# 1. Quantum mechanics essentials

- A particle's position on the real line is given by a wave function $\psi(x,t) \to \mathbb{C}$.
- Probability of finding particle in $(a,b)$ is

$$P(a,b;t) = \int_a^b |\psi(x,t)|^2 \, \mathrm{d}x$$

  Wave function is normalised so that $P(-\infty, +\infty; t) = 1$.
- Time-evolution of wave function given by **Schrodinger equation**:

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) + V(x)\psi(x,t) = \widehat{H}\psi(x,t)$$

  where $\widehat{H} = \widehat{K} + \hat{V}$ is the Hamiltonian operator, $\widehat{K}$ is kinetic energy operator, $\hat{V}$ is potential energy operator.
- Schrodinger equation is **linear**, so any linear combination of solutions is another solution (**principle of superposition**).
- An inner product is defined on the space of solutions to the Schrodinger equation:

$$\langle \psi, \varphi \rangle = \int_{-\infty}^{\infty} \psi^*(x,t)\varphi(x,t) \, \mathrm{d}x$$

- **Hilbert space**: (complex) vector space with Hermitian inner product that is also a complete metric space with metric induced by the inner product:
  - $\langle \psi, a\varphi_1 + b\varphi_2 \rangle = a\langle \psi, \varphi_1 \rangle + b\langle \psi, \varphi_2 \rangle$
  - $\langle \psi, \varphi \rangle = \langle \varphi, \psi \rangle^*$
- **Dirac notation**:
  - Write $|\psi\rangle$ (a **ket**) for vector in Hilbert space $\mathcal{H}$ corresponding to wave function $\psi$.
  - Write $\langle \varphi|$ (a **bra**) for **dual** vector in $\mathcal{H}^*$.
  - **bra-ket**:

$$\langle \varphi | \psi \rangle := \langle \varphi, \psi \rangle = \int_{-\infty}^{\infty} \varphi^*(x,t)\psi(x,t) \, \mathrm{d}x$$

- **Dual** of vector space $V$ is set of linear functionals from $V$ to $\mathbb{C}$:

$$V^* := \left\{\Phi : V \to \mathbb{C} : \forall (a,b) \in \mathbb{C}^2, \forall (z,w) \in V^2, \quad \Phi(a\boldsymbol{z} + b\boldsymbol{w}) = a\Phi(\boldsymbol{z}) + b\Phi(\boldsymbol{w})\right\}$$

  We have $\dim(V^*) = \dim(V)$.
- If $V = \mathbb{C}^n$, can think of vectors in $V$ as $n \times 1$ matrices and vectors in $V^*$ as $1 \times n$ matrices.
- A quantum mechanical system is described by a ket $|\psi\rangle$ in Hilbert space $\mathcal{H}$. For all $|\psi\rangle, |\varphi\rangle \in \mathcal{H}$:
  - $\forall (a,b) \in \mathbb{C}^2, a|\psi\rangle + b|\varphi\rangle \in \mathcal{H}$
  - Inner product of $|\psi\rangle$ with $|\varphi\rangle$ is a complex number written as $\langle \psi | \varphi \rangle$. It is Hermitian: $\langle \psi | \varphi \rangle = \langle \varphi | \psi \rangle^*$.

- Inner product is **sesquilinear** (linear in the second factor, anti-linear in the first). For $|\varphi\rangle = c_1|\varphi_1\rangle + c_2|\varphi_2\rangle$:

$$\langle\psi|\varphi\rangle = c_1\langle\psi|\varphi_1\rangle + c_2\langle\psi|\varphi_2\rangle$$
$$\langle\varphi|\psi\rangle = c_1^*\langle\varphi_1|\psi\rangle + c_2^*\langle\varphi_2|\psi\rangle$$

- $\langle\psi|\psi\rangle \geq 0$ and $\langle\psi|\psi\rangle = 0 \iff |\psi\rangle = 0$.
- States which differ by only a normalisation factor are physically equivalent:

$$\forall c \in \mathbb{C}^*, \quad |\psi\rangle \sim c|\psi\rangle$$

For this reason, pure quantum mechanical states are called **rays** in the Hilbert space, and we normally assume that a state $|\psi\rangle$ has norm 1: $\||\psi\rangle\| = 1$.

- **Physical state** condition: $\langle\psi|\psi\rangle \geq 0$ and $\langle\psi|\psi\rangle = 0 \iff |\psi\rangle = 0$.
- Note that the state labelled zero, $|0\rangle$, is not equal to the zero state (the 0 vector).
- If $\hat{A}$ is linear opearator then $\hat{A}(a|\psi\rangle + b|\varphi\rangle) = a(\hat{A}|\psi\rangle) + b(\hat{A}|\varphi\rangle)$
- Products and combinations of linear operators are also linear operators.
- **Adjoint (Hermitian conjugate)** of $\hat{A}$, $\hat{A}^\dagger$ is defined by

$$\langle\psi|\left(\hat{A}^\dagger|\varphi\rangle\right) = \left(\langle\varphi|\left(\hat{A}|\psi\rangle\right)\right)^*$$

- $\hat{A}$ is **self-adjoint (Hermitian)** if $\widehat{H}^\dagger = \widehat{H}$. Self-adjoint operators correspond to **observables** (measurable quantities) since they have real eigenvalues. Similarly, a **hermitian matrix** $H$ satisfies $H^\dagger = \left(H^T\right)^* = H$.
- $\hat{U}$ is **unitary** if $\hat{U}^\dagger\hat{U} = \hat{I}$. Unitary operators describe time-evolution in quantum mechanics. Similarly, a unitary matrix $U$ satisfies $U^\dagger U = UU^\dagger = I$.
- **Commutator** of operators $\hat{A}$ and $\hat{B}$:

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$$

- **Anti-commutator**:

$$\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$$

- **Expectation value** of observable $\hat{A}$ on state $|\psi\rangle$:

$$\langle A\rangle_\psi \coloneqq \left\langle\psi|\hat{A}|\psi\right\rangle$$

Interpreted as average outcome of many measurements of $\hat{A}$ on same state $|\psi\rangle$.

- If we have $\langle n|m\rangle = \delta_{nm}$, the basis is orthonormal.
- **Qubit system**: Hilbert space $\mathcal{H} = \operatorname{span}(|0\rangle, |1\rangle)$. Any $|\psi\rangle \in \mathcal{H}$ can be written as $a_0|0\rangle + a_1|1\rangle$. If $|\varphi\rangle = b_0|0\rangle + b_1|1\rangle$,

$$\begin{aligned}\langle\varphi|\psi\rangle &= (b_0^*\langle 0| + b_1^*|1\rangle)(a_0|0\rangle + a_1|1\rangle) \\ &= {b_0}^*a_0\langle 0|0\rangle + {b_1}^*a_1\langle 1|1\rangle + {b_0}^*a_1\langle 0|1\rangle + {b_1}^*a_0\langle 1|0\rangle = {b_0}^*a_0 + {b_1}^*a_0 \\ &= \begin{bmatrix}{b_0}^* & {b_1}^*\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix}\begin{bmatrix}a_0 \\ a_1\end{bmatrix}\end{aligned}$$

If $|0\rangle, |1\rangle$ is an energy eigenbasis, then $\widehat{H}|0\rangle = E_0|0\rangle$ and $\widehat{H}|1\rangle = E_1|1\rangle$ where $E_0, E_1$ are eigenvalues.

$\mathbb{P}(\text{measuring } E_0) = a_0^2 = |\langle 0|\psi\rangle|^2, \mathbb{P}(\text{measuring } E_1) = a_1^2 = |\langle 1|\psi\rangle|^2$. If $a_0^2 + a_1^2 = 1$, then $\langle\psi|\psi\rangle = 1$ so $\psi$ is normalised. The expected energy measurement is $\langle E\rangle = E_0\ |a_0|^2 + E_1\ |a_1|^2$.

- **Matrix form** of operator $\hat{A}$:

$$A_{nm} = \left\langle n|\hat{A}|m\right\rangle$$

For $\hat{A}^\dagger$, $\left\langle n|\hat{A}^\dagger|m\right\rangle = \left\langle m|\hat{A}|n\right\rangle^*$.

- **Change of basis**: $B = S^{-1}AS$.
- **Schrodinger equation in braket notation**:

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \widehat{H}|\psi(t)\rangle \Longrightarrow |\psi(t)\rangle = \hat{U}_t|\psi(0)\rangle$$

where $\hat{U}_t$ is unitary operator. If $\widehat{H}$ independent of $t$, then $\hat{U}_t = \exp\left(-\frac{i}{\hbar}t\widehat{H}\right)$.

- **Exponential of operator**:

$$\exp\left(\hat{A}\right) = \sum_{n=0}^{\infty}\frac{\hat{A}^n}{n!}$$

- If $\hat{A} = \text{diag}(a_1, ..., a_n)$ is diagonal, then $\exp\left(\hat{A}\right) = \text{diag}(e^{a_1}, ..., e^{a_n})$.
- If $J^2 = -I$ ($I$ is identity matrix) then

$$\exp(Jt) = \cos(t)I + \sin(t)J$$

# 2. Measurement and uncertainty

- For Hilbert space of finite dimension $N$, operator $\widehat{M}$ has $N$ eigenvalues (counting multiplicities). Eigenvalues of operator $\widehat{M}$ to possible values of the measurable quantity it represents.
- **Spectrum** of $\widehat{H}$:

$$\text{Spec}\left(\widehat{H}\right) := \left\{\lambda \in \mathbb{C} : \widehat{H} - \lambda\hat{I} \text{ non invertible}\right\}$$

For finite-dimensional Hilbert space, this is equal to the set of eigenvalues of $\widehat{H}$.

- For self-adjoint operator $\widehat{H}$, eigenstates $|n\rangle$ corresponding to different eigenvalues $\lambda_n$ are orthogonal. If eigenvalue is degenerate (multiplicity greater than one) then for each eigenspace (vector space spanned by the eigenvectors) with dimension greater than one, we can choose an orthogonal basis of eigenstates (e.g. with Gram-Schmidt).
- Only eigenvalue of identity operator is 1 with degeneracy $N$, so for any orthonormal basis of $\mathcal{H}$:

$$\hat{I} = \sum_n |n\rangle\langle n|$$

- $\hat{A}$ **diagonalisable** if $\hat{A} = \hat{S}\hat{D}\hat{S}^{-1}$ where $\hat{D}$ is diagonal and $\hat{S}$ has columns corresponding to eigenvectors of $\hat{A}$.
- For $\hat{A}$ diagonalisable,

$$\exp\left(\hat{A}\right) = \sum_{n=0}^{\infty} \frac{\left(\hat{S}\hat{D}\hat{S}^{-1}\right)^n}{n!} = \hat{S}\left(\sum_{n=0}^{\infty} \frac{\hat{D}^n}{n!}\right)\hat{S}^{-1} = \hat{S}\exp\left(\hat{D}\right)\hat{S}^{-1}$$

- **Spectral representation of operator**:

$$\hat{A} = \sum_n \lambda_n |n\rangle\langle n|$$

  for orthonomal eigenvectors $\{|n\rangle\}$ and eigenvalues $\lambda_n$. When measurement is made on state

$$|\psi\rangle = \sum_n c_n |n\rangle$$

  the result is $\lambda_n$ with probability $p_n = |\langle n|\psi\rangle|^2 = |c_n|^2$. If result is $\lambda_n$, measuring again immediately after the measurement will yield $\lambda_n$, so the state is no longer $|\psi\rangle$ but $|n\rangle$. This **collapse of the wavefunction** cannot be represented by a unitary operation, and is not reversible.
- Can describe measurement process as set of projection operators $\hat{P}_n = |n\rangle\langle n|$, then $p_n = \left\langle \psi|\hat{P}_n|\psi\right\rangle$ and resulting state $\frac{1}{\sqrt{p_n}}\hat{P}_n|\psi\rangle$ which is equal to $|n\rangle$ up to an irrelevant overall phase. $\hat{P}_n^\dagger = \hat{P}_n$ and $\hat{P}_n^2 = \hat{P}_n$. If the spectrum of $\hat{A}$ is degenerate, we can define

$$\hat{P}_\lambda := \sum_{n:\lambda_n=\lambda} |n\rangle\langle n|$$

  then we still have $p_\lambda = \left\langle \psi|\hat{P}_\lambda|\psi\right\rangle$ and resulting state is $1/\sqrt{p_\lambda}\hat{P}_\lambda|\psi\rangle$.
- $\hat{A}$ and $\hat{B}$ are **compatible** if $[\hat{A}, \hat{B}] = 0$.
- A state can only have definite values for observables $A$ and $B$ if it is a simultaneous eigenstate of both $\hat{A}$ and $\hat{B}$.
- There always exist simultaneous eigenstates for compatible operators.
- If $\hat{A}$ and $\hat{B}$ are not compatible, measuring $A$ then $B$ then $A$ again will not necessarily give the same result for the two measurements of $A$.
- We can view a function $f$ acting on real numbers as acting on $\hat{A}$ by

$$f\left(\hat{A}\right) = \sum_n f(\lambda_n)|n\rangle\langle n|$$

- A **pure state** is definite, i.e. the state of the system is completely known, and the only uncertainties are due to the uncertain nature of quantum mechanics.This is classical uncertainty rather than quantum uncertainty.
- The **density matrix** of a pure state $|\psi\rangle$ is

$$\hat{\rho} := |\psi\rangle\langle\psi|$$

- There is a bijective correspondence between density matrices and the associated pure states:

$$\begin{aligned} \widehat{M}|\psi\rangle = \lambda|\psi\rangle \quad &\longleftrightarrow \quad \widehat{M}\hat{\rho} = \lambda\hat{\rho} \\ |\psi\rangle \to \hat{U}|\psi\rangle \quad &\longleftrightarrow \quad \hat{\rho} \to \hat{U}\hat{\rho}\hat{U}^\dagger \end{aligned}$$

i.e. transforming a state $|\psi\rangle$ by unitary operator $\hat{U}$ is equivalent to transforming the density matrix $\hat{\rho}$ to $\hat{U}\hat{\rho}\hat{U}^\dagger$.

- For orthonormal basis states $|n\rangle$, **trace** of $\hat{A}$ is

$$\mathrm{tr}\left(\hat{A}\right) = \sum_n \left\langle n|\hat{A}|n\right\rangle$$

- **Cyclicity of trace**:

$$\mathrm{tr}(ABC) = \mathrm{tr}(BCA) = \mathrm{tr}(CAB)$$

- For a **density matrix** describing a **pure state** $\hat{\rho} = |\psi\rangle\langle\psi|$,

$$\begin{aligned}\mathrm{tr}(\hat{\rho}) &= \sum_n \langle n|\hat{\rho}|n\rangle = \sum_n \langle n|\psi\rangle\langle\psi|n\rangle \\ &= \sum_n \langle\psi|n\rangle\langle n|\psi\rangle = \langle\psi|\left(\sum_n |n\rangle\langle n|\right)|\psi\rangle = \left\langle\psi|\hat{I}|\psi\right\rangle = \langle\psi|\psi\rangle = 1\end{aligned}$$

Also $\mathrm{tr}(\hat{\rho}^2) = 1$ since $\hat{\rho}$ is a projector and hence $\hat{\rho}^2 = \hat{\rho}$.

- A **mixed state** is one where the state of the system is not known. It is an ensemble of pure states each with an associated probability of the system being in that state: $\{(p_i, |i\rangle)\}$, where the $|i\rangle$ are normalised (not necessarily orthogonal).
- **Density matrix** of a **mixed state** is lienar combination of density matrices for each pure state weighted by probability:

$$\hat{\rho} \coloneqq \sum_i p_i |i\rangle\langle i|$$

Can generalise definition to include possibility of ensembles containing mixed states: $\hat{\rho} = \sum_i p_i \hat{\rho}_i$ where $\hat{\rho}_i$ are mixed and/or pure density matrices.

- **Note**: generally the ensemble that gives rise to a given density matrix for a mixed state is not unique.
- **Example**: for ensemble $\left\{\left(\frac{3}{4}, |0\rangle\right), \left(\frac{1}{4}, |1\rangle\right)\right\}$,

$$\hat{\rho} = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| = \begin{bmatrix} 3/4 & 0 \\ 0 & 1/4 \end{bmatrix}$$

This ensemble is **not** unique:

$$\left\{\left(\frac{1}{2}, \sqrt{\frac{3}{4}}|0\rangle + \sqrt{\frac{1}{4}}|1\rangle\right), \left(\frac{1}{2}, \sqrt{\frac{3}{4}}|0\rangle - \sqrt{\frac{1}{4}}|1\rangle\right)\right\}$$

gives an equivalent density matrix:

$$\begin{aligned}\hat{\rho}_1 &= \left(\sqrt{\frac{3}{4}}|0\rangle + \sqrt{\frac{1}{4}}|1\rangle\right)\left(\sqrt{\frac{3}{4}}\langle 0| + \sqrt{\frac{1}{4}}\langle 1|\right) \\ &= \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| + ..., \hat{\rho}_2 = ..., \frac{1}{2}\hat{\rho}_1 + \frac{1}{2}\hat{\rho}_2 = \begin{bmatrix} 3/4 & 0 \\ 0 & 1/4 \end{bmatrix}\end{aligned}$$

- For observable $\widehat{A}$ expressed in matrix form with basis as the states $|\psi_i\rangle$, then $\langle\widehat{A}\rangle = \operatorname{tr}\left(\widehat{\rho}\widehat{A}\right)$. For mixed state, we still have $\operatorname{tr}(\widehat{\rho}) = 1$ but $\operatorname{tr}(\widehat{\rho}^2) = \sum_i p_i^2 \le 1$ with equality only when some $p_i = 1$ (i.e. a pure state). $\operatorname{tr}(\widehat{\rho}^2)$ conveys how "mixed" the state is.
- **Example**:

$$\begin{aligned}\langle E\rangle_\psi &= \left\langle\psi|\widehat{H}|I|\psi\right\rangle = \sum_n \left\langle\psi|\widehat{H}|n\right\rangle\langle n|\psi\rangle \\ &= \sum_n \langle n|\psi\rangle\left\langle\psi|\widehat{H}|n\right\rangle = \sum_n \left\langle n|\widehat{\rho}_\psi|\widehat{H}|n\right\rangle = \operatorname{tr}\left(\widehat{\rho}_\psi\widehat{H}\right)\end{aligned}$$

- Mixed states can only give a pure state when there is one pure state with probability 1.
- **Definition**: $\widehat{\rho}$ is a **density operator** on a Hilbert space if
  - **Normalised**: $\operatorname{tr}(\widehat{\rho}) = 1$
  - **Hermitian**: $\widehat{\rho}^\dagger = \widehat{\rho}$
  - **Semi-positive-definite**: for every state $|\psi\rangle$, $\langle\psi|\widehat{\rho}|\psi\rangle \ge 0$ (can be $= 0$ when $|\psi\rangle \ne 0$).
- All density matrices are density operators.
- After taking a measurement of a pure or mixed state:
  - The measurement is $\lambda$ with probability $p_\lambda = \operatorname{tr}\left(\widehat{P}_\lambda\widehat{\rho}\widehat{P}_\lambda\right) = \operatorname{tr}\left(\widehat{P}_\lambda\widehat{\rho}\right)$.
  - Density matrix after measuring value of $\lambda$ is

$$\widehat{\rho} \to \frac{1}{p_\lambda}\widehat{P}_\lambda\widehat{\rho}\widehat{P}_\lambda = \frac{1}{\operatorname{tr}\left(\widehat{P}_\lambda\widehat{\rho}\widehat{P}_\lambda\right)}\widehat{P}_\lambda\widehat{\rho}\widehat{P}_\lambda$$

- **Theorem**: let $\widehat{\rho}$ be a density operator on a Hilbert space, then $\widehat{\rho}$ corresponds to a pure state iff $\operatorname{tr}(\widehat{\rho}^2) = 1$.

# 3. Qubits and the Bloch sphere

## 3.1. Qubits

- **Definition**: a **qubit** is a state in a two-dimensional Hilbert space. Usually the **computational basis** $\{|0\rangle, |1\rangle\}$ is used to denote the basis for such a Hilbert space.
- A general pure state in a qubit system is of the form

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle, \quad 0 \le \theta \le \pi, 0 \le \varphi < 2\pi$$

  This state is normalised: $|\cos\left(\frac{\theta}{2}\right)|^2 + |e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|^2 = 1$. This gives a bijection between pure qubit states and points on $S^2$, called the **Bloch sphere**.
- Any point on the Bloch sphere can be labelled by its position vector:

$$\boldsymbol{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad x = \sin(\theta)\cos(\varphi), y = \sin(\theta)\sin(\varphi), z = \cos(\theta)$$

- There are six special states on the Bloch sphere:

$$|+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} : \quad \boldsymbol{r} = \begin{bmatrix}1\\0\\0\end{bmatrix}, \quad (\theta, \varphi) = (\pi/2, 0)$$

$$|-\rangle := \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} : \quad \boldsymbol{r} = \begin{bmatrix}-1\\0\\0\end{bmatrix}, \quad (\theta, \varphi) = (\pi/2, \pi)$$

$$|L\rangle := \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix}1\\i\end{bmatrix} : \quad \boldsymbol{r} = \begin{bmatrix}0\\1\\0\end{bmatrix}, \quad (\theta, \varphi) = (\pi/2, \pi/2)$$

$$|R\rangle := \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix} : \quad \boldsymbol{r} = \begin{bmatrix}0\\-1\\0\end{bmatrix}, \quad (\theta, \varphi) = (\pi/2, 3\pi/2)$$

$$|0\rangle \leftrightarrow \begin{bmatrix}1\\0\end{bmatrix} : \quad \boldsymbol{r} = \begin{bmatrix}0\\0\\1\end{bmatrix}, \quad (\theta, \varphi) = (0, \cdot)$$

$$|1\rangle \leftrightarrow \begin{bmatrix}0\\1\end{bmatrix} : \quad \boldsymbol{r} = \begin{bmatrix}0\\0\\-1\end{bmatrix}, \quad (\theta, \varphi) = (\pi, \cdot)$$

## 3.2. Inside the Bloch sphere

- **Definition**: **Pauli $\boldsymbol{\sigma}$-matrices** are

$$\sigma_1 := \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}, \quad \sigma_2 := \begin{bmatrix}0 & -i\\i & 0\end{bmatrix}, \quad \sigma_3 := \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$$

- Density matrix for $|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle$ is given by

$$\begin{aligned}\hat{\rho} = |\psi\rangle\langle\psi| \to \rho &= \begin{bmatrix}\cos\left(\frac{\theta}{2}\right)\\ e^{i\varphi}\sin\left(\frac{\theta}{2}\right)\end{bmatrix}\begin{bmatrix}\cos\left(\frac{\theta}{2}\right) & e^{-i\varphi}\sin\left(\frac{\theta}{2}\right)\end{bmatrix}\\ &= \frac{1}{2}\begin{bmatrix}1+\cos(\theta) & e^{-i\varphi}\sin(\theta)\\ e^{i\varphi}\sin(\theta) & 1-\cos(\theta)\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1+z & x-iy\\ x+iy & 1-z\end{bmatrix}\\ &= \frac{1}{2}(I_2 + \boldsymbol{r}\cdot\sigma)\end{aligned}$$

  where $\boldsymbol{r}\cdot\sigma = r_1\sigma_1 + r_2\sigma_2 + r_3\sigma_3 = x\sigma_1 + y\sigma_2 + z\sigma_3$.
- Density matrix for pure state is linear in the Bloch vector $\boldsymbol{r}$, so mixed states have Bloch vector given by linear combination of Bloch vectors of states in the ensemble.
- For mixed state $\{(p_i, \rho_i) : i \in [m]\}$ where $\rho_i$ are pure state density matrices defined by Bloch vectors $\boldsymbol{r}_i$, density matrix for mixed state is

$$\rho = \sum_{i=1}^{m} p_i\rho_i = \sum_{i=1}^{m} p_i\frac{1}{2}(I_2 + \boldsymbol{r}_i\cdot\sigma) = \frac{1}{2}(I_2 + \boldsymbol{r}\cdot\sigma)$$

  where $\boldsymbol{r} = \sum_{i=1}^{m} p_i r_i$. Now

$$
\begin{aligned}
|\boldsymbol{r}|^2 = |\sum_{i=1}^{m} p_i \boldsymbol{r}_i|^2 &= \sum_{(i,j)\in[m]^2} p_i p_j \boldsymbol{r}_i \cdot \boldsymbol{r}_j \\
&\leq \sum_{(i,j)\in[m]^2} p_i p_j \ |\boldsymbol{r}_i|^2 \ |\boldsymbol{r}_j|^2 = \sum_{(i,j)\in[m]^2} p_i p_j = \sum_{i=1}^{m} p_i \sum_{j=1}^{m} p_j = 1
\end{aligned}
$$

by Cauchy-Schwartz inequality. Equality holds iff all $\boldsymbol{r}_i$ are collinear, hence iff it is a pure state. So strictly mixed states are defined by a Bloch vector $\boldsymbol{r}$ with $|\boldsymbol{r}| < 1$.

- For any density matrix $\rho$,

$$\operatorname{tr}(\rho^2) = \frac{1}{2}(1 + |\boldsymbol{r}|^2)$$

## 3.3. Time evolution of a qubit

- Unitary transformations of a qubit correspond to rotations of points on/in the Bloch sphere about the origin, representing the fact that unitary transformations cannot transform pure states to mixed states
- $\operatorname{tr}(\rho^2) = \frac{1}{2}(1 + |\boldsymbol{r}|^2)$ is invariant under unitary transformations. It measures how mixed a state is: $\operatorname{tr}(\rho^2) = 1$ for pure states, $\operatorname{tr}(\rho^2) = \frac{1}{2}$ for the most mixed state (corresponds to the origin, $\boldsymbol{r} = \boldsymbol{0}$, $\rho = \frac{1}{2}I$).
- Measurements are not unitary transformations but projection operators, and can transform any state to a pure state.
- **Example**:
  - For $\boldsymbol{r_1}, \boldsymbol{r_2}$ distinct points on the Bloch sphere, density matrix corresponding to mixed state $\{(p, \boldsymbol{r_1}), (1 - p, \boldsymbol{r_2})\}$ is

$$\rho = p\rho_1 + (1-p)\rho_2 = \frac{1}{2}(I + \boldsymbol{r} \cdot \sigma), \quad \boldsymbol{r} = p\boldsymbol{r_1} + (1-p)\boldsymbol{r_2}$$

  - Geometrically, $\boldsymbol{r}$ lies in line between $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ inside the Bloch sphere (since $p \in [0, 1]$).
  - Mixing states can never produce a state further from the origin than the furthest initial state.
  - There are an infinite number of ways of writing a mixed state as an ensemble of two pure states: any line passing through the point represented by the mixed states intersects with the Bloch sphere twice - the intersection points give the pure states in the ensemble.
  - Most mixed state, with $\rho = \frac{1}{2}I_2$, corresponds to ensemble of antipodal points, each with probability $\frac{1}{2}$.
- **Definition**: **trace distance** between two density matrices:

$$D(\hat{\rho}_1, \hat{\rho}_2) = \frac{1}{2}\operatorname{tr}|\hat{\rho}_1 - \hat{\rho}_2| = \frac{1}{4}\operatorname{tr}|(\boldsymbol{r_1} - \boldsymbol{r_2}) \cdot \sigma| = \frac{1}{2}\ |\boldsymbol{r_1} - \boldsymbol{r_2}| = \sum_i |\lambda_i|$$

where $|\hat{A}| = \sqrt{\hat{A}^\dagger \hat{A}}$ and $\lambda_i$ are the eigenvalues of $\hat{\rho}_1 - \hat{\rho}_2$ (equal to sum of eigenvalues assuming that $\hat{\rho}_1 - \hat{\rho}_2$ is Hermitian).

- Trace distance defines a **metric** on set of density matrices:

- **Non-negative**: $D(\hat{\rho}_1, \hat{\rho}_2) \geq 0$.
- **Separates points**: $D(\hat{\rho}_1, \hat{\rho}_2) = 0 \iff \hat{\rho}_1 = \hat{\rho}_2$.
- **Symmetric**: $D(\hat{\rho}_1, \hat{\rho}_2) = D(\hat{\rho}_2, \hat{\rho}_2)$.
- **Triangle inequality**: $D(\hat{\rho}_1, \hat{\rho}_3) \leq D(\hat{\rho}_1, \hat{\rho}_2) + D(\hat{\rho}_2, \hat{\rho}_3)$

## 3.4. Pauli matrices

- **Definition**: **Levi-Cevita** tensor $\varepsilon_{ijk}$ satisfies:
  - $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312}$.
  - $\varepsilon_{321} = \varepsilon_{132} = \varepsilon_{213}$.
  - $\varepsilon_{ijk} = 0$ otherwise for $\{i, j, k\} \subseteq \{1, 2, 3\}$.
- Properties of Pauli matrices:
  - **Hermitian**: $\sigma_i^\dagger = \sigma_i$.
  - **Traceless**: $\operatorname{tr}(\sigma_i) = 0$.
  - $\left[\sigma_i, \sigma_j\right] = \sigma_i \sigma_j - \sigma_j \sigma_i = 2i\varepsilon_{ijk}\sigma_k$.
  - $\left\{\sigma_i, \sigma_j\right\} = \sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij} I_2$.
  - $\sigma_i \sigma_j = \delta_{ij} I_2 + i\varepsilon_{ijk}\sigma_k$.
  - Form a basis for vector space of $2 \times 2$ Hermitian traceless matrices.
- The operators

$$X = \frac{1}{2}(I_2 - \sigma_1) = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$
$$Y = \frac{1}{2}(I_2 - \sigma_2) = \frac{1}{2}\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$$
$$Z = \frac{1}{2}(I_2 - \sigma_3) = \frac{1}{2}\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

  have their eigenvectors as the six special Bloch states, with eigenvalues 0 or 1:

$$X|+\rangle = 0|+\rangle, \quad X|-\rangle = 1|-\rangle,$$
$$Y|L\rangle = 0|L\rangle, \quad Y|R\rangle = 1|R\rangle,$$
$$Z|0\rangle = 0|0\rangle, \quad Z|1\rangle = 1|1\rangle$$

- The exponentials of Pauli matrices are unitary matrices: $\forall \alpha \in \mathbb{R}$,

$$\exp(i\alpha\sigma_1) = \begin{bmatrix} \cos(\alpha) & i\sin(\alpha) \\ i\sin(\alpha) & \cos(\alpha) \end{bmatrix} = \cos(\alpha) I_2 + i\sin(\alpha)\sigma_1,$$
$$\exp(i\alpha\sigma_2) = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} = \cos(\alpha) I_2 + i\sin(\alpha)\sigma_2,$$
$$\exp(i\alpha\sigma_3) = \begin{bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{bmatrix} = \cos(\alpha) I_2 + i\sin(\alpha)\sigma_3$$

- For $\alpha \in \mathbb{R}$, $\boldsymbol{n} \in \mathbb{R}^3$, $|\boldsymbol{n}|^2 = 1$,

$$U_\alpha(\boldsymbol{n}) := \exp(i\alpha\boldsymbol{n} \cdot \sigma) = \cos(\alpha) I_2 + i\sin(\alpha)\boldsymbol{n} \cdot \sigma$$

is unitary transformation so is time evolution operator. If density mmatrix $\rho = \frac{1}{2}(I_2 + \boldsymbol{r} \cdot \sigma)$ evolves with time according to this operator, then

$$\rho \to U_\alpha(\boldsymbol{n})\rho U_\alpha(\boldsymbol{n})^\dagger = \frac{1}{2}(I_2 + (R_\alpha(\boldsymbol{n})\boldsymbol{r}) \cdot \sigma)$$

where $R_\alpha(\boldsymbol{n})$ is $3 \times 3$ orthogonal matrix corresponding to rotation of angle $2\alpha$ about axis in the direction of $\boldsymbol{n}$.

# 4. Bipartite systems

## 4.1. Tensor products

- **Tensor product** $|\varphi\rangle \otimes |\psi\rangle$ in $H_1 \otimes H_2$ satisfies:
  - **Scalar multiplication**: $c(|\varphi\rangle \otimes |\psi\rangle) = (c|\varphi\rangle) \otimes |\psi\rangle = |\varphi\rangle \otimes (c|\psi\rangle)$
  - **Linearity**:
    - $a|\psi\rangle \otimes |\varphi_1\rangle + b|\psi\rangle \otimes |\varphi_2\rangle = |\psi\rangle \otimes (a|\varphi_1\rangle + b|\varphi_2\rangle)$
    - $a|\psi_1\rangle \otimes |\varphi\rangle + b|\psi_2\rangle \otimes |\varphi\rangle = (a|\psi_1\rangle + b|\psi_2\rangle) \otimes |\varphi\rangle$
- Inner products of $H_1$ and $H_2$ induce an inner product on $H_1 \otimes H_2$: for $|\psi_1\rangle, |\psi_2\rangle \in H_1$, $|\varphi_1\rangle, |\varphi_2\rangle \in H_2$,

$$(\langle\psi_1| \otimes \langle\varphi_1|)(|\psi_2\rangle \otimes |\varphi_2\rangle) = \langle\psi_1|\psi_2\rangle\langle\varphi_1|\varphi_2\rangle$$

- For bases $\{|i\rangle\}$ for $H_1$ and $\{|j\rangle\}$ for $H_2$, $\{|i\rangle \otimes |j\rangle\}$ is basis for $H_1 \otimes H_2$: for $|\psi\rangle \in H_1$, $|\varphi\rangle \in H_2$,

$$|\psi\rangle \otimes |\varphi\rangle = \left(\sum_i a_i|i\rangle\right) \otimes \left(\sum_j b_j|j\rangle\right) = \sum_{i,j} a_i b_j |i\rangle \otimes |j\rangle$$

- The most general vector $|\psi\rangle \in H_1 \otimes H_2$ is

$$|\psi\rangle = \sum_{i,j} c_{i,j}|i\rangle \otimes |j\rangle$$

  Generally, this cannot be written as a tensor product $|\psi\rangle \otimes |\varphi\rangle$. If it can be, it is a **separable** state. If not, it is **entangled** (e.g. a linear combination of separable states is generally entangled).
- If $\{|i\rangle\}$, $\{|j\rangle\}$ orthonormal then the inner product in $H_1 \otimes H_2$ is given by

$$\begin{aligned}\langle\varphi|\psi\rangle &= \left(\sum_{i,j} d_{i,j}^* \langle i| \otimes \langle j|\right)\left(\sum_{m,n} c_{m,m}|m\rangle \otimes |n\rangle\right) \\ &= \sum_{i,j,m,n} d_{i,j}^* c_{m,n} \langle i|m\rangle\langle j|n\rangle = \sum_{i,j} d_{i,j}^* c_{i,j}\end{aligned}$$

- The Hilbert space of an $N$-qubit system is the $2^N$-dimensional Hilbert space $\mathcal{H}_N = \mathcal{H}_q^{\otimes N}$ where $\mathcal{H}_q$ is a single qubit Hilbert space.
- **Example**: let $\mathcal{H}_3 = \mathcal{H}_q \otimes \mathcal{H}_q \otimes \mathcal{H}_q$. Operator $\hat{I} \otimes \widehat{\sigma_1} \otimes \hat{I}$ acts on the second qubit and leaves the other two invariant. $\widehat{\sigma_1}|0\rangle = |1\rangle$ and $\widehat{\sigma_1}|1\rangle = |0\rangle$ so in this basis, $\sigma_1$ acts the logical NOT gate $\overline{(\cdot)}$, where $\overline{0} = 1$, $\overline{1} = 0$. So

$$(\hat{I} \otimes \hat{\sigma}_1 \otimes \hat{I})|xyz\rangle = |x\overline{y}z\rangle$$

## 4.2. Linear operators and local unitary operations

- Linear operators on $\mathcal{H}$ can be written as linear combinations of $\hat{A} \otimes \hat{B}$, where

$$\left(\hat{A} \otimes \hat{B}\right)(|\psi\rangle \otimes |\varphi\rangle) = \left(\hat{A}|\psi\rangle\right) \otimes \left(\hat{B}|\varphi\rangle\right)$$

- Properties of tensor product of linear operators:
  - $\hat{A} \otimes \hat{B} + \hat{C} \otimes \hat{B} = \left(\hat{A} + \hat{C}\right) \otimes \hat{B}$.
  - $\hat{A} \otimes \hat{B} + \hat{A} \otimes \hat{D} = \hat{A} \otimes \left(\hat{B} + \hat{D}\right)$.
  - $\left(\hat{A} \otimes \hat{B}\right)^\dagger = \hat{A}^\dagger \otimes \hat{B}^\dagger$.
  - $(\hat{A} \otimes \hat{B})(\hat{C} \otimes \hat{D}) = (\hat{A}\hat{C} \otimes \hat{B}\hat{D})$.
  - $\operatorname{tr}_{\mathcal{H}_A \otimes \mathcal{H}_B}(\hat{A} \otimes \hat{B}) = \operatorname{tr}_{\mathcal{H}_A}(\hat{A}) \operatorname{tr}_{\mathcal{H}_B}(\hat{B})$.

  So tensor product of linear operators preserves unitarity, Hermiticity, positivity, and tensor product of two projectors is a projector.
- For bipartite system, **local operations (LO)** are of the form $\hat{U}_A \otimes \hat{I}$ (for Alice) or $\hat{I} \otimes \hat{U}_B$ (for Bob). This means Alice and Bob can only act on their own qubit (assuming they have one each) where $\hat{U}_A$ and $\hat{U}_B$ are unitary operators or measurement operators.
- $\hat{U}_A \otimes \hat{I}$ and $\hat{I} \otimes \hat{U}_B$ commute: $[\hat{U}_A \otimes \hat{I}, \hat{I} \otimes \hat{U}_B] = 0$, and their product is $\hat{U}_A \otimes \hat{U}_B$.
- **Note**: any unitary transformation $\hat{U}_A \otimes \hat{U}_B$ performed using LO acting on a separable state $|\psi\rangle \otimes |\varphi\rangle$ produces another separable state: $\hat{U}_A|\psi\rangle \otimes \hat{U}_B|\varphi\rangle$. In particular, an entangled state cannot be created from a separable state.
- **Example**:

$$e^{\hat{A}\otimes\hat{I}} = \sum_{k=0}^{\infty} \frac{\hat{A} \otimes \hat{I}}{k!} = \sum_{k=0}^{\infty} \frac{\hat{A} \otimes \hat{I}}{k!} = e^{\hat{A}} \otimes \hat{I},$$

$$e^{\hat{I}\otimes\hat{B}} = \sum_{k=0}^{\infty} \frac{\left(\hat{I} \otimes \hat{B}\right)^k}{k!} = \sum_{k=0}^{\infty} \frac{\hat{I} \otimes \hat{B}^k}{k!} = \hat{I} \otimes e^{\hat{B}}$$

  Note that generally, $e^{\hat{A}} \otimes e^{\hat{B}} \neq e^{\hat{A}\otimes\hat{B}}$ since

$$e^{\hat{A}\otimes\hat{B}} = \sum_{k=0}^{\infty} \frac{\hat{A}^n \otimes \hat{B}^k}{k!},$$

$$e^{\hat{A}} \otimes e^{\hat{B}} = \left(\sum_{i=0}^{\infty} \frac{\hat{A}^i}{i!}\right) \otimes \left(\sum_{j=0}^{\infty} \frac{\hat{B}^j}{j!}\right) = \sum_{i,j=0}^{\infty} \frac{\hat{A}^i \otimes \hat{B}^j}{i!j!}$$

- **Definition**: a mixed state is **separable** iff it is an ensemble of separable states, and **entangled** otherwise.
- **Definition**: **density matrix** of separable pure state $|\Psi\rangle = |\psi\rangle \otimes |\varphi\rangle$ is

$$\hat{\rho} = |\Psi\rangle\langle\Psi| = (|\psi\rangle \otimes |\varphi\rangle)(\langle\psi| \otimes \langle\varphi|) = (|\psi\rangle\langle\psi|) \otimes (|\varphi\rangle\langle\varphi|) = \hat{\rho}_A \otimes \hat{\rho}_B$$

  where $\hat{\rho}_A = |\psi\rangle\langle\psi|$ and $\hat{\rho}_B = |\varphi\rangle\langle\varphi|$.

- **Definition**: **density matrix** of separable mixed state is

$$\hat{\rho} = \sum_i p_i \hat{\rho}_A^{(i)} \otimes \hat{\rho}_B^{(i)}$$

where $\left\{\hat{\rho}_A^{(i)}\right\}$ are mixed or pure states of first system, $\left\{\hat{\rho}_B^{(i)}\right\}$ are mixed or pure states of second system.

## 4.3. Matrix representation

- **Tensor product** of two vectors is given by e.g.

$$\begin{bmatrix}1\\2\\3\end{bmatrix} \otimes \begin{bmatrix}4\\5\end{bmatrix} = \begin{bmatrix}1\begin{bmatrix}4\\5\end{bmatrix}\\2\begin{bmatrix}4\\5\end{bmatrix}\\3\begin{bmatrix}4\\5\end{bmatrix}\end{bmatrix} = \begin{bmatrix}4\\5\\8\\10\\12\\15\end{bmatrix}$$

The expression is similar for matrices:

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} \otimes \begin{bmatrix}5 & 6\\7 & 8\end{bmatrix} = \begin{bmatrix}1\begin{bmatrix}5 & 6\\7 & 8\end{bmatrix} & 2\begin{bmatrix}5 & 6\\7 & 8\end{bmatrix}\\3\begin{bmatrix}5 & 6\\7 & 8\end{bmatrix} & 4\begin{bmatrix}5 & 6\\7 & 8\end{bmatrix}\end{bmatrix} = \begin{bmatrix}5 & 6 & 10 & 12\\7 & 8 & 14 & 16\\15 & 18 & 20 & 24\\21 & 24 & 28 & 36\end{bmatrix}$$

- If $\{|i\rangle : i \in [n]\}$ is orthonormal basis for $\mathcal{H}_A$, $\{|j\rangle : j \in [m]\}$ is orthonormal basis for $\mathcal{H}_B$, then $\{|i\rangle \otimes |j\rangle : i \in [n], j \in [m]\}$ is orthonormal basis for $\mathcal{H}_A \otimes \mathcal{H}_B$.
- **Note**: general vector in tensor product of Hilbert spaces is linear combination of tensor products (of vectors), general linear operator acting on tensor product of Hilbert spaces is linear combination of tensor products (of linear operators).
- **Controlled NOT (CNOT)** operator acts on $\mathcal{H}_2 = \mathcal{H}_q \otimes \mathcal{H}_q$ and is defined as

$$U = \frac{I_2 + \sigma_3}{2} \otimes I_2 + \frac{I_2 - \sigma_3}{2} \otimes \sigma_1$$

We have $U|00\rangle = |00\rangle$, $U|01\rangle = |01\rangle$, $U|10\rangle = |11\rangle$, $U|11\rangle = |10\rangle$.

## 4.4. Local measurements

- Alice and Bob can perform measurements on their own systems using self-adjoint operators of the form $\hat{F} = \hat{F}_A \otimes \hat{I}$ for Alice and $\hat{G} = \hat{I} \otimes \hat{G}_B$ for Bob. If $\hat{F}_A$ and $\hat{G}_B$ both have non-degenerate systems, these operators have projection operators $\hat{F}_{Ai} = |i\rangle\langle i|$ and $\hat{G}_{Bj} = |j\rangle\langle j|$.
- In the full system, $\hat{F}$ and $\hat{G}$ are degenerate, with degeneracy given by dimension of other subsystem, so $\dim(\mathcal{H}_B)$ for Alice's observable and $\dim(\mathcal{H}_A)$ for Bob's. Corresponding projection operators in full system are $\hat{F}_i = \hat{F}_{Ai} \otimes \hat{I}$ and $\hat{G}_j = \hat{I} \otimes \hat{G}_{Bj}$.
- Since $[\hat{F}, \hat{G}] = 0$, these measurements are compatible so Alice and bob can both measure, the final state is eigenstate of both $\hat{F}$ and $\hat{G}$. Probability of an outcome

occuring is not affected by whether Alice or Bob measures first (or simultaneously).

- Let $|\Psi\rangle$ be pure separable state:

$$|\Psi\rangle = |\psi\rangle \otimes |\varphi\rangle = \sum_{i,j} \alpha_i \beta_j |i\rangle \otimes |j\rangle = \sum_{i,j} \gamma_{ij} |i\rangle \otimes |j\rangle$$

where $\{|i\rangle\}$ and $\{|j\rangle\}$ are orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. If Alice measures $\hat{F}$ and obtains $f_m$ with probability $|\alpha_m|^2 = \sum_j |\gamma_{mj}|^2$, system collapses to state

$$\sum_j \beta_j |m\rangle \otimes |j\rangle = |m\rangle \otimes |\varphi\rangle$$

If Bob then measures $\hat{G}$ and obtains $g_n$ with probability $|\beta_n|^2 = \sum_i |\gamma_{in}|^2$ then final state is $|m\rangle \otimes |n\rangle$. This is the same final state as when Bob measures first, except intermediate state is $|\psi\rangle \otimes |n\rangle$. The probabiity of measuring $(f_m, g_n)$ is $|\gamma_{mn}|^2 = |\alpha_m \beta_n|^2$.

- Probability of Alice measuring $f_i$ is $|\langle i|\psi\rangle|^2 = \operatorname{tr}\left(\hat{\rho}_A \hat{F}_{Ai}\right)$ where $\hat{F}_{Ai} = |i\rangle\langle i|$. After measuring $\hat{F}_A$ and finding $f_i$, Alice's state collapses to

$$|\psi\rangle \to |i\rangle = \frac{1}{|\langle i|\psi\rangle|} \hat{F}_{Ai} |\psi\rangle = \frac{1}{\sqrt{\operatorname{tr}\left(\hat{\rho}_A \hat{F}_{Ai}\right)}} \hat{F}_{Ai} |\psi\rangle$$

$$\hat{\rho}_A \to \frac{1}{\operatorname{tr}\left(\hat{\rho}_A \hat{F}_{Ai}\right)} \hat{F}_{Ai} \hat{\rho}_A \hat{F}_{Ai}$$

- For bipartite system with separable state $|\Psi\rangle$, when Alice measures $\hat{F}_A$, she does not operate on Bob's system, so $\hat{F}_i = \hat{F}_{Ai} \otimes \hat{I}$ and density matrix is

$$\hat{\rho} = |\Psi\rangle\langle\Psi| = (|\psi\rangle\langle\psi|) \otimes (|\varphi\rangle\langle\varphi|) = \hat{\rho}_A \otimes \hat{\rho}_B$$

If Alice measures $\hat{F} = \hat{F}_A \otimes \hat{I}$, outcome is $f_i$ with probability $\operatorname{tr}(\hat{\rho}\hat{F}_i) = \operatorname{tr}(\hat{\rho}_A \hat{F}_{Ai})$ and density matrix collapses to

$$\hat{\rho} \to \frac{1}{\operatorname{tr}\left(\hat{\rho}_A \hat{F}_{Ai}\right)} \hat{F}_{Ai} \hat{\rho}_A \hat{F}_{Ai} \otimes \hat{\rho}_B = \frac{1}{\operatorname{tr}\left(\hat{\rho} \hat{F}_i\right)} = \hat{F}_i \hat{\rho} \hat{F}_i$$

Note that the eigenspace corresponding to eigenvalue $f_i$ is non-degenerate in $\mathcal{H}_A$ but any $|i\rangle \otimes |\varphi\rangle$ with $|\varphi\rangle \in \mathcal{H}_B$ is an eigenvector of $\hat{F} \otimes \hat{I}$ with eigenvalue $f_i$, so eigenspace is degenerate in $\mathcal{H}_A \otimes \mathcal{H}_B$. It does not matter if Alice or Bob measures first: if $\hat{F} = \hat{F}_A \otimes \hat{I}$ and $\hat{G} = \hat{I} \otimes \hat{G}_B$ are measured, outcome is $(f_i, g_m)$ with probability $\operatorname{tr}\left(\hat{\rho}\hat{P}_{ij}\right)$ where $\hat{P}_{ij} = \hat{F}_{Ai} \otimes \hat{G}_{Bj} = |i\rangle\langle i| \otimes |j\rangle\langle j|$, and state collapses to

$$\hat{\rho} \to \frac{1}{\operatorname{tr}\left(\hat{\rho}\hat{P}_{ij}\right)} \hat{P}_{ij} \hat{\rho} \hat{P}_{ij} = |i\rangle \otimes |j\rangle$$

- For bipartite system with entangled state $|\Psi\rangle = \sum_{i,j} \gamma_{ij} |i\rangle \otimes |j\rangle$, define coefficients

$$\alpha_m := \left(\sum_j |\gamma_{mj}|^2\right)^{1/2}, \quad \beta_n := \left(\sum_i |\gamma_{in}|^2\right)^{1/2}$$

and define auxiliary states (excluding values of $m$ and $n$ when $\beta_n = 0$ or $\alpha_m = 0$)

$$|\psi_n\rangle := \frac{1}{\beta_n}\sum_i \gamma_{in}|i\rangle \in \mathcal{H}_A,$$
$$|\varphi_m\rangle := \frac{1}{\alpha_m}\sum_j \gamma_{mj}|j\rangle \in \mathcal{H}_B$$

Then

$$|\Psi\rangle = \sum_i \alpha_i |i\rangle \otimes |\varphi_i\rangle = \sum_j \beta_j |\psi_j\rangle \otimes |j\rangle$$

If Alice measures $\hat{F}$ with $f_i$, state collapses to

$$|\Psi\rangle \to \hat{F}_i|\Psi\rangle = \left(\hat{F}_{Ai} \otimes \hat{I}\right)|\Psi\rangle \sim |i\rangle \otimes |\varphi_i\rangle$$

i.e. the entangled state collapses to a separable state. So Bob's state depends on the result of Alice's measurement.

## 4.5. Reduced density matrix

- **Definition**: for operator $\hat{C} \otimes \hat{D}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, **partial trace** over $\mathcal{H}_A$ and $\mathcal{H}_B$, $\mathrm{tr}_A : \mathrm{End}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathrm{End}(\mathcal{H}_B)$ and $\mathrm{tr}_A : \mathrm{End}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathrm{End}(\mathcal{H}_A)$, are respectively

$$\mathrm{tr}_A(\hat{C} \otimes \hat{D}) := \mathrm{tr}(\hat{C})\hat{D}, \quad \mathrm{tr}_B(\hat{C} \otimes \hat{D}) := \mathrm{tr}(\hat{D})\hat{C}$$

- **Definition**: for a given system, the **reduced density matrix** of a subsystem is partial trace of density matrix over other subsystems. So for bipartite system,

$$\hat{\rho}_A := \mathrm{tr}_B(\hat{\rho}), \quad \hat{\rho}_B := \mathrm{tr}_A(\hat{\rho})$$

- **Note**: a reduced matrix describes one subsystem, assuming no knowledge of the other system. Therefore, generally, reduced density matrices describe mixed states, even if full system is in a pure state.
- **Example**: consider state $|\beta_{00}\rangle$:

$$\hat{\rho} = |\beta_{00}\rangle\langle\beta_{00}| = \frac{1}{2}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)$$

$$\begin{aligned}\hat{\rho}_A &= \mathrm{tr}_B(\hat{\rho}) \\ &= \frac{1}{2}(|0\rangle\langle 0|\,\mathrm{tr}_B(|0\rangle\langle 0|) + |0\rangle\langle 1|\,\mathrm{tr}_B(|0\rangle\langle 1|) + |1\rangle\langle 0|\,\mathrm{tr}_B(|1\rangle\langle 0|) + |1\rangle\langle 1|\,\mathrm{tr}_B(|1\rangle\langle 1|)) \\ &= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle + \langle 1|) = \frac{1}{2}\hat{I}\end{aligned}$$

Can also obtain reduced density matrix by writing matrices:

$$
\begin{aligned}
|\beta_{00}\rangle \to \boldsymbol{v} &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\0\\1\end{bmatrix} \\
\hat{\rho} = \boldsymbol{v}\boldsymbol{v}^\dagger &= \frac{1}{2}\begin{bmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix}1&0\\0&0\end{bmatrix} \otimes \begin{bmatrix}1&0\\0&0\end{bmatrix} + \frac{1}{2}\begin{bmatrix}0&1\\0&0\end{bmatrix} \otimes \begin{bmatrix}0&1\\0&0\end{bmatrix} + \frac{1}{2}\begin{bmatrix}0&0\\1&0\end{bmatrix} \otimes \begin{bmatrix}0&0\\1&0\end{bmatrix} + \frac{1}{2}\begin{bmatrix}0&0\\0&1\end{bmatrix} \otimes \begin{bmatrix}0&0\\0&1\end{bmatrix} \\
\rho_A &= \frac{1}{2}\begin{bmatrix}1&0\\0&0\end{bmatrix} \operatorname{tr}\begin{bmatrix}1&0\\0&0\end{bmatrix} + \frac{1}{2}\begin{bmatrix}0&1\\0&0\end{bmatrix} \operatorname{tr}\begin{bmatrix}0&1\\0&0\end{bmatrix} + \frac{1}{2}\begin{bmatrix}0&0\\1&0\end{bmatrix} \operatorname{tr}\begin{bmatrix}0&0\\1&0\end{bmatrix} + \frac{1}{2}\begin{bmatrix}0&0\\0&1\end{bmatrix} \operatorname{tr}\begin{bmatrix}0&0\\0&1\end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix}1&0\\0&0\end{bmatrix} + \frac{1}{2}\begin{bmatrix}0&0\\0&1\end{bmatrix} = \frac{1}{2}I_2
\end{aligned}
$$

- Properties of reduced density matrix $\hat{\rho}_A$:
  - Invariant under all local operations in system $B$.
  - Under unitary transformations $\hat{U}$ in system $A$, $\hat{\rho}_A$ transforms as normal: $\hat{\rho}_A \to \hat{U}\hat{\rho}_A\hat{U}^\dagger$.
  - Local measurements in system $A$ can be described by $\hat{\rho}_A$ and operators acting on $\mathcal{H}_A$: $\operatorname{tr}_B(\hat{F}_i\hat{\rho}\hat{F}_i) = \hat{F}_{Ai}\hat{\rho}_A\hat{F}_{Ai}$.
- **Theorem**: if $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is pure state, then $\hat{\rho}_A$ is pure iff $|\Psi\rangle$ is separable.
- **Corollary**: if spectrum of $\hat{F}_A$ is non-degenerate then measuring $\hat{F}_A$ in system $\mathcal{H}_A$ produces separable state on system $\mathcal{H}_A \otimes \mathcal{H}_B$, i.e. measurements destroys entanglement.
- Entanglement does not violate causality (does not allow communication faster than the speed of light). I.e., if Alice makes a local measurement on an entangled system, Bob cannot detect this, even those the reduced density matrix for his system has changed.

### 4.6. Classical communication

- Alice and Bob can use classical communication (CC) to communicate results of measurements of their own subsystem. If the state was initially entangled, Bob communicating a measurement to Alice would give Alice information about her subsystem.
- If Alice and Bob can use local operations (LO) and classical communication (CC), they can use LOCC.

## 5. Entanglement applications

### 5.1. Bell states

- Checks for entangled state:
  - If $|\psi\rangle \otimes |\psi'\rangle = a|0\rangle \otimes |\varphi\rangle + b|1\rangle \otimes |\varphi\rangle$ for some $a, b \in \mathbb{C}$, $|\varphi\rangle$, then $|\psi\rangle \otimes |\psi'\rangle$ is separable, otherwise entangled.

    - If reduced density matrix of either subsystem gives a pure state ($\mathrm{tr}(\rho^2) = 1$) then state is separable. If it gives a mixed state ($\mathrm{tr}(\rho^2) < 1$), state is entangled.
    - $\mathrm{tr}(\rho_A^2) = \mathrm{tr}(\rho_B^2)$ gives measure of entanglement, with max value 1 for no entanglement, min value 1/2 (for single qubit subsystem) for maximally entangled states.
- **Bell states**: for $x, y \in \{0, 1\}$,

$$|\beta_{xy}\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |y\rangle + (-1)^x |1\rangle \otimes |\overline{y}\rangle)$$

  where $\overline{0} = 1$, $\overline{1} = 0$.
- Bell states are maximally entangled (trace of reduced density matrix of both sides is $\frac{1}{2}$) and form an orthonormal basis.
- Bell state basis is related to standard basis by unitary transformation, but Bell states can't be created from the separable standard basis by any LOCC process, since the unitary transformations between them are not of form $\hat{U}_A \otimes \hat{U}_B$ (since this preserves separability), and measurements always produce a separable state.
- Alice and Bob can individually transform any Bell state to any other Bell state by the unitary operators $\hat{U}_{xy} \otimes \hat{I}$ and $\hat{I} \otimes \hat{U}_{xy}$ respectively:

$$\left(\hat{U}_{xy} \otimes \hat{I}\right)|\beta_{00}\rangle = \left(\hat{I} \otimes \hat{U}_{xy}\right)|\beta_{00}\rangle = |\beta_{xy}\rangle$$

  where

$$\begin{aligned} U_{00} &:= I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad U_{01} := \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\ U_{10} &:= \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad U_{11} := i\sigma_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \end{aligned}$$

- **Example**: $\hat{U}_{11} = i\widehat{\sigma_2}$, so $\hat{U}_{11}|0\rangle = -|1\rangle$, $\hat{U}_{11}|1\rangle = |0\rangle$, and

$$\begin{aligned} \left(\hat{U}_{11} \otimes \hat{I}\right)|\beta_{00}\rangle &= \left(\hat{U}_{11} \otimes I\right)\left(\frac{1}{\sqrt{2}}|0\rangle \otimes |0\rangle + \frac{1}{\sqrt{2}}|1\rangle \otimes |1\rangle\right) \\ &= \left(-\frac{1}{\sqrt{2}}|1\rangle \otimes |0\rangle + \frac{1}{\sqrt{2}}|0\rangle \otimes |1\rangle\right) = |\beta_{11}\rangle \end{aligned}$$

## 5.2. Superdense coding

- Superdense coding allows one qubit to transmit two classical bits of information.
- Qubit can be used instead of classical bit: $|0\rangle$ corresponds to the bit 0, $|1\rangle$ corresponds to the bit 1. In this case, the qubit can be measured with probability 1 with the measurement operator $\frac{1}{2}(I_2 - \sigma_3)$, since

$$\frac{1}{2}(I_2 - \sigma_3)|0\rangle = 0|0\rangle, \quad \frac{1}{2}(I_2 - \sigma_3)|1\rangle = 1|1\rangle$$

  so measurement with outcome 0 means state is $|0\rangle$ with probability 1, measurement with outcome 1 means state is $|1\rangle$ with probability 1.

- Alice can prepare the qubit to represent the classical bit to send to Bob: prepare any state $|\psi\rangle$ and measure on it with operator $\frac{1}{2}(I_2 - \sigma_3)$. Outcome is 0 or 1 - if outcome is equal to the bit $x$ she wants to send, $|\psi\rangle$ has been projected to $|x\rangle$, so send this state to Bob. Otherwise, perform unitary transformation $\sigma_1|\overline{x}\rangle = |x\rangle$ and send this state to Bob.
- **Superdense coding**:
  - Alice and Bob share state $|\beta_{00}\rangle$.
  - Alice applies operation $\hat{U}_{xy} \otimes \hat{I}$ to whole system where $(xy)_2$ is the two bit message she wants to send (this just acts on her qubit). Note that this does not transmit any information to Bob, as his reduced density matrix is $\rho_B = \frac{1}{2}I$ before and after the transformation.
  - Alice sends her qubit to Bob. Then Bob has the full Bell state $|\beta_{xy}\rangle$ (he has both qubits). Bob then applies a measurement which has the four Bell states as eigenstates, which gives him the eigenvalue with probability 1, e.g. he measures

$$\hat{B} = 0|\beta_{00}\rangle\langle\beta_{00}| + 1|\beta_{01}\rangle\langle\beta_{01}| + 2|\beta_{10}\rangle\langle\beta_{10}| + 3|\beta_{11}\rangle\langle\beta_{11}|$$

# 6. Information theory

## 6.1. Classical information and Shannon entropy

- **Definition**: **Shannon entropy** is

$$H(X) := -\sum_x p(x)\log_2(p(x))$$

- **Definition**: **joint entropy** is

$$H(X,Y) := -\sum_{x,y} p(x,y)\log_2(p(x,y))$$

- **Proposition**: joint entropy obeys **subadditivity**:

$$H(X,Y) \leq H(X) + H(Y)$$

  with equality iff $X$ and $Y$ are independent variables, i.e. when $p(x,y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$.
- **Definition**: **relative entropy** is defined for two random variables which take same values but with different distributions $p(x)$ and $q(x)$:

$$\begin{aligned} H(p(x) \parallel q(x)) &:= \sum_x (p(x)\log_2(p(x)) - p(x)\log_2(q(x))) \\ &= -H(X) - \sum_x p(x)\log_2(q(x)) \end{aligned}$$

- **Proposition**: relative entropy is non-negative and

$$H(p(x) \parallel q(x)) = 0 \Longleftrightarrow \forall x, p(x) = q(x)$$

- **Remark**: relative entropy can diverge if for some $x$, $q(x) = 0$ and $p(x) \neq 0$
- **Definition**: **conditional entropy** is

$$H(X|Y) \coloneqq H(X,Y) - H(Y)$$

- **Definition**: **mutual information** of $X$ and $Y$ is

$$H(X : Y) = H(X : Y) \coloneqq H(X) + H(Y) - H(X,Y) \geq 0$$

## 6.2. Quantum entropy

- **Definition**: **von Neumann entropy** of quantum state with density operator $\hat{\rho}$ is

$$S(\hat{\rho}) = -\operatorname{tr}(\hat{\rho}\log_2(\hat{\rho})) = -\sum_i p_i \log_2(p_i)$$

  where $\log_2$ is applied entry-wise to $\hat{\rho}$ and $\hat{\rho} = \sum_i p_i |i\rangle\langle i|$.
- **Remark**: for pure state, $S(\hat{\rho}) = -1\log_2(1) = 0$.
- **Definition**: **relative entropy** is measure of distance between two mixed states:

$$S(\hat{\rho}_1 \parallel \hat{\rho}_2) \coloneqq \operatorname{tr}(\hat{\rho}_1 \log_2\left(\hat{\rho}_1\right)) - \operatorname{tr}(\hat{\rho}_1 \log_2\left(\hat{\rho}_2\right))$$

- **IGNORE SCMIDT DECOMPOSITION AND SCHMIDT NUMBER SECTION**