11. Von Neumann Entropy

1. Motivation

- *Recall*: A density operator $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ that characterizes an ensemble $\{|\psi_i\rangle, p_i\}$ of quantum vector states is the correlate of a Gibbs distribution $\rho(x, t)$ that characterizes an ensemble of classical states.

- *And*: The Gibbs entropy is given by $S_{\text{Gibbs}}(\rho) = -k\int_\Gamma \rho(x, t) \ln \rho(x, t) dx$

Recall: ρ in QM is an operator; $\rho(x, t)$ in classical Stat Mech is a function.

The ensemble average of $-k \ln \rho(x, t)$

What is the quantum correlate of S_{Gibbs}?

Def. 1 (*von Neumann entropy*). The **von Neumann entropy** $S_{\mathrm{vN}}(\rho)$ of a density operator state ρ is defined by

$$S_{\mathrm{vN}}(\rho) \equiv -\mathrm{Tr}(\rho \ln \rho)$$

John von Neumann (1903-1957)

- *Note* 1: What is "$\ln \rho$"?

We know what $\ln x$ does to a number x, but what does it do to an operator?

- Let B be an operator such that $B|\psi_i\rangle = b_i|\psi_i\rangle$.
- Let $f(x)$ be a function on the real numbers.
- Then $f(B)|\psi_i\rangle \equiv f(b_i)|\psi_i\rangle$

If B is an operator with eigenvalues b_i, then $f(B)$ is an operator with eigenvalues $f(b_i)$.

- *Recall*: Can always express ρ as $\rho = \sum_i \lambda_i |\phi_i\rangle\langle\phi_i|$, where $\rho|\phi_i\rangle = \lambda_i|\phi_i\rangle$

- ρ is an operator with eigenvalues λ_i
- $\{|\phi_i\rangle\}$ is a basis for $\mathcal{H}$

- *So*: $\ln \rho |\phi_i\rangle = \ln \lambda_i |\phi_i\rangle$

$\ln \rho$ is an operator with eigenvalues $\ln \lambda_i$.

Def. 1 (*von Neumann entropy*). The **von Neumann entropy** $S_{\mathrm{vN}}(\rho)$ of a density operator state ρ is defined by

$$S_{\mathrm{vN}}(\rho) \equiv -\mathrm{Tr}(\rho \ln \rho)$$

John von Neumann
(1903-1957)

- *Note* 2: If ρ is a density operator state on an n-dim vector space $\mathcal{H}$, then the maximum value of $S_{\mathrm{vN}}(\rho)$ is $\ln n$.

  - *Proof*: Note first:

    $\{|\phi_i\rangle\}$ is an eigenvector basis of ρ

    $$S_{\mathrm{vN}}(\rho) = -\mathrm{Tr}(\rho \ln \rho) = -\sum_{i=1}^{n} \langle \phi_i | (\rho \ln \rho) | \phi_i \rangle = -\sum_{i=1}^{n} \lambda_i \ln \lambda_i$$

  - Now recall $(d/d\lambda_i) S_{\mathrm{vN}}(\lambda_i^*) = 0$ for the λ_i^* that maximizes S_{vN}:

    $$dS_{\mathrm{vN}}(\lambda_i^*) = -\sum_{i=1}^{n} (\ln \lambda_i^* - \alpha) d\lambda_i = 0$$

    constrained by $\sum_{i=1}^{n} d\lambda_i = 0$

  - Solve for λ_i^*:

    $$\lambda_i^* = e^{\alpha} \quad \Rightarrow \quad \sum_{i=1}^{n} \lambda_i = n e^{\alpha} = 1 \quad \Rightarrow \quad \alpha = \ln(1/n) \quad \Rightarrow \quad \lambda_i^* = (1/n)$$

  - *Thus*: $S_{\mathrm{vN}}(\lambda_i^*) = -\sum_{i=1}^{n} (1/n) \ln(1/n) = -n(1/n)\ln(1/n) = \ln n$

Def. 1 (*von Neumann entropy*). The **von Neumann entropy** $S_{\text{vN}}(\rho)$ of a density operator state ρ is defined by

$$S_{\text{vN}}(\rho) \equiv -\text{Tr}(\rho \ln \rho)$$

John von Neumann
(1903-1957)

- *Note 3*: Recall $\langle O \rangle_\rho = \text{Tr}(\rho O)$ is the expectation value of the operator O in the density operator state ρ.
  - *So*: $S_{\text{vN}}(\rho)$ is the expectation value of the operator $-\ln \rho$ in the density operator state ρ.

Analogous to ensemble average in Gibbs approach!

What does $S_{\mathrm{vN}}(\rho)$ measure?

Def. 2 (*Maximally mixed density operator state*). Let ρ be a density operator on an n-dim vector space $\mathcal{H}$ with identify operator I_n. Then ρ is **maximally mixed** just when it can be expressed by $\rho = (1/n)I_n$.

- *Recall*: A mixed density operator can be expressed by $\rho = \sum_{i=1}^{n} \lambda_i |\phi_i\rangle\langle\phi_i|$ where the $|\phi_i\rangle$ form a basis for $\mathcal{H}$, and $\sum_{i=1}^{n} \lambda_i = 1$.
  - *And*: ρ is pure just when one λ_i is 1 and all the rest are 0.
- *Intuition*: The more mixed ρ is, the "farther away" it is from the pure case.
  - *And*: The maximum "distance" ρ is from the pure case occurs when all the λ_i are equal; i.e., when $\lambda_i = 1/n$.
- *So*:

$$\begin{aligned} \rho_{\max} &= \sum_{i=1}^{n} (1/n) |\phi_i\rangle\langle\phi_i| \\ &= (1/n) \sum_{i=1}^{n} |\phi_i\rangle\langle\phi_i| \\ &= (1/n) I_n \end{aligned}$$

$S_{\mathrm{vN}}(\rho)$ *is a measure of the degree to which* ρ *is mixed!*

Claim 1. Let ρ be a density operator on an n-dim vector space $\mathcal{H}$. Then $S_{\mathrm{vN}}(\rho)$ varies from zero, if ρ is a pure density operator state, to $\ln n$, if ρ is a maximally mixed density operator state.

- *Proof*: Suppose ρ is a pure density operator state.

  - *Then*: $S_{\mathrm{vN}}(\rho) = -\sum_{i=1}^{n} \lambda_i \ln \lambda_i = -\ln(1) = 0$

- Now Suppose ρ is a maximally mixed density operator state.

  - *Then*: $S_{\mathrm{vN}}(\rho) = -\sum_{i=1}^{n} \lambda_i \ln \lambda_i = -\sum_{i=1}^{n} (1/n) \ln (1/n)$

$$= -\ln(1/n)$$

$$= \ln n$$

The maximum value of S_{vN}*!*

2. Comparison with S_{Shan}

$S_{\text{vN}}(\rho)$ as a measure of information compression?

- The *Shannon Entropy*:

$$S_{\text{Shan}}(X) = -\sum_i p_i \log_2 p_i$$

Specifies the minimal number of bits required to encode the output of a classical information source.

  - $X = \{x_1, ..., x_n\}$, where x_i is a state produced by a classical information source, and p_i is a probability distribution over such states.

- The *von Neumann Entropy*:

$$S_{\text{vN}}(\rho) = -\sum_i p_i \ln p_i$$

Specifies the minimal number of qubits required to encode the output of a quantum information source.

  - $\rho = -\sum_i p_i |\psi_i\rangle\langle\psi_i|$,where $|\psi_i\rangle$ is a vector state produced by a quantum information source, and p_i is a probability distribution over such states.

$S_{\text{vN}}(\rho)$ *as a measure of uncertainty?*

- The *Shannon Entropy*:

$$S_{\text{Shan}}(X) = -\sum_i p_i \log_2 p_i$$

- $X = \{x_1, ..., x_n\}$, where x_i is a state produced by a classical information source, and p_i is a probability distribution over such states.

Let $-\log p_i$ be "info gained" upon finding a system to be in a state drawn from a set of states with probabilites p_i.

Expected value of information gained upon measurement of X with outcome x_i.

- The *von Neumann Entropy*:

$$S_{\text{vN}}(\rho) = -\sum_i p_i \ln p_i$$

- $\rho = -\sum_i p_i |\psi_i\rangle\langle\psi_i|$,where $|\psi_i\rangle$ is a vector state produced by a quantum information source, and p_i is a probability distribution over such states.

Expected value of information gained upon measurement of ρ with outcome $|\psi_i\rangle$.

But: $S_{\text{vN}}(\rho)$ measures mixedness of ρ.

And: Mixedness does not necessarily entail uncertainty...

3. Entanglement Entropy

Def. 3 (*Partial trace*). Let $O_A \otimes O_B$ be an operator on $\mathcal{H}_A \otimes \mathcal{H}_B$, and let $\{|w_{Bi}\rangle\}$ be a basis of $\mathcal{H}_B$. The **partial trace** $\mathrm{Tr}_B(O_A \otimes O_B)$ of $O_A \otimes O_B$ over $\mathcal{H}_B$ is defined by

$$\mathrm{Tr}_B(O_A \otimes O_B) \equiv \sum_i \langle w_{Bi}|O_A \otimes O_B|w_{Bi}\rangle$$

Traces out the degrees of freedom of subsystem B.

- <u>*Note*</u>. $\mathrm{Tr}_B(O_A \otimes O_B) = \sum_i \langle w_{Bi}|O_A \otimes O_B|w_{Bi}\rangle$

$$= \sum_i O_A \langle w_{Bi}|O_B|w_{Bi}\rangle$$

$$= O_A \mathrm{Tr}(O_B)$$

$\mathrm{Tr}(O_B)$ is a number, so $O_A \mathrm{Tr}(O_B)$ is an operator on $\mathcal{H}_A$!

Def. 4 (*Reduced density operator*). Let ρ_{AB} be a density operator for a bipartite system AB with subsystems A and B. The **reduced density operator** for subsystem A is defined by $\rho_A \equiv \mathrm{Tr}_B(\rho_{AB})$.

ρ_A is the A part of ρ_{AB}!

Example. Consider the entangled vector state $|\psi_{AB}\rangle = \sqrt{½}\{|0_A 0_B\rangle + |1_A 1_B\rangle\}$.

$$\begin{aligned}\rho_{AB} &= |\psi_{AB}\rangle\langle\psi_{AB}| \\ &= ½\{|0_A 0_B\rangle\langle 0_A 0_B| + |1_A 1_B\rangle\langle 0_A 0_B| + |0_A 0_B\rangle\langle 1_A 1_B| + |1_A 1_B\rangle\langle 1_A 1_B|\}\end{aligned}$$

$$\begin{aligned}\rho_A &= \mathrm{Tr}_B(\rho_{AB}) \\ &= \sum_{i=1}^{2}\langle w_{B_i}|\rho_{AB}|w_{B_i}\rangle \\ &= ½\langle 0_B|\{|0_A 0_B\rangle\langle 0_A 0_B| + |1_A 1_B\rangle\langle 0_A 0_B| + |0_A 0_B\rangle\langle 1_A 1_B| + |1_A 1_B\rangle\langle 1_A 1_B|\}|0_B\rangle \\ &\quad + ½\langle 1_B|\{|0_A 0_B\rangle\langle 0_A 0_B| + |1_A 1_B\rangle\langle 0_A 0_B| + |0_A 0_B\rangle\langle 1_A 1_B| + |1_A 1_B\rangle\langle 1_A 1_B|\}|1_B\rangle \\ &= ½\{|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|\} \\ &= ½ I_A\end{aligned}$$

Let $\{|w_{B_i}\rangle\}$ be the basis $\{|0_B\rangle, |1_B\rangle\}$

Def. 4 (*Reduced density operator*). Let ρ_{AB} be a density operator for a bipartite system AB with subsystems A and B. The **reduced density operator** for subsystem A is defined by $\rho_A \equiv \mathrm{Tr}_B(\rho_{AB})$.

ρ_A is the A part of ρ_{AB}!

Example. Consider the entangled vector state $|\psi_{AB}\rangle = \sqrt{½}\{|0_A 0_B\rangle + |1_A 1_B\rangle\}$.

$$\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$$
$$= ½\{|0_A 0_B\rangle\langle 0_A 0_B| + |1_A 1_B\rangle\langle 0_A 0_B| + |0_A 0_B\rangle\langle 1_A 1_B| + |1_A 1_B\rangle\langle 1_A 1_B|\}$$

pure state

$$\rho_A = \mathrm{Tr}_B(\rho_{AB})$$
$$= ½\{|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|\} = ½ I_A$$

mixed states

$$\rho_B = \mathrm{Tr}_A(\rho_{AB})$$
$$= ½\{|0_B\rangle\langle 0_B| + |1_B\rangle\langle 1_B|\} = ½ I_B$$

No ignorance interpretation of ρ_A or ρ_B:

- *An ignorance interpretation of ρ_A, ρ_B suggests subsystems A, B are either in vector states $|0\rangle$ or $|1\rangle$.*
- *But: This would entail ρ_{AB} is $½\{|0_A 0_B\rangle\langle 0_A 0_B| + |1_A 1_B\rangle\langle 1_A 1_B|\}$.*

Entanglement quantified!

Def. 5 (*Entanglement entropy*). For a bipartite system AB with density operator ρ_{AB}, the **entanglement entropy** S_A of subsystem A is defined by:

$$S_A \equiv S_{\mathrm{vN}}(\rho_A) = -\mathrm{Tr}(\rho_A \log \rho_A)$$

S_A is a measure of the degree to which the density operator ρ_A is mixed.

But what does this have to do with entanglement?...

Claim 2. Let $\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$ be a pure density operator state on a product vector space $\mathcal{H}_A \otimes \mathcal{H}_B$. Then $|\psi_{AB}\rangle$ is an entangled vector state *if and only if* $S_A > 0$ (i.e., ρ_A is mixed).

- *Proof*: "$\Leftarrow$". Suppose $|\psi_{AB}\rangle = |\varphi_A \phi_B\rangle$ is a product vector state, where $|\varphi_A\rangle \in \mathcal{H}_A$ and $|\phi_B\rangle \in \mathcal{H}_B$, and let $\{|w_{B_i}\rangle\}$ be a basis of $\mathcal{H}_B$.
  - *Then*: $\rho_A = \mathrm{Tr}_B(\rho_{AB}) = \sum_i \langle w_{B_i}|\rho_{AB}|w_{B_i}\rangle$

$$= \sum_i \langle w_{B_i}|\varphi_A \phi_B\rangle\langle \varphi_A \phi_B|w_{B_i}\rangle$$

$$= |\varphi_A\rangle\langle\varphi_A| \sum_i \langle\phi_B|w_{B_i}\rangle\langle w_{B_i}|\phi_B\rangle \quad \leftarrow \sum_i |w_{B_i}\rangle\langle w_{B_i}| = I_n$$

$$= |\varphi_A\rangle\langle\varphi_A|\langle\phi_B|\phi_B\rangle = |\varphi_A\rangle\langle\varphi_A|$$

  - *Thus*: ρ_A is pure (i.e., not mixed).

So: If ρ_A is mixed, then $|\psi_{AB}\rangle$ is a non-product (i.e., entangled) vector state.

Claim 2. Let $\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$ be a pure density operator state on a product vector space $\mathcal{H}_A \otimes \mathcal{H}_B$. Then $|\psi_{AB}\rangle$ is an entangled vector state *if and only if* $S_A > 0$ (i.e., ρ_A is mixed).

- <u>*Proof*</u>: "⇒". Let $\{|w_{A_i}\rangle\}, \{|w_{B_i}\rangle\}$ be bases for $\mathcal{H}_A$ and $\mathcal{H}_B$.

  Then $|\psi_{AB}\rangle$ can be written as $|\psi_{AB}\rangle = \sum_i \alpha_i |w_{A_i} w_{B_i}\rangle$.

  We want to show that if ρ_A is pure, then there's only one term in this "biorthogonal expansion".

  - <u>*Strategy*</u>: Let $\{|w_{A_i}\rangle\}$ be an eigenbasis of ρ_A.
  - <u>*Then*</u>: $\rho_A = \mathrm{Tr}_B(\rho_{AB}) = \sum_i \langle w_{B_i}|\rho_{AB}|w_{B_i}\rangle$

$$= \sum_i \langle w_{B_i}| \sum_j \alpha_j |w_{A_j} w_{B_j}\rangle \sum_k \alpha_k^* \langle w_{A_k} w_{B_k}|w_{B_i}\rangle$$

$$= \sum_{i,j,k} \alpha_j \alpha_k^* |w_{A_j}\rangle\langle w_{A_k}| \langle w_{B_i}|w_{B_j}\rangle\langle w_{B_k}|w_{B_i}\rangle$$

$$= \sum_{i,j} \alpha_j \alpha_i^* |w_{A_j}\rangle\langle w_{A_i}| \langle w_{B_i}|w_{B_j}\rangle$$

$$= \sum_i \alpha_i \alpha_i^* |w_{A_i}\rangle\langle w_{A_i}|$$

  - If ρ_A is pure, then all the α_i are zero except for one; and this entails $|\psi_{AB}\rangle$ is a product vector state!

<u>*So*</u>*: If $|\psi_{AB}\rangle$ is a non-product (entangled) vector state, then ρ_A is mixed.*

Claim 2. Let $\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$ be a pure density operator state on a product vector space $\mathcal{H}_A \otimes \mathcal{H}_B$. Then $|\psi_{AB}\rangle$ is an entangled vector state *if and only if* $S_A > 0$ (i.e., ρ_A is mixed).

- *So*: *If* a *bipartite* system is in a *pure* state, then S_A is a measure of the degree to which the vector state of the system is entangled.
  - *More provocatively*: S_A *is a measure of the degree to which subsystem* A *is entangled with subsystem* B.
  - *The more mixed* ρ_A *is, the greater the entanglement between* A *and* B.

Side note: What if ρ_{AB} *is mixed and/or the composite system has more than two subsystems?...*