Chapter 3

Reduced and mixed quantum states

So far we have represented quantum states as a vector $|\psi\rangle$. Density operators, whereby a quantum state is represented by a matrix ρ, is an alternative formalism for representing quantum states. In particular, this perspective will allow us to- i. handle classical uncertainty as well as quantum uncertainty in a single formalism and ii. extract the state of part of a quantum system from knowledge of its composite system.

3.1 Density operators

The *density operator* corresponding to a state $|\psi\rangle$ is given by the matrix $\rho = |\psi\rangle\langle\psi|$. The average value of an observable O in the state ρ is then given by:

$$\langle O\rangle = \mathrm{Tr}(\rho O) = \mathrm{Tr}(|\psi\rangle\langle\psi|O) = \langle\psi|0|\psi\rangle \tag{3.1}$$

Where we have applied the cyclicity of the trace directly to Eq. (3.1) to see that this does indeed give the same expectation value as the standard state vector formalism.

Examples 3.1.1. 1. The state $|\psi\rangle = |1\rangle$ is a pure state of the system, and the corresponding density operator ρ is given by:

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

2. The state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ is written in density matrix form as:

$$\rho = \begin{pmatrix} \frac{1}{2} & \pm\frac{1}{2} \\ \pm\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

3. <u>Bell States:</u> The density operator corresponding to the Bell state $|\Psi_-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ is given by:

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

Density operators open up a new perspective on the Bloch sphere. To see this first note that density operator of a single qubit $\cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$ can be written as

$$\rho = |\psi\rangle\langle\psi| = \begin{pmatrix} \cos(\theta/2)^2 & \cos(\theta/2)\sin(\theta/2)e^{-i\phi} \\ \cos(\theta/2)\sin(\theta/2)e^{i\phi} & \sin(\theta/2)^2 \end{pmatrix} \tag{3.2}$$

Next we note that any 2×2 matrix can be written as a weighted sum of the Pauli matrices,

$$\{\sigma_i\}_{i=0}^3 := \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \tag{3.3}$$

because Pauli matrices form an orthogonal basis with the following helpful properties

$$\mathrm{Tr}[\sigma_0] = 1, \mathrm{Tr}[\sigma_k] = 0 \ \text{ for } \ k \neq 0 \tag{3.4}$$

$$\mathrm{Tr}[\sigma_j \sigma_k] = 2\delta_{k,j} \ . \tag{3.5}$$

It follows that we can write

$$\rho = \frac{1}{2}\sigma_0 + \frac{1}{2}\sum_{i=1}^{3} v_i \sigma_i \tag{3.6}$$

where the factor of 2 is to account for the factor of 2 in $\mathrm{Tr}[\sigma_j \sigma_k] = 2\delta_{k,j}$.

Next we ask, what is the significance of the vector $\mathbf{v} = (v_1, v_2, v_3)$ in Eq. (3.6). To answer this - we first note that it follows from the properties of the Pauli matrices (namely, $\mathrm{Tr}[\sigma_j \sigma_k] = 2\delta_{k,j}$) that $v_i = \mathrm{Tr}[\rho\sigma_i]$. It follows that the vector $\mathbf{v}$ is a vector of the expectation values of the Pauli observables:

$$\mathbf{v} = \begin{pmatrix} \langle\sigma_x\rangle \\ \langle\sigma_y\rangle \\ \langle\sigma_z\rangle \end{pmatrix} . \tag{3.7}$$

Alternatively, one can verify by direct comparison of Eq. (3.2) and Eq. (3.6) (check this for yourself!) that

$$\mathbf{v} = \begin{pmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{pmatrix} . \tag{3.8}$$

That is, the vector $\mathbf{v}$ is the unit Bloch vector which can be used to represent a quantum state on the Bloch sphere.

Thus far this switch in representation may seem rather arbitrary. We have provided a different perspective on quantum states but not done anything more. The real power of this formalism will be made clearer in the following two sections.

3.1.1 Pure states and mixed states

Suppose someone prepares a system S in the state $|\psi\rangle$ with probability p and state $|\phi\rangle$ with probability $1-p$ by tossing a biased coin, how would we mathematically represent the state of the system S? We we want a mathematical entity that allows us to correctly compute the expectation value of any observable O. Now we know from basic probability that the expectation of O should be

$$\begin{aligned} \langle O\rangle &= p\langle O\rangle_\psi + (1-p)\langle O\rangle_\phi \\ &= p\,\langle\psi|\,O\,|\psi\rangle + (1-p)\,\langle\phi|\,O\,|\phi\rangle \\ &= p\,\mathrm{Tr}(|\psi\rangle\langle\psi|O) + (1-p)\,\mathrm{Tr}(|\phi\rangle\langle\phi|O) \\ &:= \mathrm{Tr}(\rho O) \end{aligned}$$

where we have used $\mathrm{Tr}(|\psi\rangle\langle\psi|O) = \langle\psi| O |\psi\rangle$ and in the final line defined

$$\rho := p|\psi\rangle\langle\psi| + (1-p)|\phi\rangle\langle\phi| \,. \tag{3.9}$$

That is, the density operator ρ allows us to compute any expectation value for the system described above where the system was prepared in the state $|\psi\rangle$ with probability p and state $|\phi\rangle$ with probability $1-p$.

More generally, if a system is prepared in state $|\psi_k\rangle$ with probability p_k it can be described by the density operator

$$\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k| \,.$$

Such states are known as *statistical mixtures* or as *mixed* states. In contrast a state where the exact quantum state is known (i.e. all states studied until now) are known as *pure* quantum states.

How does a generic single qubit mixed state look on the Bloch sphere? To study this we start by recalling Eq. (3.6) and writing

$$\begin{aligned}
|\psi\rangle\langle\psi| &= \frac{1}{2}\sigma_0 + \frac{1}{2}\sum_{i=1}^{3} v_i\sigma_i \\
|\phi\rangle\langle\phi| &= \frac{1}{2}\sigma_0 + \frac{1}{2}\sum_{i=1}^{3} u_i\sigma_i \,.
\end{aligned}$$

Then we note that the mixed state

$$\begin{aligned}
\rho &= p|\psi\rangle\langle\psi| + (1-p)|\phi\rangle\langle\phi| \\
&= \frac{1}{2}\sigma_0 + \frac{1}{2}\sum_{i=1}^{3}(pv_i + (1-p)u_i)\sigma_i \,.
\end{aligned}$$

That is, the mixed state has a Bloch vector

$$\mathbf{w} = p\mathbf{v} + (1-p)\mathbf{u} \tag{3.10}$$

composed from the weighted convex combination of the Bloch vectors of the original pure state Bloch vectors. This is when the geometric representation provided by the Bloch sphere really comes into its own. If one already knows the original Bloch vectors, it is basic geometry to sketch the new Bloch vector for the corresponding mixed state (see Fig. 3.1).

While pure states have a Bloch vector of norm 1 and sit on the outside of the Bloch sphere, mixed states fall within the Bloch sphere. This follows immediately from the observation that $\mathbf{w} = p\mathbf{v} + (1-p)\mathbf{u}$. Unless $p = 0$, $p = 1$ or $\mathbf{v} = \mathbf{u}$ (which correspond to pure states), the vector $\mathbf{w}$ will point to some point in the interior of the Bloch sphere with $|\mathbf{w}|^2 = p^2 + (1-p)^2 + 2p(1-p)\mathbf{u}.\mathbf{v} = 1 - 2p(1-p)(1-\mathbf{u}.\mathbf{v}) \leq 1$.

A good physical example of a mixed state is that of a thermal state. A thermal state of a Hamiltonian H at inverse temperature $\beta = 1/k_B T$ can be written as

$$\rho = \frac{e^{-\beta H}}{Z} \tag{3.11}$$

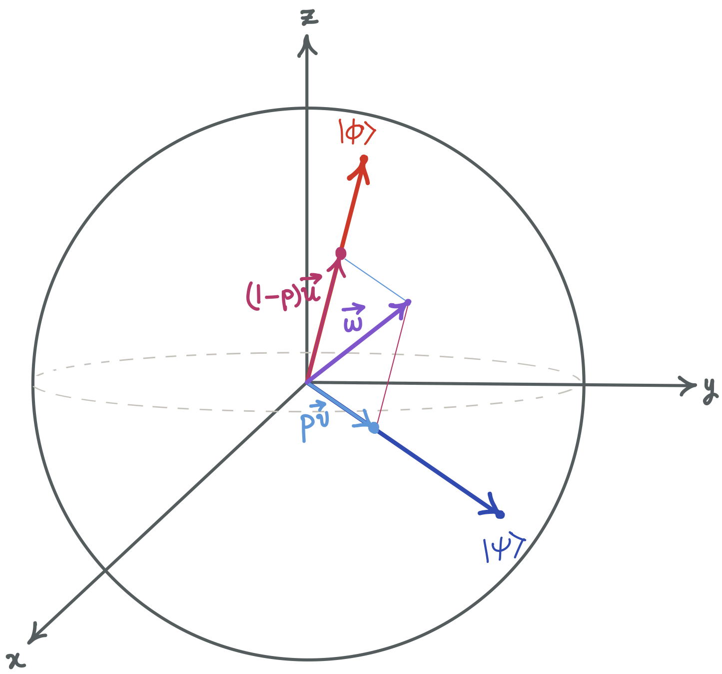

Figure 3.1: **Mixed State.**

where Z is the partition function of the system $Z = \mathrm{Tr}[e^{-\beta H}]$. To see that this reduces to more familiar notions of the thermal state let us expand it in the eigenbasis of $H = \sum_k E_k |E_k\rangle\langle E_k|$. Using the standard definition of the matrix exponential, in this basis we have

$$\begin{aligned} \rho &= \frac{1}{Z}\sum_k e^{-\beta E_k}|E_k\rangle\langle E_k| \\ Z &= \sum_k e^{-\beta E_k}\,. \end{aligned} \tag{3.12}$$

That is, ρ corresponds to a mixed state where the energy eigenstate $|E_k\rangle$ is prepared with the probability $p_k = e^{-\beta E_k}/Z$ which should look familiar as the standard Boltzmann distribution from your statistical mechanics courses. The state $\rho = \frac{e^{-\beta H}}{Z}$ can be treated as any quantum state - you can combine it with other quantum states, evolve it unitarily, perform quantum measurements etc etc. Thus we see that the density matrix formalism allows one to combine classical statistical mechanics and quantum mechanics.

3.1.2 Reduced states

In this course so far we have constructed the state of a composite system from the states of the individual systems using the tensor product. But what if one wants to go in the other direction? Say you are given the state $|\Psi\rangle$ of a 4-dimensional system corresponding to two qubits - how could you describe the state of just one of the qubits? If the state of the composite system is a product state, i.e. $|\Psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, this is straightforward, i.e. the state of A is just $|\psi_A\rangle$. But what if $|\Psi\rangle$ is entangled? For example, what if it's the Bell state $|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$? Now it's no longer clear how to describe the state of the system A alone. Here we show how this question can be addressed using density operators.

Consider a system composed of two subsystems, A and B, and corresponding Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. The core idea is to introduce an operator ρ_A (to be defined!) that one can compute from $|\Psi\rangle$ that will allow one to compute all properties of system A alone. That is, given any measurement operator $O \otimes \mathbb{1}_B$ that acts non-trivially on A alone, we want to define an operator ρ_A, defined on the Hilbert space of $\mathcal{H}_A$ alone, that will allow one to compute the expectation value of O.

To identify such an operator let us first write the operator O in its eigenbasis as $O = \sum_{j=1}^{d_A} \lambda_j |\lambda_j\rangle\langle\lambda|_j$. Now we note that the average value of O is given by

$$\langle O\rangle = \sum_{j=1}^{d_A} \lambda_j P_A(\lambda_j) \tag{3.13}$$

where $P_A(\lambda_j)$ is the probability of getting λ_j when measuring system A. Now this probability can be rewritten in terms of

$$P_{AB}(\lambda_j, k) = \langle\lambda_j k|\rho_{AB}|\lambda_j k\rangle\,, \tag{3.14}$$

the joint probability of finding A to be in $|\lambda_j\rangle$ and system B to be in the state computational basis state[1] $|k\rangle$. Concretely, we have

$$P_A(\lambda_j) = \sum_{k=1}^{d_B} P_{AB}(\lambda_j, k) = \sum_{k=1}^{d_B} \langle\lambda_j k|\rho_{AB}|\lambda_j k\rangle\,. \tag{3.15}$$

Thus we have

$$\begin{aligned}
\langle O\rangle &= \sum_{j=1}^{d_A} \lambda_j \sum_{k=1}^{d_B} \langle\lambda_j k|\rho_{AB}|\lambda_j k\rangle \\
&= \sum_{j=1}^{d_A} \lambda_j \langle\lambda_j| \left(\sum_{k=1}^{d_B} (\mathbb{I}_A \otimes \langle k|)\rho_{AB}(\mathbb{I}_A \otimes |k\rangle) \right) |\lambda_j\rangle \\
&= \sum_{j=1}^{d_A} \lambda_j \langle\lambda_j|\rho_A|\lambda_j\rangle \\
&= \mathrm{Tr}[\rho_A O]
\end{aligned} \tag{3.16}$$

where we have defined

$$\rho_A := \sum_{k=1}^{d_B} (\mathbb{I}_A \otimes \langle k|)\rho_{AB}(\mathbb{I}_A \otimes |k\rangle) \equiv \mathrm{Tr}_B[\rho_{AB}] \tag{3.17}$$

This operator is known as a *reduced state* and is another type of density operator. Note that since the trace of an operator is invariant under a change of basis, the use of a density operator to calculate the average value of O does not depend on the choice of the basis used to define this operator.

It is worthwhile becoming fluent at taking the partial trace of a quantum state. This is usually easiest to do using braket notation rather than working with the explicit matrix forms. To do so, it's helpful to note (prove this to yourself!) that:

$$\mathrm{Tr}_B[|ij\rangle\langle kl|] = |i\rangle\langle k|\,\mathrm{Tr}[|j\rangle\langle l|] \tag{3.18}$$

from which point you can make use of the standard properties of the trace (e.g. cyclicity).

[1] This choice in basis is arbitrary. Any orthogonal basis will do.

Example 3.1.2. The reduced state of $|\psi_A\psi_B\rangle$ is given by the density operator $\rho_A = |\psi_A\rangle\langle\psi_A|$ as one would expect from our arguments at the start:

$$\begin{aligned}\rho_A &= \mathrm{Tr}_B[|\psi_A\psi_B\rangle\langle\psi_A\psi_B|]\\ &= |\psi_A\rangle\langle\psi_A|\mathrm{Tr}_B[|\psi_B\rangle\langle\psi_B|]\\ &= |\psi_A\rangle\langle\psi_A|\langle\psi_B|\psi_B\rangle\\ &= |\psi_A\rangle\langle\psi_A|\,.\end{aligned} \tag{3.19}$$

Example 3.1.3. Consider the Bell state $|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle$. The reduced state on qubit A is given by

$$\begin{aligned}\rho_A &= \mathrm{Tr}_B[|\Phi^+\rangle\langle\Phi^+|]\\ &= \frac{1}{2}\mathrm{Tr}_B[|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|]\\ &= \frac{1}{2}(|0\rangle\langle 0|\mathrm{Tr}[|0\rangle\langle 0|] + |0\rangle\langle 1|\mathrm{Tr}[|0\rangle\langle 1|] + |1\rangle\langle 0|\mathrm{Tr}[|1\rangle\langle 0|] + |1\rangle\langle 1|\mathrm{Tr}[|1\rangle\langle 1|])\\ &= \frac{1}{2}(|0\rangle\langle 0|\langle 0|0\rangle + |0\rangle\langle 1|\langle 0|1\rangle + |1\rangle\langle 0|\langle 0|1\rangle + |1\rangle\langle 1|\langle 1|1\rangle)\\ &= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)\\ &= \frac{1}{2}\mathbb{1}\end{aligned} \tag{3.20}$$

That is, the reduced state on qubit A is the *maximally mixed* state where with equal probability the qubit is in state 0 or state 1. Similarly, $\rho_B = \frac{1}{2}\mathbb{1}$. Crucially we note that

$$\begin{aligned}\rho_A \otimes \rho_B &= \frac{1}{2}\left(|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| + |11\rangle\langle 11|\right)\\ &\neq |\Phi^+\rangle\langle\Phi^+|\,.\end{aligned} \tag{3.21}$$

Thus we see that if you look at only one half of a Bell state the statistical outcomes are no different to tossing a fair coin. But the state of two fair coins is not the same as a Bell state. The interesting behaviour of a Bell state can only be captured by studying the correlations between both systems and captured by the pure state $|\Phi^+\rangle$.

Exercise: Use the notion of a reduced state to argue that entanglement cannot be used for faster than light signalling.

3.1.3 General properties of density operators

Above we have presented two different ways of obtaining mixed state density operators: by direct construction or as the reduced state of a larger system. More generally, density operators can be introduced more abstractly as any operator with the following properties.

Property 3.1.4.

1. *The density operator is self-adjoint, that is to say,* $\rho_A^\dagger = \rho_A$*,*
2. $\mathrm{Tr}(\rho_A) = \sum_i \rho_{ii} = \sum_{i,\mu} |\alpha_{i,\mu}|^2 = \|\psi\|^2 = 1$,
3. *The density operator is positive semidefinite, i.e.* $\langle\phi|\rho_A|\phi\rangle \geq 0$ *for all* $|\phi\rangle \in A$.

It is straightforward to show that the reduced states introduced above satisfy these properties.

Demo. 1. We have:

$$\rho_{ij} = \sum_\mu \alpha^*_{i,\mu}\alpha_{j,\mu}$$
$$\rho_{ji} = \sum_\mu \alpha^*_{j,\mu}\alpha_{i,\mu}$$

One should see that

$$\rho_{ij} = \overline{\rho_{ji}}$$

2. We compute:

$$\begin{aligned}\sum_i \rho_{ii} = \sum_i\sum_\mu \alpha^*_{i,\mu}\alpha_{i,\mu} &= \sum_i\sum_\mu \langle i\mu|\psi\rangle\langle\psi|i\mu\rangle \\ &= \sum_{i,\mu} |\langle i\mu|\psi\rangle|^2\end{aligned}$$

The $|i\rangle$ and $|\mu\rangle$ form a basis of A and B, respectively. Thus, the sum over i and μ give the norm of $|\psi\rangle$, which is by definition normalized to 1.

3. We compute :

$$\begin{aligned}\langle\phi|\rho_A|\phi\rangle &= \sum_{i,j}\sum_\mu \langle\phi|i\rangle\langle j|\phi\rangle\langle i\mu|\psi\rangle\langle\psi|j\mu\rangle \\ &= \sum_\mu \beta_\mu\beta^*_\mu \\ &= \|\beta\|^2 \geq 0,\end{aligned}$$

where $\beta_\mu = \langle\phi|i\rangle\langle i\mu|\psi\rangle$

□

Notice that these properties imply, in particular:

- There exists a basis in which ρ_A is diagonal (from point 1),
- Furthermore, points 2 and 3 impose a particular form on the diagonal representation of the operator ρ_A:

$$\rho_A = \sum_j p_j\,|j\rangle\langle j|\,,$$

where $p_j \geq 0$ and $\sum p_j = 1$. Thus,

$$\langle O\rangle = \mathrm{Tr}(\rho_A O) = \sum_j p_j\,\langle j|O|j\rangle = \sum_j p_j\,\langle O\rangle_{|j\rangle}\,,$$

where $\langle O\rangle_{|j\rangle}$ denotes the average value of O for the subsystem consisting of state $|j\rangle$.

We note that if a density operator describes a pure state, then it is a projector, i.e., $\rho^2 = \rho$. In fact, the two properties are equivalent: if $\rho^2 = \rho$, the eigenvalues of the density operator must necessarily be 0 or 1. But since the sum of the eigenvalues of a density operator must be equal to 1, there must be a single eigenvalue of the density operator that equals 1, and it is unique.

On the other hand, if ρ is not pure then we have $\rho^2 \neq \rho$ and $\mathrm{Tr}[\rho^2] \leq 1$. To see this we consider writing in its eigenbasis as $\rho = \sum_k \lambda_k|\psi_k\rangle\langle\psi_k|$. It follows that $\rho^2 = \sum_k \lambda_k^2|\psi_k\rangle\langle\psi_k| \neq \rho$ and

$\text{Tr}[\rho^2] = \sum_k \lambda_k^2$ and this is less than 1 unless $\{\lambda_k\} = \{1, 0\}$ which again reduces to the case where $\rho = |\psi_0\rangle\langle\psi_0|$ is a pure state. The quantity $\text{Tr}[\rho^2]$ is known as the *purity* of a state - it takes its maximal value of 1 for a pure state and is less than 1 otherwise. An alternative way of showing that mixed states live within the interior of the Bloch sphere is to establish that the condition that $\text{Tr}[\rho^2] \le 1$ implies that the norm of the Bloch vector is less that 1, i.e. $|\mathbf{w}| \le 1$. We leave this as an exercise for the reader.

3.1.4 Evolution of density operators

Let's consider a density operator in diagonal form at $t = 0$:

$$\rho(t = 0) = \sum_j \alpha_j |\psi_j(0)\rangle \langle\psi_j(0)|$$

We are interested in determining the laws governing its time evolution. We assume that the statistical mixture does not change over time. In other words, α_i does not depend on t, and

$$\rho(t) = \sum_j \alpha_j |\psi_j(t)\rangle \langle\psi_j(t)| \,.$$

The time evolution of a state has already been characterized as:

$$|\psi_j(t)\rangle = e^{-iHt} |\psi_j(0)\rangle$$

Using these two equations, we obtain:

$$\rho(t) = \sum_j \alpha_j e^{-iHt} |\psi_j(0)\rangle \langle\psi_j(0)| \, e^{iHt}$$

We differentiate:

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= \sum_j \alpha_j \, (-iH) \, e^{-iHt} |\psi_j(0)\rangle \langle\psi_j(0)| \, e^{-iHt} \\ &+ \sum_j \alpha_j e^{-iHt} |\psi_j(0)\rangle \langle\psi_j(0)| \, (iH) \, e^{iHt} \\ &= (-iH) \, \rho + \rho \, (iH)\end{aligned}$$

which leads to the equation:

$$i\frac{\partial\rho}{\partial t} = [\hat{H}, \rho], \tag{3.22}$$

describing the time evolution of the density operator. Note that while this may look like the Heisenberg equation, ρ does not define an observable physical quantity!