Chapter 2

Density matrices and composite systems

In this chapter we will take a step further in our description of quantum systems. First we will show that quantum information can be combined with classical information in a new object ρ, called the density matrix, which generalizes the idea of a ket. Then we show how to describe systems composed of multiple parts. At first these two things will seem unrelated. But in Sec. **??** we will connect the dots and show that there is an intimate relation between the two. The remainder of the chapter is then dedicated to the basic toolbox of Quantum Information; that is, the basic tools used to quantify information-theoretic measures in quantum systems.

2.1 The density matrix

A ket $|\psi\rangle$ is actually not the most general way of defining a quantum state. To motivate this, consider the state $|\boldsymbol{n}_+\rangle$ in Eq. (1.49) and the corresponding expectation values computed in Eq. (1.50). This state always points *somewhere*: it points at the direction $\boldsymbol{n}$ of the Bloch sphere. It is never possible to find a quantum ket $|\psi\rangle$ where the state doesn't point somewhere specific; that is, where it is isotropic. That sounds strange since, if we put the spin in a high temperature oven without any magnetic fields, then we certainly expect that it will never have a preferred magnetization direction. The solution to this paradox is that, when we put a spin in an oven, we are actually adding a *classical uncertainty* to the problem, whereas kets are only able to encompass quantum uncertainty.

The most general representation of a quantum system is written in terms of an operator ρ called the *density operator*, or *density matrix*. It is built in such a way that it naturally encompasses both quantum and classical probabilities. In this section I want to introduce the idea of density matrix by looking at a system which mixes quantum and classical probabilities. There is also another cool way of introducing the density matrix as stemming from the **entanglement** between two sub-systems. We will deal with that in Sec. **??**.

Suppose we have an apparatus which prepares quantum systems in certain states. For instance, this could be an oven producing spin 1/2 particles, or a quantum optics setup producing photons. But suppose that this apparatus is imperfect, so it does not always produces the same state. That is, suppose that it produces a state $|\psi_1\rangle$ with a certian probability q_1 or a state $|\psi_2\rangle$ with a certain probability q_2 and so on. Notice how we are introducing here a *classical uncertainty.* We can have as many q's as we want. All we assume is that they behave like classical probabilities:

$$q_i \in [0,1], \qquad \text{and} \quad \sum_i q_i = 1 \tag{2.1}$$

Now let A be an observable. If the state is $|\psi_1\rangle$, then the expectation value of A will be $\langle\psi_1|A|\psi_1\rangle$. But if it is $|\psi_2\rangle$ then it will be $\langle\psi_2|A|\psi_2\rangle$. To compute the actual expectation value of A we must therefore perform an average of quantum averages:

$$\langle A\rangle = \sum_i q_i \langle\psi_i|A|\psi_i\rangle \tag{2.2}$$

What is important to realize is that this type of average *cannot* be writen as $\langle\phi|A|\phi\rangle$ for some ket $|\phi\rangle$. If we want to attribute a "state" to our system, then we must generalize the idea of ket. To do that, we use Eq. (1.77) to write

$$\langle\psi_i|A|\psi_i\rangle = \text{tr}\Big[A|\psi_i\rangle\langle\psi_i|\Big]$$

Then Eq. (2.2) may be written as

$$\langle A\rangle = \sum_i q_i\, \text{tr}\Big[A|\psi_i\rangle\langle\psi_i|\Big] = \text{tr}\Big\{A\sum_i q_i|\psi_i\rangle\langle\psi_i|\Big\}$$

This motivates us to define the **density matrix** as

$$\boxed{\rho = \sum_i q_i |\psi_i\rangle\langle\psi_i|} \tag{2.3}$$

Then we may finally write Eq. (2.2) as

$$\boxed{\langle A\rangle = \text{tr}(A\rho)} \tag{2.4}$$

which, by the way, is the same as $\text{tr}(\rho A)$ since the trace is cyclic [Eq. (1.75)].

With this idea, we may now recast *all* of quantum mechanics in terms of density matrices, instead of kets. If it happens that a density matrix can be written as $\rho = |\psi\rangle\langle\psi|$, we say we have a **pure state**. And in this case it is not necessary to use ρ at all. One may simply continue to use $|\psi\rangle$. For instance, Eq. (2.4) reduces to the usual result: $\mathrm{tr}(A\rho) = \langle\psi|A|\psi\rangle$. A state which is not pure is usually called a **mixed state**. In this case kets won't do us no good and we must use ρ.

Examples

To start, suppose a machine tries to produce qubits in the state $|0\rangle$. But it is not very good so it only produces $|0\rangle$ with probability q. And, with probability $1-q$ it produces a state $|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle$, where θ may be some small angle (just for the sake of example). The density matrix for this system will then be

$$\rho = q|0\rangle\langle 0| + (1-q)|\psi\rangle\langle\psi| = \begin{pmatrix} q + (1-q)\cos^2\frac{\theta}{2} & (1-q)\sin\frac{\theta}{2}\cos\frac{\theta}{2} \\ (1-q)\sin\frac{\theta}{2}\cos\frac{\theta}{2} & (1-q)\sin^2\frac{\theta}{2} \end{pmatrix} \tag{2.5}$$

which is a just a very ugly guy. We could generalize it as we wish For instance, we can have a machine that produces $|0\rangle$, $|\psi_1\rangle$ and $|\psi_2\rangle$, and so on. Of course, the more terms we add, the uglier will the result be, but the idea is the same.

Next I want to discuss with you something called the **ambiguity of mixtures**. The idea is quite simple: if you mix stuff, you generally loose information, so you don't always know where you started at. To see what I mean, consider a state which is a 50-50 mixture of $|0\rangle$ and $|1\rangle$. The corresponding density matrix will then be

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Alternatively, consider a 50-50 mixture of the states $|\pm\rangle$ in Eq. (1.12). In this case we get

$$\rho = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -| = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We see that both are identical. Hence, we have *no way to tell* if we began with a 50-50 mixture of $|0\rangle$ and $|1\rangle$ or of $|+\rangle$ and $|-\rangle$. By mixing stuff, we have lost information.

Properties of the density matrix

The density matrix satisfies a bunch of very special properties. We can figure them out using only the definition (2.3) and recalling that $q_i \in [0,1]$ and $\sum_i q_i = 1$ [Eq. (2.1)]. First, the density matrix is a Hermitian operator:

$$\rho^\dagger = \rho. \tag{2.6}$$

Second,

$$\mathrm{tr}(\rho) = \sum_i q_i\,\mathrm{tr}(|\psi_i\rangle\langle\psi_i|) = \sum_i q_i\langle\psi_i|\psi_i\rangle = \sum_i q_i = 1. \tag{2.7}$$

This is the *normalization* condition of the density matrix. Another way to see this is from Eq. (2.4) by choosing $A = 1$. Then, since $\langle 1 \rangle = 1$ we again get $\text{tr}(\rho) = 1$.

We also see from Eq. (2.9) that $\langle \phi|\rho|\phi\rangle$ is a sum of quantum probabilities $|\langle\phi|\psi_i\rangle|^2$ averaged by classical probabilities q_i. This entails the following interpretation: for an arbitrary state $|\phi\rangle$,

$$\boxed{\langle\phi|\rho|\phi\rangle = \text{Prob. of finding the system at state } |\phi\rangle \text{ given that it's state is } \rho} \tag{2.8}$$

Besides normalization, the other big property of a density matrix is that it is **positive semi-definite**, which we write symbolically as $\rho \geq 0$. What this means is that *its sandwich in any quantum state is always non-negative.* In symbols, if $|\phi\rangle$ is an arbitrary quantum state then

$$\langle\phi|\rho|\phi\rangle = \sum_i q_i |\langle\phi|\psi_i\rangle|^2 \geq 0. \tag{2.9}$$

Of course, this makes sense in view of the probabilistic interpretation of Eq. (2.8). Please note that this does *not* mean that all entries of ρ are non-negative. Some of them may be negative. It does mean, however, that the diagonal entries are always non-negative, no matter which basis you use.

Another equivalent definition of a positive semi-definite operator is one whose **eigenvalues are always non-negative**. In Eq. (2.3) it already looks as if ρ is in diagonal form. However, we need to be a bit careful because the $|\psi_i\rangle$ are arbitrary states and do not necessarily form a basis [which can be seen explicitly in the example in Eq. (2.5)]. Thus, in general, the diagonal structure of ρ will be different. Notwithstanding, ρ is Hermitian and may therefore be diagonalized by some orthonormal basis $|k\rangle$ as

$$\rho = \sum_k p_k |k\rangle\langle k|, \tag{2.10}$$

for certain eigenvalues p_k. Since Eq. (2.9) must be true for any state $|\phi\rangle$ we may choose, in particular, $|\phi\rangle = |k\rangle$, which gives

$$p_k = \langle k|\rho|k\rangle \geq 0.$$

Thus, we see that the statement of positive semi-definiteness is equivalent to saying that the eigenvalues are non-negative. In addition to this, we also have that $\text{tr}(\rho) = 1$, which implies that $\sum_k p_k = 1$. Thus we conclude that the eigenvalues of ρ behave like probabilities:

$$p_k \in [0,1], \qquad \sum_k p_k = 1. \tag{2.11}$$

But they are not the same probabilities q_i. They just behave like a set of probabilities, that is all.

For future reference, let me summarize what we learned in a big box: the basic properties of a density matrix are

$$\boxed{\text{Defining properties of a density matrix:} \qquad \operatorname{tr}(\rho) = 1 \quad \text{and} \quad \rho \geq 0.} \tag{2.12}$$

Any normalized positive semi-definite matrix is a valid candidate for a density matrix.

Purity

Next let us look at ρ^2. The eigenvalues of this matrix are p_k^2 so

$$\operatorname{tr}(\rho^2) = \sum_k p_k^2 \leq 1 \tag{2.13}$$

The only case when $\operatorname{tr}(\rho^2) = 1$ is when ρ is a pure state. In that case it can be written as $\rho = |\psi\rangle\langle\psi|$ so it will have one eigenvalue $p_1 = 1$ and all other eigenvalues equal to zero. Hence, the quantity $\operatorname{tr}(\rho^2)$ represents the **purity** of the quantum state. When it is 1 the state is pure. Otherwise, it will be smaller than 1:

$$\boxed{\text{Purity} = \mathcal{P} := \operatorname{tr}(\rho^2) \leq 1} \tag{2.14}$$

As a side note, when the dimension of the Hilbert space d is finite, it also follows that $\operatorname{tr}(\rho^2)$ will have a lower bound:

$$\frac{1}{d} \leq \operatorname{tr}(\rho^2) \leq 1 \tag{2.15}$$

This lower bound occurs when ρ is the **maximally disordered state**

$$\rho = \frac{\mathbb{I}_d}{d} \tag{2.16}$$

where $\mathbb{I}_d$ is the identity matrix of dimension d.

The von Neumann equation

The time evolution of any ket $|\psi\rangle$ under unitary dynamics is given by Eq. (1.78): $|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle$. Any density operator may be written in the form (2.3) so its time evolution will be

$$\rho(t) = \sum_i q_i e^{-iHt}|\psi_i(0)\rangle\langle\psi_i(0)|e^{iHt} = e^{-iHt}\rho(0)e^{iHt}.$$

Differentiating with respect to t we then get

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = (-iH)e^{-iHt}\rho(0)e^{iHt} + e^{-iHt}\rho(0)e^{iHt}(iH) = -iH\rho(t) + i\rho(t)H$$

Thus, we reach von Neumann's equation:

$$\boxed{\frac{\mathrm{d}\rho}{\mathrm{d}t} = -i[H,\rho], \qquad \rho(t) = e^{-iHt}\rho(0)e^{iHt}} \tag{2.17}$$

This is equivalent to Schrödinger's equation, but written in terms of the density matrix. Thus, in this sense, it is more general than Schrödinger's equation. But, of course, that is a bit of an exaggeration since if we solve one, we solve the other. Also, in practice, the best thing is just solve for U. That is, to compute $U = e^{-iHt}$. Then it doesn't matter if we apply this to a ket or a density matrix.

2.2 Bloch's sphere and coherence

The density matrix for a qubit will be 2×2 and may therefore be parametrized as

$$\rho = \begin{pmatrix} p_+ & q \\ q^* & p_- \end{pmatrix}. \tag{2.18}$$

where normalization requires that $p_+ + p_- = 1$. If the state is pure then it can be written as $|\psi\rangle = a|0\rangle + b|1\rangle$, in which case the density matrix becomes

$$\rho = |\psi\rangle\langle\psi| = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}. \tag{2.19}$$

This is the density matrix of a system which is in a *superposition* of $|0\rangle$ and $|1\rangle$. Conversely, we could construct a state which can be in $|0\rangle$ **or** $|1\rangle$ with different probabilities. According to the very definition of the density matrix in Eq. (2.3), this state would be

$$\rho = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1| = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}. \tag{2.20}$$

This is a classical state, obtained from classical probability theory. The examples in Eqs. (2.19) and (2.20) reflect well the difference between quantum superpositions and classical probability distributions.

We can also make this more operational by defining the idea of **coherence**. Coherence is a basis dependent concept. Given a certain basis (here the computational basis), we say a state is *incoherent* if it is diagonal in that basis:

$$\rho = \sum_i p_i |i\rangle\langle i|. \tag{2.21}$$

Any state that is not diagonal contains a certain amount of coherence (with respect to the chosen basis). Coherence plays a central role in quantum information and has seen a boom of interest in the last few years. It is seen as the ultimate **quantum resource**, from which other resources such as entanglement (which will be discussed below) can be extracted. It is also a central concept in the transition from quantum to classical.

The origin of this transition lies in the *interaction of a system with its environment*, a process known as **decoherence**. Decoherence is the process through which the system looses the coherence (off-diagonal entries) in pure states such as (2.19) to end up in mixed states such as (2.20). All system are in contact with their surroundings, so this process is really unavoidable. In fact, even if the system is in the middle of space, with nothing around it, it will still be in contact with the **electromagnetic vacuum**. For this reason, isolating quantum systems turns out to be quite a difficult task and constitutes an active topic of research nowadays. In fact, decoherence is not something that can be avoided. All we can hope to do is to delay it long enough to be able to do something interesting with our state.

When we first learn about quantum mechanics, we use Schrödinger's equation to solve, for instance, for the Hydrogen atom, from which we find the allowed energy levels for the electron. In principle quantum mechanics would allow us to place an electron into a superposition of these eigenstates. However, we *never* do that: I am sure you never heard someone saying "the electron is in a superposition of 1s and 3p" [which would represent a state of the form (2.19)]. Instead, we always assume a situation similar to Eq. (2.20). For instance, when we use the Hydrogen energy levels to construct (approximately) the periodic table, we simply "put" the electrons into the energy levels, stacking them one on top of each other. The reason why this works is because for electrons the decoherence rate is usually really high, so that even if we were to prepare the electron in a pure superposition state, it would quickly decohere to a mixture.

You may also be wondering, if coherence is a basis dependent concept, then how can we talk about decoherence as a physical effect, independent of *our* choice of basis? I mean, what makes one basis more important than another. The answer, again, is the environment. The environment and, in particular, the system-environment interaction, singles out a specific basis.

Models of decoherence are usually based on **master equations**. This name which, admittedly, is a bit weird, corresponds to modifications of the von Neumann Eq. (2.17) to include also the effects of the environment. There is an infinite number of different such models and we will explore them in this course in quite some detail. Here I want to study the simplest model of a master equation, called the **dephasing noise**. We consider a qubit with some Hamiltonian $H = \frac{\Omega}{2}\sigma_z$ and subject it to the equation

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -i[H,\rho] + \frac{\gamma}{2}(\sigma_z\rho\sigma_z - \rho) \tag{2.22}$$

The first part is Eq. (2.17) and the second part describes the action of the environment. In this case γ is a constant specifying the *coupling strength* to the environment: if you set $\gamma \to 0$ you isolate your system and recover the unitary dynamics of the von Neumann Eq. (2.17).

The solution of Eq. (2.22) is really easy. We assume that the density matrix can be parametrized as in Eq. (2.18), with $p_\pm(t)$ and $q(t)$ being just time-dependent functions. Plugging this into Eq. (2.22) and doing all the silly 2×2 matrix multiplications we get the following system of equations:

$$\frac{\mathrm{d}p_\pm}{\mathrm{d}t} = 0 \tag{2.23}$$

$$\frac{\mathrm{d}q}{\mathrm{d}t} = -(i\Omega + \gamma)q \tag{2.24}$$

Thus, we see that the dephasing noise does not change the populations, but only affects the off-diagonal elements. In particular, the solution for q is

$$q(t) = e^{-(i\Omega+\gamma)t} q(0) \tag{2.25}$$

so that $|q|$ decays exponentially. Thus, the larger the coupling strength γ the faster the system looses coherence. If we start in a pure state such as (2.19) then after a time t the state will be

$$\rho(t) = \begin{pmatrix} |a|^2 & ab^* e^{-(i\Omega+\gamma)t} \\ a^* b e^{-(-i\Omega+\gamma)t} & |b|^2 \end{pmatrix}. \tag{2.26}$$

Thus, after a sufficiently long time, coherence will be completely lost and eventually the system will reach

$$\lim_{t\to\infty} \rho(t) = \begin{pmatrix} |a|^2 & 0 \\ 0 & |b|^2 \end{pmatrix}. \tag{2.27}$$

This is no longer a superposition, but simply a classical statistical mixture, just like (2.20). The action of the environment therefore changed quantum probabilities into classical probabilities.

The above example illustrates well I think, how the environment acts to destroy quantum features. Of course, that is not the only way that the bath can act. The dephasing model is very special in that it does not change the populations. In general, the interaction with the environment will induce both decoherence *and* changes in populations. A qubit model which captures that is the **amplitude damping** described by the master equation

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -i[H, \rho] + \gamma\left[\sigma_- \rho \sigma_+ - \frac{1}{2}\{\sigma_+\sigma_-, \rho\}\right] \tag{2.28}$$

where $\{A, B\} = AB - BA$ is the anti-commutator. I will leave the study of this model for you as an exercise. As you will find, in this case $q(t)$ continues to decay exponentially, but $p_\pm$ change as well.

Let us now go back to Bloch's sphere. So far we have seen that a pure state can be described as point on the surface of a unit sphere. Now let us see what happens for mixed states. Another convenient way to write the state (2.18) is as

$$\boxed{\rho = \frac{1}{2}(1 + \boldsymbol{s}\cdot\boldsymbol{\sigma}) = \frac{1}{2}\begin{pmatrix} 1+s_z & s_x - i s_y \\ s_x + i s_y & 1 - s_z \end{pmatrix}.} \tag{2.29}$$

where $\boldsymbol{s} = (s_x, s_y, s_z)$ is a vector. The physical interpretation of $\boldsymbol{s}$ becomes evident from the following relation:

$$s_i = \langle \sigma_i \rangle = \text{tr}(\sigma_i \rho). \tag{2.30}$$

The relation between these parameters and the parametrization in Eq. (2.18) is

$$\begin{aligned}
\langle \sigma_x \rangle &= q + q^*, \\
\langle \sigma_y \rangle &= i(q - q^*), \\
\langle \sigma_z \rangle &= p_+ - p_-.
\end{aligned}$$

Next we look at the purity of a qubit density matrix. From Eq. (2.29) one also readily finds that

$$\text{tr}(\rho^2) = \frac{1}{2}(1 + \boldsymbol{s}^2). \tag{2.31}$$

Thus, due to Eq. (2.13), it also follows that

$$\boldsymbol{s}^2 = s_x^2 + s_y^2 + s_z^2 \leq 1. \tag{2.32}$$

When $\boldsymbol{s}^2 = 1$ we are in a pure state. In this case the vector $\boldsymbol{s}$ lays on the surface of the Bloch sphere. For mixed states $\boldsymbol{s}^2 < 1$ and the vector is inside the Bloch sphere. Thus, we see that the purity can be directly associated with the radius in Bloch's sphere. The smaller the radius, the more mixed is the state. In particular, the maximally disordered state occurs when $\boldsymbol{s} = 0$ and reads

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.33}$$

In this case the state lies in the center of the sphere. A graphical representation of pure and mixed states in the Bloch sphere is shown in Fig. 2.1.

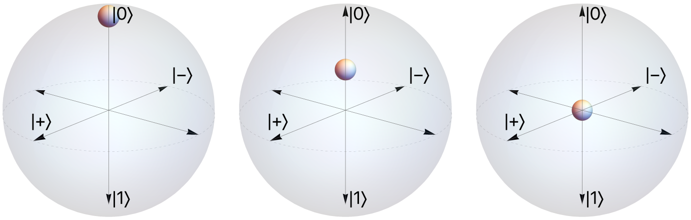

Figure 2.1: Examples of pure and mixed states in the z axis. Left: a pure state. Center: an arbitrary mixed state. Right: the maximally mixed state (2.33).

2.3 Composite systems and the almighty kron

So far we have considered only a single quantum system described by a basis $|i\rangle$. Now we turn to the question of how to describe mathematically a system