

Title: Quantum Information Lecture

Speakers: Eduardo Martin-Martinez

Collection: Quantum Information 2023/24

Date: March 01, 2024 - 9:00 AM

URL: <https://pirsa.org/24030051>

as:

$$e^{\lambda x + \mu p} \longrightarrow e^{\lambda \hat{x} + \mu \hat{p}}. \quad (14)$$

This is just a choice and it is not unique!

2.2 Postulates of QM

Let \mathcal{H} be the Hilbert space associated with a quantum mechanical physical system.

- *Postulate 1:* To each physical state corresponds a unique state operator, called the *density operator*, which has to satisfy:
 1. **Positivity** $\langle \psi | \hat{\rho} | \psi \rangle \geq 0$ for $|\psi\rangle \in \mathcal{H}$ (all eigenvalues are positive semi-definite).
 2. **Trace normalisation:** $\text{Tr}(\hat{\rho}) = \sum_i \langle e_i | \hat{\rho} | e_i \rangle = 1$, where $\{|e_i\rangle\}$ is an orthonormal basis of \mathcal{H} .
 3. It is **self-adjoint**, $\hat{\rho}^\dagger = \hat{\rho}$
- *Postulate 2:* Physical observables (in principle measurable properties of the system) are represented by (possibly unbounded) self-adjoint operators on the Hilbert space \mathcal{H} . The result of a measurement of an observable that produces a definite value will always be in the spectrum of the operator that represented it.
- *Postulate 3: (Born's Rule)* The probability of measuring an outcome α_i when measuring the observable represented by the self-adjoint operator \hat{A} on a state $\hat{\rho}$ is given by: $\text{Tr}(\hat{\rho} |\alpha_i\rangle \langle \alpha_i|)$, where $\hat{A} |\alpha_i\rangle = \alpha_i |\alpha_i\rangle$, therefore: $\langle \hat{A} \rangle = \text{Tr}(\hat{\rho} \hat{A})$

2.3 The density matrix

There is a difference between classical probabilistic systems and quantum systems: in the classical case the probabilistic nature exists because we do not have access to the microstates, but in the quantum case even if we know the microscopic states we may not be certain about the outcome of a measurement of an observable. Nevertheless, in quantum systems we may have classical ignorance of what exact state we have, of course on top of the intrinsic probabilistic nature of the predictions of outcomes of measurements.

Our knowledge about the state of a quantum system is described by the density matrix $\hat{\rho}$. We can always think of the density matrix as a classical probability distribution over “*pure states*”. One way to see this is to recall the fact that the density matrix is always diagonalisable in some basis:

$$\hat{\rho} = \sum_i p_i |e_i\rangle\langle e_i|, \quad (15)$$

where

$$\sum_{i=1} p_i = 1, p_i \geq 0 \Rightarrow \text{Tr}(\hat{\rho}) = 1. \quad (16)$$

The set $\{|e_i\rangle\}$ is an orthonormal basis of the Hilbert space \mathcal{H} , i.e. $\sum_i |e_i\rangle\langle e_i| = \mathbb{1}$.

Notice an important point: the interpretation of $\hat{\rho}$ as probability distribution over states is not unique:

observable. Nevertheless, in quantum systems we may have classical ignorance of what exact state we have, of course on top of the intrinsic probabilistic nature of the predictions of outcomes of measurements.

Our knowledge about the state of a quantum system is described by the density matrix $\hat{\rho}$. We can always think of the density matrix as a classical probability distribution over “pure states”. One way to see this is to recall the fact that the density matrix is always diagonalisable in some basis:

$$\hat{\rho} = \sum_i p_i |e_i\rangle\langle e_i|, \quad (15)$$

where

$$\sum_{i=1} p_i = 1, p_i \geq 0 \Rightarrow \text{Tr}(\hat{\rho}) = 1. \quad (16)$$

The set $\{|e_i\rangle\}$ is an orthonormal basis of the Hilbert space \mathcal{H} , i.e. $\sum_i |e_i\rangle\langle e_i| = \mathbf{1}$.

Notice an important point: the interpretation of $\hat{\rho}$ as probability distribution over states is not unique:

Example 2.1. *A system of one qubit be in state that is a probability distribution of the states $|0\rangle$ and $|1\rangle$, thus the density matrix is:*

$$\hat{\rho} = \frac{1}{4} |0\rangle\langle 0| + \frac{3}{4} |1\rangle\langle 1|. \quad (17)$$

We want to show that we can rewrite the density matrix as a probability distribution over different states. For example, consider the states $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Then, we can rewrite the density matrix as:

$$\hat{\rho} = \left(\frac{1}{4}(|0\rangle\langle 0| + |1\rangle\langle 1|) + \frac{1}{8}(|0\rangle\langle 1| - |0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 0|) \right) + \frac{1}{2} |1\rangle\langle 1| \quad (18)$$

$$\hat{\rho} = \frac{1}{8}(|0\rangle + |1\rangle)(\langle 0| + \langle 1|) + \frac{1}{8}(|0\rangle - |1\rangle)(\langle 0| - \langle 1|) + \frac{1}{2} |1\rangle\langle 1| \quad (19)$$

$$\hat{\rho} = \frac{1}{4} |+\rangle\langle +| + \frac{1}{4} |-\rangle\langle -| + \frac{1}{2} |1\rangle\langle 1|, \quad (20)$$

which is a statistical mixture of the states $|1\rangle$, $|+\rangle$ and $|-\rangle$. The density matrix does not contain information about how a state is prepared.

In general, for arbitrary states and probability distributions, we can always

Example 2.1. A system of one qubit is in state that is a probability distribution of the states $|0\rangle$ and $|1\rangle$, thus the density matrix is:

$$\hat{\rho} = \frac{1}{4} |0\rangle\langle 0| + \frac{3}{4} |1\rangle\langle 1|. \quad (17)$$

We want to show that we can rewrite the density matrix as a probability distribution over different states. For example, consider the states $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Then, we can rewrite the density matrix as:

$$\hat{\rho} = \left(\frac{1}{4}(|0\rangle\langle 0| + |1\rangle\langle 1|) + \frac{1}{8}(|0\rangle\langle 1| - |0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 0|) \right) + \frac{1}{2} |1\rangle\langle 1| \quad (18)$$

$$\hat{\rho} = \frac{1}{8}(|0\rangle + |1\rangle)(\langle 0| + \langle 1|) + \frac{1}{8}(|0\rangle - |1\rangle)(\langle 0| - \langle 1|) + \frac{1}{2} |1\rangle\langle 1| \quad (19)$$

$$\hat{\rho} = \frac{1}{4} |+\rangle\langle +| + \frac{1}{4} |-\rangle\langle -| + \frac{1}{2} |1\rangle\langle 1|, \quad (20)$$

which is a statistical mixture of the states $|1\rangle$, $|+\rangle$ and $|-\rangle$. The density matrix does not contain information about how a state is prepared.

In general, for arbitrary states and probability distributions, we can always find infinitely many sets of states that one can use for probabilistic preparation/interpretation. More formally, if I consider one set $\{|\psi_i\rangle\}$, we can find infinitely many sets $\{|\phi_i\rangle\}$ so that:

$$\hat{\rho} = \sum_j p_j |\psi_j\rangle\langle \psi_j| = \sum_j q_j |\phi_j\rangle\langle \phi_j|, \quad (21)$$

Expectation values of operators We can express the expectation value of an observable \hat{O} in a state of density matrix $\hat{\rho}$ as

$$\langle \hat{O} \rangle = \text{Tr}(\hat{\rho}\hat{O}) \quad (22)$$

The expectation value of an observable \hat{O} , can also be thought as an expectation of a classical distribution. Assume $\{e_j\}$ to be the basis that diagonalises the density matrix: $\hat{\rho} = \sum p_i |e_i\rangle\langle e_i|$. By definition the trace is:

$$\langle \hat{O} \rangle = \text{Tr}(\hat{\rho}\hat{O}) = \sum_j \langle e_j | \hat{O} | e_j \rangle \quad (23)$$

We write the density matrix in its diagonal form in the basis $\{|e_j\rangle\}$:

$$\begin{aligned} \langle \hat{O} \rangle &= \sum_i \sum_j p_i \langle e_j | e_i \rangle \langle e_i | \hat{O} | e_j \rangle \\ &= \sum_i \sum_j p_i \langle e_i | \hat{O} | e_j \rangle \delta_{ij} = \sum_i p_i \langle e_i | \hat{O} | e_i \rangle. \end{aligned} \quad (24)$$

So, it is the average of expectation values of \hat{O} for all the possible pure states, with the probabilities being those of the pure states.

2.4 Pure and non-pure states

If we know exactly ~~which vector in \mathcal{H} represents the state~~ of the system, then

So, it is the average of expectation values of U for all the possible pure states, with the probabilities being those of the pure states.

2.4 Pure and non-pure states

If we know exactly which vector in \mathcal{H} represents the state of the system, then the density matrix is $\hat{\rho} = |\psi\rangle\langle\psi|$. Such a rank-1 density matrix $\hat{\rho}$ is said to represent a pure state. If $\hat{\rho}$ is pure, then

$$\hat{\rho}^2 = \hat{\rho} \Rightarrow \text{Tr}(\hat{\rho}^2) = \text{Tr}(\hat{\rho}) = 1 \quad (25)$$

and, since $\hat{\rho}$ is of the form (15), we have that

$$\text{Tr}(\hat{\rho}^2) = 1 \Rightarrow \hat{\rho}^2 = \hat{\rho}. \quad (26)$$

Thus, if the trace of $\hat{\rho}^2$ equals 1 then the state is pure, and the function $\hat{\rho}^2$ can be thought of as a measure of *purity*.

Definition 2.1 (Purity). *We define as purity of a state $\hat{\rho}$:*

$$\mathcal{P}(\hat{\rho}) = \text{Tr}(\hat{\rho}^2) \quad (27)$$

Purity is bounded from above and below:

$$1/d \leq \mathcal{P}(\hat{\rho}) \leq 1, \quad (28)$$

where d is the dimension of the Hilbert space \mathcal{H} .

It is easy to check the bounds of purity: We know that the maximally mixed state (the state with the highest Shannon entropy) is a multiple of the identity (homogeneous distribution over all possible orthogonal sets of pure states) $\hat{\rho}_{\max} = \frac{1}{d} \mathbb{1}$. Therefore the minimum possible purity will be the purity of $\hat{\rho}_{\max}$

$$P\left(\frac{1}{d} \mathbb{1}\right) = \text{Tr}\left(\left(\frac{1}{d} \mathbb{1}\right)^2\right) = \frac{1}{d^2} \text{Tr}(\mathbb{1}) = \frac{1}{d} \quad (29)$$

Additionally, one needs to prove that the purity is monotonically increasing as one gets from a mixed state to a pure state, consider the following family of states

$$\hat{\rho} = a |\psi\rangle\langle\psi| + (1-a) \frac{1}{d} \mathbb{1} \quad (30)$$

For $a = 1$ we have a pure state and as a decreases, the state becomes more and more mixed until it becomes maximally mixed for $a = 0$. Then,

$$\hat{\rho}^2 = \left(a^2 + a(1-a) \frac{2}{d} |\psi\rangle\langle\psi| + (1-a)^2 \frac{1}{d^2} \mathbb{1} \right) \Rightarrow \quad (31)$$

$$\text{Tr} \hat{\rho}^2 = \left(a^2 + a \frac{2}{d} (1-a) + (1-a)^2 \frac{1}{d^2} d \right) \Rightarrow \quad (32)$$

$$\text{Tr} \hat{\rho}^2 = a^2 + \frac{1}{d} (1-a^2) = \left(1 - \frac{1}{d} \right) a^2 + \frac{1}{d} \quad (33)$$

which is a monotonically increasing function of a .

So, we have shown two things. Most importantly that purity actually quantifying the purity of a state, as it monotonically increasing with the purity of a state. Secondly, this monotonicity yields the inequality (28).

2.5 Von Neumann Entropy

$$\rho = a |\psi\rangle\langle\psi| + (1-a) \frac{1}{d} \mathbb{1} \quad (30)$$

For $a = 1$ we have a pure state and as a decreases, the state becomes more and more mixed until it becomes maximally mixed for $a = 0$. Then,

$$\hat{\rho}^2 = \left(a^2 + a(1-a) \frac{2}{d} |\psi\rangle\langle\psi| + (1-a)^2 \frac{1}{d^2} \mathbb{1} \right) \Rightarrow \quad (31)$$

$$\text{Tr} \hat{\rho}^2 = \left(a^2 + a \frac{2}{d} (1-a) + (1-a)^2 \frac{1}{d^2} d \right) \Rightarrow \quad (32)$$

$$\text{Tr} \hat{\rho}^2 = a^2 + \frac{1}{d} (1-a^2) = \left(1 - \frac{1}{d} \right) a^2 + \frac{1}{d} \quad (33)$$

which is a monotonically increasing function of a .

So, we have shown two things. Most importantly that purity actually quantifying the purity of a state, as it monotonically increasing with the purity of a state. Secondly, this monotonicity yields the inequality (28).

2.5 Von Neumann Entropy

The *von Neumann entropy* of a state is defined as the Shannon entropy associated with the probability distribution over distinguishable pure states that the density operator $\hat{\rho}$ represents.

Definition 2.2 (von Neumann entropy). *Let a system be in a state $\hat{\rho}$. The von Neumann entropy of that state is:*

$$S(\hat{\rho}) := -\text{Tr}(\hat{\rho} \log \hat{\rho}) = -\sum_i p_i \log p_i \quad (34)$$

The von Neumann entropy is bounded by $0 \leq S(\hat{\rho}) \leq \log d$. Multiplying by the coefficient $k_B T$ we get the entropy dimensions that are compatible with thermodynamics.

$$f(\hat{H}) = \sum_{\alpha} f(h_{\alpha}) |h_{\alpha}\rangle \langle h_{\alpha}|$$

$$\hat{H} |h_{\alpha}\rangle = h_{\alpha} |h_{\alpha}\rangle$$

A note on Thermodynamics and Information

Gibbs Thermality: The state of equilibrium corresponds to

Max of Von Neumann entropy at "constant energy": there is

Maximize entropy.

$$\frac{\partial}{\partial \hat{\rho}} \left[S - \beta (\text{Tr}(\hat{\rho} \hat{H}) - E) - \mu (\text{tr}(\hat{\rho}) - 1) \right]$$

leads to a maximum of thermodynamic entropy, $\text{tr}(\rho \hat{H})$

there is a Hamiltonian \hat{H} . Constant energy $\langle \hat{H} \rangle = E = \text{constant}$
Trace 1 constant $\text{tr} \rho = 1$
 $(\rho) - 1 \Big] = 0$

$$S(\hat{\rho}) = -\text{tr} \hat{\rho} \text{Log}(\hat{\rho}) = \sum_i n_i p_i \text{Log}(p_i)$$

$$\text{tr}(\hat{\rho}) = \sum_i n_i p_i = 1$$

$$\langle \hat{H} \rangle = \sum_i n_i p_i E_i = \bar{E}$$

$$[\hat{\rho}, \hat{H}] = 0 \quad \text{Stabilität}$$

$$\hat{\rho} = \sum_i p_i \hat{P}_i$$

$$\hat{H} = \sum_i E_i \hat{P}_i$$

$$\text{Rank}(\hat{P}_i) = n_i$$

$$\hat{P}_i = 1/\psi_i \langle \psi_i | \dots$$

$$\frac{\partial}{\partial p_i} \left(-\sum_i n_i p_i \log(p_i) - \beta \left(\sum_i n_i E_i p_i - E \right) - \mu \left(\sum_i n_i p_i - 1 \right) \right) = 0 \Rightarrow$$

$$-\log(p_i) - 1 - \beta E_i - \mu = 0$$

multiply both sides by \hat{p}_i
and sum them all

$$-\log(\hat{p}) - \mathbb{1} - \beta \hat{H} - \mu \mathbb{1} = 0 \Rightarrow \hat{p} = e^{-(\mu+1)\mathbb{1}} e^{-\beta \hat{H}} = \frac{1}{Z(\beta)} e^{-\beta \hat{H}}$$

$$\text{tr}(\hat{p}) = 1 \Rightarrow Z(\beta) = \text{tr}(e^{-\beta \hat{H}})$$

$$) = 0 \Rightarrow$$

$$\sum \hat{P}_i = 1$$

$$f(\hat{H}) = \sum_k f(h_k) |h_k\rangle\langle h_k|$$

$$\sum |e_i\rangle\langle e_i| = 1$$

$$H|h_i\rangle = h_i|h_i\rangle$$

$$\frac{1}{Z(\beta)} e^{-\beta H}$$

$$(M+1)_i = \text{Log}(Z(\beta))$$

$$\text{tr}(\hat{\rho}) = 1 \Rightarrow Z(\beta) = \text{tr}(e^{-\beta \hat{H}})$$

$$\text{tr}(\beta \hat{H}) = E \Rightarrow E = \frac{1}{Z(\beta)} \text{tr}(\hat{H} e^{-\beta \hat{H}}) = -\frac{\partial}{\partial \beta} \ln(Z(\beta)) = E$$

$$\hat{\rho}_{\beta} = \frac{1}{Z(\beta)} \sum_i e^{-\beta E_i} \hat{\rho}_{E_i}$$

$$\beta = (k_B T)^{-1}$$

$$\hat{P} = \sum_i p_i \hat{P}_i$$

$$\text{Rank}(\hat{P}_i) = n_i$$

$$\hat{P}_i = 1/\psi_i \langle \psi_i |$$

$$\hat{H} = \sum_i n_i E_i \hat{P}_i$$

$$\hat{P} \propto e^{-\beta \hat{H}}$$

$$\hat{O} \propto e^{i t \hat{H}}$$

$$t \rightarrow i\beta$$