(A)$ of qubit A (as a function of α).

(ii) Use the observables defined in Eq. (310) and Eq. (311) to evaluate $\langle\psi|\,\boldsymbol{\sigma}_A\,|\psi\rangle$, $\langle\psi|\,\boldsymbol{\sigma}_B\,|\psi\rangle$ and $\langle\psi|\,\boldsymbol{\sigma}_A\otimes\boldsymbol{\sigma}_B\,|\psi\rangle$.
(iii) Let $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{n}_3$ be three unit vectors 120° to each other in the xz plane, evaluate the left-hand-side of the Bell inequality
$p(\boldsymbol{n}_1\cdot\boldsymbol{\sigma}_A=\boldsymbol{n}_2\cdot\boldsymbol{\sigma}_B)+p(\boldsymbol{n}_2\cdot\boldsymbol{\sigma}_A=\boldsymbol{n}_3\cdot\boldsymbol{\sigma}_B)+p(\boldsymbol{n}_3\cdot\boldsymbol{\sigma}_A=\boldsymbol{n}_1\cdot\boldsymbol{\sigma}_B)$
as a function of α.

We can plot the l.h.s. of the Bell inequality v.s. the 2nd Rényi entanglement entropy for different α:



- For *pure* state, such as $|\psi\rangle$ in the above example, entanglement entropy $S^{(2)}(A)>0\Leftrightarrow$ the state is *entangled*. But the Bell inequality is *not always* violated. ⇒ It is an **entanglement witness**.
- For *mixed* state, entropy no longer provides a good measure of quantum entanglement. We had to rely on Bell inequalities and other entanglement witness.

Solution (HW 15)

Quantum Many-Body Systems*

Combining Systems

Axiom 5 (Composition): The *Hilbert space* of a *combined* quantum system is the **direct product** of the *Hilbert space* of each *subsystem*.

Suppose systems A and B are associated with the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively,

$$\mathcal{H}_A=\operatorname{span}\{|i\rangle_A\},\ \mathcal{H}_B=\operatorname{span}\{|j\rangle_B\},\tag{336}$$

the composite system $A\bigcup B$ will be associated with the Hilbert space

$$\mathcal{H}_{A\bigcup B}=\mathcal{H}_A\otimes\mathcal{H}_B=\operatorname{span}\{|i\rangle_A\otimes|j\rangle_B\}=\operatorname{span}\{|ij\rangle\}.\tag{337}$$

- Hilbert space **tensor product** ⇒ Hilbert space *dimension* **multiplies**

$$\dim\mathcal{H}_{A\bigcup B}=\dim\mathcal{H}_A\dim\mathcal{H}_B.\tag{338}$$

- **Generic states** in $\mathcal{H}_{A\bigcup B}$

$$|\psi\rangle = \sum_{i,j} \psi_{ij}\, |ij\rangle. \tag{339}$$

- **Generic operators** in $\mathcal{H}_{A\cup B}$

$$L = \sum_{i,j,k,l} |ij\rangle\, L_{ij,kl}\, \langle kl|, \tag{340}$$

where the matrix (tensor) element

$$L_{ij,kl} = \langle ij|\, L\, |kl\rangle. \tag{341}$$

- **Tensor product** of **states**. Suppose $|\psi\rangle = \sum_i \psi_i\, |i\rangle_A$, $|\phi\rangle = \sum_j \phi_j\, |j\rangle_B$

$$|\psi\rangle \otimes |\phi\rangle = \sum_{i,j} \psi_i\, \phi_j\, |i\rangle_A \otimes |j\rangle_B = \sum_{i,j} \psi_i\, \phi_j\, |ij\rangle. \tag{342}$$

  - Note: the **double index** ij labels a **single state** $|ij\rangle$.
  - Rule of **inner product**.

$$\langle ij \mid kl\rangle = \langle j|_B \otimes \langle i|_A\, |k\rangle_A \otimes |l\rangle_B = \langle i \mid k\rangle_A\, \langle j \mid l\rangle_B = \delta_{ik}\, \delta_{jl}. \tag{343}$$

- **Tensor product** of **operators**. Suppose $A = \sum_{i,j} |i\rangle_A\, A_{ij}\, \langle j|_A$, $B = \sum_{k,l} |k\rangle_B\, B_{kl}\, \langle l|_B$,

$$\begin{aligned} A \otimes B &= \sum_{i,j,k,l} A_{ij}\, B_{kl}\, |i\rangle_A\, |k\rangle_B\, \langle j|_A\, \langle l|_B \\ &= \sum_{i,j,k,l} A_{ij}\, B_{kl}\, |ik\rangle\, \langle jl|. \end{aligned} \tag{344}$$

▪ Tensor Network and Quantum Circuit

Complicated quantum systems can be built out of qubits.

- **Many-body Hilbert space** $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3 \otimes \ldots$
- **States** in $\mathcal{H}$

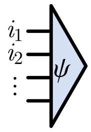
$$|\psi\rangle = \sum_{i_1\, i_2\, \ldots} \psi_{i_1\, i_2\, \ldots}\, |i_1\, i_2\, \ldots\rangle = \sum_{[i]} \psi_{[i]}\, |[i]\rangle. \tag{345}$$

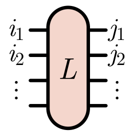
Notation: bundled index $[i] = i_1\, i_2\, \ldots$.

- **Operators** in $\mathcal{H}$

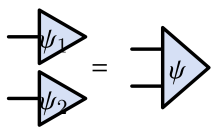
$$\begin{aligned} L &= \sum_{i_1\, i_2\, \ldots} \sum_{j_1\, j_2\, \ldots} |i_1\, i_2\, \ldots\rangle\, L_{i_1\, i_2\, \ldots, j_1\, j_2\, \ldots}\, \langle j_1\, j_2\, \ldots| \\ &= \sum_{[i],[j]} |[i]\rangle\, L_{[i][j]}\, \langle [j]|. \end{aligned} \tag{346}$$

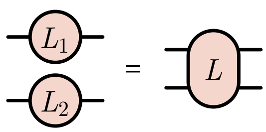
States and operators are both represented as **tensors** in general. Note: the *tensor* here is just a *multi-dimensional array*, without the requirement of covariance as in general relativity.

$\psi_{[i]} = \psi_{i_1\, i_2\, \ldots} =$ 

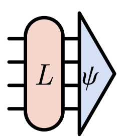
$L_{[i][j]} =$ 

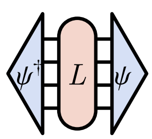
- **Tensor product**: simply put the tensors together.

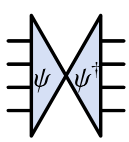
$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle =$ 

$L = L_1 \otimes L_2 =$ 

- **Tensor Contraction**: indices on *internal legs* are automatically *summed* over.

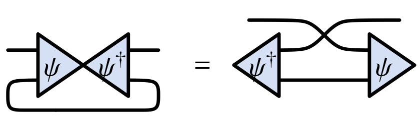
$L\,|\psi\rangle =$ 

$\langle\psi|\, L\, |\psi\rangle =$ 

$\rho = |\psi\rangle\,\langle\psi| =$ 

- **(Partial) Trace**: connect the pair of legs to be traced.

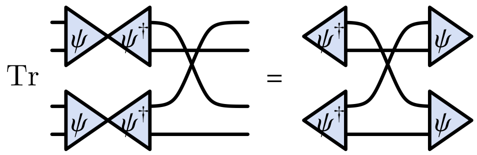
$$\rho_A = \mathrm{Tr}_{\bar{A}}\, |\psi\rangle\, \langle\psi|. \tag{347}$$



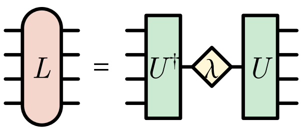
Let us try to express the 2nd Rényi entropy

$$\begin{aligned} e^{-S^{(2)}(A)} &= \mathrm{Tr}_A\, \rho_A^2 \\ &= \mathrm{Tr}(|\psi\rangle\,\langle\psi|)^{\otimes 2}\left(X_A \bigotimes \mathbb{1}_{\bar{A}}\right) \\ &= \langle\psi|^{\otimes 2}\left(X_A \bigotimes \mathbb{1}_{\bar{A}}\right)|\psi\rangle^{\otimes 2}. \end{aligned} \tag{348}$$

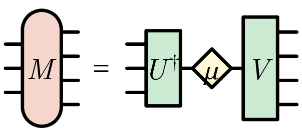
The $\otimes$ and $\otimes$ tensor products have *different* meanings. This ambiguity can be resolved by the tensor network.



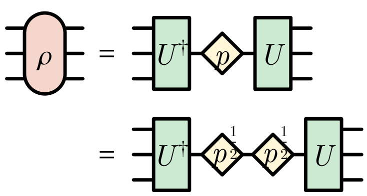
- **Diagonalization** or **Singular Value Decomposition (SVD)**.
    - For Hermitian operator, decompose by matrix diagonalization,



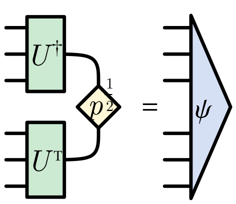
    - For more general tensors, decompose by SVD.



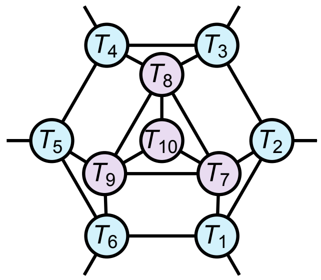
Mix state purification: given a *mixed* state density matrix ρ, find a *pure* state $|\psi\rangle$ (in a larger Hilbert space), such that its *reduced* density matrix reproduces ρ. The procedure is to first diagonalize ρ and split its eigenvalues in square roots $p_i = p_i^{1/2}\, p_i^{1/2}$.



Take *one* square root and bend around the unitary ⇒ the purified state $|\psi\rangle$. It is also called the **thermal field double** state, if p follows the *thermal equilibrium* distribution, i.e. $p_i \propto e^{-\beta E_i}$.



Tensor network: a collection of *tensors* connected by *contractions*.

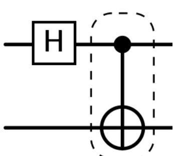


- Efficient representation of big tensors ⇒ numerical method to solve *quantum many-body problems*.
- Conceptual tools to visualize the entanglement structures and symmetry properties ⇒ tensor network *holography*, tensor network formulation of *topological order*.

Quantum circuits are a subclass of *tensor networks*.

- Each wire: a qubit.
- Each block: a unitary operator, also called a **quantum gate**.

Example: a simple quantum circuit that prepares *Bell states*.



It consists of two tensors: a **Hadamard gate** (H) and a **controlled NOT gate** (CNOT, in the dashed region)

$$\mathsf{H} \doteq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

$$\mathsf{CNOT} = e^{\frac{i\pi}{4}(\mathbb{1}-\sigma_1^z)(\mathbb{1}-\sigma_2^x)} \doteq \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right). \tag{349}$$

Convention: for quantum circuit, the state *enters* from *left*, and *exits* form *right*.

HW 16 What are the resulting states when the above quantum circuit acts on (i) the state $|\uparrow\uparrow\rangle$ and (ii) the state $|\downarrow\downarrow\rangle$?

Solution (HW 16)

■ Quantum Decoherence

Consider a qubit coupled to a bath.

- System A: a **qubit** → two-dimensional Hilbert space

$$\mathcal{H}_A = \text{span}\{|\uparrow\rangle, |\downarrow\rangle\}, \tag{352}$$

- System B: a **bath** → d-dimensional Hilbert space (d is supposed to be large)

$$\mathcal{H}_B = \text{span}\{|i\rangle\}_{i=1,\ldots,d} \tag{353}$$

The Hilbert space of the combined system

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B = \text{span}\{|\uparrow\rangle \otimes |i\rangle, |\downarrow\rangle \otimes |i\rangle\}_{i=1,\ldots,d}. \tag{354}$$

Suppose the interaction between the *qubit* and the *bath* is described by the **Hamiltonian**

$$H = \sigma^z \otimes M, \tag{355}$$

where M is a *Hermitian* operator acting on $\mathcal{H}_B$ (or represented as a $d \times d$ Hermitian matrix).

- Initial state: a **product state** of qubit ρ_A and bath ρ_B

$$\rho(0) = \rho_A(0) \otimes \rho_B(0). \tag{356}$$

Evolve the system with H by time t,

$$\rho(t) = U(t)\,\rho(0)\,U(t)^\dagger, \tag{357}$$

where $U(t) = e^{-i\,H\,t} = e^{-i\,\sigma^z \otimes M\,t}$.

- Goal: *trace out* the bath and focus on the **reduced density matrix** of the qubit

$$\rho_A(t) = \text{Tr}_B\,\rho(t). \tag{358}$$

In general, recall Eq. (213), $\rho_A(t)$ takes the form

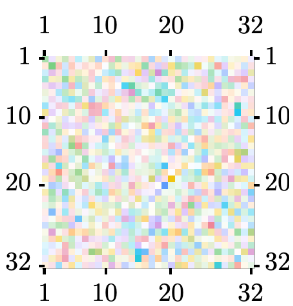
$$\rho_A(t) = \frac{1}{2}\,(\mathbb{1} + \langle \boldsymbol{\sigma}\,(t)\rangle \cdot \boldsymbol{\sigma}). \tag{359}$$

Alternatively, we just need to determine $\langle \boldsymbol{\sigma}\,(t)\rangle$, which is directly related to physical observables.

▫ Numerics

Start by setting up a $d \times d$ *random* Hermitian matrix M

```
d = 32;
M = (# + ConjugateTranspose[#]) / 2 &[
   RandomVariate[NormalDistribution[0, 1 / Sqrt[d]], {d, d, 2}] . {1, I}];
ComplexMatrixPlot@M
```



Construct the Hamiltonian and then define the unitary operator

```
H = KroneckerProduct[PauliMatrix[3], M];
U[t_] := MatrixExp[-ⅈ H t];
```

Prepare a initial state

$$\rho(0) = |\psi(0)\rangle\langle\psi(0)|,$$
$$|\psi(0)\rangle = \frac{3|\uparrow\rangle + 4|\downarrow\rangle}{5} \otimes |1\rangle. \tag{360}$$

```
ψ0 = Flatten[({3, 4} / 5)⊗SparseArray[{1 → 1}, d]];
```

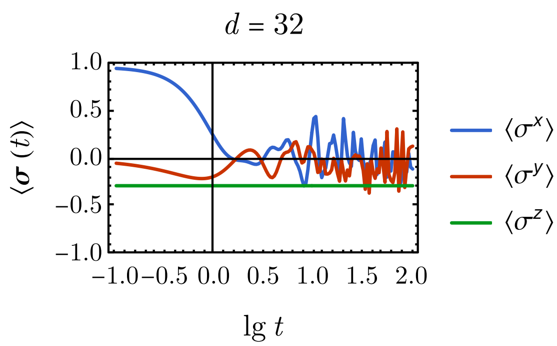
Evolve the state and measure spin expectation values of the qubit

```
ψ[t_] := U[t].ψ0;
spin[t_] :=
  Re@Table[Conjugate[#].KroneckerProduct[PauliMatrix[a], IdentityMatrix[d]].#,
      {a, 3}] &@ψ[t];
spin[0]
spin[10]
```

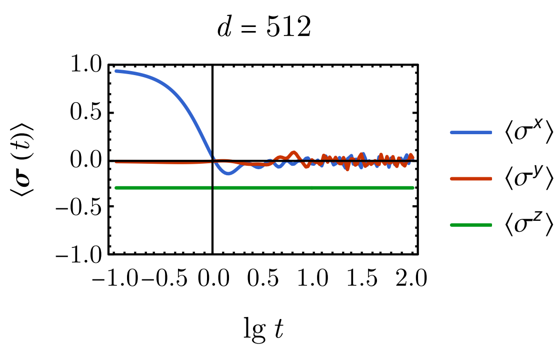
```
{0.96, 0., -0.28}
```

```
{0.425254, 0.0619321, -0.28}
```

Plot the spin expectation values (in log time scale)



With a larger bath (5 qubits → 9 qubits), the fluctuation is quickly suppressed.



Observations:

- As time evolves, $\langle\sigma^x\rangle$, $\langle\sigma^y\rangle$ decays to zero (+ fluctuations) in $O(1)$ time.
- $\langle\sigma^z\rangle$ is conserved (since $[\sigma^z, H] = 0$).

The consequence is that the *off-diagonal* elements of ρ_A decays with time ⇒ **decoherence** of the qubit under the **interaction** with a bath. Note: the *diagonal basis* is set by how the qubit is coupled to the bath (if $H = \sigma^x \otimes M$ then $\langle\sigma^y\rangle$ and $\langle\sigma^z\rangle$ will decay, and the eigenbasis of σ^x is the diagonal basis).

- In general, if a qubit couples to a bath via a spin operator $\boldsymbol{n}\cdot\boldsymbol{\sigma}$ as

  $$H = \boldsymbol{n}\cdot\boldsymbol{\sigma}\otimes M, \tag{361}$$

  under *unitary* time evolution $e^{-i H t}$ of the combined system, its spin expectation value $\langle\boldsymbol{\sigma}\rangle$ will decay to $(\langle\boldsymbol{\sigma}\rangle\cdot\boldsymbol{n})\,\boldsymbol{n}$, i.e.

  $$\langle\boldsymbol{\sigma}\,(t)\rangle \xrightarrow{t\gg 1} (\langle\boldsymbol{\sigma}\,(0)\rangle\cdot\boldsymbol{n})\,\boldsymbol{n}, \tag{362}$$

  which effectively *projects* the initial spin expectation value $\langle\boldsymbol{\sigma}\,(0)\rangle$ to the direction $\boldsymbol{n}$.
- More generally, if there are more than one coupling channels in the Hamiltonian

  $$H = \boldsymbol{n}_1\cdot\boldsymbol{\sigma}\otimes M_1 + \boldsymbol{n}_2\cdot\boldsymbol{\sigma}\otimes M_2, \tag{363}$$

  where M_1 and M_2 are *independently* random, then the spin expectation value $\langle\boldsymbol{\sigma}\rangle$ will decay to *zero* as long as $|\boldsymbol{n}_1\cdot\boldsymbol{n}_2| < 1$. An intuitive understanding: $\langle\boldsymbol{\sigma}\rangle$ is projected towards $\boldsymbol{n}_1$ or $\boldsymbol{n}_2$ repeatedly, ⇒ Its magnitude $|\langle\boldsymbol{\sigma}\rangle|$ always decays. ⇒ Eventually, it will decay to zero, i.e.

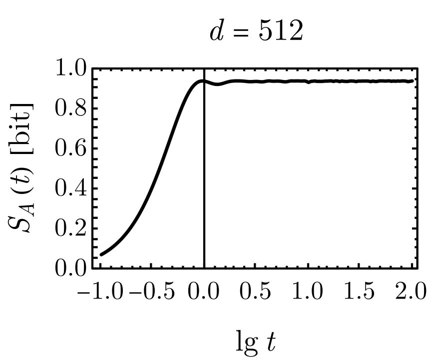
  $$\langle\boldsymbol{\sigma}\,(t)\rangle \xrightarrow{t\gg 1} \mathbf{0}. \tag{364}$$

  In this case, the qubit decohere to the *maximally mixed state*.

  The decoherence of the qubit is also reflected in the **growth** of its **entanglement entropy**.

  $$S_A(t) = -\operatorname{Tr}\rho_A(t)\ln\rho_A(t). \tag{365}$$

- ρ_A evolves from a *pure* state ($S_A = 0$) to a *mixed* state ($S_A > 0$).
- The qubit-bath coupling **entangles** the qubit with the bath under unitary time evolution. ⇒ *Quantum information* of the qubit (partially) spread into the bath via the **quantum entanglement**. For the qubit itself, as if the information is lost ⇒ entropy must grow.



▫ **Theory**

■ Quantum Darwinism

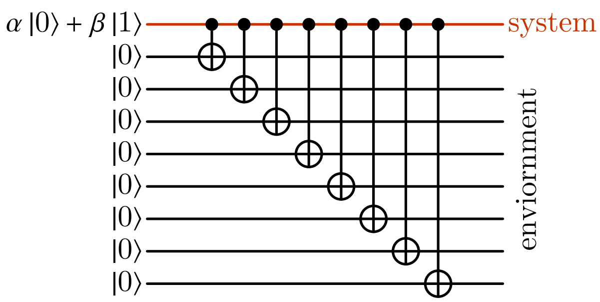
Although quantum decoherence explains how a set of measurement basis can emerge by interacting with the environment, it has not explained how the quantum system collapses to a definite classical reality.

Question: where is the quantum-classical boundary?

- There is *no sharp boundary* between quantum and classical physics.
- Classical reality is an *emergent phenomenon* in quantum many-body systems.

Quantum Darwinism: the *classical* observables are *selected* from *quantum* observables in a process loosely analogous to natural selection in evolution, i.e. classical observables are the *fittest* observables in the sense that *many* independent observers can *agree* on its outcome via *local* measurements.

A simplified model in terms of a quantum circuit



A single-qubit system couples to an environment of N qubits. The unitary time evolution is model by a sequence of CNOT gates that flip the environment qubits controlled by the system qubit:

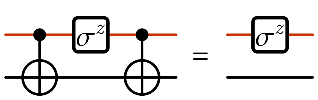
$$U = \prod_{i=1}^{N} e^{\frac{i\pi}{4}(\mathbb{1}-\sigma_0^z)(\mathbb{1}-\sigma_i^x)}. \tag{383}$$

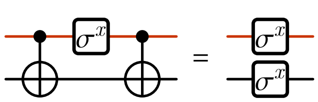
- **State evolution**. $|\psi_{\text{ini}}\rangle \to |\psi_{\text{fin}}\rangle$

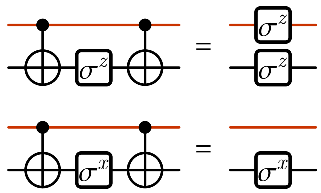
$$\begin{aligned} |\psi_{\text{ini}}\rangle &= \alpha\,|0000\,\ldots\rangle + \beta\,|1000\,\ldots\rangle, \\ |\psi_{\text{fin}}\rangle &= U\,|\psi_{\text{ini}}\rangle = \alpha\,|0000\,\ldots\rangle + \beta\,|1111\,\ldots\rangle. \end{aligned} \tag{384}$$

Quantum entanglement established between system and environment.

- **Operator evolution**. Basic rules of operator evolution under the CNOT gate:







- Observables of the *system* evolve *forward* in time: σ_0^z operator remains unchanged, σ_0^x operator spreads to the environment and grows into a non-local operator (becomes locally intractable).

  $$\begin{aligned} &\sigma_0^z \to U\,\sigma_0^z\,U^\dagger = \sigma_0^z \\ &\sigma_0^x \to U\,\sigma_0^x\,U^\dagger = \sigma_0^x\,\sigma_1^x\,\sigma_2^x\,\sigma_3^x \ldots \end{aligned} \tag{385}$$

  The reduced density matrix *decohere* to the σ_0^z basis.

  $$\begin{aligned} &\rho_{0,\text{ini}} \doteq \begin{pmatrix} |\alpha|^2 & \beta^*\,\alpha \\ \alpha^*\,\beta & |\beta|^2 \end{pmatrix}, \\ &\rho_{\text{ini}} = \rho_{0,\text{ini}} \otimes (|0\rangle\,\langle 0|)^{\otimes N} = |\psi_{\text{ini}}\rangle\,\langle\psi_{\text{ini}}|, \\ &\rho_{\text{fin}} = U\,\rho_{\text{ini}}\,U^\dagger = |\psi_{\text{fin}}\rangle\,\langle\psi_{\text{fin}}|, \\ &\rho_{0,\text{fin}} = \underset{1,2,\ldots,N}{\text{Tr}}\,\rho_{\text{fin}} \doteq \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}. \end{aligned} \tag{386}$$

- (Local) observables of the environment evolve *backward* in time: every σ_i^z observable encodes the information about σ_0^z.

  $$\sigma_i^z \to U^\dagger\,\sigma_i^z\,U = \sigma_0^z\,\sigma_i^z, \tag{387}$$

  Given the initial state for every environment qubit is $|0\rangle$, measuring σ_i^z in the final state is effectively measuring σ_0^z in the initial state,

  $$\begin{aligned} &\langle\psi_{\text{fin}}|\,\sigma_i^z\,|\psi_{\text{fin}}\rangle \\ &= \langle\psi_{\text{ini}}|\,U^\dagger\,\sigma_i^z\,U\,|\psi_{\text{ini}}\rangle \\ &= \langle\psi_{\text{ini}}|\,\sigma_0^z\,\sigma_i^z\,|\psi_{\text{ini}}\rangle \\ &= \langle\psi_{\text{ini}}|\,\sigma_0^z\,|\psi_{\text{ini}}\rangle. \end{aligned} \tag{388}$$

  The information of $\langle\psi_{\text{ini}}|\,\sigma_0^z\,|\psi_{\text{ini}}\rangle$ is copied many times and stores separately in every environment qubit $\Rightarrow$ σ_0^z survives the selection and becomes the classical reality.

  On the contrary, no *local* observables in the environment can tell $\langle\psi_{\text{ini}}|\,\sigma_0^x\,|\psi_{\text{ini}}\rangle$ or $\langle\psi_{\text{ini}}|\,\sigma_0^y\,|\psi_{\text{ini}}\rangle$ $\Rightarrow$ σ_0^x and σ_0^y remains quantum.

The *fittest observable* gets imprinted in the environment many times $\Rightarrow$ such that different observers can *agree* on the observation by independent measuring a small fraction of the environment -- a hallmark of classical behavior $\Rightarrow$ **classical observable** *emerges* as the fittest observable under “natural selection” in quantum Darwinism.

Two essential steps of **quantum state collapse**:

- a set of *measurement basis* emerges under *decoherence*,

- copies of *classical* information *proliferate* in the environment.

They are closely related: the very same process that is responsible for decoherence should inscribe multiple copies of the classical information in the environment.

A testable prediction: **information overload** effect - Any *small fraction* of the environment is enough to provide the *maximal classical information* about the observed system, such that the information one can gather about the system *quickly saturates*. This can be quantified by the mutual information $I(\text{system} : \text{partial environment})$

$$I(\{0\} : \{1, 2, \ldots, n\}) = S(\{0\}) + S(\{1, 2, \ldots, n\}) - S(\{0, 1, 2, \ldots, n\}), \tag{389}$$

evaluated in the final state (assuming $\alpha = \beta = \frac{1}{\sqrt{2}}$)

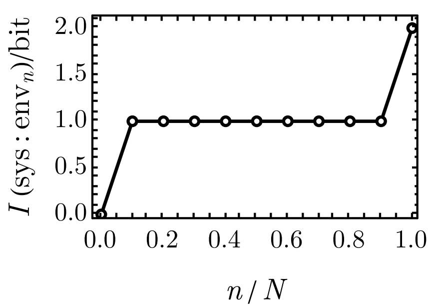
$$|\psi_{\text{fin}}\rangle = \frac{|0000 \ldots\rangle + |1111 \ldots\rangle}{\sqrt{2}}. \tag{390}$$

The entanglement entropy follows

$$\begin{aligned} &S(\text{empty set}) = 0, \\ &S(\text{full set}) = 0, \\ &S(\text{any subset}) = \log 2, \end{aligned} \tag{391}$$

therefore

$$I(\{0\} : \{1, 2, \ldots, n\}) = \begin{cases} 0 & n = 0 \\ \log 2 & 0 < n < N \\ 2 \log 2 & n = N \end{cases}. \tag{392}$$



Quantum Error Correction

Quantum decoherence posts a serious threat to *quantum information processing*.

- Qubits *couple* to the environment and *decohere* inevitably.
- In the extreme case, a qubit can become **maximally mixed** ⇒ An **erasure error**: the quantum information of the qubit is fully *scrambled* with the *environment*, *as if* the information has been *erased*.

Quantum error correction: protecting the quantum information from errors by *spreading* the information into a *highly entangled* quantum many-body state (which we have access to).

$$\textit{one}\text{ logical qubit} \underset{\text{decoded to}}{\overset{\text{encoded in}}{\rightleftharpoons}} \textit{many}\text{ physical qubits.} \tag{393}$$

- **Logical qubit**: the information theoretic qubit (software level), whose basis states are denoted as $|\uparrow\_\rangle$, $|\downarrow\_\rangle$ (with a *underline*).
- **Physical qubit**: the actual qubit implemented on quantum devices (hardware level).

Even if some physical qubits are corrupted or erased, one can still retrieve the logical qubit from the rest of the physical qubits.

Five-qubit code: a quantum error correction code that encodes **one** *logical qubit* into **five** *physical qubits*, where the logical qubit is protected against the *erasure* of *any* **two** physical qubits.

- The *logical qubit* states $|\uparrow\_\rangle$, $|\downarrow\_\rangle$ span a **code subspace** in the physical qubit Hilbert space.
- The *code subspace* is specified by four **commuting Pauli operators** on the physical qubits:

$$\begin{aligned} M_1 &\doteq \sigma^{13\,310}, \\ M_2 &\doteq \sigma^{01\,331}, \\ M_3 &\doteq \sigma^{10\,133}, \\ M_4 &\doteq \sigma^{31\,013}. \end{aligned} \tag{394}$$

- These operators are called **stabilizers**, as they stabilize the *logical qubit* as their *common eigenstates* of eigenvalue $+1$, i.e.

$$\forall\, i: M_i\,|\underline{\uparrow}\rangle = |\underline{\uparrow}\rangle,\;\; M_i\,|\underline{\downarrow}\rangle = |\underline{\downarrow}\rangle. \tag{395}$$

Recall that we can simultaneously diagonalize *commuting* operators by constructing a **many-body Hamiltonian**, e.g.

$$\begin{aligned} H &= -M_1 - M_2 - M_3 - M_4 \\ &\doteq -\sigma^{13\,310} - \sigma^{01\,331} - \sigma^{10\,133} - \sigma^{31\,013}. \end{aligned} \tag{396}$$

  - The **code subspace** = the **common eigenspace** of *stabilizers* that $\forall\, i: M_i = +1$ = the **ground state subspace** of the *Hamiltonian* H.
  - The code subspace is *two-dimensional* ⇒ can encode a **logical qubit**. How do we know? 5 qubits, 4 stabilizers: each stabilizer **halves** the Hilbert space ⇒ the remaining space dimension:

$$\frac{2^5}{2^4} = 2. \tag{397}$$

  - Within the code subspace, a choice of the basis is (can be obtained by diagonalize H)

$$\begin{aligned} |\underline{\uparrow}\rangle = \frac{1}{4}\,\big(&-|\downarrow\downarrow\downarrow\downarrow\uparrow\rangle - |\downarrow\downarrow\downarrow\uparrow\downarrow\rangle - |\downarrow\downarrow\uparrow\downarrow\downarrow\rangle - |\downarrow\downarrow\uparrow\uparrow\uparrow\rangle - \\ &|\downarrow\uparrow\downarrow\downarrow\downarrow\rangle + |\downarrow\uparrow\downarrow\uparrow\uparrow\rangle + |\downarrow\uparrow\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\uparrow\downarrow\rangle - |\uparrow\downarrow\downarrow\downarrow\downarrow\rangle - |\uparrow\downarrow\downarrow\uparrow\uparrow\rangle + \\ &|\uparrow\downarrow\uparrow\downarrow\uparrow\rangle + |\uparrow\downarrow\uparrow\uparrow\downarrow\rangle - |\uparrow\uparrow\downarrow\downarrow\uparrow\rangle + |\uparrow\uparrow\downarrow\uparrow\downarrow\rangle - |\uparrow\uparrow\uparrow\downarrow\downarrow\rangle + |\uparrow\uparrow\uparrow\uparrow\uparrow\rangle\big), \end{aligned} \tag{398}$$

$$
\begin{aligned}
|\underline{\downarrow}\rangle = \frac{1}{4}\, \big(& |{\downarrow\downarrow\downarrow\downarrow\downarrow}\rangle - |{\downarrow\downarrow\downarrow\uparrow\uparrow}\rangle + |{\downarrow\downarrow\uparrow\downarrow\uparrow}\rangle - |{\downarrow\downarrow\uparrow\uparrow\downarrow}\rangle + \\
& |{\downarrow\uparrow\downarrow\downarrow\uparrow}\rangle + |{\downarrow\uparrow\downarrow\uparrow\downarrow}\rangle - |{\downarrow\uparrow\uparrow\downarrow\downarrow}\rangle - |{\downarrow\uparrow\uparrow\uparrow\uparrow}\rangle - |{\uparrow\downarrow\downarrow\downarrow\uparrow}\rangle + |{\uparrow\downarrow\downarrow\uparrow\downarrow}\rangle + \\
& |{\uparrow\downarrow\uparrow\downarrow\downarrow}\rangle - |{\uparrow\downarrow\uparrow\uparrow\uparrow}\rangle - |{\uparrow\uparrow\downarrow\downarrow\downarrow}\rangle - |{\uparrow\uparrow\downarrow\uparrow\uparrow}\rangle - |{\uparrow\uparrow\uparrow\downarrow\uparrow}\rangle - |{\uparrow\uparrow\uparrow\uparrow\downarrow}\rangle \big)
\end{aligned}
$$

- **Logical gates**: quantum gates that effectively operate on the logical qubit

$$
\begin{aligned}
&\underline{Z}\,|\underline{\uparrow}\rangle = |\underline{\uparrow}\rangle,\ \underline{Z}\,|\underline{\downarrow}\rangle = -|\underline{\downarrow}\rangle,\\
&\underline{X}\,|\underline{\uparrow}\rangle = |\underline{\downarrow}\rangle,\ \underline{X}\,|\underline{\downarrow}\rangle = |\underline{\uparrow}\rangle.
\end{aligned}
\tag{399}
$$

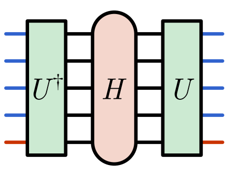
  - $\underline{Z}$ and $\underline{X}$ must *commute* with all stabilizers (to remain in the code subspace), yet not any product of stabilizers (to be nontrivial). One canonical choice is

$$
\underline{Z} \doteq \sigma^{33\,333},\ \underline{X} \doteq \sigma^{11\,111},\ \underline{Y} = i\, \underline{X}\, \underline{Z} = \sigma^{22\,222}. \tag{400}
$$

  - It is hard to decohere the logical qubit, because $\underline{X}$, $\underline{Y}$, $\underline{Z}$ are non-local. ⇒ Their couplings to the environment are typically weak.

A **diagrammatic** understanding: the *unitary matrix* U that diagonalize the Hamiltonian H can be viewed as a *quantum circuit*,

$$
U^\dagger\, H\, U = E. \tag{401}
$$



The *quantum circuit* U should also simultaneously diagonalize all the stabilizers. With a proper basis choice, one can find U

$$
U \doteq i\, e^{-\frac{i\pi}{4}\sigma^{23\,310}}\, e^{-\frac{i\pi}{4}\sigma^{02\,331}}\, e^{\frac{i\pi}{4}\sigma^{33\,123}}\, e^{\frac{i\pi}{4}\sigma^{33\,312}}\, e^{\frac{i\pi}{4}\sigma^{33\,331}}\, e^{\frac{i\pi}{4}\sigma^{30\,302}}\, e^{-\frac{i\pi}{4}\sigma^{00\,001}}\, e^{-\frac{i\pi}{4}\sigma^{00\,003}}, \tag{402}
$$

such that

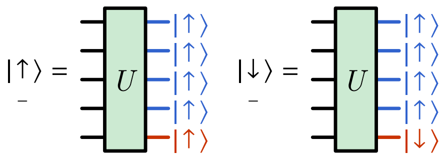
$$
\begin{aligned}
&U^\dagger\, M_1\, U \doteq \sigma^{30\,000},\\
&U^\dagger\, M_2\, U \doteq \sigma^{03\,000},\\
&U^\dagger\, M_3\, U \doteq \sigma^{00\,300},\\
&U^\dagger\, M_4\, U \doteq \sigma^{00\,030}.
\end{aligned}
\tag{403}
$$

As a result, the Hamiltonian transforms to

$$
U^\dagger\, H\, U \doteq -\sigma^{30\,000} - \sigma^{03\,000} - \sigma^{00\,300} - \sigma^{00\,030}. \tag{404}
$$

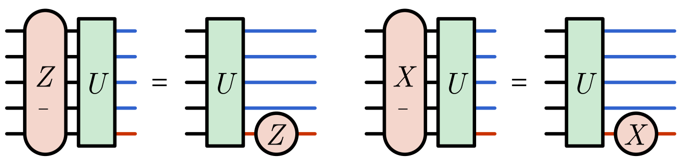
- The first four qubits are *pinned* by the Hamiltonian to $|{\uparrow\uparrow\uparrow\uparrow}\rangle$ to lower the energy ⇒ **syndrome qubits**.
- The last qubit is free ⇒ **logical qubit**.

The quantum circuit *encodes* the *logical* qubit into five *physical* qubits, given the *syndrome* qubits pinned to $|\uparrow\uparrow\uparrow\uparrow\rangle$. This is how Eq. (398) was obtained.



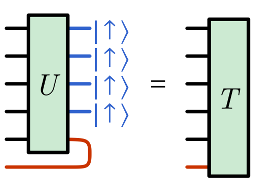
In addition, the *logical gates* do acts on the *logical qubit* as expected,

$$
\begin{aligned}
&U^\dagger\, \underline{Z}\, U \doteq \sigma^{00\,003},\\
&U^\dagger\, \underline{X}\, U \doteq \sigma^{00\,001}.
\end{aligned}
\tag{405}
$$

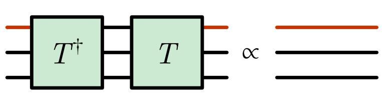


- Logical gates will not take the system out of the *code subspace*, as they will not touch the syndrome qubit.
- If any of the syndrome qubit is *flipped*. ⇒ The system is carried out of the *code subspace* (**excitation** created). ⇒ Signals an error. ⇒ Correct the error by applying appropriate unitary operations based on the syndrome.

How well is the logical qubit protected? Take the *unitary circuit*, pin the *syndrome qubits* and bend around the *logical qubit* → a six-leg tensor T describing how the *logical* and the *physical* qubits are related



It is a **perfect tensor**, because of an amazing property: T is proportional to a **unitary** matrix from *any half* of legs to the rest *half* of legs.



Treat T as a many-body state (after normalization) ⇒ it describes a *pure state* of *six* qubits, where *any* set of three qubits is *maximally entangled* with the complementary set of three qubits. Such states have been called **absolutely maximally entangled** states.

HW 17

(i) Use the perfect tensor property to show that the nth Rényi entanglement entropy of any m qubits in the six-qubit state $|T\rangle$ is $S^{(n)}(m) = \min(m, 6-m) \ln 2$.
(ii) Use the above result to show Eq. (406).

The **mutual information** between the *logical qubit* and any *m physical qubits* is given by

$$I^{(n)}(1:m) = \begin{cases} 0 & m \leq 2, \\ 2\ln 2 & m \geq 3. \end{cases} \tag{406}$$

The *five-qubit code* has the property that

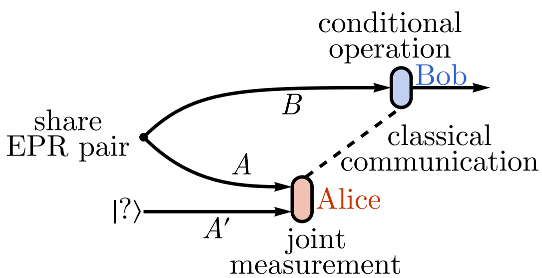
- any *two* qubits have *no* information about the logical qubit.
- any *three* qubits have *complete* information about the logical qubit.

Therefore the logical qubit is protected against *erasure error* up to *two* qubits.

Solution (HW 17)

Quantum Teleportation

Quantum teleportation transfers (unknown) quantum states from one location to another using *quantum entanglement* resource and *classical communication.*



- The model involves three qubits: A', A and B.
  - Establish **quantum entanglement**: A and B qubits are prepared in an EPR state, and shared between Alice and Bob respectively

$$|\text{EPR}\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|\uparrow_A \uparrow_B\rangle + |\downarrow_A \downarrow_B\rangle\right) \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}. \tag{409}$$

  - A' qubit is in an *unknown* quantum state in Alice's possession

$$|\psi\rangle_{A'} = \alpha\,|\uparrow_{A'}\rangle + \beta\,|\downarrow_{A'}\rangle \doteq \begin{pmatrix}\alpha\\ \beta\end{pmatrix}, \tag{410}$$

(Alice knowns that the single-qubit pure state must take this form, but does not need to know what α, β are.)

  - The three-qubit system is in a joint state

$$|\Psi\rangle = |\psi\rangle_{A'} \otimes |\text{EPR}\rangle_{AB} \doteq \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{411}$$

- Goal: to teleport the quantum state $|\psi\rangle$ from A' to B without handing the qubit A' to Bob.

Protocol:

- Alice makes a **joint measurement** of A' and A, by observing $\sigma_{A'}^x \sigma_A^x$ and $\sigma_{A'}^z \sigma_A^z$ (note that they are commuting observables that can be measured simultaneously). There are four possible outcomes

$$\begin{array}{c|cccc} \sigma_{A'}^x \sigma_A^x & +1 & +1 & -1 & -1 \\ \sigma_{A'}^z \sigma_A^z & +1 & -1 & +1 & -1 \\ P_{ab} & P_{++} & P_{+-} & P_{-+} & P_{--} \end{array}, \tag{412}$$

each corresponds to a projection operator (labeled by the measurement outcomes $a := \sigma_{A'}^x \sigma_A^x = \pm 1,\ b := \sigma_{A'}^z \sigma_A^z = \pm 1$)

$$\begin{aligned} P_{ab} &= \frac{\mathbb{1} + a\,\sigma_{A'}^x \sigma_A^x}{2} \frac{\mathbb{1} + b\,\sigma_{A'}^z \sigma_A^z}{2} \\ &\doteq \frac{1}{4} \begin{pmatrix} 1+b & 0 & 0 & a\,(1+b) \\ 0 & 1-b & a(1-b) & 0 \\ 0 & a(1-b) & 1-b & 0 \\ a\,(1+b) & 0 & 0 & 1+b \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{413}$$

- After the measurement, the three-qubit state will collapse to

$$|\Psi\rangle \xrightarrow{\sigma_{A'}^x \sigma_A^x = a,\, \sigma_{A'}^z \sigma_A^z = b} \frac{P_{ab}\,|\Psi\rangle}{\text{normalization} \ldots}, \tag{414}$$

more explicitly as

$$P_{ab}\,|\Psi\rangle \doteq \frac{1}{4\sqrt{2}} \begin{pmatrix} \alpha\,(1+b) \\ \beta\,a\,(1+b) \\ \beta\,a\,(1-b) \\ \alpha\,(1-b) \\ \beta\,(1-b) \\ \alpha\,a\,(1-b) \\ \alpha\,a\,(1+b) \\ \beta\,(1+b) \end{pmatrix}. \tag{415}$$

Let us enumerate all four cases

$$\begin{array}{c|cccc}
\sigma^x_{A'}\,\sigma^x_A & +1 & +1 & -1 & -1 \\
\sigma^z_{A'}\,\sigma^z_A & +1 & -1 & +1 & -1 \\
P_{ab}\,|\Psi\rangle & \begin{pmatrix}\alpha\\ \beta\\ 0\\ 0\\ 0\\ 0\\ \alpha\\ \beta\end{pmatrix} & \begin{pmatrix}0\\ 0\\ \beta\\ \alpha\\ \beta\\ \alpha\\ 0\\ 0\end{pmatrix} & \begin{pmatrix}\alpha\\ -\beta\\ 0\\ 0\\ 0\\ 0\\ -\alpha\\ \beta\end{pmatrix} & \begin{pmatrix}0\\ 0\\ -\beta\\ \alpha\\ \beta\\ -\alpha\\ 0\\ 0\end{pmatrix}
\end{array},$$

It turns out that they can all be written as the tensor product state between $A'\,A$ and B as

$$\begin{array}{ccc}
\sigma^x_{A'}\,\sigma^x_A & \sigma^z_{A'}\,\sigma^z_A & P_{ab}\,|\Psi\rangle \\
\hline
+1 & +1 & \begin{pmatrix}1\\ 0\\ 0\\ 1\end{pmatrix}\otimes\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \\
+1 & -1 & \begin{pmatrix}0\\ 1\\ 1\\ 0\end{pmatrix}\otimes\begin{pmatrix}\beta\\ \alpha\end{pmatrix} \\
-1 & +1 & \begin{pmatrix}1\\ 0\\ 0\\ -1\end{pmatrix}\otimes\begin{pmatrix}\alpha\\ -\beta\end{pmatrix} \\
-1 & -1 & \begin{pmatrix}0\\ 1\\ -1\\ 0\end{pmatrix}\otimes\begin{pmatrix}-\beta\\ \alpha\end{pmatrix}
\end{array} \tag{417}$$

or in terms of ket state notation as

$$\begin{aligned}
P_{++}\,|\Psi\rangle &= \frac{1}{\sqrt{2}}\,(|\uparrow_{A'}\uparrow_A\rangle + |\downarrow_{A'}\downarrow_A\rangle)\otimes(\alpha\,|\uparrow_B\rangle + \beta\,|\downarrow_B\rangle),\\
P_{+-}\,|\Psi\rangle &= \frac{1}{\sqrt{2}}\,(|\uparrow_{A'}\downarrow_A\rangle + |\downarrow_{A'}\uparrow_A\rangle)\otimes(\beta\,|\uparrow_B\rangle + \alpha\,|\downarrow_B\rangle),\\
P_{-+}\,|\Psi\rangle &= \frac{1}{\sqrt{2}}\,(|\uparrow_{A'}\uparrow_A\rangle - |\downarrow_{A'}\downarrow_A\rangle)\otimes(\alpha\,|\uparrow_B\rangle - \beta\,|\downarrow_B\rangle),\\
P_{--}\,|\Psi\rangle &= \frac{1}{\sqrt{2}}\,(|\uparrow_{A'}\downarrow_A\rangle - |\downarrow_{A'}\uparrow_A\rangle)\otimes(-\beta\,|\uparrow_B\rangle + \alpha\,|\downarrow_B\rangle).
\end{aligned} \tag{418}$$

- Alice will tell Bob her measurement outcome

$$(\sigma^x_{A'}\,\sigma^x_A = a,\, \sigma^z_{A'}\,\sigma^z_A = b), \tag{419}$$

via a **classical communication channel** (e.g. by making a phone call).

- If $(a, b) = ++$, the qubit B is in the state

  $$\alpha\,|\uparrow_B\rangle + \beta\,|\downarrow_B\rangle = |\psi\rangle_B. \tag{420}$$

  There is nothing more Bob needs to do. The state $|\psi\rangle$ has been teleported to B successfully.

- If $(a, b) = +-$, the qubit B is in the state

  $$\beta\,|\uparrow_B\rangle + \alpha\,|\downarrow_B\rangle, \tag{421}$$

  Bob will apply a σ_B^x operator to the qubit B

  $$\sigma_B^x(\beta\,|\uparrow_B\rangle + \alpha\,|\downarrow_B\rangle) = \alpha\,|\uparrow_B\rangle + \beta\,|\downarrow_B\rangle = |\psi\rangle_B, \tag{422}$$

  then the qubit B is converted to the state $|\psi\rangle$.

- If $(a, b) = -+$, the qubit B is in the state

  $$\alpha\,|\uparrow_B\rangle - \beta\,|\downarrow_B\rangle, \tag{423}$$

  Bob will apply a σ_B^z operator to the qubit B

  $$\sigma_B^z(\alpha\,|\uparrow_B\rangle - \beta\,|\downarrow_B\rangle) = \alpha\,|\uparrow_B\rangle + \beta\,|\downarrow_B\rangle = |\psi\rangle_B, \tag{424}$$

  then the qubit B is converted to the state $|\psi\rangle$.

- If $(a, b) = +-$, the qubit B is in the state

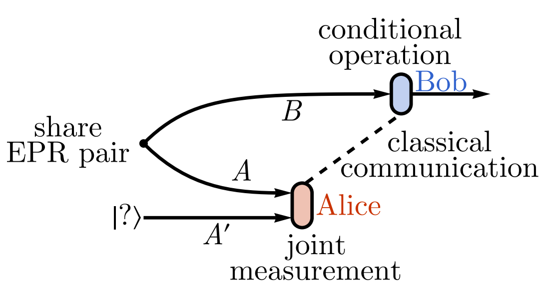
  $$-\beta\,|\uparrow_B\rangle + \alpha\,|\downarrow_B\rangle, \tag{425}$$

  Bob will apply the composite operator $\sigma_B^z\,\sigma_B^x$ to the qubit B

  $$\sigma_B^z\,\sigma_B^x(-\beta\,|\uparrow_B\rangle + \alpha\,|\downarrow_B\rangle) = \alpha\,|\uparrow_B\rangle + \beta\,|\downarrow_B\rangle = |\psi\rangle_B, \tag{426}$$

  then the qubit B is converted to the state $|\psi\rangle$.

Summary:



- Alice and Bob establish shared a *entanglement resource* (e.g. a EPR pair).
- For any unknown state $|\psi\rangle_{A'}$ handed to Alice, she measures $\sigma_{A'}^x\,\sigma_A^x$ and $\sigma_{A'}^z\,\sigma_A^z$.
- Alice tells Bob her the measurement outcomes by *classical communication.*
- Depending on the information, Bob performs conditional operation on his qubit

$\sigma^x_{A'}\sigma^x_A$	$\sigma^z_{A'}\sigma^z_A$	Bob's operation
$+1$	$+1$	$\mathbb{1}_B$
$+1$	-1	σ^x_B
-1	$+1$	σ^z_B
-1	-1	$\sigma^z_B\sigma^x_B$

.

- After the operation, the state $|\psi\rangle$ will appear on Bob's qubit.

Comments:

- The original state $|\psi\rangle$ on qubit A' is destroyed in the process → **No-Cloning Theorem**: it is impossible to create an identical copy of an arbitrary unknown state. You can only teleport an unknown state but not duplicate it.
- **Traversable wormhole** ⇔ interstellar **quantum teleportation**.
  - Entanglement resource: wormhole ⇔ entangled pairs of black holes
  - Classical communication: classical interaction between two black holes

Quantum Search

What is a search problem?

- Given a **bit string** x and a **query function** $Q(x)$ that tells if x is the target.

$$\begin{array}{ccccc} x & & Q(x) & \\ 0000 & \to & 0 & \text{no} \\ 0001 & \to & 0 & \text{no} \\ \vdots & & \vdots & \\ 0110 & \to & 1 & \text{yes!} \\ \vdots & & \vdots & \end{array} \tag{428}$$

- Assuming there is only one target bit string to look for, the goal is to find it!

Complexity of unstructured search (over N entities)

- Classical computer ~ $\mathcal{O}(N)$,
- Quantum computer (Grover's algorithm) ~ $\mathcal{O}\left(N^{1/2}\right)$. The quantum speedup can be significant when N is large.

Key idea: *quantum superposition* allows search to be carried out in *parallel*.

$$\begin{aligned} |s\rangle &= \frac{1}{\sqrt{N}}\left(|0000\rangle + |0001\rangle + \ldots + |0110\rangle + \ldots\right) \\ &= \frac{1}{\sqrt{N}}\sum_x |x\rangle. \end{aligned} \tag{429}$$

- Initially, all bit string states $|x\rangle$ have *equal* probability $1/N$ in a uniform superposition state $|s\rangle$.
- Grover's algorithm iteratively *enhance* the probability of the *target* state by *quantum interference*.

- After about $N^{1/2}$ steps, the target state emerges as the probability ~ 1 state.

The **Grover's algorithm** involves two steps in each iteration. Both steps are implemented as *unitary* operations.

- **Quantum query** (in parallel)

$$U_Q\,|x\rangle = (-1)^{Q(x)}\,|x\rangle. \tag{430}$$

It simply marks the target state with a minus sign. Let $|t\rangle$ be the *target* state, Eq. (430) can also be written as

$$U_Q\,|x\rangle = |x\rangle - 2\,|t\rangle\,\langle t|x\rangle. \tag{431}$$

- **Grover diffusion**

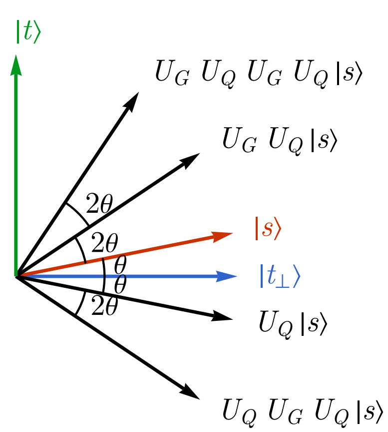
$$U_G\,|x\rangle = 2\,|s\rangle\,\langle s|x\rangle - |x\rangle, \tag{432}$$

which reflect any state about the *source* state $|s\rangle$.

In the two-dimensional Hilbert space spanned by $|t\rangle$, $|s\rangle$, we can define the state $|t_\perp\rangle$ orthogonal to $|t\rangle$

$$|t_\perp\rangle = \frac{1}{\sqrt{N-1}} \sum_{x \neq t} |x\rangle. \tag{433}$$

- In each iteration, the quantum query $U_Q = \mathbb{1} - 2\,|t\rangle\,\langle t|$ reflects any state vector about $|t_\perp\rangle$, and the Grover diffusion $U_G = 2\,|s\rangle\,\langle s| - \mathbb{1}$ reflects any state vector about $|s\rangle$.



- Two reflections make a rotation of $2\,\theta$ angle, where the θ denotes the angle between the two reflection axes (i.e. the angle between $|s\rangle$ and $|t_\perp\rangle$) and is given by

$$\sin\theta = \langle s|t\rangle = \frac{1}{\sqrt{N}}, \tag{434}$$

- After k steps of Grover iterations, the source state $|s\rangle$ is rotated away from $|t_\perp\rangle$ by $(2\,k+1)\,\theta$.
- To rotate the state from $|s\rangle$ to $|t\rangle$ (which is $\pi/2$ from $|t_\perp\rangle$), the total rotation angle must accumulates to $\pi/2$

$$\frac{\pi}{2} = (2\,k+1)\,\theta, \tag{435}$$

such that the solution for k is given by

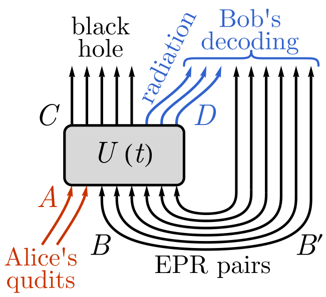
$$k = \frac{1}{2}\left(\frac{\pi}{2\,\theta} - 1\right) = \frac{1}{2}\left(\frac{\pi}{2\arcsin N^{-1/2}} - 1\right). \tag{436}$$

- In the large N limit (when there are many bit strings to search),

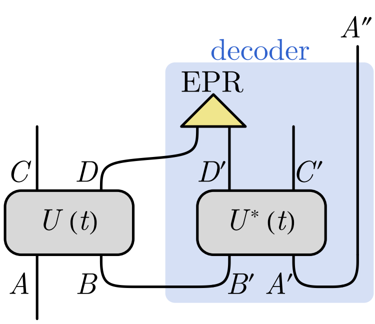
$$k \approx \frac{\pi}{4}\sqrt{N}, \tag{437}$$

which is indeed of order $N^{1/2}$ as mentioned.

Hayden-Preskill problem: can we recover the object that has fallen into a black hole? Yes. Combining quantum search and quantum teleportation (Yoshida, Kitaev, arXiv:1710.03363).



- Resources needed:
  - **Entanglement** resource: the black hole (on Alice's side) must be *entangled* with another black hole (on Bob's side) in the lab (effectively forming a wormhole).
  - **Quantum computation** resource: a strong enough *quantum computer* to simulate the quantum dynamics of black hole and to perform quantum search.
- Basic idea: *Hawking radiation* is an efficient *encoding* of the object that falls into the black hole (black hole is very much like a **quantum hash function**).
  - Collect the Hawking radiation a few moments after the object has fallen into the black hole on Alice's side.
  - Try to decode the Hawking radiation to recover the quantum information of the object ⇔ quantum teleportation of the object through the wormhole.



- The key idea to search for a *collision* of Hawking radiations (hash keys) between the two entangled black holes. This relies on the quantum search algorithm.
- Once the Hawking radiations from both black holes are made the same (same = perfectly entangled), the object will reemerge from Bob's black hole, as a result of quantum teleportation.