Module I, Lecture 10: Density Operator Formalism

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

$$\frac{\partial \psi}{\partial t} = -i\hat{H}\psi \qquad \Leftrightarrow \qquad \frac{\partial}{\partial t}|\psi\rangle = -i\hat{H}|\psi\rangle \tag{1}$$

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

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

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

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

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

$$\hat{\rho}|\varphi\rangle = |\psi\rangle\langle\psi|\varphi\rangle = a|\psi\rangle, \qquad a = \langle\psi|\varphi\rangle \tag{4}$$

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

$$\hat{\rho}^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \hat{\rho} \tag{5}$$

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:

$$\frac{\partial}{\partial t}\hat{\rho}(t) = \frac{\partial}{\partial t}\left(|\psi(t)\rangle\langle\psi(t)|\right) = \left(\frac{\partial}{\partial t}|\psi(t)\rangle\right)\langle\psi(t)| + |\psi(t)\rangle\left(\frac{\partial}{\partial t}\langle\psi(t)|\right) = \left\{\begin{array}{c}\text{use Schrodinger}\\\text{equation}\end{array}\right\}$$

$$= -i\hat{H}(t)|\psi(t)\rangle\langle\psi(t)| + i|\psi(t)\rangle\langle\psi(t)|\hat{H}(t) = -i\hat{H}\hat{\rho} + i\hat{\rho}\hat{H} = -i\left[\hat{H}(t),\hat{\rho}(t)\right] \tag{6}$$

$$\frac{\partial}{\partial t}\hat{\rho}(t) = -i\left[\hat{H}(t),\hat{\rho}(t)\right]$$

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

$$|\psi(t)\rangle = \exp\left[-i\hat{H}t\right]|\psi(0)\rangle \qquad \Rightarrow \qquad |\psi(t)\rangle\langle\psi(t)| = \exp\left[-i\hat{H}t\right]|\psi(0)\rangle\langle\psi(0)|\exp\left[+i\hat{H}t\right]$$

$$\hat{\rho}(t) = \exp\left[-i\hat{H}t\right]\hat{\rho}(0)\exp\left[+i\hat{H}t\right] \tag{7}$$

$$\hat{\rho}(t) = \hat{\vec{T}}\exp\left(-i\int_0^t \hat{H}(t)dt\right)\hat{\rho}(0)\hat{\vec{T}}\exp\left(+i\int_0^t \hat{H}(t)dt\right)$$

The calculation of observables similarly undergoes only a cosmetic alteration:

$$\langle O\rangle = \langle\psi|\hat{O}|\psi\rangle = \mathrm{Tr}\left(\langle\psi|\hat{O}|\psi\rangle\right) = \mathrm{Tr}\left(\hat{O}|\psi\rangle\langle\psi|\right) = \mathrm{Tr}\left(\hat{O}\hat{\rho}\right) \tag{8}$$

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:

$$\frac{\partial}{\partial t}\langle O\rangle = \mathrm{Tr}\left(\hat{O}\frac{\partial}{\partial t}\hat{\rho}\right) = \mathrm{Tr}\left(-i\hat{O}\left[\hat{H},\hat{\rho}\right]\right) = -i\mathrm{Tr}\left(\hat{O}\hat{H}\hat{\rho} - \hat{O}\hat{\rho}\hat{H}\right) =$$

$$= -i\mathrm{Tr}\left(\hat{O}\hat{H}\hat{\rho} - \hat{H}\hat{O}\hat{\rho}\right) = i\mathrm{Tr}\left(\left[\hat{H},\hat{O}\right]\hat{\rho}\right) = i\left\langle\left[\hat{H},\hat{O}\right]\right\rangle \tag{9}$$

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:

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

2. $|\psi\rangle$ becomes unacceptably large for ~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

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

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:

$$|\psi\rangle = ... + c_n|n\rangle + c_k|k\rangle + ...$$
$$\hat{\rho} = |\psi\rangle\langle\psi| = \left(... + c_n|n\rangle + c_k|k\rangle + ...\right)\left(... + c_n^*\langle n| + c_k^*\langle k| + ...\right) =$$
$$= ...|c_n|^2|n\rangle\langle n| + c_n c_k^*|n\rangle\langle k| + c_k c_n^*|k\rangle\langle n| + |c_k|^2|k\rangle\langle k| + ...$$
$$\langle n|\hat{\rho}|k\rangle = c_n c_k^* \qquad (12)$$

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:

$$\hat{\sigma} = \hat{\rho} - k\hat{E} \qquad k = 1/\dim(\hat{\rho}) \qquad (13)$$

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:

$$\frac{\partial}{\partial t}\left[\hat{\sigma}(t) + k\hat{E}\right] = \frac{\partial}{\partial t}\hat{\sigma}(t) = -i\left[\hat{H}(t), \hat{\sigma}(t) + k\hat{E}\right] = -i\left[\hat{H}(t), \hat{\sigma}(t)\right] \qquad (14)$$

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:

$$\langle O\rangle = \mathrm{Tr}(\hat{O}\hat{\rho}) = \mathrm{Tr}\left(\hat{O}\left[\hat{\sigma} + k\hat{E}\right]\right) = \mathrm{Tr}(\hat{O}\hat{\sigma}) + k\mathrm{Tr}(\hat{O}\hat{E}) = \mathrm{Tr}(\hat{O}\hat{\sigma}) \qquad (15)$$

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.

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-½ particle is completely polarized along the Z axis, that is, $|\psi\rangle = |\alpha\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}^{\mathrm{T}}$, then the corresponding density matrix is:

$$\hat{\rho} = |\psi\rangle\langle\psi| = |\alpha\rangle\langle\alpha| = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \left\{ \begin{matrix} kill\ the \\ trace \end{matrix} \right\} = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} = \hat{L}_{\mathrm{Z}} \tag{16}$$

and so the density matrix happens to be equal to the $\hat{L}_{\mathrm{Z}}$ operator. It can be similarly shown that $\hat{\rho}$ for the spin oriented along X and Y axes is likewise equal to $\hat{L}_{\mathrm{X}}$ and $\hat{L}_{\mathrm{Y}}$ 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 $\hat{L}_{\mathrm{Z}}$ state", this means that its density operator is equal to $\hat{L}_{\mathrm{Z}}$, 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:

$$\hat{\hat{H}}\hat{\rho} = \left[\hat{H}, \hat{\rho}\right] = \hat{H}\hat{\rho} - \hat{\rho}\hat{H} \tag{17}$$

Clearly, commutation with $\hat{H}$ is a linear transformation on the space of the density operators. Equation (17) allows us to re-cast the fairly cumbersome Equations (6) and (7) in a slightly neater form:

$$\frac{\partial}{\partial t}\hat{\rho}(t) = -i\hat{\hat{H}}\hat{\rho}(t) \quad \Rightarrow \quad \hat{\rho}(t) = \exp\left[-i\hat{\hat{H}}t\right]\hat{\rho}(0)$$

$$\frac{\partial}{\partial t}\hat{\rho}(t) = -i\hat{\hat{H}}(t)\hat{\rho}(t) \quad \Rightarrow \quad \hat{\rho}(t) = \hat{T}\exp\left(-i\int_0^t \hat{\hat{H}}(t)dt\right)\hat{\rho}(0) \tag{18}$$

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.

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:

$$|\psi_{\mathrm{AB}}\rangle = |\psi_{\mathrm{A}}\rangle \otimes |\psi_{\mathrm{B}}\rangle \tag{19}$$

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

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD) \qquad \mathrm{Tr}(A \otimes B) = \mathrm{Tr}(A)\mathrm{Tr}(B) \tag{20}$$

For the density matrices we therefore have:

$$\hat{\rho}_{\mathrm{AB}} = |\psi_{\mathrm{AB}}\rangle\langle\psi_{\mathrm{AB}}| = \left(|\psi_{\mathrm{A}}\rangle \otimes |\psi_{\mathrm{B}}\rangle\right)\left(\langle\psi_{\mathrm{A}}| \otimes \langle\psi_{\mathrm{B}}|\right) = \left(|\psi_{\mathrm{A}}\rangle\langle\psi_{\mathrm{A}}|\right) \otimes \left(|\psi_{\mathrm{B}}\rangle\langle\psi_{\mathrm{B}}|\right) = \hat{\rho}_{\mathrm{A}} \otimes \hat{\rho}_{\mathrm{B}} \tag{21}$$

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:

$$\hat{H}_{AB} = \hat{H}_A \otimes \hat{E}_B + \hat{E}_A \otimes \hat{H}_B \qquad (22)$$

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 (21)-(22) used to update the operator matrices at each step of the procedure.

Writing your own simulations – an example

We can now continue building the simulation from the previous lecture. Our first stage was to define the Pauli matrices:

```
% Define Pauli matrices
sigma_x=[0, 1/2; 1/2, 0];
sigma_y=[0, -1i/2; 1i/2, 0];
sigma_z=[1/2, 0; 0, -1/2];
unit=[1 0; 0 1];
```

The second stage was to use the Kronecker product procedure to generate two-spin operators:

```
% Build two-spin operators
Lx=kron(sigma_x,unit); Sx=kron(unit,sigma_x);
Ly=kron(sigma_y,unit); Sy=kron(unit,sigma_y);
Lz=kron(sigma_z,unit); Sz=kron(unit,sigma_z);
```

And the third stage was to build the Hamiltonian:

```
% Build the Hamiltonian
omega_L=2*pi*200; % 200 Hz
omega_S=2*pi*400; % 400 Hz
omega_J=2*pi*40;  % 40 Hz
H=omega_L*Lz+omega_S*Sz+omega_J*(Lx*Sx+Ly*Sy+Lz*Sz);
```

We can now proceed with this simulation – let us take the initial condition to be $\hat{\rho}(0) = \hat{L}_X + \hat{S}_X$, this corresponds to both spins sitting on the X axis, as they would as a result of a 90-degree radiofrequency pulse. For the observable operator, we will pick the total X magnetization for the real part and the total Y magnetization for the imaginary part, that is $\hat{A} = \left(\hat{L}_X + \hat{S}_X\right) + i\left(\hat{L}_Y + \hat{S}_Y\right)$. In Matlab language, this would simply be:

```
% Initial and detection state
rho=Lx+Sx; coil=(Lx+Sx)+1i*(Ly+Sy);
```

where the term "coil" was used for the observable operator. Before computing the propagator, we need to decide the time step in the simulation. To observe the Nyquist condition, we need two points per period of the fastest oscillation in the system. This means that we need to calculate the highest frequency present in the system. That is actually the definition of the Hamiltonian norm, and therefore:

```
% Decide the time step
time_step=1/norm(H);
```

We can now compute the step propagator:

```
% Build the step propagator
P_evol=expm(-1i*H*time_step);
```

The following needs to happen from this point onwards: the system should be propagated forward by one step, then the observable should be calculated and recorded, then the system should be propagated forward by another step, and so on. This is accomplished by the use of the "for" loop:

```
% Time evolution with for 2048 steps
for n=1:2048
    fid(n)=trace(coil.'*rho);
```

```
            rho=P_evol*rho*P_evol';
        end
```

Our free induction decay is now recorded. The steps from this point essentially amount to NMR data processing. The signal needs to be multiplied by a decaying function (this is called apodization):

```
% Apodization
window_function=exp(-5*linspace(0,1,2048))';
fid=fid.*window_function;
```

The statement above used several advanced features of Matlab syntax. The `linspace` function returns an array of 2048 numbers spaced equally between 0 and 1. Calling an exponential of this array computes the exponential of every element. The element-by-element multiplication operation (denoted by dot-star) then multiplies every element of the fid array by the corresponding element of the window function array. This compact notation avoids quite a few loops. Finally, we need to run a Fourier transform (with zerofilling to 8196 points) and ask Matlab to plot the result:

```
% Fourier transform with zerofill
spectrum=fftshift(fft(fid,8196));

% Plotting
plot(real(spectrum));
```

The following picture is produced, with the instantly recognizable two-spin NMR spectrum: