any observable $\mathbf{L}$ in state $|\psi\rangle$ is given by

$$\langle\psi|\mathbf{L}|\psi\rangle = Tr\ |\psi\rangle\langle\psi|\ \mathbf{L}. \tag{7.12}$$

Here are the steps to prove it. Pick any basis $|i\rangle$. Then, using the definition of *trace*, write

$$Tr\ |\psi\rangle\langle\psi|\ \mathbf{L} = \sum_i \langle i|\psi\rangle\langle\psi|\mathbf{L}|i\rangle.$$

The two factors in the summation are just numbers, so we can reverse their ordering,

$$Tr\ |\psi\rangle\langle\psi|\ \mathbf{L} = \sum_i \langle\psi|\mathbf{L}|i\rangle\langle i|\psi\rangle.$$

Carrying out the sum and using $\sum|i\rangle\langle i| = I$, we get

$$Tr\ |\psi\rangle\langle\psi|\ \mathbf{L} = \langle\psi|\mathbf{L}|\psi\rangle.$$

The right side is just the expectation value of $\mathbf{L}$.

7.3 Density Matrices: A New Tool

Up to now, we have learned how to make predictions about a system when we know the system's exact quantum state. But more often than not, we don't have complete knowledge of the state. For example, suppose Alice has prepared a spin using an apparatus oriented along some axis. She gives the spin to Bob but doesn't tell him the axis along which the apparatus was oriented. Perhaps she gives him some partial

information, such as the fact that the axis was either along the z axis or the x axis, but she refuses to tell him more than that. What does Bob do? How does he use this information to make predictions?

Bob reasons as follows: If Alice prepared the spin in the state $|\psi\rangle$, then the expectation value of any observable $\mathbf{L}$ is

$$Tr\ |\psi\rangle\langle\psi|\mathbf{L} = \langle\psi|\mathbf{L}|\psi\rangle.$$

On the other hand, if Alice prepared the spin in state $|\phi\rangle$, then the expectation value of $\mathbf{L}$ is

$$Tr\ |\phi\rangle\langle\phi|\mathbf{L} = \langle\phi|\mathbf{L}|\phi\rangle.$$

What if there is a 50 percent probability that she prepared $|\psi\rangle$ and a 50 percent probability that she prepared $|\phi\rangle$? Obviously, the expectation value is

$$\langle\mathbf{L}\rangle = \frac{1}{2}Tr\ |\psi\rangle\langle\psi|\mathbf{L} + \frac{1}{2}Tr\ |\phi\rangle\langle\phi|\mathbf{L}.$$

All we are doing is averaging over Bob's ignorance of the state prepared by Alice.

But now we can combine the terms into a single expression by defining a density matrix ρ that encodes Bob's knowledge. In this case the density matrix is half the projection operator onto $|\phi\rangle$ plus half the projection operator onto $|\psi\rangle$,

$$\rho = \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\phi\rangle\langle\phi|.$$

We've now packaged all of Bob's knowledge of the system into a single operator ρ. At this point, the rule to compute expectation values becomes very simple:

$$\langle L \rangle = Tr\ \rho \mathbf{L}. \tag{7.13}$$

We can generalize this. Suppose that Alice tells Bob that she has prepared one of several states—call them $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$, and so on. Moreover, she specifies probabilities P_1, P_2, $P_3, \ldots$ for each of these states. Bob can still package all his knowledge into a density matrix:

$$\rho = P_1|\phi_1\rangle\langle\phi_1| + P_2|\phi_2\rangle\langle\phi_2| + P_3|\phi_3\rangle\langle\phi_3| + \ldots.$$

Furthermore, he can use exactly the same rule, Eq. 7.13, to compute the expectation value.

When the density matrix corresponds to a single state, it is a projection operator that projects onto that state. In this case, we say that the state is *pure*. A pure state represents the maximum amount of knowledge that Bob can have of a quantum system. But in the more general case, the density matrix is a mix of several projection operators. We then say that the density matrix represents a *mixed* state.

I have used the term *density matrix*, but strictly speaking, ρ is an operator. It only becomes a matrix when a basis is chosen. Suppose we choose the basis $|a\rangle$. The density matrix is just the matrix representation of ρ with respect to this basis:

Don't understand how this equation comes up!

$$\rho_{aa'} = \langle a|\rho|a'\rangle.$$

If the matrix representation of $\mathbf{L}$ is $L_{a',a}$ then 7.13 takes the form

And also don't understand how this equation comes up!

$$\langle \mathbf{L} \rangle = \sum_{a,a'} L_{a',a}\rho_{a,a'}. \tag{7.14}$$

7.4 Entanglement and Density Matrices

Classical physics also has its notion of pure and mixed states, although they are not called by those names. Just to illustrate, let's consider a system of two particles moving along a line. According to the rules of classical mechanics, we can calculate the orbits of the particles if we know the values of their positions (x_1 and x_2) and momenta (p_1 and p_2) at a certain instant in time. The state of the system is thus specified by four numbers: x_1, x_2, p_1, and p_2. If we know these four numbers, we have as complete a description of the two-particle system as it is possible to have: there is no more to know. We can call this a pure classical state.

Often, however, we don't know the exact state, but only some probabilistic information. That information can be encoded in a probability density

$$\rho(x_1, x_2, p_1, p_2).$$

A classical pure state is just a special case of a probability