A note on quantum formalism: density operators

Elliot Cheung

Definition. *A density matrix, or density* operator*, is an operator* $\rho : \mathcal{H} \to \mathcal{H}'$ *satisfying;*

- $\rho^\dagger = \rho$
- *For any* $|\psi\rangle \in \mathcal{H}'$*, we have that* $\langle\psi| \rho |\psi\rangle \geqslant 0$.
- $\mathrm{Tr}(\rho) = 1$

The above is the *mathematical definition* of a density operator. As a matter of convenience, we will assume in this note that $\mathcal{H} = \mathcal{H}'$. In terms of physics, density operators provide an alternative framework for describing the axioms of quantum mechanics that is more flexible than the standard description of quantum mechanics via state vectors (and Hilbert spaces of states). In particular, while the state vector formalism provides a good framework for describing *closed* quantum mechanical systems, the language of density operators tends to be more convenient if one wishes to also describe *open* quantum mechanical systems. That is, the density operator formalism can easily describe both closed systems, and systems that may possibly interact with the environment (i.e. incorporating quantum noise).

When describing in what density operator formalism can "replace" state vector formalism, it is instructive to consider in what sense can a state vector "describe" a quantum system. Recall that in the state vector formalism: given an observable $\mathbf{A}$ (i.e. a self-adjoint operator), with eigenvalue m, if a quantum system $\mathcal{H}$ (i.e. a Hilbert space of states) is in state $|\psi\rangle$, then the *probability* of measuring $\mathbf{A}$ with outcome m is given by the formula:

$$\underline{\mathrm{Prob}}\Big(m \,\big|\, |\psi\rangle\Big) = \langle\psi| \Pi_m |\psi\rangle = \langle \Pi_m \psi| \, |\psi\rangle = \| \Pi_m |\psi\rangle \|^2 \tag{1}$$

where Π_m denotes the (orthogonal) projection operator, onto the eigenspace of $\mathbf{A}$ corresponding to eigenvalue m.

Furthermore, the post measurement state of the system just after the measurement is given by

$$\frac{\Pi_m |\psi\rangle}{\| \Pi_m |\psi\rangle \|}$$

Now, as we can write $\mathbf{A} = \sum m \Pi_m$, we can compute the expectation value of the measurement outcome to be

$$\begin{aligned}\langle \mathbf{A} \rangle &= \sum_m \underline{\text{Prob}}\Big(m \,\big|\, |\psi\rangle\Big) \\ &= \sum_m m \cdot \Big(\langle\psi| \,\Pi_m\, |\psi\rangle\Big) \\ &= m \cdot \Big(\langle\psi| \,\Pi_m\, |\psi\rangle\Big) \\ &= \langle\psi| \,\mathbf{A}\, |\psi\rangle\end{aligned}$$

Given $\mathbf{A}$ and ψ, a *complete* set of outcomes $\{m\}$ for ψ is a collection of possible outcomes such that

$$\sum \underline{\text{Prob}}\Big(m\big| \,|\psi\rangle\Big) = 1 \tag{2}$$

In general, a complete set of outcomes $\{m\}$ is a collection of possible outcomes, for an observable U, such that 2 holds for *all* possible state vectors $\psi \in \mathcal{H}$. Given the definitions above, it is easy to see that a complete set of outcomes $\{m\}$ for an observable U may be equivalently described as a collection $\{M_m\}$ of unitary operators (called *measurement operators*), where M_m is an operator with eigenvalue m (i.e. the outcome m), such that

$$\sum M_m^\dagger M_m = \text{Id}$$

As $\mathbf{A}$ is self-adjoint, we have that $\mathbf{A}$ is diagonalizable and so we may write

$$\mathbf{A} = \sum m\Pi_m$$

where m is an eigenvalue for U and Π_m is the projection operator onto the eigenspace of U corresponding to eigenvalue m. By virtue of diagonalizability, the collection $\{\Pi_m\}$ forms a complete set of measurement operators (with respect to $\mathbf{A}$) for $\mathcal{H}$. That is, $\sum \Pi_m^\dagger \Pi_m = \sum \Pi_m = \text{Id}$.

Then, one can calculate the total expected value of the observable $\mathbf{A}$, given the state $|\psi\rangle$

$$\begin{aligned}\underline{\text{Prob}}\Big(\mathbf{A}\big| \,|\psi\rangle\Big) &= \sum m\underline{\text{Prob}}\Big(m \,\big|\, |\psi\rangle\Big) \\ &= \sum m\langle\psi| \,\Pi_m\, |\psi\rangle \\ &= \sum \langle\psi| \,m\Pi_m\, |\psi\rangle \\ &= \langle\psi| \,\mathbf{A}\, |\psi\rangle\end{aligned}$$

Ultimately, this is to say that **for any observable A, state vectors of a system describe the measurement statistics for the observable**. Furthermore, we also have a description of the resulting state vector the system is in post-measurement. Note that as far as physics goes, one may consider measurement statistics of observables (*any observable!*) to be the totality of

physically relevant information one can discern from a physical (quantum) system. This begs the question: is it possible to extract these measurement statistics (for all observables), without necessarily having knowledge of a fixed state vector of a system?

The answer is **yes** – and this brings us to the density operator formalism and an alternative description of a quantum system via density operators. To motivate the introduction of the density operator, let us consider the following scenario:

Suppose that instead of having a fixed state vector $|\psi\rangle$ of a quantum system $\mathcal{H}$, we have an *ensemble* of possible states. That is, suppose that we have a *probability distribution* of possible state vectors, $\{p_i, |\psi_i\rangle\}_{i=0,\cdots N}$ – meaning that $0 \leqslant p_i \leqslant 1$, $\sum p_i = 1$ where p_i represents the probability that $\mathcal{H}$ is in the state $|\psi_i\rangle$. Note that in the scenario before, when describing the state vector formalism, our underlying assumption was always that $\mathcal{H}$ was in a definite state $|\psi\rangle$. The situation now is a slightly different – we do not suppose that $\mathcal{H}$ is necessarily in a definite state $|\psi\rangle$, but there is a fixed *complete* set of probabilities for a collection of possible states $|\psi_i\rangle$ for $\mathcal{H}$. By "complete" here, we of course mean that $\sum p_i = 1$. To put it formally:

Definition. *An ensemble of states $\{p_i, |\psi_i\rangle\}_{i=0,\cdots N}$, is a finite collection of state vectors $|\psi_i\rangle$ with associated probabilities p_i (i.e. satisfying $0 \leqslant p_i \leqslant 1$ and $\sum p_i = 1$).*

Now, the claim is that while we know that measurement statistics can be obtained from a statevector, for any observable – we can also obtain measurement statistics for any observable, given a fixed (complete) ensemble of states. Note that a definite state vector $|\psi\rangle$ of a system $\mathcal{H}$ is also a valid ensemble of states – this is of course the situation where the ensemble is simply given by ψ with probability 1.

Terminology: one usually refers to an ensemble (or density operator) given by a definite state vector as a *pure state* – otherwise, it is referred to as a *mixed state*.

The idea here is: given an ensemble $\{p_i, |\psi_i\rangle\}$ and an observable U, we would like to "measure U" with respect to the ensemble. To this end, we define the density matrix of the ensemble $\{p_i, |\psi_i\rangle\}$ to be the operator $\rho : \mathcal{H} \to \mathcal{H}$ given by

$$\rho = \sum p_i \, |\psi_i\rangle\langle\psi_i|$$

Then, we define the expectation value of $\mathbf{A}$ with respect to ρ to be

$$\mathrm{E}_\rho(\mathbf{A}) = \mathrm{Tr}(\mathbf{A}\rho)$$

if Π_m is the projection onto the eigenspace associated to an eigenvalue m of U, we define

$$\underline{\mathrm{Prob}}_\rho(m) = \mathrm{Tr}(\Pi_m \rho)$$

Then, where $\mathbf{A} = \sum m\Pi_m$ with $\Pi_m \perp \Pi_{m'}$ for $m \neq m'$, we have that

$$\mathrm{E}_\rho(\mathbf{A}) = \mathrm{Tr}(\mathbf{A}\rho) = \sum m \cdot \underline{\mathrm{Prob}}_\rho(m)$$

Recall from definition 1, that given the initial state $|\psi_i\rangle$ for the system $\mathcal{H}$, the probability of the measurement of $\mathbf{A}$ resulting in m is given by

$$\begin{aligned}\underline{\mathrm{Prob}}\Big(m\big|\,|\psi_i\rangle\Big) &= \langle\psi_i|\,\Pi_m\,|\psi_i\rangle \\ &= \mathrm{Tr}\left(\Pi_m\,|\psi_i\rangle\langle\psi_i|\right)\end{aligned}$$

Now, as $\{p_i, |\psi_i\rangle\}$ is a *complete* ensemble, one has that $\sum p_i = 1$, and

$$\begin{aligned}&\underline{\mathrm{Prob}}(m) = \sum_i \underline{\mathrm{Prob}}\Big(m\big|\,|\psi_i\rangle\Big)\cdot p_i \\ &= \sum_i p_i\,\mathrm{Tr}(\Pi_m\,|\psi_i\rangle\langle\psi_i|) \\ &= \mathrm{Tr}(\Pi_m\cdot\rho) = \underline{\mathrm{Prob}}_\rho(m)\end{aligned}$$

Therefore, if are not given a definite state $|\psi\rangle$ of our system $\mathcal{H}$, but rather an ensemble of possible states (of which there are finitely many), then the total probability of measuring the outcome m from our observable $\mathbf{A}$ is equal to $\mathrm{Tr}(\Pi_m \cdot \rho)$, given by the density operator. Similarly, an expression for the total expectation value of the observable $\mathbf{A}$ can be calculated in terms of the density operator associated to the ensemble $\{p_i, |\psi_i\rangle\}$:

$$\mathrm{E}(\mathbf{A}) = \sum m\underline{\mathrm{Prob}}(m) = \sum m\,\mathrm{Tr}(\Pi_m\cdot\rho) = \mathrm{Tr}(\mathbf{A}\cdot\rho) = \mathrm{E}_\rho(\mathbf{A})$$

At this stage, let us point out that there is a difference between a state in equal superposition and an ensemble of states with uniformly distributed probability. For example, let us consider an ensemble given by $|0\rangle$ with probability $\frac{1}{2}$ and $|1\rangle$ with probability $\frac{1}{2}$. The corresponding density operator is given by $\rho = \frac{1}{2}\,|0\rangle\langle 0| + \frac{1}{2}\,|1\rangle\langle 1| = \frac{1}{2}\,\mathrm{Id}$, as $\{|0\rangle, |1\rangle\}$ forms a basis for $\mathcal{H}$. We may also consider the superposition state $|+\rangle = \frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}\,|1\rangle$. First, we point out the obvious difference: a system $\mathcal{H}$ in state $|+\rangle$ would be in a pure state, while on the other hand if the system $\mathcal{H}$ described by ρ would be a system in a mixed state. In other words, the density operator associated to $|\psi\rangle$ is given by $\rho_+ = |+\rangle\langle+|$. The idea is that the density operators ρ and ρ_+ are *physically distinguishable*. That is, there exists a choice of observable $\mathbf{A}$ for which ρ and ρ_+ produce different measurement statistics.

Indeed, we set $\mathbf{A}$ to be the projection operator onto the subspace spanned by $|+\rangle$. Then, $\underline{\mathrm{Prob}}_{\rho_+}(1) = 1$, while $\underline{\mathrm{Prob}}_\rho(1) = \frac{1}{2}$.

From ensembles to density operators

It is easy to see that given an ensemble $\{p_i, |\psi_i\rangle\}$, the corresponding operator

$$\rho = \sum p_i \, |\psi_i\rangle\langle\psi_i| \tag{3}$$

satisfies: $\rho^\dagger = \rho$, $\mathrm{Tr}(\rho) = 1$, and for any $|\psi\rangle \in \mathcal{H}$, we have that $\langle\psi| \, \rho \, |\psi\rangle \geqslant 0$ (see 1). On the other hand, it is easy to see that given any density operator ρ (as defined in 1), one can express ρ in the form 3 for some state vectors $|\psi_i\rangle$. However, describing ρ as a density operator is a mathematically convenient way to express the operator ρ without committing to a choice of ensemble. Given a density operator ρ, one says that 3 is a choice of an *ensemble presentation* of ρ. It is clear that the choice of ensemble presentation is not uniquely determined by ρ.

(Reduced) density operators and open systems: the partial trace

As mentioned as the beginning of this note, density operator formalism allows us to describe measurement statistics for observables acting on a Hilbert space $\mathcal{H}_0$, where $\mathcal{H}_0$ is in fact part of a larger *composite* system $\mathcal{H}_0 \hookrightarrow \mathcal{H}$. The axioms of quantum formalism tell us that such a composite system may be represented as $\mathcal{H} = \mathcal{H}_0 \otimes \mathcal{H}_0'$. An important example to keep in mind of such a set-up, is the example of a quantum system $\mathcal{H}_0$ which may interact with the external environment, but one considers the measurement of observables that only act on the system $\mathcal{H}_0$ and not the external environment.

The important thing is realize is that the usual Hilbert space of states formalism for quantum systems is applicable when one has a *closed* quantum system. In this case, the Hilbert space $\mathcal{H}$ models all possible states of a given system, and evolution of the system is determined by unitary operators acting on $\mathcal{H}$. Of course, in practise, it may prove to be an intractible problem to completely describe a closed system – and one may instead consider a more tractible subsystem. For example, one may consider the two dimensional qubit system determined by the spin of a particle – where in reality, the particle may be involved in a much larger, more complex state space. In restricting to a subsystem, it is no longer true that states (in the subsystem) are necessarily described by rays in a Hilbert space, and evolution of the (sub)system is no longer necessarily described by unitary operators on the subsystem. We refer to such a subsystem $\mathcal{H}_0$ as an *open system.*

Therefore, the question is as follows: given an open system $\mathcal{H}_0$, which is a subsystem of a larger closed system $\mathcal{H}$, how can one get a handle on measurement statistics of physical observables $\mathbf{A}$ that act on on $\mathcal{H}_0$, without knowledge of the larger closed system $\mathcal{H}$ containing $\mathcal{H}_0$?

With this question, we may already sense the utility of the density operator formalism. In terms of states, to determine measurement statistics of $\mathbf{A}$ acting on $\mathcal{H}_0$, one can consider a state $|\psi_0\rangle$ in $\mathcal{H}_0$ as a truncation of a state $|\psi\rangle$ in $\mathcal{H}$ and then measure the observable $\mathbf{A}$ in state $|\psi\rangle$. However, this of course requires one to consider a state $|\psi\rangle$ in the larger closed system $\mathcal{H}$. On the other hand, recall that the density operator allows us to describe measurement statistics in terms of *mixed states*. That is, in lieu of a definite state vector, we can still define measurement statistics via the formula $\mathrm{E}_\rho(\mathbf{A}) = \mathrm{Tr}(\mathbf{A} \cdot \rho)$. Then, the idea is that the larger closed subsystem $\mathcal{H}$ induces a density operator ρ for $\mathcal{H}_0$. The resulting density operator for $\mathcal{H}_0$ is known as the *reduced density operator* for the subsystem $\mathcal{H}_0$ in $\mathcal{H}$.

To obtain a reduced density operator for $\mathcal{H}_0$, one takes a density operator on $\mathcal{H}$ and takes the *partial trace* of it over the "complementary subsystem" $\mathcal{H}_0'$ to $\mathcal{H}_0$ inside $\mathcal{H}$.

Definition. *Suppose that $\mathcal{H} = \mathcal{H}_0 \otimes \mathcal{H}_0'$ is a composite system, and $\mathbf{A} = \sum_i (A_0)_i \otimes (A_0')_i$ is a linear operator on $\mathcal{H}$. Then, we define the partial trace of $\mathbf{A}$ over the subsystem $\mathcal{H}_0'$ to be*

$$\mathrm{Tr}_{\mathcal{H}_0'}(\mathbf{A}) = \sum_i (\mathbf{A_0})_i \,\mathrm{Tr}\left((\mathbf{A_0'})_i\right) : \mathcal{H}_0 \to \mathcal{H}_0$$

Proposition. *For any Hilbert space $\mathcal{H}_0$ with density operator $\rho : \mathcal{H}_0 \to \mathcal{H}_0$, there exists a larger system $\mathcal{H}$ such that $\mathcal{H} = \mathcal{H}_0 \otimes \mathcal{H}_0'$, such that ρ can be represented as the partial trace $\mathrm{Tr}_{\mathcal{H}_0'}(|\psi\rangle\langle\psi|)$ of a density operator corresponding to a* pure state *$|\psi\rangle \in \mathcal{H}$. Such a $|\psi\rangle$ is referred to a* purification *of ρ (which may be a* mixed *state)*