

Lecture 9: density operator formalism

General formalism

We saw in the previous lecture that the Schrödinger equation for the wavefunction

$$\frac{\partial \psi}{\partial t} = -i\hat{H}\psi \quad \Leftrightarrow \quad \frac{\partial}{\partial t}|\psi\rangle = -i\hat{H}|\psi\rangle \quad (1)$$

is relatively straightforward to solve for both static and time-dependent Hamiltonians:

$$|\psi(t)\rangle = \exp[-i\hat{H}t]|\psi(0)\rangle \quad |\psi(t)\rangle = \hat{T} \exp\left(-i\int_0^t \hat{H}(t) dt\right)|\psi(0)\rangle \quad (2)$$

We will now build an equivalent formalism based on the dynamics of the corresponding projection operator, commonly called the *density operator*

$$\hat{\rho}(t) = |\psi(t)\rangle\langle\psi(t)| \quad (3)$$

It is an operator because it can act upon a wavefunction and return another wavefunction:

$$\hat{\rho}|\varphi\rangle = |\psi\rangle\langle\psi|\varphi\rangle = a|\psi\rangle, \quad a = \langle\psi|\varphi\rangle \quad (4)$$

As defined in Equation (3), it is also an idempotent operator, that is:

$$\hat{\rho}^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \hat{\rho} \quad (5)$$

The equation of motion for $\hat{\rho}(t)$ (called *Liouville – von Neumann equation*) is easily obtained from the definition in Equation (3) and the Schrödinger equation:

$$\begin{aligned} \frac{\partial}{\partial t}\hat{\rho}(t) &= \frac{\partial}{\partial t}(|\psi(t)\rangle\langle\psi(t)|) = \left(\frac{\partial}{\partial t}|\psi(t)\rangle\right)\langle\psi(t)| + |\psi(t)\rangle\left(\frac{\partial}{\partial t}\langle\psi(t)|\right) = \left\{ \begin{array}{l} \text{use Schrodinger} \\ \text{equation} \end{array} \right\} \\ &= -i\hat{H}(t)|\psi(t)\rangle\langle\psi(t)| + i|\psi(t)\rangle\langle\psi(t)|\hat{H}(t) = -i\hat{H}\hat{\rho} + i\hat{\rho}\hat{H} = -i[\hat{H}(t), \hat{\rho}(t)] \\ &\frac{\partial}{\partial t}\hat{\rho}(t) = -i[\hat{H}(t), \hat{\rho}(t)] \end{aligned} \quad (6)$$

This equation of motion inherits its solutions from the Schrödinger equation:

$$\begin{aligned} |\psi(t)\rangle &= \exp[-i\hat{H}t]|\psi(0)\rangle \quad \Rightarrow \quad |\psi(t)\rangle\langle\psi(t)| = \exp[-i\hat{H}t]|\psi(0)\rangle\langle\psi(0)|\exp[+i\hat{H}t] \\ \hat{\rho}(t) &= \exp[-i\hat{H}t]\hat{\rho}(0)\exp[+i\hat{H}t] \\ \hat{\rho}(t) &= \hat{T} \exp\left(-i\int_0^t \hat{H}(t) dt\right)\hat{\rho}(0)\hat{T} \exp\left(+i\int_0^t \hat{H}(t) dt\right) \end{aligned} \quad (7)$$

The calculation of observables similarly undergoes only a cosmetic alteration (*N.B.* the observable operators in spin dynamics are often non-Hermitian):

$$\begin{aligned} \langle A \rangle &= \langle\psi|\hat{A}|\psi\rangle = \sum_{ij} \psi_i^* A_{ij} \psi_j = \sum_{ij} (\psi_i^* \psi_j) A_{ij} = \sum_{ij} \rho_{ij} A_{ij} = \\ &= \sum_{ij} \rho_{ij} A_{ji}^* = \text{Tr}(\hat{\rho}\hat{A}^\dagger) = \text{Tr}(\hat{A}^\dagger\hat{\rho}) \\ &\langle A \rangle = \text{Tr}(\hat{A}^\dagger\hat{\rho}) \end{aligned} \quad (8)$$

Computing the dynamics of a specific observable leads to the equation that is identical to the one we had already derived a few lectures ago from the Schrödinger equation:

$$\begin{aligned}\frac{\partial}{\partial t}\langle A \rangle &= \text{Tr}\left(\hat{A}^\dagger \frac{\partial}{\partial t} \hat{\rho}\right) = \text{Tr}\left(-i\hat{A}^\dagger [\hat{H}, \hat{\rho}]\right) = -i\text{Tr}\left(\hat{A}^\dagger \hat{H} \hat{\rho} - \hat{A}^\dagger \hat{\rho} \hat{H}\right) = \\ &= -i\text{Tr}\left(\hat{A}^\dagger \hat{H} \hat{\rho} - \hat{H} \hat{A}^\dagger \hat{\rho}\right) = i\text{Tr}\left([\hat{H}, \hat{A}]^\dagger \hat{\rho}\right) = i\langle [\hat{H}, \hat{A}] \rangle\end{aligned}\quad (9)$$

It does therefore appear that the formalism around the density operator defined in Equation (3) is in every respect equivalent to the original wavefunction formalism. There are two primary reasons why this description is often preferred in spin dynamics simulations:

1. $|\psi\rangle$ does not survive the ensemble average, whereas $|\psi\rangle\langle\psi|$ does:

$$\overline{|\psi\rangle} = \frac{1}{2\pi} \int_0^{2\pi} e^{i\varphi} |\psi\rangle d\varphi = 0, \quad \overline{|\psi\rangle\langle\psi|} = \frac{1}{2\pi} \int_0^{2\pi} e^{i\varphi} |\psi\rangle\langle\psi| e^{-i\varphi} d\varphi = |\psi\rangle\langle\psi| \quad (10)$$

2. $|\psi\rangle$ becomes unacceptably large for $\sim 10^{23}$ spins, whereas the above mentioned ensemble average provides a serviceable substitute to the explicit description for $|\psi\rangle\langle\psi|$.

For the moment, however, we will not carry out any ensemble averaging and will treat the density operator formalism as an exact substitute for the wavefunction formalism developed in the previous lecture. The physical meaning of $\hat{\rho}(t)$ can be glimpsed from the expressions for its matrix elements. The diagonal elements

$$\langle n|\hat{\rho}|n\rangle = \langle n|\psi\rangle\langle\psi|n\rangle = |\langle\psi|n\rangle|^2 = |c_n|^2 = p_n \quad (11)$$

correspond to the probability of finding the system in a state $|n\rangle$ and the off-diagonal elements indicate the presence of a superposition in the wavefunction, the corresponding matrix element being equal to the product of the coefficients in the superposition:

$$\begin{aligned}|\psi\rangle &= \dots + c_n |n\rangle + c_k |k\rangle + \dots \\ \hat{\rho} = |\psi\rangle\langle\psi| &= (\dots + c_n |n\rangle + c_k |k\rangle + \dots)(\dots + c_n^* \langle n| + c_k^* \langle k| + \dots) = \\ &= \dots |c_n|^2 |n\rangle\langle n| + c_n c_k^* |n\rangle\langle k| + c_k c_n^* |k\rangle\langle n| + |c_k|^2 |k\rangle\langle k| + \dots \\ &\quad \langle n|\hat{\rho}|k\rangle = c_n c_k^*\end{aligned}\quad (12)$$

Equation (11) suggests that the trace of the density matrix as defined by Equation (3) should always be unity. This is indeed the case, but this is often inconvenient for numerical reasons, and a scaled unit matrix is commonly subtracted out to leave a traceless operator. The new definition therefore is:

$$\hat{\sigma} = \hat{\rho} - k\hat{1} \quad \text{Tr}(k\hat{1}) = 1 \quad (13)$$

This does not affect the system evolution under Equation (6), or any calculation results, because the unit operator is time-independent and commutes with the Hamiltonian:

$$\frac{\partial}{\partial t} [\hat{\sigma}(t) + k\hat{1}] = \frac{\partial}{\partial t} \hat{\sigma}(t) = -i[\hat{H}(t), \hat{\sigma}(t) + k\hat{1}] = -i[\hat{H}(t), \hat{\sigma}(t)] \quad (14)$$

and because all observable operators (as we saw in the previous lectures) are traceless, and therefore the $k\hat{1}$ term does not contribute to any observables either:

$$\langle A \rangle = \text{Tr}(\hat{A}^\dagger \hat{\rho}) = \text{Tr}(\hat{A}^\dagger [\hat{\sigma} + k\hat{1}]) = \text{Tr}(\hat{A}^\dagger \hat{\sigma}) + k\text{Tr}(\hat{A}^\dagger \hat{1}) = \text{Tr}(\hat{A}^\dagger \hat{\sigma}) \quad (15)$$

This unit trace term will become important when we get to relaxation theory, but for now we will forget about it and consider the density matrix traceless.

Treatment of composite systems

We know that for a system composed of two non-interacting uncorrelated subsystems the total wavefunction is a direct product of the wavefunctions of the individual subsystems:

$$|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle \quad (16)$$

We can also easily prove two useful properties of direct products:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD) \quad \text{Tr}(A \otimes B) = \text{Tr}(A)\text{Tr}(B) \quad (17)$$

For the density matrices we therefore have:

$$\hat{\rho}_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}| = (|\psi_A\rangle \otimes |\psi_B\rangle)(\langle\psi_A| \otimes \langle\psi_B|) = (|\psi_A\rangle\langle\psi_A|) \otimes (|\psi_B\rangle\langle\psi_B|) = \hat{\rho}_A \otimes \hat{\rho}_B \quad (18)$$

The Hamiltonian built differently – every part must affect its own subsystem, but must leave the other subsystem intact (*i.e.* perform unit operator action). The Hamiltonian of a two-spin system is therefore built as:

$$\hat{H}_{AB} = \hat{H}_A \otimes \hat{E}_B + \hat{E}_A \otimes \hat{H}_B \quad (19)$$

where \hat{E} is the identity operator. With this in place, we only need to know the density matrix description of a single spin, and the rest of the formalism can be built up from there using direct products – spins can be added to the system one by one and Equations (18)-(19) used to update the operator matrices at each step of the procedure.

Operator-state duality

Equation (8) for the expectation value of the observables is, in fact, a scalar product on the space of spin operators. This gives the density operator formalism a convenient property – the density matrix of a system with specific expectation values of orthogonal physical observables is a linear combination of the corresponding operators. For example, if a spin- $\frac{1}{2}$ particle is completely polarized along the Z axis, that is, $|\psi\rangle = |\alpha\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$, then the corresponding density matrix is:

$$\hat{\rho} = |\psi\rangle\langle\psi| = |\alpha\rangle\langle\alpha| = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \left\{ \begin{array}{l} \text{kill the} \\ \text{trace} \end{array} \right\} = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} = \hat{L}_Z \quad (20)$$

and so the density matrix happens to be equal to the \hat{L}_Z operator. It can be similarly shown that $\hat{\rho}$ for the spin oriented along X and Y axes is likewise equal to \hat{L}_X and \hat{L}_Y operators respectively. This is the explanation of the common practice of associating these spin operators with the state of the spin system. When we say that the spin system is “in the \hat{L}_Z state”, this means that its density operator is equal to \hat{L}_Z , and it so happens that this corresponds to the Z orientation of the actual spin.

Superoperators

Various linear transformations of the density operator may be thought of as operations defined on the space of operators. The set of all such operations (which are distinguished by a double hat) is called the *superoperator space*. We will consider it in much greater detail in Term II – here we will only mention one specific type of superoperators called *commutation superoperators*, defined as:

$$\hat{\hat{H}}\hat{\rho} = [\hat{H}, \hat{\rho}] = \hat{H}\hat{\rho} - \hat{\rho}\hat{H} \quad (21)$$

Clearly, commutation with \hat{H} is a linear transformation on the space of the density operators. Equation (21) allows us to re-cast the fairly cumbersome Equations (6) and (7) in a slightly neater form:

$$\begin{aligned} \frac{\partial}{\partial t}\hat{\rho}(t) = -i\hat{\hat{H}}\hat{\rho}(t) &\Rightarrow \hat{\rho}(t) = \exp\left[-i\hat{\hat{H}}t\right]\hat{\rho}(0) \\ \frac{\partial}{\partial t}\hat{\rho}(t) = -i\hat{\hat{H}}(t)\hat{\rho}(t) &\Rightarrow \hat{\rho}(t) = \hat{T}\exp\left(-i\int_0^t \hat{\hat{H}}(t) dt\right)\hat{\rho}(0) \end{aligned} \quad (22)$$

This is useful in practice because the double-sided multiplication in Equation (7) tends to be numerically inconvenient. Some important processes (notably relaxation, more on this in Term III) cannot be expressed as commutation operations, necessitating the use of the superoperator formalism.