

Contents

4	Density operators	3
4.1	Motivation	3
4.2	Definition of a density operator	4
4.2.1	Traces	4
4.2.2	Density operators of pure states	5
4.2.3	General density operators	6
4.2.4	Examples	8
4.2.5	Non-uniqueness of ensemble representations	9
4.3	Von-Neumann equation	9
4.4	Measurement for density operators	10
4.5	Mixing and states of composite systems	11
4.5.1	Mixing quantum states	11
4.5.2	Partial traces	12
4.5.3	Entangled states	13

Chapter 4

Density operators

This will be a relatively short chapter: We will merely introduce here the concept of a density operator, which incorporates the concept of classical probabilities into quantum theory. Density operators are quantum states and hence generalize the concept of state vectors.

4.1 Motivation

Let us imagine we have a single spin, associated with a Hilbert space $\mathcal{H} \simeq \mathbb{C}^2$. We now throw a coin. In case of heads, we prepare the spin in $|0\rangle$, in case of tails, we prepare it in $|1\rangle$. That is to say, with the classical probability $1/2$ we have $|0\rangle$, and with classical probability $1/2$ we get $|1\rangle$. How do we capture this situation? Can we describe the system by a state vector

$$|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}? \quad (4.1)$$

Not quite. This is easy to see: In case of a σ_x measurement, we would always get the same outcome. But this is different from the situation we encounter here. In fact, when we make a measurement of σ_x , we would get both outcomes with equal probability. Or

$$|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}? \quad (4.2)$$

Again, this will not work, for the same reason. In fact, no state vector is associated with such a situation, and for that, we need to generalize our concept of a quantum state slightly: to density operators. This is, however, the most general quantum state in standard quantum mechanics, and we will not have to generalize it any further.

In fact, the above situation is an instance of the situation where we prepare with probability p_j , $j = 1, \dots, n$, a system in a state vector $|\psi_j\rangle$. Since we encounter a probability distribution, we have

$$\sum_{j=1}^n p_j = 1. \quad (4.3)$$

Such a situation is sometimes referred to as a mixed ensemble. How do we incorporate that?

4.2 Definition of a density operator

4.2.1 Traces

A hint we have already available: Since all operations we can apply to state vectors act linearly (time evolution and measurement), we already know the following: Let us assume that we initially have the situation that with probability p_j the state vector $|\psi_j(t_0)\rangle$ is prepared. Then we, say, evolve the system in time, which means that we apply the unitary $U(t, t_0)$ to each state vector,

$$|\psi_j(t)\rangle = U(t, t_0)|\psi_j(t_0)\rangle. \quad (4.4)$$

In order to compute the expectation value of some observable A at time t , we merely need to compute the expectation value for each of the initial states, to obtain

$$\sum_{j=1}^n p_j \langle \psi_j(t) | A | \psi_j(t) \rangle. \quad (4.5)$$

Before we can define the density operator, based on this intuition, we quickly remind ourselves again of the trace of a matrix that we saw in the second chapter:

Trace of a matrix: The trace of a $d \times d$ -matrix A is defined as

$$\text{tr}(A) = \sum_{j=0}^{d-1} \langle j | A | j \rangle. \quad (4.6)$$

The trace has a number of interesting properties:

Properties of the trace: For any matrix A and any unitary U ,

$$\text{tr}(UAU^\dagger) = \text{tr}(A). \quad (4.7)$$

For two matrices A, B ,

$$\text{tr}(AB) = \text{tr}(BA) \quad (4.8)$$

(cyclic interchange) and

$$\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B). \quad (4.9)$$

Also

$$\text{tr}(A^*) = \text{tr}(A)^* \quad (4.10)$$

and

$$\text{tr}(A^T) = \text{tr}(A). \quad (4.11)$$

For numbers $\alpha \in \mathbb{C}$,

$$\text{tr}(\alpha A) = \alpha \text{tr}(A). \quad (4.12)$$

The unitary invariance also means that the basis that we pick in order to compute the trace does not matter. These properties also imply that the trace is a linear functional.

4.2.2 Density operators of pure states

We are now in the position to introduce a density operator for a so-called pure state:

Density operator of a pure state: A pure state associated with a state vector $|\psi\rangle \in \mathcal{H}$ from some Hilbert space \mathcal{H} is given by the density operator

$$\rho = |\psi\rangle\langle\psi|. \quad (4.13)$$

We immediately find some properties of such an operator: We obviously have that

$$\rho = \rho^\dagger. \quad (4.14)$$

Then,

$$\text{tr}(\rho) = 1. \quad (4.15)$$

Finally, we have that

$$\rho \geq 0, \quad (4.16)$$

which means that all of its eigenvalues are non-negative, which is clearly the case, as all the eigenvalues are given by 0 or 1, clearly non-negative numbers. We also have the

property that

$$\text{tr}(\rho^2) = \text{tr}(|\psi\rangle\langle\psi|\psi\rangle\langle\psi|) = \text{tr}(|\psi\rangle\langle\psi|) = \text{tr}(\rho) = 1. \quad (4.17)$$

How to we compute expectation values from such a density operator? Well, we know that for an observable

$$\langle A \rangle = \langle \psi | A | \psi \rangle. \quad (4.18)$$

This we can equally well (although at this instance in time this may look unnecessarily complicated, but we will see the point in a second) write as

$$\langle A \rangle = \langle \psi | A | \psi \rangle = \text{tr}(A | \psi \rangle \langle \psi |) = \text{tr}(A \rho). \quad (4.19)$$

We have hence made use of one of the above rules, and have written expectation values as a trace of the observable, multiplied with the density operator.

4.2.3 General density operators

A general density operator is just extended by linearity from this definition.

Density operator of a mixed ensemble: Consider the situation of preparing $|\psi_j\rangle$, $j = 1, \dots, n$ with probability p_j . This is associated with a density operator

$$\rho = \sum_{j=1}^n p_j |\psi_j\rangle\langle\psi_j|. \quad (4.20)$$

How do we compute expectation values from that? We have for observables A that

$$\langle A \rangle = \sum_{j=1}^n p_j \langle \psi_j | A | \psi_j \rangle, \quad (4.21)$$

from the very definition of a mixed ensemble. This we can, however, also write as

$$\begin{aligned} \langle A \rangle &= \sum_{j=1}^n p_j \text{tr}(A |\psi_j\rangle\langle\psi_j|) \\ &= \text{tr} \left(A \sum_{j=1}^n p_j |\psi_j\rangle\langle\psi_j| \right) \\ &= \text{tr}(A \rho). \end{aligned} \quad (4.22)$$

So again, expectation values are just computable as the trace of the density operator multiplied with the observable.

We now once more investigate properties of such a density operator: We find that again,

$$\rho = \rho^\dagger. \quad (4.23)$$

In the same fashion as before, since now

$$\sum_{j=1}^n p_j = 1, \quad (4.24)$$

we also have that

$$\text{tr}(\rho) = 1. \quad (4.25)$$

Finally, we have that

$$\rho \geq 0, \quad (4.26)$$

since all of the probabilities are positive, and a sum of positive operators is positive. These are exactly the same properties as above, except from one that is now missing. We no longer have $\text{tr}(\rho^2) = 1$. In fact, this property is replaced by

$$\begin{aligned} \text{tr}(\rho^2) &= \text{tr} \left(\sum_{j=1}^n p_j |\psi_j\rangle\langle\psi_j| \right) \left(\sum_{k=1}^n p_k |\psi_k\rangle\langle\psi_k| \right) \\ &= \sum_{j,k=1}^n p_j p_k \text{tr}(|\psi_j\rangle\langle\psi_j| |\psi_k\rangle\langle\psi_k|) \\ &\leq 1, \end{aligned} \quad (4.27)$$

where we have bounded the scalar products between two arbitrary state vectors. We have now arrived at the most general concept of a state in (standard) quantum mechanics. This is surely worth a box:

Density operators: General states of quantum systems with Hilbert space \mathcal{H} are given by density operators ρ . Their properties are

$$\rho = \rho^\dagger \text{ (Hermicity)}, \quad (4.28)$$

$$\rho \geq 0 \text{ (Positivity)}, \quad (4.29)$$

$$\text{tr}(\rho) = 1 \text{ (Normalization)}. \quad (4.30)$$

Pure states are those density operators for which

$$\text{tr}(\rho) = 1, \quad (4.31)$$

those can be represented by state vectors $|\psi\rangle \in \mathcal{H}$ as

$$\rho = |\psi\rangle\langle\psi|. \quad (4.32)$$

Otherwise, if $\text{tr}(\rho^2) < 1$, the state is called *mixed*. For observables, expectation values are computed as

$$\langle A \rangle = \text{tr}(A\rho). \quad (4.33)$$

4.2.4 Examples

This is a good moment to discuss a number of examples. Let us go back to our initial situation discussed at the beginning of the chapter, of preparing $|0\rangle$ or $|1\rangle$ with equal probability. We can now easily associate this with a density operator

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|. \quad (4.34)$$

We can write this in matrix form – remember that operators and their matrix representation are identified with each other throughout the script

$$\rho = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}. \quad (4.35)$$

We have that

$$\text{tr}(\rho^2) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} < 1. \quad (4.36)$$

This in fact the minimum value $\text{tr}(\rho^2)$ can take for a system with $\mathcal{H} \simeq \mathbb{C}^2$. The pure state $\rho = |0\rangle\langle 0|$ in turn is represented as

$$\rho = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad (4.37)$$

obviously satisfying $\text{tr}(\rho) = 1$. Generally, if we have probabilities p_0 and p_1 to prepare $|0\rangle$ and $|1\rangle$, we have the density operator

$$\rho = \begin{bmatrix} p_0 & 0 \\ 0 & p_1 \end{bmatrix}. \quad (4.38)$$

But of course, we are not forced to take the standard basis. The situation of having prepared $|+\rangle$ and $|-\rangle$ with equal probabilities is captured as

$$\rho = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|. \quad (4.39)$$

This is

$$\begin{aligned} \rho &= \frac{1}{4}((|0\rangle + |1\rangle)(\langle 0| + \langle 1|)) + \frac{1}{4}((|0\rangle - |1\rangle)(\langle 0| - \langle 1|)) \\ &= \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|, \end{aligned} \quad (4.40)$$

with matrix representation

$$\rho = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}. \quad (4.41)$$

4.2.5 Non-uniqueness of ensemble representations

Eh, wait a minute? Was this not the matrix representation of the ensemble consisting of $|0\rangle$ and $|1\rangle$? So we are faced here with the irritating situation that two different mixed ensembles are described by the same density operator. Let us face it: There are many different ways of preparing the same density operator! Since all expectation values of observables are computed as

$$\langle A \rangle = \text{tr}(A\rho), \quad (4.42)$$

we get exactly same same value for all observables in case of

$$\rho = \sum_{j=1}^n p_j |\psi_j\rangle\langle\psi_j| = \sum_{k=1}^m q_k |\phi_k\rangle\langle\phi_k|, \quad (4.43)$$

even if all of the probabilities $\{p_j\}$ and $\{q_k\}$ as well as all state vectors $\{|\psi_j\rangle\}$ and $\{|\phi_k\rangle\}$ are different. In fact, now even $n = m$ has to hold. What matters for all outcomes in all experiments is the density operator, not the mixed ensemble we have started with.

The situation is hence quite subtle: Yes, a density operator is a concept that allows to introduce the concept of classical probability distributions into quantum mechanics. We have to have a way of incorporating probabilistic preparation, where we do one thing with some probability and another with another. In statistical physics, we will encounter such a situation frequently, where density operators are ubiquitous, to say the least.

But no, once we arrive at a given density operator, there is no way to reconstruct the mixed ensemble that can be held responsible for the density operator. In retrospect, there always would have been infinitely many other ways of preparing the same density operator (unless it is a pure state). Sometimes, people use notions of the kind, “the system *is* in some pure state vector $|\psi_j\rangle$, $j = 1, \dots, n$, we simply do not know which one”. Such reasoning is not quite precise and can be plain wrong, in which case it is referred to as *preferred ensemble fallacy*.

4.3 Von-Neumann equation

Now that we have understood what a density operator is, the rest will be a piece of cake. How do density operators evolve in time? Well, this equation is just inherited from the Schroedinger equation by linearity. Since it was von-Neumann who first described this situation well, it is called von-Neumann equation. But really, it is just the ordinary Schroedinger equation written for density operators. There is no new physics happening here.

Von Neumann equation: Density operators of physical systems described by Hamiltonians H evolve in time according to

$$i\hbar \frac{d}{dt} \rho(t) = [H, \rho(t)]. \quad (4.44)$$

So together with the initial condition at $\rho(t_0)$, this differential equation fully specifies the density operator at a later time, exactly as the Schroedinger equation does for state vectors. Is this equation mysterious? Not at all: Let us write

$$\rho(t) = \sum_{j=1}^n p_j |\psi_j(t)\rangle \langle \psi_j(t)|, \quad (4.45)$$

then

$$\begin{aligned} i\hbar \frac{d}{dt} \rho(t) &= i\hbar \sum_{j=1}^n p_j \frac{d}{dt} |\psi_j(t)\rangle \langle \psi_j(t)| \\ &= i\hbar \sum_{j=1}^n p_j \left(\left(\frac{d}{dt} |\psi_j(t)\rangle \right) \langle \psi_j(t)| + |\psi_j(t)\rangle \left(\frac{d}{dt} \langle \psi_j(t)| \right) \right) \\ &= \sum_{j=1}^n p_j (H |\psi_j(t)\rangle \langle \psi_j(t)| - |\psi_j(t)\rangle \langle \psi_j(t)| H) \\ &= [H, \rho(t)]. \end{aligned} \quad (4.46)$$

Of course we can again write the non-differential form of time evolution:

Time evolution in terms of the time evolution operator: We have that

$$\rho(t) = U(t, t_0) \rho U(t, t_0)^\dagger. \quad (4.47)$$

4.4 Measurement for density operators

The last little piece that is missing is how to capture measurements in terms of density operators. But this is again just inherited from the previous formalism by linearity. So again, there is no new physics associated with that.

Measurement in terms of density operators: Consider a measurement of the observable A of a system prepared in the state ρ , and let π_k with

$$A = \sum_{k=0}^{D-1} \lambda_k \pi_k, \quad (4.48)$$

and

$$\pi_l \pi_k = \delta_{k,l} \pi_k \quad (4.49)$$

the projectors defined in the projection postulate. Then the probability of obtaining the k -th outcome is given by

$$p_k = \text{tr}(\pi_k \rho). \quad (4.50)$$

The state immediately after the measurement is given by

$$\rho_k = \frac{\pi_k \rho \pi_k}{\text{tr}(\pi_k \rho \pi_k)} = \frac{\pi_k \rho \pi_k}{p_k}. \quad (4.51)$$

It is easy to see that for pure states with $\rho = |\psi\rangle\langle\psi|$ this reduces to the same projection postulate that we have seen above.

4.5 Mixing and states of composite systems

4.5.1 Mixing quantum states

If we have a machine that prepares either ρ_1 or ρ_2 with probability p and $1 - p$, how do we describe this probabilistic situation? Well, this we already know: The new state is given by the density operator

$$\rho = p\rho_1 + (1 - p)\rho_2. \quad (4.52)$$

Such an operation is called *mixing*. Mathematically, this is called a convex combination of ρ_1 and ρ_2 . Physically, it means that we create a new state out of the probabilistic situation of having either of the previous two states. Maybe unsurprisingly, mixing mixes states: So even if $\rho_1 = |0\rangle\langle 0|$ and $\rho_2 = |1\rangle\langle 1|$ are both pure states,

$$\rho = p|0\rangle\langle 0| + (1 - p)|1\rangle\langle 1| \quad (4.53)$$

is a mixed state, unless $p = 0$ or $p = 1$. Such a mixing reflects ignorance of the exact preparation procedure. This is why in statistical mechanics, such mixed density operators are ubiquitous, to say the least (one there usually never knows the exact preparation of macroscopic bodies).

We also now know how to describe the initial state of the first example of the first day: This initial preparation is simply

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|. \quad (4.54)$$

Such a state is called *maximally mixed*. It captures the situation that we have no a-priori knowledge of the spin of the atom coming out of the oven.

4.5.2 Partial traces

We will now learn something quite remarkable, an insight the significance of which may not be easy to grasp at this point, but which has quite profound consequences. Let us first define the partial trace.

Partial trace: Consider a composite quantum system consisting of parts equipped with Hilbert spaces $\mathcal{H}_A \sim \mathbb{C}^2$ and $\mathcal{H}_B \sim \mathbb{C}^2$ each, so

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B. \quad (4.55)$$

The partial trace of a quantum state ρ over the second system is then given by

$$\rho_A = \text{tr}_B(\rho) = \sum_{j=0}^{d-1} \langle j | \rho | j \rangle, \quad (4.56)$$

where $\{|j\rangle : j = 0, \dots, d-1\}$ is a basis of \mathcal{H}_B . Similarly, one can define ρ_B . The new states ρ_A and ρ_B are states on \mathcal{H}_A and \mathcal{H}_B , respectively, and are called *reduced states*.

What is the significance of such a reduced state? Well, let us consider a composite quantum system and consider an observable A that is supported on \mathcal{H}_A only: So on both Hilbert spaces, this is

$$A \otimes \mathbb{1}. \quad (4.57)$$

If we have a state ρ , what is the expectation of this observable in this state? It is clearly given by

$$\langle A \otimes \mathbb{1} \rangle = \text{tr}((A \otimes \mathbb{1})\rho). \quad (4.58)$$

But this is nothing but

$$\text{tr}((A \otimes \mathbb{1})\rho) = \text{tr}(A\rho_A). \quad (4.59)$$

So if A is supported on \mathcal{H}_A only, the expectation values with respect to the full ρ and the reduced ρ_A are the same.

This operation of a partial trace is hence the one related to “disregarding a part” of a system. Colloquially speaking, it reflects the situation that we only consider a part of a composite subsystem and are no longer interested in the entire system. Clearly, expectation values of observables that act only on this part are not affected by this. This is conceptually an important operation, as in physics, one never speaks of the “wave function of the universe”: Obviously, we have systems in the laboratory, and we want to speak about their quantum state. This means that we disregard the rest of the world when speaking about what we have in the lab. And this “disregarding” is reflected by the partial trace.

Let us mention that a similar concept of course also exists in classical physics. Here one simply integrates or sums over the degrees of freedom that one would like to disregard. If we have a joint probability distribution, of course it makes sense to speak of a probability distribution of a part.

4.5.3 Entangled states

To come back to quantum mechanics, fortunately, if we have a product state, there is nothing to worry about when it comes to the partial trace:

Product state: A product state of a composite quantum system is a state of the form

$$\rho = \rho_A \otimes \rho_B. \quad (4.60)$$

One indeed has that

$$\text{tr}_B(\rho) = \rho_A. \quad (4.61)$$

If there are no correlations between a system (for example, the lab) and an environment (for example, the rest of the building), the partial trace of the joint state is just the state of the lab itself. This is reassuring: This is the very reason why state vectors of objects in labs make any sense, and why quantum mechanics courses usually start with that concept.

However, let us have a look at the following situation. Let us consider a joint system that is prepared in a pure state $\rho = |\psi\rangle\langle\psi|$ with

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0, 0\rangle + |1, 1\rangle). \quad (4.62)$$

Such a state vector makes perfect sense, it is contained in the Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. But let us now have a look at its partial trace (here the vectors $|k\rangle$ act on the second Hilbert space \mathcal{H}_B):

$$\rho_A = \frac{1}{2} \sum_{k=0}^1 \langle k| ((|0, 0\rangle + |1, 1\rangle) (\langle 0, 0| + \langle 1, 1|)) |k\rangle = \frac{1}{2} (|0\rangle\langle 0| + |1\rangle\langle 1|). \quad (4.63)$$

This is a mixed state! In fact, it is a maximally mixed state.

Entangled states: Pure states of composite systems which are not product states are called entangled states. Their reduced states are mixed.

This is radically different from the situation in classical physics: If I look at a marginal distribution in classical physics, if the joint distribution was extremal, this is still true for the marginal distribution. However, here we encounter that reduced states

of pure entangled states are mixed. This means that if the state of the system in the lab and its environment are entangled, there is *no way* we can associate it with a state vector: It simply does not make any sense, since the state is mixed. In such a situation, there is no “ $|\psi\rangle$ of the system itself”.

The concept of entanglement is related to several remarkable phenomena in quantum physics: In fact, it plays a key role in condensed-matter physics, in quantum information theory, and in quantum many-body theory. In fact, in a subtle way it is even responsible for the classical limit of quantum theory: One needs entanglement to explain how classical properties (such as trajectories of particles) emerge in quantum theory.