Mixed States

Now we're ready to start extending the axioms. The principal "problem" with the axioms is that they describe a closed system, one completely isolated from its environment or any other physical system.

Consider two systems A and B in an arbitrary state $|\psi\rangle_{AB}$ and measurement of an observable M_A on system A. The expectation value of M_A is

$$\langle M_A\rangle = {}_{AB}\langle\psi|M_A\otimes\mathbb{1}_B|\psi\rangle_{AB} = \sum_{kk'\ell\ell'}\psi^*_{k'\ell'}\,{}_A\langle k'|M_A|k\rangle_A\,{}_B\langle\ell'|\ell\rangle_B\,\psi_{k\ell}$$

$$= \sum_{kk'}{}_A\langle k'|M_A|k\rangle_A\sum_\ell\psi^*_{k'\ell}\psi_{k\ell} = \mathrm{Tr}\Big[M_A\sum_{kk'}|k\rangle\langle k'|\sum_\ell\psi^*_{k'\ell}\psi_{k\ell}\Big]$$

$$= \mathrm{Tr}[M_A\rho_A] \qquad \rho_A = \sum_{kk'}|k\rangle\langle k'|\sum_\ell\psi^*_{k'\ell}\psi_{k\ell} = \mathrm{Tr}_B\big[|\psi\rangle\langle\psi|_{AB}\big]$$

Thus we can define a quantity pertaining only to system A which allows us to calculate all probabilities and expectation values on A.

$\mathrm{Tr}_B[\cdot]$ is the partial trace — just the trace over the B system:

$\mathrm{Tr}[M_{AB}] = \sum_i {}_B\langle i|M_{AB}|i\rangle_B$ for any ON basis $\{|i\rangle_B\}$. So we have

$$\rho_A = \mathrm{Tr}_B[|\psi\rangle\langle\psi|] = \mathrm{Tr}_B\Big[\sum_{\substack{kk'\\ \ell\ell'}}\psi^*_{k'\ell'}\psi_{k\ell}|k\rangle\langle k'|_A\otimes|\ell\rangle\langle\ell'|_B\Big]$$

$$= \sum_{\substack{kk'\\ \ell\ell'}}\psi^*_{k'\ell'}\psi_{k\ell}|k\rangle\langle k'|_A\otimes\underbrace{\mathrm{Tr}_B\big[|\ell\rangle\langle\ell'|_B\big]}_{\delta_{\ell\ell'}} = \sum_{kk'}|k\rangle\langle k'|\sum_\ell\psi^*_{k'\ell}\psi_{k\ell}$$

ρ_A is called the density operator or density matrix and has the following properties:

1. $\rho_A = \rho_A^\dagger$

2. $\rho_A \geq 0$, that is $\langle\varphi|\rho_A|\varphi\rangle \geq 0$ for any $|\varphi\rangle$

3. $\mathrm{Tr}[\rho_A] = 1$

1. $\rho_A^\dagger = \sum_{kk'} |k\rangle\langle k'|^\dagger \left(\sum_\ell \psi_{k'\ell}^* \psi_{k\ell}\right)^* = \sum_{kk'} |k'\rangle\langle k| \sum_\ell \psi_{k'\ell}\psi_{k\ell}^* = \rho_A$

2. $\langle\varphi|\rho_A|\varphi\rangle = \sum_{kk'} \langle\varphi|k\rangle\langle k'|\varphi\rangle \psi_{k'\ell}^* \psi_{k\ell} = \sum_k |\psi_{k\ell}\langle\varphi|k\rangle|^2 \geq 0$

3. $\mathrm{Tr}[\rho_A] = \sum_{k\ell} |\psi_{k\ell}|^2 = \langle\psi|\psi\rangle = 1.$

States as we defined them originally have $\rho_A = |\psi\rangle\langle\psi|_A$ and are called pure states. We may also think of them as subsystems of a separate state, since if $|\psi\rangle_{AB} = |\xi\rangle_A |\eta\rangle_B$, then $\rho_A = \mathrm{Tr}_B\left[|\xi\rangle\langle\xi|_A \otimes |\eta\rangle\langle\eta|_B\right] = |\xi\rangle\langle\xi|_A$.

Pure states have only nonzero eigenvalue, and satisfy $\mathrm{Tr}[\rho_A^2] = 1$. Density operators with more than one nonzero eigenvalue are called mixed states, since they are mixtures of their eigenvectors: $\rho_A = \sum_\ell p_\ell |\varphi_\ell\rangle\langle\varphi_\ell|$

This is also called an incoherent superposition, since unlike in a coherent superposition, there's no phase relation between the constituent states.

For instance, compare the qubit states $\rho_1 = \frac{1}{2}\mathbb{1}$ and $\rho_2 = |\hat{x}\rangle\langle\hat{x}|$

Both have $\Pr(\pm\hat{z}) = \frac{1}{2}$ and can be thought of as superpositions of the states $|\pm\hat{z}\rangle$. However $\Pr(\hat{x}|\rho_1) = \frac{1}{2}$ while $\Pr(\hat{x}|\rho_2) = 1$; ρ_1 is an incoherent superposition, while ρ_2 is a coherent superposition.

We can think of a mixed state ρ as an <u>ensemble</u> of possible pure states $|\varphi_\ell\rangle$, each occurring with probability p_ℓ, using the eigendecomposition $\rho = \sum_\ell p_\ell |\varphi_\ell\rangle\langle\varphi_\ell|$ mentioned above. This works because the probability rule is linear. Thus for any expectation value

$$\langle M\rangle = \mathrm{Tr}[M\rho] = \sum_\ell p_\ell \langle\varphi_\ell|M|\varphi_\ell\rangle = \sum_\ell p_\ell \langle M\rangle_{\varphi_\ell}$$

So we can regard $\langle M\rangle$ as the weighted average of the $\langle M\rangle_{\varphi_\ell}$ we would obtain for state $|\varphi_\ell\rangle$.

Looking back to the analogy with classical probability theory it becomes apparent that density operators are the proper quantum version of $\vec{p}$. This holds for two reasons. First, just as $\vec{p}$ can be regarded as a convex combination of sharp distributions, so too are density operators mixtures of pure states. Pure states are pure, and sharp distributions sharp, because they cannot be expressed as a nontrivial convex combination of other states/distributions.

Secondly, neither for unsharp $\vec{p}$ nor for mixed ρ can we find an event which is certain to occur, i.e. $\Pr(E) = 1$.

Returning to the qubit case for a moment, we can immediately see that the possible density operators lie inside the Bloch sphere. First, any ρ is Hermitian, so we can express it as $\rho = r_0(\mathbb{1} + \vec{r}\cdot\vec{\sigma})$ But $\mathrm{Tr}[\rho]=1$ so $r_0 = \frac{1}{2}$. And $\rho \geq 0$, meaning both eigenvalues must be positive, and therefore $1 \pm |\vec{r}| \geq 0$. In other words, the set of possible $\vec{r}$, which are in 1-1 correspondence with allowed qubit density operators, forms a ball in $\mathbb{R}^3$. It's usually called the Bloch sphere even though the sphere is just the surface, which corresponds to the pure states.

Although density operators form a convex set like classical distributions, they do not form a <u>simplex.</u> That is, the convex decomposition of a given mixed state is not unique. This is easy to see on the Bloch sphere (ball): pick an $\vec{r}$ with $|\vec{r}|<1$ corresponding to ρ and then choose any line segment connecting two points on the surface which includes $\vec{r}$. Calling the end points $\vec{s}_1$ and $\vec{s}_2$, it follows that $\vec{r} = \lambda\vec{s}_1 + (1-\lambda)\vec{s}_2$ for some λ with $0 \leq \lambda \leq 1$. Then $\rho = \lambda\rho_1 + (1-\lambda)\rho_2$ Since

$$\lambda\rho_1 + (1-\lambda)\rho_2 = \frac{1}{2}\left(\mathbb{1} + (\lambda\vec{s}_1 + (1-\lambda)\vec{s}_2)\cdot\vec{\sigma}\right) = \frac{1}{2}(\mathbb{1} + \vec{r}\cdot\vec{\sigma}) = \rho.$$

As there are many such lines, the decomposition of ρ is not unique.

For a general density operator ρ on a d-dimensional state space, there are many ensembles $\{p_k, |\theta_k\rangle\}_{k=1}^n$ such that

$$\rho = \sum_{k=1}^{n} p_k |\theta_k\rangle\langle\theta_k|$$

In general, two ensemble decompositions $\{p_k, |\varphi_k\rangle\}$ and $\{q_k, |\psi_k\rangle\}$ of a state ρ are related by a unitary matrix U_{jk}:

$$\sqrt{q_j}\,|\psi_j\rangle = \sum_k U_{jk}\sqrt{p_k}\,|\varphi_k\rangle \quad \text{for all } j.$$

This statement is known as the HJW (Houghston-Jozsa-Wootters) theorem, was already known to Schrödinger, and can be derived in a straightforward manner.

We'll take a more informative route here and see that it all goes back to the singular value decomposition, by way of the <u>purification</u> of a mixed state.

The notion of a density operator was derived by starting with a bipartite pure state and using the partial trace to construct the state of only one subsystem. Can we turn this around and regard any density operator as the partial trace of a larger pure state? The answer is yes. Such a state is called a purification of ρ.

It's easy to construct purifications from ensembles: Given an ensemble description of ρ_A as $\{p_k, |\theta_k\rangle_A\}_{k=1}^n$, i.e.

$$\rho_A = \sum_{k=1}^n p_k |\theta_k\rangle\langle\theta_k|, \quad \text{we can}$$

simply invent an additional system B, having dimension n, and define $|\psi\rangle_{AB} = \sum_{k=1}^n \sqrt{p_k}\,|\theta_k\rangle_A |k\rangle_B$. Clearly $\mathrm{Tr}_B[\psi_{AB}] = \rho_A$.

Regarding a mixed state as part of a larger pure state is done all the time in quantum information theory and is called "going to the Church of the Larger Hilbert Space".

The Schmidt decomposition of a purification $|\psi\rangle_{AB}$ of ρ_A is in fact related to the eigendecomposition of ρ_A itself. Consider $|\psi\rangle_{AB} = \sum_k \sqrt{\lambda_k}\,|\xi_k\rangle_A |\eta_k\rangle_B$ such that $\rho_A = \mathrm{Tr}_B\left[|\psi\rangle\langle\psi|_{AB}\right]$

But $\mathrm{Tr}_B\left[|\psi\rangle\langle\psi|_{AB}\right] = \sum_k \lambda_k |\xi_k\rangle\langle\xi_k|$ since $\{|\eta_k\rangle\}$ is an ON basis, which we may use to compute the partial trace. Then, since $\{|\xi_k\rangle\}$ is an ON basis, $\{\lambda_k, |\xi_k\rangle\}$ must be the eigenvalues and eigenvectors of ρ_A.

The Schmidt decomposition then implies that any two purifications of a state ρ_A must be related by a unitary operation on the purifying system B. Suppose $|\psi\rangle_{AB}$ and $|\psi'\rangle_{AB}$ are two purifications of ρ_A. Their Schmidt forms must be

$|\psi\rangle_{AB} = \sum_k \sqrt{\lambda_k}\,|\xi_k\rangle_A |\eta_k\rangle_B$ $\quad |\psi'\rangle_{AB} = \sum_k \sqrt{\lambda_k}\,|\xi_k\rangle_A |\eta_k'\rangle_B$ by the above.

Both $\{|\eta_k\rangle\}$ and $\{|\eta_k'\rangle\}$ are ON bases, so they must be related by some unitary transformation U : $U|\eta_k\rangle = |\eta_k'\rangle$. Thus it follows immediately that $|\psi'\rangle_{AB} = U_B|\psi\rangle_{AB}$.

Now we'd like to prove the HJW theorem. From two ensemble decompositions $\rho_A = \sum_k p_k |\varphi_k\rangle\langle\varphi_k| = \sum_k q_k |\psi_k\rangle\langle\psi_k|$

we can construct two purifications

$$|\theta_1\rangle_{AB} = \sum_k \sqrt{p_k}\, |\varphi_k\rangle_A |k\rangle_B \quad \text{and} \quad |\theta_2\rangle_{AB} = \sum_k \sqrt{q_k}\, |\psi_k\rangle_A |k\rangle_B .$$

By the above argument, $|\theta_2\rangle_{AB} = U_B |\theta_1\rangle_{AB}$ for some unitary U_B. But we then have

$$_B\langle j|\theta_2\rangle_{AB} = \sum_k \sqrt{q_k}\, |\psi_k\rangle \underbrace{\langle j|k\rangle}_{\delta_{jk}} = \sqrt{q_j}\, |\psi_j\rangle \quad , \quad \text{and}$$

$$_B\langle j|\theta_2\rangle_{AB} = \sum_k \sqrt{p_k}\, |\varphi_k\rangle \langle j|U_B|k\rangle_B = \sum_k U_{jk} \sqrt{p_k}\, |\varphi_k\rangle,$$

which proves the theorem.

Do purifications exist classically? That is, given a distribution $\vec{p}_A$, can we find a sharp joint distribution on AB whose marginal is $\vec{p}_A$? No. A sharp distribution $\vec{p}_{AB}$ has components $(p_{AB})_{jk} = \delta_{jj'}\delta_{kk'}$ for some j' and k', where j refers to A and k to B. The marginal is clearly $p_j = \delta_{jj'}$, which is itself sharp. Another, more instructive way to see this is to regard classical probability theory as the special case of quantum probability theory in which the density matrix is always diagonal in some preordained basis. $\vec{p}$ goes over to ρ via $\rho = \sum_k p_k |k\rangle\langle k|$. Purifying ρ, we see that it is not diagonal in the $|jk\rangle_{AB}$ basis, and therefore doesn't exist in classical probability theory.