III. Ensembles & Density Operators

Often, we only have partial information about the state of a system.
Important to be able to describe situation in which we only know that system has some probability of being in a certain quantum state.

Definition (quantum ensemble)
The ensemble $\{ p_i, |\psi_i\rangle \}$ describes a system which has probability p_i of being in state $|\psi_i\rangle$, where:
$|\psi_i\rangle \in \mathcal{H}$, $0 \le p_i \le 1$, $\sum_i p_i = 1$.

Proposition 1
Measurement of the ensemble $\{ p_i, |\psi_i\rangle \}$ can be completely described in terms of the operator:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$$

ρ is called the density operator (or the density matrix) for the ensemble.

To prove this, we need:

Lemma (Cyclic property of trace)

For any state $|\psi\rangle \in \mathcal{H}$ and operator $X \in \mathcal{B}(\mathcal{H})$,

$$\langle\psi|X|\psi\rangle = \operatorname{tr}[X\,|\psi\rangle\langle\psi|].$$

Proof (for finite dimensional Hilbert spaces)

Let $\{|i\rangle\}$ be orthonormal basis for $\mathcal{H}$, so that $|\psi\rangle = \sum_i \psi_i\,|i\rangle$, $X = \sum_{i,j} x_{ij}\,|i\rangle\langle j|$. Then

$$\begin{aligned}
\langle\psi|X|\psi\rangle &= \Big(\sum_i \bar{\psi}_i \langle i|\Big)\Big(\sum_{j,k} x_{jk}\,|j\rangle\langle k|\Big)\Big(\sum_\ell \psi_\ell\,|\ell\rangle\Big) \\
&= \sum_{i,j,k,\ell} \bar{\psi}_i\, x_{jk}\, \psi_\ell\, \underbrace{\langle i|j\rangle}_{=\delta_{ij}}\underbrace{\langle k|\ell\rangle}_{=\delta_{k\ell}} \\
&= \sum_{ik} \bar{\psi}_i\, x_{ik}\, \psi_k \\
&= \sum_{ik} x_{ik}\, \bar{\psi}_i\, \psi_k
\end{aligned}$$

$$\begin{aligned}
\operatorname{tr}[X\,|\psi\rangle\langle\psi|] &= \operatorname{tr}\Big[\Big(\sum_{ij} x_{ij}\,|i\rangle\langle j|\Big)\Big(\sum_k \psi_k\,|k\rangle\Big)\Big(\sum_\ell \bar{\psi}_\ell \langle\ell|\Big)\Big] \\
&= \sum_{ijk\ell} x_{ij}\,\psi_k\,\bar{\psi}_\ell\,\delta_{jk}\,\underbrace{\operatorname{tr}[\,|i\rangle\langle\ell|\,]}_{=\delta_{i\ell}} \\
&= \sum_{ik} x_{ik}\,\psi_k\,\bar{\psi}_i \\
&= \langle\psi|X|\psi\rangle
\end{aligned}$$

Lemma (Further cyclic property of trace)

$\operatorname{tr}[AB] = \operatorname{tr}[BA]$

Proof: Exercise!

Proof of Proposition 1

Consider measurement $\{M_m\}$ on ensemble $\{p_i, |\psi_i\rangle\}$.

$$\Pr(\text{outcome } m) = \sum_i p_i \Pr(m \mid |\psi_i\rangle)$$

$$= \sum_i p_i \langle\psi_i|M_m|\psi_i\rangle \quad \text{Postulate 3}$$

$$= \sum_i p_i \operatorname{tr}[M_m |\psi_i\rangle\langle\psi_i|] \quad \text{Lemma}$$

$$= \operatorname{tr}[\textstyle\sum_i p_i |\psi_i\rangle\langle\psi_i|] \quad \text{Linearity of tr}$$

So probability of <u>any</u> measurement outcome is given by $\operatorname{tr}[M_m \rho]$ where $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ is the density operator.

Remains to show that we can determine post-measurement density operator from ρ alone (i.e. don't need full description $\{p_i, |\psi_i\rangle\}$ of ensemble).

With probability p_i, initial system state was $|\psi_i\rangle$, so post-measurement state given outcome m is:

$$|\psi_i'\rangle = \frac{M_m |\psi_i\rangle}{\| M_m |\psi_i\rangle \|} \quad \text{by Postulate 3.}$$

$$= \frac{M_m |\psi_i\rangle}{\sqrt{\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle}}$$

But be careful!

Probability of state $|\psi_i'\rangle$ given outcom m is <u>not</u> p_i. Measurement outcome gives us some information about which state $|\psi_i\rangle$ system was in initially.

E.g. for ensemble $\left\{ \left(p_0 = \tfrac{1}{2}, |0\rangle\right), \left(p_1 = \tfrac{1}{2}, \frac{|1\rangle + \varepsilon|0\rangle}{\sqrt{1+\varepsilon}}\right) \right\}$

and measurement $\{ |0\rangle\langle 0|, |1\rangle\langle 1| \}$

if get outcome 0, probability initial state was $|0\rangle$ is very high for small ε.

Post-measurement density operator given outcome m is:

$$\rho' = \sum_i Pr\left(|\psi_i\rangle \,|\, m\right) \, |\psi_i'\rangle\langle\psi_i'|$$

$$= \sum_i \frac{Pr(m \,|\, |\psi_i\rangle) \, Pr(|\psi_i\rangle)}{Pr(m)} \cdot |\psi_i'\rangle\langle\psi_i'| \qquad \text{Bayes' rule}$$

$$= \sum_i \frac{\cancel{\langle\psi_i| M_m^\dagger M_m |\psi_i\rangle} \, p_i}{\sum_j p_j \langle\psi_j| M_m^\dagger M_m |\psi_j\rangle} \cdot \frac{M_m |\psi_i\rangle\langle\psi_i| M_m^\dagger}{\cancel{\langle\psi_i| M_m^\dagger M_m |\psi_i\rangle}}$$

$$= \frac{M_m \left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) M_m^\dagger}{\sum_j p_j \, tr\left[M_m |\psi_j\rangle\langle\psi_j| M_m^\dagger\right]} \qquad \text{linearity + Lemmas}$$

$$= \frac{M_m \, \rho \, M_m^\dagger}{tr\left[M_m \rho M_m^\dagger\right]}$$

so we've recovered ρ' from ρ, as required. $\square$

Proposition 2

If system evolves according to unitary evolution operator U, the density matrix ρ for the ensemble $\{p_i, |\psi_i\rangle\}$ transforms as

$$\rho \longrightarrow \rho' = U \rho U.$$

Proof

With probability p_i, system is initially in state $|\psi_i\rangle$, which evolves to $U|\psi_i\rangle$ (Postulate 2) so ensemble evolves to $\{p_i, U|\psi_i\rangle\}$ giving

$$\rho' = \sum_i p_i \, U \, |\psi_i\rangle\langle\psi_i| \, U^\dagger$$

$$= U \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \, U^\dagger = U \rho U^\dagger \quad \square$$

The only way we can gain information about a quantum system is to measure it. So the only physical properties of a system are measurement probabilities.

Since these are completely described by the density operator, and the effect of both evolution and measurement on the density operator can be described in terms of the density operator alone, the density operator <u>is</u> a <u>full</u> physical description of the state of the system.

We can therefore reformulate the postulates of QM equivalently in terms of density operators.

Definition (density operator)

A density operator ρ is any operator in $\mathcal{B}(\mathcal{H})$ satisfying:

(i) $\rho \geqslant 0$ (positive $\Rightarrow$ implicitly $\rho^\dagger = \rho$)

(ii) $\mathrm{tr}(\rho) = 1$ (trace 1)

Postulate 1'

The state of a quantum system is described by a density operator $\rho \in \mathcal{B}(\mathcal{H})$.

Postulate 2'

The density operator evolves according to

$$\rho(t) = U_t \, \rho \, U_t^\dagger$$

where $U_t = e^{-iHt}$ is unitary.

Postulate 3'

A measurement is described a set of operators $\{M_m\}$ satisfying completeness: $\sum_m M_m^\dagger M_m = \mathbb{1}$ (as before).

The probability of obtaining outcome m on state ρ is

$$\Pr(m) = \operatorname{tr}[M_m \rho]$$

and the post-measurement state given outcome m is

$$\rho' = \frac{M_m \rho M_m^\dagger}{\operatorname{tr}[M_m \rho M_m^\dagger]}.$$

Note that if $\rho \geq 0$, $\operatorname{tr}[\rho] = 1$, then

$U \rho U^\dagger \geq 0$ (U unitary, so preserves spectrum)

$\operatorname{tr}[U \rho U^\dagger] = \operatorname{tr}[\rho U^\dagger U] = \operatorname{tr}[\rho \cdot \mathbb{1}] = 1.$

so evolution transforms density operators into density operators, as required, and

$$\operatorname{tr}\left[\frac{M_m \rho M_m^\dagger}{\operatorname{tr}[M_m \rho M_m^\dagger]}\right] = \frac{\operatorname{tr}[M_m \rho M_m^\dagger]}{\operatorname{tr}[M_m \rho M_m^\dagger]} = 1$$

[What if $\operatorname{tr}[M_m \rho M_m^\dagger] = 0$? Exercise: think about it!]

$\forall |\psi\rangle: \langle \psi | M_m \rho M_m^\dagger | \psi \rangle = \langle \psi' | \rho | \psi' \rangle \geq 0$

so $\dfrac{M_m \rho M_m^\dagger}{\operatorname{tr}[M_m \rho M_m^\dagger]} \geq 0,$

thus measurement also transforms density ops. to density ops., as required.

We have already seen that any ensemble $\{p_i, |\psi_i\rangle\}$ defines a density operator $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Conversely,

Lemma

Any density operator ρ can be realised by some ensemble.

Proof

Consider spectral decomposition $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$.

$$1 = \mathrm{tr}\,\rho = \mathrm{tr}\left[\sum_i \lambda_i |\psi_i\rangle\langle\psi_i|\right]$$

$$= \sum_i \lambda_i \,\mathrm{tr}\left[|\psi_i\rangle\langle\psi_i|\right] = \sum_i \lambda_i \langle\psi_i|\psi_i\rangle$$

$$= \sum_i \lambda_i$$

$\therefore$ λ_i form a probability distribution, and ensemble $\{\lambda_i, |\psi_i\rangle\}$ reaslises ρ. $\square$

It is very important to note that this ensemble is not the only one realising ρ, and there is nothing special about this particular ensemble!

E.g. $\{(\frac{3}{4}, |0\rangle), (\frac{1}{4}, |1\rangle)\}$ and $\{(\frac{1}{2}, |\psi\rangle), (\frac{1}{2}, |\varphi\rangle)\}$

where $|\psi\rangle = \sqrt{\frac{3}{4}}|0\rangle + \sqrt{\frac{1}{4}}|1\rangle$

$|\varphi\rangle = \sqrt{\frac{3}{4}}|0\rangle - \sqrt{\frac{1}{4}}|1\rangle$

both have the same density matrix (Exercise: check!)

Definition (pure state, mixed state)
We say ρ is a **pure state** if $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle$.
A state that is not pure is a **mixed state**.

A pure state $\rho = |\psi\rangle\langle\psi|$ **does** correspond to a unique ensemble, namely $\{p=1, |\psi\rangle\}$.
(This is obvious. Exercise: prove it!)

Lemma
$tr(\rho^2) \leq 1$, with equality iff ρ is pure

Proof. Exercise

IV. Taking systems apart: Reduced states & the partial trace

If we have a state ρ_{AB} on a joint Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, have seen that measurement on A alone is described by measurement operators $\{ M_m \otimes \mathbb{1} \}$.

Can we describe measurement in terms of the A system alone?

Note that ρ_{AB} can be expanded in a basis as $\rho = \rho_{ijkl} |i\rangle\langle k| \otimes |j\rangle\langle l|$. Then

$$\begin{aligned}
Pr(m) &= tr\left[M_m \otimes \mathbb{1} \, \rho_{AB} \right] \\
&= tr\left[M_m \otimes \mathbb{1} \sum_{ijkl} \rho_{ijkl} |i\rangle\langle k| \otimes |j\rangle\langle l| \right] \\
&= \sum_{ijkl} \rho_{ijkl} \, tr\left[M_m |i\rangle\langle k| \otimes |j\rangle\langle l| \right] \\
&= \sum \rho_{ijkl} \, tr\left[M_m |i\rangle\langle k| \right] tr\left[|j\rangle\langle l| \right] \\
&= tr\left[M_m \underbrace{\sum_{ijkl} \rho_{ijkl} |i\rangle\langle k| \, tr\left[|j\rangle\langle l| \right]}_{:= \rho_A} \right]
\end{aligned}$$

Motivates...

Definition (partial trace)

tr_B is linear operator defined by

$$tr_B\left[|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2| \right] := |a_1\rangle\langle a_2| \, tr\left[|b_1\rangle\langle b_2| \right]$$

(i.e. extend by linearity to any operator on $\mathcal{H}_A \otimes \mathcal{H}_B$)

Is this unique?

<u>Theorem</u> (uniqueness of partial trace)

$\mathrm{tr}_B : \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B) \longrightarrow \mathcal{B}(\mathcal{H}_A)$

is the unique linear operator satisfying

$$\mathrm{tr}[M \, \mathrm{tr}_B(X_{AB})] = \mathrm{tr}[M \otimes \mathbb{1} \cdot X_{AB}]$$

To prove this, we first need the following very import fact: $\mathcal{B}(\mathcal{H})$ can itself be made into a Hilbert space.

<u>Definition</u> (Hilbert - Schmidt inner product)

The Hilbert - Schmidt inner product $\langle X, Y \rangle$ on $\mathcal{B}(\mathcal{H})$ is defined by

$$\langle X, Y \rangle := \mathrm{tr}[X^\dagger Y]$$

Exercise: verify that this does indeed define an inner product.

The Hilbert - Schmidt inner product turns the vector space $\mathcal{B}(\mathcal{H})$ into a Hilbert space.

Thus there exist orthogonal (with respect to H-S inner product) bases $\{A_i\}$ of operators, and any operator $X \in \mathcal{B}(\mathcal{H})$ can be expanded

$$X = A_i \; \mathrm{tr}[A_i^\dagger X]$$

Proof of uniqueness theorem

Let f, g be two maps

$$f, g : \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B) \longrightarrow \mathcal{B}(\mathcal{H}_A)$$

satisfying the property in the theorem.

Expanding in an orthonormal operator basis $\{A_i\}$ for $\mathcal{B}(\mathcal{H}_A)$, we have

$$\begin{aligned} f(X_{AB}) &= \sum_i A_i \; \mathrm{tr}[A_i^\dagger f(X_{AB})] \\ &= \sum_i A_i \; \mathrm{tr}[A_i^\dagger \otimes \mathbb{1} \; X_{AB}] \qquad \text{by assumption} \\ &= \sum_i A_i \; \mathrm{tr}[A_i^\dagger \; g(X_{AB})] \qquad \text{"} \\ &= g(X_{AB}) \qquad \square \end{aligned}$$

Definition (reduced density operator)

$\rho_A := \mathrm{tr}_B(\rho_{AB})$ is the <u>reduced density operator</u> on A of the state ρ_{AB}.

Lemma

Let $\{M_m \otimes \mathbb{1}\}$ be a measurement on A alone
ρ_{AB} be the state of a joint system AB.
The reduced state ρ_A' of a post-measurement state is fully determined by the reduced state $\rho_A := \mathrm{tr}_B[\rho_{AB}]$:

$$\rho_A' = \frac{M_m \rho_A M_m^\dagger}{\mathrm{tr}[M_m^\dagger M_m \rho_A]}$$

Proof: Exercise

Thus the reduced state ρ_A fully determines the probabilities of measurement outcomes for all measurements of the A system alone.

Example:

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|0\rangle|0\rangle + |1\rangle|1\rangle\right)$$

$$\rho_{AB} = |\psi\rangle\langle\psi| \qquad \text{(a } \underline{\text{pure state}}\text{)}$$

$$\rho_A = \frac{1}{2}\,\mathrm{tr}_B\left[\left(|0\rangle|0\rangle + |1\rangle|1\rangle\right)\left(\langle 0|\langle 0| + \langle 1|\langle 1|\right)\right]$$

$$= \frac{1}{2}\Big(|0\rangle\langle 0|\,\overbrace{\mathrm{tr}(|0\rangle\langle 0|)}^{=\langle 0|0\rangle = 1} + |1\rangle\langle 0|\,\overbrace{\mathrm{tr}(|1\rangle\langle 0|)}^{=\langle 0|1\rangle = 0}$$

$$+ |0\rangle\langle 1|\,\underbrace{\mathrm{tr}(|0\rangle\langle 1|)}_{=0} + |1\rangle\langle 1|\,\underbrace{\mathrm{tr}(|1\rangle\langle 1|)}_{=1}\Big)$$

$$= \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \mathbb{1}/2 \qquad \text{(a } \underline{\text{mixed state}}\text{)}$$

Even if the joint system is in a <u>pure state</u> $\rho_{AB} = |\psi\rangle\langle\psi|$, so that the state is known to be exactly $|\psi\rangle$ with certainty, the reduced states can nonetheless be <u>mixed</u>, implying that the of the states of the individual subsystems are <u>not</u> known with certainty!

Another manifestation of the strange properties of entanglement...

V. A brief introduction to entropy

If we have an ensemble $\{p_i, |\psi_i\rangle\}$, and someone tells us the system is in fact in state $|\psi_i\rangle$, how much information have we gained?

More generally, if an event E occurs in a probabilistic experiment, how much information does this provide?

Want to quantify this using an "information content" function $I(E)$.

Some reasonable requirements on $I(E)$ are:

(i) $I(E)$ is a function only of the probability that E occurs.

(ii) I is a smooth function of probability.

(iii) If two independent events E & F occur, the total information gained should be the sum of the information gained from each event.

(iv) The more likely E is, the less information we gain by discovering that it occurred.
(If $Pr(E)=1$, then the fact that E occurred tells us nothing new!)
so I should be a decreasing function of $Pr(E)$.

Lemma

If $I(E)$ satisfies (i) – (iv), then

$$I(E) = I(p) = -k \log p$$

for some constant $k > 0$, where $p = Pr(E)$.

Proof (sketch)

From (i), must have $I(E) = I(Pr(E)) \equiv I(p)$.

From (iii),

$$\begin{array}{ccc} I(E \cap F) & = & I(E) + I(F) \\ \| & & \| \\ I(pq) & & I(p) + I(q) \end{array}$$

so $I(pq) = I(p) + I(q)$.

Together with (ii), this implies $I(b^x) = x\, I(b)$.

Let $f(p) = \dfrac{I(p)}{I(b)}$ for fixed b.

Then $f(b^x) = \dfrac{I(b^x)}{I(b)} = \dfrac{x\, I(b)}{I(b)} = x$

so $f(p) = \log_b p$.

Thus $I(p) = I(b) \log_b p = c \log p$.

From (iv), $c < 0$. $\square$

Let some event occur from a set of mutually exclusive events $\{E_i\}$ with probabilities $\{p_i\}$ ($\sum_i p_i = 1$ since they're mutually exclusive).

Average information gain is then:

$$\sum_i p_i \; I(p_i) \;=\; -k \sum_i p_i \log p_i$$

This motivates the following definition of the information content of a random variable:

<u>Definition</u> (Shannon entropy)

The <u>Shannon entropy</u> of a probability distribution $X = \{p_1, p_2, \ldots, p_n\}$ is

$$H(X) := -\sum_i p_i \log p_i \, .$$

(where we define $0 \log 0 := 0$).

For a quantum ensemble $X = \{p_i, |\psi_i\rangle\}$, we have $H(X) = -\sum_i p_i \log p_i$.

How can we extend this to density matrices? Recall that the density matrix $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ (spectral decomposition) is realised by the ensemble $\{\lambda_i, |\psi_i\rangle\}$.

This ensemble is not unique, but it <u>is</u> the unique ensemble for which all $|\psi_i\rangle$ are mutually orthogonal, which is significant in this context because...

Lemma

Two states $|\psi_1\rangle, |\psi_2\rangle$ can be distinguished perfectly (i.e. $\exists$ measurement $\{M_1, M_2\}$ such that

$$\Pr(\text{outcome } i) = \begin{cases} 1 & \text{if state is } |\psi_i\rangle \\ 0 & \text{otherwise} \end{cases} \quad)$$

iff $|\psi_1\rangle, |\psi_2\rangle$ are orthogonal (i.e. $\langle\psi_1|\psi_2\rangle = 0$).

Proof: Exercise.

This motivates the following definition of the entropy of a density operator:

Definition (von Neumann entropy)

The <u>von Neumann entropy</u> of density operator $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ is

$$S(\rho) := -\sum_i \lambda_i \log \lambda_i = -\operatorname{tr}(\rho \log \rho).$$

Note: there are much better justifications for these definitions of entropy, in terms of data compression, but sadly that's beyond the scope of this course :-(

A nice application of von Neumann entropy:

How can we tell if a bipartite state is entangled?

E.g. is $\frac{1}{2}(|00\rangle + i|01\rangle + i|10\rangle + |11\rangle)$ entangled or not?

Lemma

$|\psi\rangle_{AB}$ is entangled iff $S(\rho_A), S(\rho_B) > 0$, where the reduced states $\rho_{A,B} = tr_{B,A}(|\psi\rangle\langle\psi|)$.

Proof

Recall $|\psi\rangle$ is a product state (i.e. not entangled) if it can be written

$|\psi\rangle = |a\rangle|b\rangle$.

In this case,

$\rho_A = |a\rangle\langle a| \, tr[|b\rangle\langle b|] = |a\rangle\langle a|$

only has one non-zero eigval $\lambda = 1$, and

$S(\rho_A) = \sum_i \lambda_i \log \lambda_i = 1 \cdot \log 1 = 0$.

Conversely, if $S(\rho_A) = 0$ then ρ_A can have only one non-zero eigval $\lambda = 1$ (Exercise: convince yourself of this), so

$\rho_A = |a\rangle\langle a|$.

Similarly, $\rho_B = |b\rangle\langle b|$.

But $|\psi\rangle = |a\rangle|b\rangle$ is the unique state with these reduced states (Exercise: prove this!) $\square$