

Introduction to NMR and NMR based qubits

October 21, 2018

1 Spin systems

In this section, we recall some of the basic phenomena associated with spin operators. To start, let us first recall the spin operators in the most basic conceivable non-trivial quantum system - a quantum system described by a two-dimensional Hilbert space, which is at the same time the smallest non-trivial representation of the angular momentum algebra.

The (complexified) Lie algebra $\mathfrak{su}(2)$ of the three-dimensional rotation group is the space of all trace-free and anti-hermitian complex 2×2 matrices (this is a well known fact, see for instance chapter 16 of [4]). In physics, it is custom to use the set of **Pauli spin matrices**

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (1)$$

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (2)$$

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3)$$

which form a basis of the space of trace-free and hermitian matrices so that a basis of the Lie algebra $\mathfrak{su}(2)$ is given by $i\sigma_x, i\sigma_y$ and $i\sigma_z$. These matrices have a few remarkable properties. First, the square of each of these matrices is 1. They all have eigenvalues 1 and -1 and therefore determinant -1 . For their products, we have the relation

$$\sigma_x \sigma_y = i\sigma_z$$

and cyclic extensions. In addition, a short calculation shows that any two different Pauli matrices anti-commute. It follows that the commutation relations are given by

$$[\sigma_x, \sigma_y] = 2i\sigma_z$$

plus again cyclic extensions. A set of matrices that obey the usual commutation relations that physicists use to characterize angular momenta is thus given by

$$I_a = \frac{\hbar}{2} \sigma_a$$

where a denotes one of the indices x, y, z . It is obvious that then

$$[I_x, I_y] = i\hbar I_z$$

with cyclic extensions.

Let us now consider the rotation operators associated with these matrices. Different conventions exist for the signs in these equations, but we follow [7] and define the rotation around the z-axis by the angle ϕ as

$$R_z(\phi) = e^{-\frac{i}{\hbar}\phi I_z} = e^{-\frac{1}{2}i\phi\sigma_z} \quad (4)$$

and similarly for the other axes. These rotation operators are easily computed using the fact that the square of the Pauli matrices is one and by splitting the series for the exponential into the cosine (even powers) and sine (odd powers) part. The cosine of a multiple of a Pauli matrix σ is given by

$$\cos a\sigma = \sum_{k=0}^{\infty} (-1)^k \frac{a^{2k}}{(2k)!} \sigma^{2k} = 1 \cdot \cos a$$

and the sine is given by

$$\sin a\sigma = \sum_{k=0}^{\infty} (-1)^k \frac{a^{2k+1}}{(2k+1)!} \sigma^{2k+1} = \sigma \cdot \sin a$$

This makes it very easy to calculate the rotation matrices based on the Euler identity (which is an identity of power series). For instance

$$\begin{aligned} R_y(\phi) &= e^{-i\frac{\phi}{2}\sigma_y} \\ &= \cos\left[\frac{\phi}{2}\sigma_y\right] - i\sin\left[\frac{\phi}{2}\sigma_y\right] \\ &= 1 \cdot \cos\frac{\phi}{2} - i\sigma_y \sin\frac{\phi}{2} \end{aligned}$$

or, as a matrix

$$R_y(\phi) = \begin{pmatrix} \cos\frac{\phi}{2} & -\sin\frac{\phi}{2} \\ -\sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{pmatrix}$$

When working with two-dimensional Hilbert spaces, it is often useful to be able to visualize the physical states of the system. We know that the space of physical states is the space of rays in the Hilbert space, i.e. the projective space of \mathbb{C}^2 which is topologically equivalent to a 2-sphere S^2 . Let us now introduce a specific parametrization that is known as the **Bloch sphere**.

Suppose we are given a non-zero state $|\psi\rangle$ in our two-dimensional Hilbert space. We can write this state as

$$|\psi\rangle = a|0\rangle + b|1\rangle$$

with complex numbers a and b . Let us also assume that we have normalized this state. As every complex number, both, a and b , can be written as a phase times a non-negative real number. By multiplying the entire state by a phase (which does not change the ray), we can assume that the phase of a is one, i.e. that

$$|\psi\rangle = \alpha|0\rangle + e^{i\varphi}\beta|1\rangle$$

with non-negative real numbers α, β which fulfill the relation

$$\alpha^2 + \beta^2 = 1$$

It is also clear that this representation is unique. Now, due to this relation, (α, β) is a point on the unit circle in the real two-dimensional plane. It can therefore be written as

$$(\alpha, \beta) = \left(\cos \frac{\Theta}{2}, -\sin \frac{\Theta}{2} \right)$$

with an angle Θ . In other words, we can write the vector $|\psi\rangle$ as

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\varphi} \end{pmatrix} R_y(\Theta)|0\rangle = e^{-\frac{1}{2}i\phi} R_z(\Phi) R_y(\Theta)|0\rangle$$

We have therefore parametrized the space of rays by the two angles Φ and Θ , i.e. with spherical coordinates, which gives us the desired identification. Specifically, the state $|0\rangle$ is the north pole. The presentation in terms of rotations shows that a general state is then obtained by first rotating by the angle Θ around the y-axis - so that Θ is the polar angle - and then by the angle ϕ around the z-axis, i.e. ϕ is the azimuth. The vectors of the form

$$|0\rangle + e^{i\phi}|1\rangle$$

are then the points on the equator of the sphere. The angle $\Theta = \pi$ corresponds to $R_y(\pi) = \sigma_x$ and therefore to $|1\rangle$, i.e. the south pole is the vector $|1\rangle$. Finally, applying R_z only changes the angle Φ and therefore corresponds to a rotation around the z-axis (which fixes exactly the north pole $|0\rangle$ and the south pole $|1\rangle$ as expected).

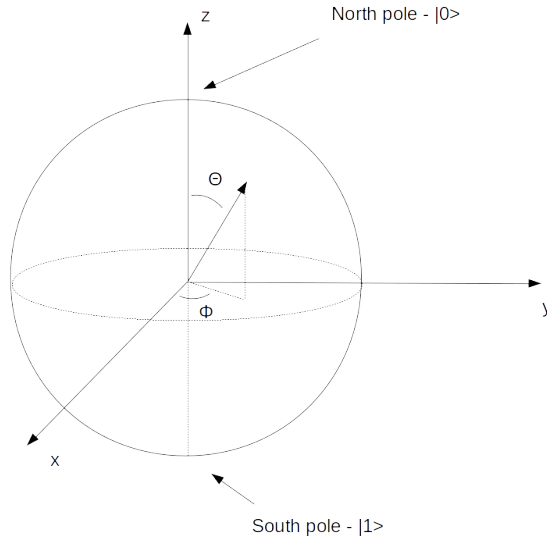


Figure 1: Representation of qubit states on the Bloch sphere

Let us now turn to a physically relevant example that illustrates some of the typical behavior of spin systems. Let V denote the 2-dimensional Hilbert space spanned by the standard basis $|0\rangle$ and $|1\rangle$. Let V_0 denote the Hilbert space of

a physical system into which spin degrees of freedom are not yet incorporated. Then we can add spin to this model by passing to the tensor product

$$V_0 \otimes V$$

If the Hamiltonian of the original system is H_0 and H is a Hamiltonian acting on the spin degrees of freedom alone, i.e. a Hamiltonian on V , then we can use the sum

$$H_0 + H$$

as the Hamiltonian for the composite system. This model is of course only an approximation that ignores interactions between the original system and the spin, but an approximation which works in many cases.

To make things concrete, we can imagine that H_0 is the Hamiltonian describing the spatial degrees of freedom of a hydrogen atom. In this case, V_0 can be taken to be $L^2(\mathbb{R}^3)$ or a suitable Sobolev space. The elements of this space are the familiar wave functions. The new space

$$V_0 \otimes V = L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$$

can then be represented as two-component wave functions

$$\begin{pmatrix} \psi_+(x) \\ \psi_-(x) \end{pmatrix} = \psi_+|0\rangle + \psi_-|1\rangle$$

This is what physicists call a (2-component) **spinor**. The Hamiltonian will then turn into a (generally coupled) system of partial differential equations for the components of the spinor.

Now assume that we expose the system - i.e. in our example the hydrogen atom - to some external magnetic field B . This will add a term

$$H = -\gamma(B_x I_x + B_y I_y + B_z I_z)$$

to the Hamiltonian which is conveniently written as

$$H = -\gamma B \cdot I$$

with a "vector" $I = (I_x, I_y, I_z)$ (it will add other terms as well, like a corresponding term for the rotational angular momentum, but let us ignore this for a moment). For simplicity, let us assume that B is oriented along the z-axis, so that our extra term becomes

$$H = -\gamma B_z I_z = -\frac{\hbar}{2} \gamma B_z \sigma_z$$

In these expressions, γ is a constant (depending on the system) that is called the **gyromagnetic ratio**. The term γI is often called the **magnetic moment**. Thus the statement is that the spin generates a magnetic moment and that this magnetic moment interacts with the external magnetic field.

Let us now see how this term affects the eigenstates and eigenvalues of the Hamiltonian assuming a static field (so that modified the Hamiltonian is again time-independent). First, the matrix σ_z is already diagonal in the standard

basis. If therefore $|\psi_0\rangle$ is an eigenstate of the original system with energy E , we obtain two eigenstates $|\psi_0\rangle|0\rangle$ and $|\psi_0\rangle|1\rangle$ of the new system with energies

$$E \pm -\frac{\hbar}{2}\gamma B_z$$

Thus we see a split of the energy levels, known as the **Zeeman split** in the NMR literature (there is a similar effect for the interaction with the rotational angular momentum, see also [5], chapter 5.2 for a full treatment and a short history).

We can also derive directly how a state changes over time. Let us first calculate the time evolution operator acting on the spin degree of freedom alone. We have

$$e^{-\frac{i}{\hbar}Ht} = e^{-\frac{i}{\hbar}\gamma t B_z I_z} = R_z(\gamma B_z t)$$

As H and H_0 act on different spaces, they commute and the entire time evolution operator is given by

$$U(t)R_z(\gamma B_z t)$$

where $U(t)$ is the time evolution of the original system acting on V_0 and the rotation acts on V . We therefore see that time evolution is described on the Bloch sphere as a precession around the z-axis with frequency

$$\omega = \gamma B_z$$

This is called the **Larmor precession** and the frequency is called the **Larmor frequency**. The energy gap caused by the Zeeman effect is then just $\hbar\omega$.

In some textbooks and introductory material, this fact is described by saying that "the spin axis precesses around the magnetic field direction". This is of course not true - spin is an intrinsic property of a particle and no rotation around some axis in three-dimensional space. Still, this intuition is useful, and the visualization on the Bloch sphere appears to be a more precise version of that statement. The point on the Bloch sphere respectively the vector from the origin to that point is often called the **spin polarization vector**.

2 The basic mechanism of NMR

After these preparations, we are now ready to understand the basic idea of NMR spectrography (see for instance chapter 35 of [1] or the introductory chapters of [7]). We first describe the approach heuristically, and then show later how the machinery of quantum statistics can be used to make this more precise.

NRM spectrography is concerned with the interaction between **nuclear spins** and magnetic fields. Recall that the nucleus of an atom consists of protons and neutrons. Both particle types have spin $1/2$. In a nucleus, the individual spins combine to give an overall spin of the nucleus.

Of course, this spin depends on the state of the nucleus. However, it turns out that the energy difference between the ground state and first excited state tends to be very large. Therefore, in practical applications, we can treat the nucleus as a system with a fixed spin I which is the spin in the ground state, and ignore excited states with different total spin quantum numbers completely. The value of this spin is a property of the element in question. For instance, the

ground state spin of a deuterium nucleus (made up of a proton and a neutron) is 1, the ground state nuclear spin of the main phosphorus isotope ^{31}P is $1/2$ and the ground state spin of ^{12}C is zero.

Now suppose that we place a molecular probe in a strong, static magnetic field along the z-axis - for simplicity, assume that $I = \frac{1}{2}$. If we ignore spatial and other degrees of freedoms for a moment and only focus on the degree of freedom associated with the nuclear spin, we can apply our results from the previous section. The nuclear spin will be subject to a Hamiltonian proportional to σ_z , and the energy levels will split.

Initially, some nuclei will be in "spin up" state and some nuclei will be in the "spin down" state (this is of course not exactly true, most spins will be in random superposition, but we will see that as far as macroscopic observables are involved, we can treat the system as if this were true). If all the nuclei would be fully isolated, each of the states would evolve along a Larmor precession and the z-component of the spin would not change. Now, in reality, the individual systems are not fully isolated, but will exchange energy with other nuclei and the environment. Thus, after some time, the system will reach thermal equilibrium, and the number of spin up states and spin down states are determined by the Boltzmann distribution. Thus the number of nuclei that are in the lower energy state will be slightly higher than the number of nuclei in the higher energy state. This leads to a very small **net magnetization** of the probe.

This magnetization itself is very difficult to measure. But consider what happens if we direct a microwave pulse at the probe which has frequency ω , i.e. equal to the Larmor frequency. Then the photons that constitute the pulse will have energy $\hbar\omega$, which is exactly the energy gap between the low and high energy eigenstates. Thus, some nuclei will change state, and we will observe an **absorption** of energy by the probe. This absorption can be shown to exist only at pulse frequencies very close to the Larmor frequency of the nuclei. If we are able to measure this absorption, we can determine the Larmor frequency and obtain information on the constituents and structure of the probe. This is the basic idea of nuclear magnetic resonance experiments.

3 NMR spectography for a single particle

Let us now try to make this "hand-waving" argument a bit more precise. First, let us study the behavior of a spin-1/2 system under the influence of a short electromagnetic pulse. There are different ways to do this - we can treat the electromagnetic field classically or quantize it - but it turns out that a classical treatment is sufficient to give good results.

To simplify the necessary calculations, it is useful to pass to a different reference frame. We know on general grounds that if we pass from our initial frame of reference to a reference frame rotated by an angle Θ around the z-axis, a physical state that is described by the vector $|\psi\rangle$ in our initial reference frame will now be described by the vector

$$|\widetilde{\psi}\rangle = R_z(-\Theta)|\psi\rangle$$

where the rotation R_z is as above - in fact, this is more or less the definition of a spin-1/2 particle (of course, the sign is subject to conventions). Let us

now choose a specific frame that is rotating around the z-axis in the laboratory frame with a fixed frequency ω_{ref} , so that our transformation rule is

$$|\widetilde{\psi}\rangle = R_z(-\omega_{ref}t)|\psi\rangle = e^{i\frac{\omega_{ref}}{\hbar}tI_z}|\psi\rangle$$

To obtain the equation of motion in this new frame, let us take the time derivative of this equation. Applying the product rule and the Schrödinger equation, we find that

$$\begin{aligned} i\hbar \frac{d}{dt}|\widetilde{\psi}\rangle &= -\hbar \frac{\omega_{ref}}{\hbar} I_z |\widetilde{\psi}\rangle + R_z(-\omega_{ref}t)H|\psi\rangle \\ &= R_z(-\omega_{ref}t)HR_z(\omega_{ref}t)|\widetilde{\psi}\rangle - \omega_{ref}I_z|\widetilde{\psi}\rangle \end{aligned}$$

If we now define a transformed Hamilton operator - which now has an explicit time-dependency - by the formula

$$\widetilde{H} = R_z(-\omega_{ref}t)HR_z(\omega_{ref}t) - \omega_{ref}I_z$$

we can write this as a modified Schrödinger equation

$$i\hbar \frac{d}{dt}|\widetilde{\psi}\rangle = \widetilde{H}|\widetilde{\psi}\rangle$$

This modified Hamiltonian is usually called the **rotating frame Hamiltonian**. Note that the extra term $-\omega_{ref}I_z$ is very similar to the additional forces that appear in classical mechanics when moving into a rotating reference frame - this is formally just an additional fictitious magnetic field similar to a Coriolis force.

There are a few interesting observations that we can directly make. First, let us again specialize to the case that the Hamiltonian is a multiple of I_z , i.e.

$$H = -\omega I_z$$

with the Larmor frequency ω . Then, the Hamiltonian commutes with all rotations around the z-axis, and we obtain the new Hamiltonian

$$\widetilde{H} = -(\omega - \omega_{ref})I_z$$

This is formally again a spin Hamiltonian, corresponding to a Larmor frequency

$$\Omega^0 = \omega - \omega_{ref}$$

So we can, without any further calculations, derive immediately that in the rotating frame, the vector $|\widetilde{\psi}\rangle$ will experience a precession around the z-axis with this frequency. If $\omega = \omega_{ref}$, then the vector is stable. This is not surprising - after all, the transformation was made to achieve this exactly this, note the similarity of this approach to the way how we pass into the interaction picture in scattering theory.

Let us now apply this machinery to see what happens if we add an additional magnetic field that is rotating with frequency ω_{ref} in the x-y-plane. In the laboratory frame, such a field is given by

$$B_{RF} [\cos(\omega_{ref}t + \phi_p)e_x + \sin(\omega_{ref}t + \phi_p)e_y]$$

Here Φ_p is the angle of the rotation at time $t = 0$ (which is, as we will see, relevant). We assume that this field is only applied for some time τ . During this time, the Hamiltonian is then modified by adding the additional term

$$H_{RF} = -\gamma B_{RF} [\cos(\omega_{ref}t + \phi_p)I_x + \sin(\omega_{ref}t + \phi_p)I_y]$$

Now the relations between rotations in append A imply that

$$\cos(\omega_{ref}t + \phi_p)I_x + \sin(\omega_{ref}t + \phi_p)I_y = R_z(\omega_{ref}t + \Phi_p)I_x R_z(-\omega_{ref}t - \Phi_p)$$

Using this, we can calculate the form of the additional in the rotation frame and find that this will change the Hamiltonian in the rotating frame by adding a term

$$\widetilde{H_{RF}} = -\omega_{nut} R_z(\Phi_p) I_x R_z(-\Phi_p)$$

where

$$\omega_{nut} = \gamma B_{RF}$$

is the **nutaton frequency**. Using again the relations from appendix A, we could write this equally well as

$$\widetilde{H_{RF}} = -\omega_{nut} (\cos \Phi_p I_x + \sin \Phi_p I_y)$$

In both cases, the explicit dependence has been removed by passing to the rotating frame representation - again not really surprising. The full Hamiltonian is then given by

$$\widetilde{H} = \Omega^0 I_z - \omega_{nut} (\cos \Phi_p I_x + \sin \Phi_p I_y)$$

Let us now consider a few special cases to see what this implies. First, let us suppose that we create a pulse with frequency ω_{ref} being exactly the Larmor frequency ω , so that $\Omega^0 = 0$, and with $\Phi_p = 0$ - this is called an **x-pulse** in [7]. The Hamiltonian in the rotating frame is then simply

$$\widetilde{H} - \omega_{nut} I_x$$

Thus, the time evolution operator for a duration τ of the pulse is

$$e^{\omega_{nut}\tau \frac{i}{\hbar} I_x} = R_x(-\omega_{nut}\tau)$$

We see that applying an x-pulse for a duration τ rotates the spin polarization axis by an angle $-\omega_{nut}\tau$ around the x-axis on the Bloch sphere. Similarly, for other values of Φ_p , we obtain a rotation around an axis in the x-y-plane, spanning the angle Φ_p with the x-axis. In particular, a pulse of length $\tau = \pi\omega_{nut}^{-1}$ exchanges the states $|0\rangle$ and $|1\rangle$ and *a pulse of half that length rotates the spin polarization axis by 90 degree around the x-axis.*

Let us now study the impact of a pulse with a frequency "off resonance", i.e. for which $\Omega^0 \neq 0$. To describe this, let us introduce the vector

$$\Omega = -\omega_{nut} \cos \Phi_p e_x - \omega_{nut} \sin \Phi_p e_y + \Omega^0 e_z$$

so that

$$\widetilde{H} = \Omega \cdot I$$

The length of this vector is called the effective frequency ω_{eff} and is given by

$$\omega_{eff} = |\Omega| = \sqrt{\omega_{nut}^2 + (\omega_{ref} - \omega)^2}$$

so that we can write

$$\tilde{H} = \omega_{eff} \hat{\Omega} \cdot I$$

where $\hat{\Omega}$ is the unit vector in the direction of Ω . If we apply the pulse again for the time τ , the time evolution operator can be written - again using the formulas in appendix A - as

$$e^{-\frac{i}{\hbar} \tilde{H} \tau} = e^{-\frac{i}{\hbar} \tau \omega_{eff} \hat{\Omega} \cdot I} = \cos \omega_{eff} \frac{\tau}{2} - i(\hat{\Omega} \cdot I) \sin \frac{\tau}{2}$$

If we apply the pulse for the time it would take to flip the state if we were exactly on resonance, i.e. for the time $\tau = \pi \omega_{nut}^{-1}$ and with $\Phi_p = 0$, we get

$$\cos \frac{\omega_{eff} \pi}{2 \omega_{nut}} - i(\hat{\Omega} \cdot \sigma) \sin \frac{\omega_{eff} \pi}{2 \omega_{nut}}$$

Now let $\lambda = \frac{\Omega^0}{\omega_{nut}}$. We can then write

$$\frac{\omega_{eff}}{\omega_{nut}} = \sqrt{1 + \lambda^2}$$

and

$$\hat{\Omega} \cdot \sigma = -\frac{\omega_{nut}}{\omega_{eff}} (\sigma_x - \lambda \sigma_z) = \frac{-1}{\sqrt{1 + \lambda^2}} (\sigma_x - \lambda \sigma_z)$$

Let us now calculate the transition probability for the transition $|0\rangle \rightarrow |1\rangle$. Intuitively, we expect that off resonance, when the energy of the photons that make up the magnetic field is different from the energy gap between the two states, this probability is small. To verify this, we can form the matrix element

$$\begin{aligned} \langle 0 | e^{-\frac{i}{\hbar} \tilde{H} \tau} | 1 \rangle &= \frac{i}{\sqrt{1 + \lambda^2}} \langle 0 | \sigma_x | 1 \rangle \sin \frac{\omega_{eff} \pi}{2 \omega_{nut}} \\ &= \frac{i}{\sqrt{1 + \lambda^2}} \sin \left[\frac{\pi}{2} \sqrt{1 + \lambda^2} \right] \end{aligned}$$

where we have used that there is no contribution from σ_y as $\Phi_p = 0$ and no contribution from σ_z as σ_z is diagonal. The the probability amplitude then becomes

$$P(|0\rangle \rightarrow |1\rangle) = |\langle 0 | e^{-\frac{i}{\hbar} \tilde{H} \tau} | 1 \rangle|^2 = \frac{1}{1 + \lambda^2} \sin^2 \left[\frac{\pi}{2} \sqrt{1 + \lambda^2} \right]$$

Clearly, this function is symmetric, has a maximum with value 1 at $\lambda = 0$ and decays quadratically with growing λ .

It turns out that the width of this peak is related to the relation between the strength of the vertical magnetic field B and the amplitude B_{RF} of the rotating field. In fact, we can write

$$\lambda = \frac{\omega - \omega_{ref}}{\gamma B_{RF}} = \frac{\omega - \omega_{ref}}{\omega} \frac{\omega}{\gamma B_{RF}} = \Delta \frac{\gamma B}{\gamma B_{RF}} = \Delta \frac{B}{B_{RF}}$$

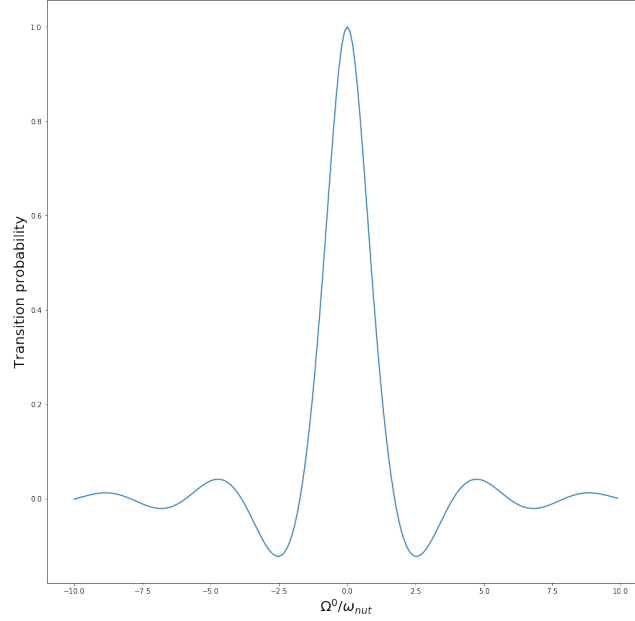


Figure 2: Resonance peak around the Larmor frequency

where Δ is the relative deviation of the rotation frequency from the Larmor frequency. Thus, in the usual setup where the rotating field is much weaker than the horizontal field, then even very small relative deviations will lead to significant values of λ and hence to small transition probabilities. This implies that the time evolution operator is approximately diagonal, with phase factors in the diagonal, and the time evolution will not change the physical states. Thus we will have a very sharp peak around the Larmor frequency.

How is the rotating pulse created in practice? Typical Larmor frequencies for strong static fields are in the order of a few hundred MHz. Obviously, it is almost impossible to create a rotating magnetic by rotating a magnet or coil at that frequency. Instead, one usually generates an oscillating field of the form

$$B \cos(\omega_{ref}t + \Phi_p)e_x$$

Obviously, we can write this as

$$\begin{aligned} & \frac{B}{2} [\cos(\omega_{ref}t + \Phi_p)e_x + \sin(\omega_{ref}t + \Phi_p)e_y] + \\ & \frac{B}{2} [\cos(-\omega_{ref}t - \Phi_p)e_x + \sin(-\omega_{ref}t - \Phi_p)e_y] \end{aligned}$$

This is a superposition of two rotating magnetic fields, one rotating with frequency ω_{ref} , the other one rotating with frequency ω_{ref} , i.e. in the reverse direction. If we now transform this into the rotating frame Hamiltonian, we will get two terms. The first term is the time-independent field that we have seen before, the second one is a term that will oscillate rapidly (with frequency $2\omega_{ref}$).

Now during a typical NMR experiment, the duration of the pulse will be much longer than the time it takes to complete one precession, because the

nutration frequency is much smaller than the Larmor frequency. Therefore the second term will oscillate forth and back many times over the duration of a typical pulse. Now, quite generally,

$$\left| \int_a^b e^{i\omega t} dt \right| = \left| \frac{-i}{\omega} [e^{i\omega t}]_a^b \right| = \frac{|a - b|}{\omega}$$

i.e. when we integrate over a time period covering a large number of oscillations, the integral over a phase factor becomes small - rapid oscillations tend to average out over longer time periods. Thus, we can hope to obtain a reasonable approximation if we ignore the term rotating with frequency $-\omega_{ref}$ - this is called the **rotating wave approximation**. In other words, we hope to obtain a very good approximation to a rotating magnetic field by simply applying a strong oscillating magnetic field, and this is how a real world NMR experiment usually works.

Supported by the calculations done in this section, we are now in a slightly better position to understand what happens during an NMR experiment. Initially, we apply the static magnetic field in the direction of the z-axis. For a nucleus that had a spin polarization vector close to the z-axis, the spin polarization vector will then start to precess with the Larmor frequency around the z-axis. After some time, we apply an oscillating magnetic field with frequency ω_{ref} with length

$$\tau = \frac{\pi}{2} \omega_{nut}^{-1}$$

If ω_{ref} is close to the Larmor frequency, this will turn the spin polarization almost into the x-y-plane. When we turn off the pulse, the precession will resume, but now the spin polarization axis will precess in a plane almost identical to the x-y-plane.

In a semi-classical view, this rotation will again create a rotating magnetic field. This field will induce a current in a nearby coil, and it is this current that can be measured and produce sharp peaks at the Larmor frequency.

However, there is still a problem with this argument. We have assumed that the initial spin polarization is almost along the z-axis. However, there is no reason why this should be the case. If our probe consists of a large number of nuclei, we should expect that the initial spin polarization vectors will be randomly oriented. This will then still be the case after applying the pulse, and it is not clear that the magnetization caused by the precession of each nucleus does not cancel out in the average, but leaves a non-zero net magnetization. To understand how this works, we need to extend our model to describe a large number of nuclei with statistical behavior.

4 NMR spectography for an ensemble

In this section, we will generalize our considerations for a single spin-1/2 particle to the more realistic case that we are working with an entire probe, consisting of a large number N of nuclei. Instead of the wave function of the individual systems given by each nucleus, we follow the usual approach in statistical quantum physics and use the density operator formalism (see appendix B for a short introduction).

First, let us try to determine the state - represented by a density matrix ρ - after turning on the strong magnetic field and waiting for some time until the system has reached some equilibrium state. Here we make the assumption that the quantum systems describing the individual nuclei are of course not fully isolated. Instead, they interact - and they interact with the environment. To model this situation, let us use the model of a canonical ensemble at a constant temperature T , so that the initial state is given by

$$\rho(t=0) = \frac{1}{Z} e^{-\frac{H}{kT}}$$

where again H is the Hamiltonian of the system. In our case, the Hamiltonian (in the laboratory frame) is again

$$H = -\omega I_z = \omega \frac{\hbar}{2} \sigma_z$$

where $\omega = \gamma B$ is the Larmor frequency. Therefore we find that

$$\rho(t=0) = \frac{1}{Z} \exp\left(\frac{1}{2} \frac{\hbar \gamma B}{kT} \sigma_z\right)$$

Let us now introduce the ratio

$$\beta = \frac{\hbar \gamma B}{kT}$$

This is a ratio of two energies. The first energy is the energy of a photon with the Larmor frequency. In the example of a proton at a magnetic field of a few Tesla, this energy is in the order of 10^{-25} Joule. The thermal energy at room temperature (300 K), on the other hand, is roughly $4 \cdot 10^{-21}$ Joule. Thus if the magnetic field is sufficiently strong, the ratio β will be in the order of 10^{-4} . We can therefore use the **high temperature approximation** and replace the exponential series by its first two terms. With this approximation, we have

$$\rho(t=0) = \frac{1}{Z} \left[1 + \frac{1}{2} \beta \sigma_z\right]$$

To determine the partition function Z , we can calculate the trace. As the trace of σ_z vanishes, we find that

$$Z = \frac{1}{2}$$

and obtain

$$\rho(t=0) = \frac{1}{2} + \frac{1}{4} \beta \sigma_z = \begin{pmatrix} \frac{1}{2} + \frac{\beta}{4} & 0 \\ 0 & \frac{1}{2} - \frac{\beta}{4} \end{pmatrix} = \frac{1}{2} + \frac{1}{2} \frac{\beta}{\hbar} I_z$$

Let us now pause for a moment and look at this matrix. The first diagonal element ρ_{00} can be interpreted as the probability to be in the ground state (or, more precisely, as the average expectation value of the operator $|0\rangle\langle 0|$). Assuming the standard case of positive γ , this probability is slightly higher than the probability ρ_{11} to be in the first excited state. This corresponds to the fact that the Hamiltonian - which is proportional to **minus** ωI_z - takes a negative value on this state, i.e. this is the state of lowest energy. We note,

however, that the difference between the probabilities is very small, as we have assumed that β is very small.

We also see that the density matrix is diagonal. This is what we expect for an equilibrium state. In fact, the Liouville-von Neumann equation tells us that for a stationary state ρ , the commutator $[\rho, H]$ vanishes, so the energy eigenstates $|0\rangle$ and $|1\rangle$ are also eigenvectors of ρ . Thus our density matrix is the matrix that we would obtain from a statistical ensemble in which all nuclei are either in a spin-down or a spin-up energy eigenstate. This is quite remarkable - in a real ensemble, we would expect that the spin polarization axis is randomly distributed and almost no nucleus is in an exact eigenstate. The fact that the density matrix in thermal equilibrium is diagonal tells us that as far as macroscopic observables are involved, we can treat the system **as if** all nuclei were either spin up or spin down.

Next, let us compute this state in the rotating frame basis. If a state vector transforms with a unitary transformation U , the density matrix transforms as

$$\rho \mapsto U\rho U^{-1}$$

When we pass from the laboratory reference frame to the rotating frame coordinate system, the transformation U is a rotation around the z-axis. However, these rotations clearly commute with the density operator (which corresponds to the fact that the thermal equilibrium state is invariant with respect to rotations around the z-axis), so that the density operator in the rotating frame is again given by

$$\tilde{\rho}(t=0) = \frac{1}{2} + \frac{1}{2} \frac{\beta}{\hbar} I_z$$

Now let us try to compute the net magnetization of our probe. Each individual spin has a magnetic moment $\mu = \gamma I$. The average magnetic moment can therefore be computed as

$$\langle \mu \rangle = \gamma \text{tr}(\rho I)$$

Now the Pauli matrices and therefore I_x, I_y and I_z are traceless. Therefore, almost all terms in ρI do not contribute to the trace. The only contribution comes from

$$\frac{1}{2} \frac{\beta}{\hbar} I_z^2 = \frac{\beta \hbar}{8}$$

Thus we find that

$$\langle \mu \rangle = \frac{\gamma \beta \hbar}{4} e_z$$

To calculate the resulting magnetic field, we can treat the probe as a magnetic dipole. In general, the relation between the magnetic vector potential produced by a magnetic dipole at the coordinate origin and the magnetic moment μ of the dipole is given by

$$A(r) = \frac{\mu_0}{4\pi} \frac{\mu \times r}{|r|^3}$$

so that the magnetic field is given by

$$B(r) = \frac{\mu_0}{4\pi} \left[\frac{3r(\mu \cdot r)}{|r|^5} - \frac{\mu}{|r|^3} \right]$$

Let us now suppose that we have placed a coil in the x-y-plane, say at a point on the x-axis, at a distance d from our probe. Then $r = de_x$. As the magnetic

moment is parallel to e_z , we will observe a magnetic field in the z-direction with strength

$$\frac{\mu_0 |\mu|}{4\pi d^3} = N \frac{\mu_0 \gamma \beta \hbar}{4\pi 4d^3}$$

where N is the number of particles in the probe. This is a very weak field, many orders of magnitude smaller than the external field but in the same direction and static, so it is very hard to measure.

But there is hope - based on our previous results, we expect that we can rotate this field into the x-y-plane by applying a short pulse. To see how this works, let us now study the evolution of the density matrix. We will again switch into the rotating reference frame and study the time evolution under Hamiltonians of the form

$$\tilde{H} = -\omega_{eff} n \cdot I$$

for a unit vector n - in fact, all Hamiltonians that we need in this section are of this form. We know that $tr(\rho) = 1$, so that the matrix $\rho - \frac{1}{2}$ is traceless and hermitian and therefore be written as a linear combination of Pauli matrices with real coefficients. In other words, we can write

$$\rho(t) = \frac{1}{2} + \frac{1}{2} f(t) \cdot I$$

with a time dependent vector valued function f . Then, using the relations from the appendix, we find that

$$[H, \rho] = [-\omega_{eff} n \cdot I, \frac{1}{2} f(t) \cdot I] = -i\hbar \frac{\omega_{eff}}{2} (n \times f(t)) \cdot I$$

Plugging this into the Liouville-von Neumann equation yields

$$i\hbar \frac{1}{2} \dot{f} \cdot I = -i\hbar \frac{\omega_{eff}}{2} (n \times f(t)) \cdot I$$

i.e.

$$\dot{f} = \omega_{eff} [f \times n]$$

Before we proceed to find solutions for this equation, let us quickly clarify the physical interpretation of the vector f . To do this, let us try to express the expectation value of the magnetic moment in terms of f . For the x-component, we have

$$\begin{aligned} \langle \mu_x \rangle &= \gamma tr(\rho I_x) \\ &= \gamma tr\left(\frac{1}{2} I_x + \frac{1}{2} (f \cdot I) I_x\right) \\ &= \frac{\gamma}{2} tr(f_x I_x^2) = \frac{\gamma \hbar^2}{4} f_x \end{aligned}$$

where we have again used that all angular momentum operators are traceless. Similar calculations work for all other components, so that we obtain

$$\langle \mu \rangle = \frac{\gamma \hbar^2}{4} f$$

Thus, up to a constant, the vector f is nothing but the average expectation value of the magnetic moment. We can therefore multiply our differential equation for the magnetic moment by this constant and obtain

$$\dot{\mu} = \omega_{eff} [\mu \times n]$$

for the macroscopically observed magnetic moment μ . This equation is called the **Bloch equation** (the actual Bloch equation also takes relaxation into account by adding an additional term that we ignore for the time being).

We claim that this equation describes a precession of the magnetic moment around the axis n . To see this, let us split the magnetic moment into a part $\mu_{||}$ parallel to n and a part μ_{\perp} perpendicular to n . Then

$$\dot{\mu}_{||} + \dot{\mu}_{\perp} = \omega_{eff} f(\mu_{\perp} \times n)$$

which gives two equations

$$\begin{aligned}\dot{\mu}_{||} &= 0 \\ \dot{\mu}_{\perp} &= \omega_{eff} f(\mu_{\perp} \times n)\end{aligned}$$

Let now v_x, v_y be two orthogonal vectors in the plane perpendicular to n , oriented such that $v_x \times v_y = n$. We can then write

$$\mu_{\perp}(t) = a(t)v_x + b(t)v_y$$

Let us also assume that v_x and v_y are chosen such that $\mu_{\perp}(t=0) = v_x$. Then

$$\mu_{\perp} \times n = a(t)v_x \times n + b(t)v_y \times n = b(t)v_x - a(t)v_y$$

so that we obtain

$$\begin{aligned}\dot{a} &= \omega_{eff} b(t) \\ \dot{b} &= -\omega_{eff} a(t)\end{aligned}$$

which is solved by

$$\mu_{\perp}(t) = v_x \cos \omega_{eff} t + v_y \sin \omega_{eff} t$$

Thus we see that the vector μ_{\perp} rotates around n with the frequency ω_{eff} , as claimed.

We are now in a position to go through the individual steps of a typical NMR experiment and compute the resulting magnetic moment. First, we initialize the system, i.e. we establish a strong static magnetic field B along the z-axis and wait until the system has settled in thermal equilibrium. We have seen above that in thermal equilibrium, the net magnetic moment of the probe is in the direction of the z-axis, with magnitude

$$\frac{\gamma \beta \hbar}{4} = \frac{\gamma^2 \hbar^2 B}{4kT}$$

Now we turn on a pulse with frequency ω_{ref} . Let us assume that we are exactly on-resonance, i.e. that $\omega_{ref} = \omega$, the Larmor frequency. As we have seen, this results in a Hamiltonian

$$\tilde{H} = -\omega_{nut}(\cos \Phi_p I_x + \sin \Phi_p I_y)$$

This is of the form considered above, with $n = \cos \Phi_p e_x + \sin \Phi_p e_y$ and $\omega_{eff} = \omega_{nut}$. For simplicity, let us assume that $\Phi_p = 0$. Then we find that the effect of the pulse is *a rotation of the net magnetic moment around the x-axis* with frequency ω_{nut} . If we do this for a time τ such that

$$\omega_{nut}\tau = \frac{\pi}{2}$$

the net magnetic moment after applying the pulse will be in the direction of the x-axis, with unchanged magnitude, i.e. the density matrix in the rotating frame will be

$$\tilde{\rho} = \frac{1}{2} + \frac{1}{2} \frac{\beta}{\hbar} I_y$$

and the net magnetization will be

$$\tilde{\mu} = \frac{\gamma^2 \hbar^2 B}{4kT} e_y$$

Note that this is the net magnetic moment in the rotating frame. When we now switch off the pulse again, the Hamiltonian in the rotating frame will become zero again. Thus this new state is again a stable state in the rotating frame. In the laboratory frame, the magnetic moment will therefore be in the direction of the rotating frame, i.e. it will rotate around the z-axis with frequency ω_{ref} . Thus, after the pulse, the magnetic moment in the laboratory frame will be

$$\mu(t) = \frac{\gamma^2 \hbar^2 B}{4kT} [\cos \omega_{ref} t e_x + \sin \omega_{ref} t e_y] + const.$$

where the constant reflects the offset between the rotating frame and the laboratory frame at the time when the pulse is switched off. This will now result in a time-dependent magnetic field

$$B(r) = \frac{\mu_0}{4\pi} \left[\frac{3r(\mu \cdot r)}{|r|^5} - \frac{\mu}{|r|^3} \right]$$

Let us now calculate this for the special case $r = d e_x$. Then

$$r(\mu \cdot r) = r(d\mu_x) = d^2 \mu_x e_x$$

and we find that the x-component of the field is given by

$$B_x(r) = \frac{\mu_0}{4\pi} \frac{2\mu_x}{d^3} = \frac{\mu_0}{4\pi} \frac{1}{d^3} \frac{\gamma^2 \hbar^2 B}{2kT} \cos \omega_{ref} t$$

If we now place a receiver coil at the point $r = d e_x$ along the x-axis, the change of the flux through the coil created by all nuclei in the probe will be

$$-\frac{\mu_0}{4\pi} \frac{1}{d^3} \frac{\gamma^2 \hbar^2 B}{2kT} A \omega_{ref} N \sin \omega_{ref} t$$

where A is the area enclosed by the coil and N is the number of particles in the probe. If the receiver coil has M turns, this will produce a voltage

$$V = \frac{\mu_0}{4\pi} \frac{1}{d^3} \frac{\gamma^2 \hbar^2 B}{2kT} A \omega_{ref} N M \sin \omega_{ref} t$$

in the receiver coil. Using our previous notation and the fact that we are on-resonance, we can also write this as

$$V = \left[\frac{\mu_0}{4\pi} \frac{1}{d^3} \frac{\gamma\beta\hbar}{4} N \right] 2A\omega M \sin \omega_{ref} t$$

The term in the parentheses is the static magnetic field that we have already found to be present in the static state (but directed along the z-axis). We see that the voltage that is induced is many orders of magnitudes higher than the strength of that field, mainly caused by the very rapid oscillations, i.e. by the high Larmor frequency. This voltage is therefore measurable. Thus by applying the pulse to turn the magnetic moment into the x-axis where it then precesses with the very high Larmor frequency, we have turned the very weak net magnetic moment into a measurable quantity!

To get an idea of the magnitudes, let us do an example. Assume that the static magnetic field has 11.74 Tesla, and that our substance is water. Then the Larmor frequency of the nucleus (a proton) will be

$$\omega \approx 500 \text{ MHz}$$

For room temperature, i.e. $T = 300 \text{ K}$, the energy ratio β is then

$$\beta \approx 8 \cdot 10^{-5}$$

The static magnetic field in the thermal equilibrium caused by the difference between the populations of the ground state and the excited state at a distance $d = 0.1$ from the probe is then

$$B_z \approx 3.4 \cdot 10^{-11} \text{ T}$$

per mol of substance. If our receiver coil has 10 turns and an area of $A = 10^{-4} \text{ m}^2$, this will lead to an induced voltage with an amplitude of

$$V \approx 0.2 \text{ mV}$$

oscillating with 500 MHz, again per mol of substance in the probe (here we count the number of protons, which is twice the number of water molecules, so that 6g of water corresponds to roughly N_A protons, i.e. one mol of substance in this sense). This is small, but measurable.

In our calculations, the magnetic moment will precess around the z-axis forever, corresponding to a constant density matrix in the rotating frame of reference. In reality, this is of course not true. Semi-classically, the rotating magnetic dipole emits energy as a electromagnetic field, so the rotation will slow down. Quantum mechanically, we will again have an interaction with the environment (remember that we modelled the probe as a canonical ensemble, i.e. as being in contact with a heat reservoir). Thus the density matrix will change, and the system will slowly return to equilibrium. This process is called **relaxation**. This is typically modelled by introducing two **relaxation times**. The first time of relaxation is called **transverse relaxation** and denoted by T_2 . This relaxation time determines how fast the off-diagonal elements of the density matrix return to zero (these off-diagonal matrix elements are often called **coherences**). Recall that right after the pulse, the density matrix is

$$\frac{1}{2} + \frac{1}{2} \frac{\beta}{\hbar} I_y$$

Thus the off-diagonal elements are

$$\begin{aligned}\rho_{12} &= \frac{-i}{2} \frac{\beta}{\hbar} \\ \rho_{21} &= \frac{i}{2} \frac{\beta}{\hbar}\end{aligned}$$

To model the transversal relaxation, this equation is modified to become

$$\begin{aligned}\rho_{12}(t) &= \frac{-i}{2} \frac{\beta}{\hbar} e^{-T_2^{-1}t} \\ \rho_{21}(t) &= \frac{i}{2} \frac{\beta}{\hbar} e^{-T_2^{-1}t}\end{aligned}$$

Thus the off-diagonal elements are assumed to decay exponentially, where T_2 determines the speed of the decay. This amounts to a decay of the magnetization with the same rate.

A second type of relaxation is called **longitudinal relaxation**. This process models the fact that the populations of the density matrix (the elements on the diagonal) slowly return to the values predicted by the thermal equilibrium state. Again, the assumption is that the difference between the populations after the pulse and the populations in the thermal equilibrium happens exponentially at a rate T_1^{-1} , where the time T_1 is called the **longitudinal relaxation time**. This process models the fact that the nuclei can exchange energy with the environment to return to thermal equilibrium. We refer to [7], section 11.9 or chapter 20 for more details.

5 Interactions

A Some further properties of rotations

In this appendix, we collect a few additional properties of the rotation matrices defined in equation (4). First, let us calculate the conjugate of a Pauli matrix with a rotation. For every real number a , we have

$$\begin{aligned}e^{ia\sigma_x}\sigma_y e^{-ia\sigma_x} &= (\cos a + i \sin a \sigma_x) \sigma_y (\cos a - i \sin a \sigma_x) \\ &= \cos^2 a \sigma_y - i \cos a \sin a \sigma_y \sigma_x + i \sin a \cos a \sigma_x \sigma_y + \sin^2 a \sigma_x \sigma_y \sigma_x \\ &= [\cos^2 - \sin^2] \sigma_y + 2i \sin a \cos a \sigma_x \sigma_y \\ &= [\cos^2 - \sin^2] \sigma_y - 2 \sin a \cos a \sigma_z \\ &= \cos 2a \sigma_y - \sin 2a \sigma_z\end{aligned}$$

where we have used that any two Pauli matrices anti-commute and that $\sigma_x \sigma_y = i \sigma_z$ as well as the usual addition theorems for sine and cosine. If we combine this with the definition of the rotations, we obtain

$$R_x(\phi) I_y R_x(-\phi) = \cos \phi I_y + \sin \phi I_z \quad (5)$$

This obviously implies that

$$R_x(\phi) I_y^k R_x(-\phi) = \cos \phi I_y + \sin \phi I_z$$

for all integers k and therefore

$$R_x(\phi)R_y(\Theta)R_x(-\phi) = \exp \left[-i \frac{\Theta}{\hbar} (\cos \phi I_y + \sin \phi I_z) \right] \quad (6)$$

This equation has an obvious geometrical interpretation - a rotation round x , followed by a rotation around y , followed by reversing the rotation around x is the same as a rotation around the axis given by $\cos \phi e_y + \sin \phi e_z$. Obviously the same relations hold for any cyclic combination of indices. For instance

$$R_z(\phi)I_xR_z(-\phi) = \cos \phi I_x + \sin \phi I_y$$

Let us now investigate a second type of relations which is very useful. For a three-vector n , we can formally define a dot-product

$$n \cdot \sigma = n_x \sigma_x + n_y \sigma_y + n_z \sigma_z$$

A short calculation using the commutation relations then shows that

$$(n \cdot \sigma)^2 = |n|^2$$

In fact, the product contains each combination of σ -matrices twice, but in different order, so all combinations except the squares that equal to one cancel each other. Using this, we immediately obtain from the power series expansion that for a unit vector n ,

$$e^{ian \cdot \sigma} = \cos a + in \cdot \sigma \sin a$$

Now let us assume that a and b are three vectors. We can then form the dot product

$$a \cdot I = a_x I_x + a_y I_y + a_z I_z$$

and similarly for b . Then $a \cdot I$ and $b \cdot I$ are matrices. Let us calculate their commutator.

$$\begin{aligned} [a \cdot I, b \cdot I] &= [a_x I_x + a_y I_y + a_z I_z, b_x I_x + b_y I_y + b_z I_z] \\ &= i\hbar [(a_x b_y - a_y b_x) I_z + (a_z b_x - b_z a_x) I_y + (a_y b_z - a_z b_y) I_x] \end{aligned}$$

This looks complicated, but can be summarized nicely in the equation

$$[a \cdot I, b \cdot I] = i\hbar(a \times b) \cdot I$$

B The density matrix formalism

In this appendix, we give a quick introduction into the density matrix formalism, which is very common in quantum statistics. For the sake of simplicity, we will restrict ourselves to the case of finite dimensional Hilbert spaces. We refer the reader to section 3.3 in [5] or section 2.4 in [6] for more details.

To motivate the introduction of the density matrix, let us first assume that we are given a quantum mechanical system in a normalized state $|\psi\rangle$. Suppose that we want to calculate the expectation value of an observable described by a hermitian operator A with an eigenbasis $|a\rangle$. We know that the expectation value is simply

$$\langle A \rangle = \langle \psi | A | \psi \rangle$$

To calculate this, let us expand $|\psi\rangle$ into an eigenbasis of A . We have

$$|\psi\rangle = \sum_a \langle a|\psi\rangle |a\rangle$$

Plugging this into the expression for the expectation value, we find that

$$\langle A \rangle = \sum_a a \langle \psi|a\rangle \langle a|\psi\rangle = \sum_a a |\langle \psi|a\rangle|^2 \quad (7)$$

So far this is nothing new and not surprising - this is just the average across all values of a , weighted by the probability to find the system in the state $|a\rangle$ after a measurement. In this way, we usually express the expectation values of all observables in terms of $|\psi\rangle$ and then argue that, because $|\psi\rangle$ in this way determines the expectation values of all observables, the vector (or rather the ray of the vector in the underlying projective space) can be considered to obtain the full information on the physical state of the system.

However, it turns out that there is a different object that we can use instead of $|\psi\rangle$ to express these expectation values - and that we can then, with the same rationale, accept as being a representative of the state of the physical system. To see this, note that we can express the individual summands in equation (7) as matrix elements of an operator. In fact,

$$a \langle \psi|a\rangle |\psi\rangle = a [|\psi\rangle \langle \psi|] |a\rangle = [|\psi\rangle \langle \psi| A] |a\rangle$$

so that

$$a \langle \psi|a\rangle \langle a|\psi\rangle = \langle a| [|\psi\rangle \langle \psi| A] |a\rangle$$

Thus the expression for the expectation value in (7) is the sum over the elements along the diagonal of the operator

$$|\psi\rangle \langle \psi| A$$

in other words, the trace of this operator. The operator

$$\rho = |\psi\rangle \langle \psi|$$

is called the **density matrix** associated with the state $|\psi\rangle$, and our calculation shows that

$$\langle A \rangle = \text{tr}(\rho A)$$

Thus instead of expressing the expectation values of all observables in terms of the vector $|\psi\rangle$, we can as well express all these expectation values in terms of the operator ρ and can use this object as the primary description of the physical state.

Let us collect a few properties of the density matrix. First, it is obvious that the density matrix is hermitian. Further, the trace of the density matrix itself is one, and the operator is non-negative, i.e. all eigenvalues are non-negative. It is now common to make the following definition.

DEFINITION B.1. *An non-negative operator ρ which is hermitian and has trace one is called a **density operator**.*

We have seen that every normalized state vector gives rise to a density operator. However, the converse is not true. In fact, the following holds.

PROPOSITION B.1. *For a density operator ρ , the following conditions are equivalent.*

1. ρ is a projection, i.e. $\rho^2 = \rho$
2. ρ can be written as $|\psi\rangle\langle\psi|$ for a normed state $|\psi\rangle$

Proof. It is obvious that every operator of the form $|\psi\rangle\langle\psi|$ is a projection. Conversely, suppose we are given a density operator ρ . As ρ is hermitian, we can diagonalize ρ , i.e. we can write ρ as

$$\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$$

with normed eigenvectors $|\psi_i\rangle$. As ρ is non-negative, all eigenvalues λ_i are positive. The condition that the trace of ρ is one then translates into

$$1 = \sum_i \lambda_i$$

In particular, all λ_i are real numbers between 0 and 1. Now the square of ρ is given by

$$\rho^2 = \sum_i \lambda_i^2 |\psi_i\rangle\langle\psi_i|$$

If $\rho^2 = \rho$, this implies that $\lambda_i^2 = \lambda_i$ for all i . This is only possible if all λ_i are either zero or one, i.e. if we can write

$$\rho = |\psi_i\rangle\langle\psi_i|$$

for one of the indices i , as claimed. \square

Thus only those density operators that are projections correspond to normed vectors and therefore to physical states of our quantum system. These density operators are called **pure states**. A density operator that does not fulfill this condition is called a **mixed state**. We will learn more about mixed states and their physical interpretation later.

Let us now try to understand how the time evolution of a system can be expressed in terms of the density operator. The time evolution of the vector representing a given state is

$$\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}H|\psi(t)\rangle$$

Using the product rule, we can derive a differential equation for the corresponding density operator.

$$\begin{aligned} \frac{d}{dt}\rho(t) &= \frac{d}{dt}|\psi\rangle\langle\psi| \\ &= \left|\frac{d}{dt}\psi\right\rangle\langle\psi| + |\psi\rangle\left\langle\frac{d}{dt}\psi\right| \\ &= -\frac{i}{\hbar}|H\psi\rangle\langle\psi| + \frac{i}{\hbar}|\psi\rangle\langle H\psi| \\ &= -\frac{i}{\hbar}[H|\psi\rangle\langle\psi| - |\psi\rangle\langle\psi|H] \end{aligned}$$

We can rewrite this equation as

$$i\hbar \frac{d}{dt} \rho(t) = [H, \rho]$$

This equation of motion for the density operator is often called the **Liouville-von Neumann equation**.

Let us now make the relation between the description of states in terms of rays and the description of states in terms of density operators a bit more precise. Let \mathcal{H} denote the unit sphere in the Hilbert space \mathcal{H} . We have a well-defined map

$$\begin{aligned} S\mathcal{H} &\rightarrow \mathcal{B}(\mathcal{H}) \\ |\psi\rangle &\mapsto |\psi\rangle\langle\psi| \end{aligned}$$

Now it is obvious that for a scalar a , we have

$$|a\psi\rangle\langle a\psi| = aa^* |\psi\rangle\langle\psi|$$

This shows that this map factors to provide a map

$$\mathbb{P}\mathcal{H} \rightarrow \mathcal{B}(\mathcal{H})$$

from the projective space, i.e. the space of all rays which we identify with the set of all physical states, into the set of linear operators on \mathcal{H} . We have also seen that the image of this mapping is exactly the set of all density operators that are idempotent, i.e. are projections. Clearly, this map is one-to-one - the ray can be reconstructed from ρ as it is simply the image of ρ . Thus we find that in fact, we have bijections

$$\{\text{all physical states}\} \leftrightarrow \mathbb{P}\mathcal{H} \leftrightarrow \{\text{idempotent density operators}\}$$

Consequently, the set of rays in the Hilbert space and the set of pure states in the density matrix formalism are two fully equivalent formulations of quantum mechanics.

So far all we have done is to re-express physical states described by rays in a Hilbert space in terms of density operators. The density matrix formalism does not seem to bring a real advantage in this case. This changes as soon as we consider **statistical ensembles** of quantum systems.

To prepare for this, let us note the following interesting property of density operators.

LEMMA B.2. *A convex combination of density operators is again a density operator.*

Proof. Suppose we are given a set A_i of density operators and real, non-negative coefficients a_i with $\sum_i a_i = 1$. We want to show that

$$A = \sum_i a_i A_i$$

is again a density operator. Clearly, A is again hermitian. The trace is

$$\text{tr}(A) = \text{tr}\left(\sum_i a_i A_i\right) = \sum_i a_i \text{tr}(A_i) = \sum_i a_i = 1$$

To prove that A is non-negative, recall that in general, a hermitian operator A is non-negative if and only if $\langle x|A|x\rangle \geq 0$ for all vectors x (a fact which is not difficult to prove using a basis in which A is diagonal). Now in our case

$$\langle x|A|x\rangle = \sum_i a_i \langle x|A_i|x\rangle$$

which is non-negative as this is true for each A_i and the coefficients a_i are non-negative. \square

Now suppose we are given an ensemble of N quantum systems with identical Hilbert spaces and Hamiltonians, i.e. an ensemble of N copies of the same system. Suppose that at a given point in time, each system is in a state $|\psi_i\rangle$, which we can equivalently describe by a density matrix ρ_i . By the lemma above, the operator

$$\rho = \frac{1}{N} \sum_i \rho_i$$

is again a density operator. If now A is an observable on each of the (identical) systems, the expectation value of A is the average of the expectation values of A on each of the individual Hilbert spaces, i.e. given by

$$\langle A \rangle = \frac{1}{N} \sum_i \langle A_i \rangle$$

where A_i is the copy of A acting on the Hilbert space of system i . Now we know how to compute each $\langle A_i \rangle$, this is just the trace of $\rho_i A$. As the trace is linear, we obtain

$$\langle A \rangle = \frac{1}{N} \sum_i \text{tr}(\rho_i A) = \text{tr}\left(\frac{1}{N} \sum_i \rho_i A\right) = \text{tr}(\rho A)$$

Thus we find the same relation between expectation value and density matrix as for pure states. This becomes even more apparent if we let n_i denote the number of systems which are in state ρ_i (which can be one). Then the fraction

$$p_i \frac{n_i}{N}$$

is the probability for one of the individual systems to be in state ρ_i (note that here the probability does not refer to the stochastic interpretation of a measurement, but simply refers to the frequentist's interpretation of probability as a ratio). The density matrix is then

$$\rho = \sum_i p_i \rho_i = \sum_i p_i |\psi_i\rangle \langle \psi_i|$$

where $|\psi_i\rangle$ is the state of system i in the vector formalism. Thus we see that the overall density matrix is simply the weighted sum of the density matrices of the individual systems, weighted with the frequency of occurrence, i.e. the probability. Our formula for the expectation value then becomes

$$\langle A \rangle = \sum_i p_i \text{tr}(\rho_i A)$$

which is just the weighted sum of the expectation values of the individual system.

The fact that we can compute the (average) expectation value of an observable once we know the density matrix is remarkable. After all, the full state of the system is described by all N state vectors $|\psi_i\rangle$ or - equivalently - by all density operators ρ_i . If we build the overall density matrix by summing up these contributions, we lose a huge part of the information, but the information which is left is still sufficient to calculate expectation values.

In fact, two completely different statistical ensembles, consisting of systems in states $|\psi_i\rangle$ with frequencies p_i can give rise to the same density matrix. To see an obvious example, consider a two state system and the density matrix

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$$

This matrix would be the result of starting with an ensemble of a large number of copies of that system, in which half of the copies are in state $|0\rangle$ and half of the systems are in state $|1\rangle$. Now consider the alternative basis of the Hilbert space given by the two vectors

$$|+\rangle = \frac{1}{\sqrt{2}}[|0\rangle + |1\rangle]$$

and

$$|-\rangle = \frac{1}{\sqrt{2}}[|0\rangle - |1\rangle]$$

A short calculation shows that

$$\frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -| = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$$

So we have two very different ensembles giving the same density matrix. The representation on the right represents an ensemble where all systems are in an eigenstate of σ_z , whereas the left side represents an ensemble where all systems are an eigenstate of σ_z . Still, the density matrices are the same. If we accept that all macroscopically observable quantities are averages across the ensemble and can therefore be derived from the density matrix, we have to accept these two systems as macroscopically equivalent.

References

- [1] R. Feynman, The Feynman Lectures on Physics, Vol. II, Addison Wesley 1964, available online at <http://www.feynmanlectures.caltech.edu/>
- [2] R. Feynman, The Feynman Lectures on Physics, Vol. III, Addison Wesley 1964, available online at <http://www.feynmanlectures.caltech.edu/>
- [3] W.S. Warren et. al., The Usefulness of NMR Quantum Computing, Science 277 (5332), 1688-1689, available at <http://science.sciencemag.org/content/277/5332/1688>
- [4] B.C. Hall, Quantum theory for mathematicians, Springer, New York 2013
- [5] S. Weinberg, Lectures on quantum mechanics, Cambridge University Press, Cambridge 2013

- [6] M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2010
- [7] M.H. Levitt, *Spin dynamics - basics of nuclear magnetic resonance*, John Wiley & Sons 2008