

Introduction to NMR and NMR based qubits

October 29, 2018

1 Introduction

In the standard approach to quantum computation, the central logical unit that is considered and manipulated is a **qubit**. Conceptionally, a qubit is a two-state quantum mechanical system, i.e. a system that is described by a two-dimensional Hilbert space.

Now the best known two-state quantum system (and in fact, up to unitary equivalence, the only one) is the spin degree of freedom of a spin-1/2 particle. Thus spin systems like a single electron or certain atomic nuclei are a very obvious candidate for a physical implementation of quantum computers.

Unfortunately, manipulating individual spins is experimentally very hard. Ion traps that are capable of confining individual charged particles to a small region of space and then cool them down to the point that the energy in the spatial degrees of freedom becomes very small is technically possible (and quantum computation on that basis has been experimentally demonstrated for a small number of qubits), but remains challenging.

At the same time, a technology to control states of an **ensemble** of spins has been available since the second half of the last century - **Nuclear magnetic resonance (NMR) spectroscopy**. In NMR spectroscopy, a probe, for instance a liquid, containing a large number of identical molecules is considered. The nuclei of these molecules have a spin that creates a magnetic moment. By applying properly designed static and oscillating magnetic fields, the spin states can be manipulated and read out. Even though there are some challenges due to the fact that we cannot control an individual spin, but only a statistical ensemble of spins, NMR techniques have successfully been applied to realize quantum computations with a small number of qubits - for instance, the factorization of small integers using Shor's algorithm. NMR experiments require strong (often superconducting) magnets, but can otherwise be executed at room temperature, and highly developed NMR devices are available on the market.

In these notes, we present the basics of quantum computation with NMR methods. We first recall some basic properties of spin systems before we describe the basic mechanics of an NMR experiment on a heuristic level. We then work out the formalism for a single particle and a full ensemble, before we discuss how quantum logic gates and quantum algorithms, including state preparation and read-out, can be implemented.

2 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. 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

$$R_z(\phi) = \exp\left(-\frac{i}{\hbar} \phi I_z\right) = \exp\left(-i \frac{\phi}{2} \sigma_z\right) \quad (4)$$

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

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

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

$$R_y(\phi) = \exp\left(-i\frac{\phi}{2}\sigma_y\right) = \cos\frac{\phi}{2} - i\sigma_y \sin\frac{\phi}{2}$$

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

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

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

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

We have therefore parametrized the space of rays by the two angles Φ and Θ , i.e. by 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 $R_y(\pi) = -i\sigma_y$ 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).

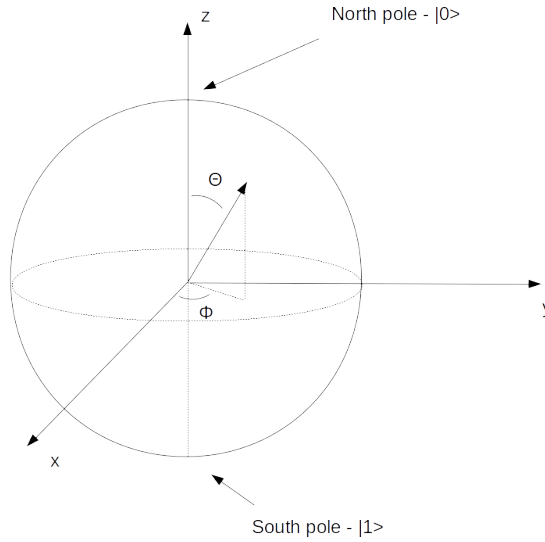


Figure 1: Representation of qubit states on the Bloch sphere

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 elements of 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) and a constant γ that is known as the **gyromagnetic ratio**. For simplicity, let us assume that B is oriented along the z-axis, so that our extra term becomes

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

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) of size $\hbar\gamma B$.

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

$$\exp\left(-\frac{i}{\hbar} H t\right) = \exp\left(\frac{i}{\hbar} \gamma t B_z I_z\right) = R_z(-\gamma B_z t)$$

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

$$U_0(t) \otimes R_z(-\gamma B_z t)$$

where $U_0(t)$ 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**.

3 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, while ^{13}C has spin $1/2$.

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 a 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 were 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. According to this distribution, 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 , 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.

4 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

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

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 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 appendix 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

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

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

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

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

$$\tilde{H} = \omega_{nut} I_x$$

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

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

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

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

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 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 .

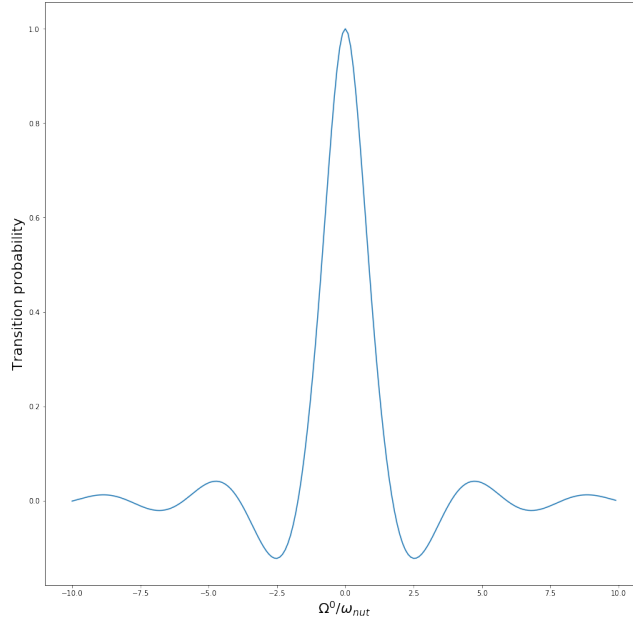


Figure 2: Resonance peak around the Larmor frequency

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

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

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, 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 not easy to create a rotating magnetic field by really 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 nutation 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,

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

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} for the time

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

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.

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 magnetizations caused by the precession of each individual nucleus do not cancel out in the average. To understand how this works, we need to extend our model to describe a large number of nuclei with statistical behavior.

5 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 reflect 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 temperate (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** $\gamma B I_z$ - takes a negative value on this state, i.e. this is the state of lowest energy (this is true when $\gamma > 0$, which is more common, however there are also nuclei with $\gamma < 0$). 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

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

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

$$\tilde{H} = \omega_{ef} n \cdot I$$

for a unit vector n - in fact, all Hamiltonians that we need in this section are of this form. We know that $\text{tr}(\rho) = 1$, so that the matrix $\rho - \frac{1}{2}$ is traceless and hermitian and can 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

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

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). By splitting the magnetization into a component perpendicular to n and a component parallel to n , it is easy to see that this equation describes a precession of the magnetic moment around the axis n , as expected.

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 y-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

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

plus a phase that reflects the offset between the rotating frame and the laboratory frame at the time when the pulse is switched off and that we ignore for simplicity (but which will become important in section 8 when we discuss measurements in more detail). 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 $r = de_x$. 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 $r = de_x$ 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!

In NMR terminology, this signal is known as **free induction decay**, as we let the system involve freely while we measure the signal. In reality, one observes a decay of this signal over time, caused by the fact that the system will slowly return to thermal equilibrium and the signal fades out. This is called the **relaxation** and not yet reflected in our simple model.

This is a good point in time to visualize what is happening. Given the explicit formulas for the density matrix derived in this section, and the formulas describing the measurement process that we will derive in a later section, it is not difficult to numerically simulate the state changes and NMR signals during an experiment as the one that we have just described.

Diagram 3 shows the result of such a simulation. Here, we have simulated a carbon nucleus in a TCE (Trichloroethylene) molecule which, at 11.74 Tesla, has a Larmor precession frequency of 125 MHz. At the start of the simulation, the system was put into a thermal state. Then, an RF pulse was applied to flip the magnetization in the direction of the x-axis, and then a sample was taken over 0.1 seconds.

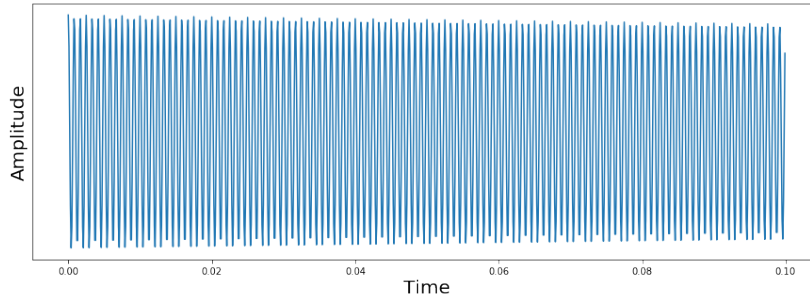


Figure 3: FID (Free induction decay) signal of a single nucleus

The signal that we see looks at the first glance as expected. We see an oscillating signal with an amplitude that is slowly decaying. However, you might notice that the frequency of the oscillation is clearly not 125 MHz. Instead, the period is roughly 0.001 seconds, corresponding to a frequency of 1200 Hz. The reason for this is that in an NMR spectrometer, the circuit processing the received signal will typically apply a combination of a mixer and a low pass filter to effectively shift the frequency by an adjustable reference frequency (see

section 8 or [7] for details). In our case, the reference frequency was adjusted to be 1200 Hz above the Larmor frequency of the carbon nucleus, so the signal will oscillate with a frequency of 1200 Hz. In practice, the reference frequency determines a window in the frequency space in which we can detect signals, and all frequencies outside this window will be suppressed by the low pass filter.

Now let us take a look at a more complicated signal. We again place the system in the thermal equilibrium state first, but then apply RF pulses to flip the spin of both carbon nuclei into the x-axis (in a simulation, this is easy, in a real experiment, this requires some thought, as the Larmor frequencies of these two carbon nuclei differ only by a small chemical shift of 900 Hz). We then again take a sample and plot the signal. The result is shown in diagram

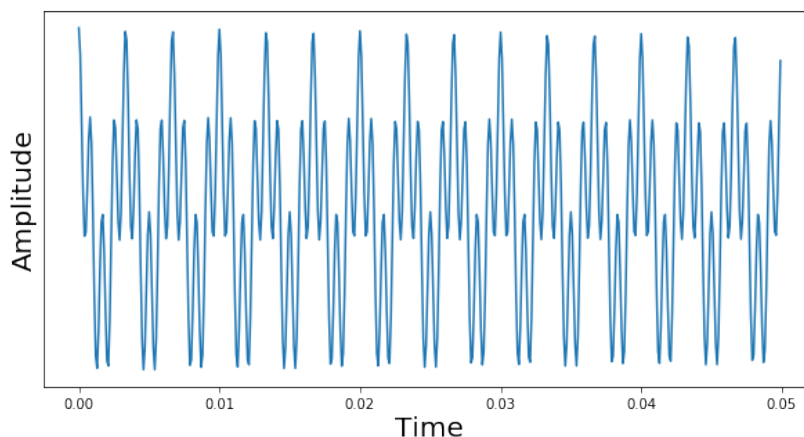


Figure 4: FID signal of a pair of nuclei

This time, we see a superposition of two oscillations. The first oscillations is what we have already seen - an oscillation with 1200 Hz, which is the difference of the chosen reference frequency and the Larmor frequency of the first carbon. The second oscillation corresponds to a frequency of roughly 300 Hz. This is the signal caused by the Larmor precession of the second spin. As we again measure the difference between the real frequency and the reference frequency, we can conclude that this frequency differs from the Larmor frequency of the first spin by 900 Hz - the chemical shift between the two nuclei.

In reality, the signal that we observe is the superposition of many different oscillations and is not easy to interpret - even with a few oscillations, it soon becomes impossible to extract the frequencies by a graphical analysis as we have done it so far. Instead, one usually digitizes the signal and applies a Fourier transform (or, more precisely, a discrete Fourier transform). The following diagram shows the result of applying such a DFT to the signal above.

Here, we have shifted the x-axis by the difference between the Larmor frequency ω_0 of the first nucleus and the reference frequency, so that the value zero corresponds to ω_0 . We clearly see two peaks. The first peak at zero, i.e. at Larmor frequency ω_0 , is the signal emitted by the first nucleus. The second peak is shifted by 900 Hz and is emitted by the second nucleus. In general, each nucleus will result in one peak (ignoring couplings that we will study in a later section) and the differences between the peaks belonging to nuclei of the same

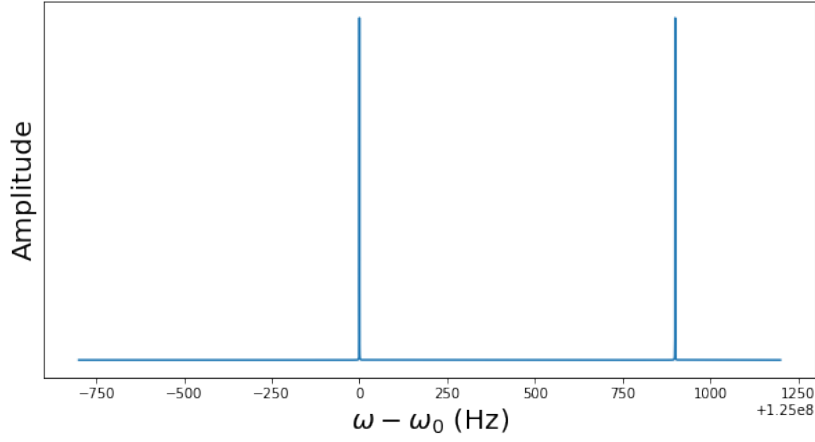


Figure 5: FFT of the FID for a pair of nuclei

isotope are the chemical shifts.

To get an idea of the magnitudes of the voltages that we need to measure, let calculate the field strength for 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 $A = 10^{-4} \text{ m}^2$, and if we set $N = N_A$, this will lead to an induced voltage with an amplitude of

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

oscillating with 500 MHz per mol of protons in the probe. 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

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

and 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.

6 Interactions

So far, we have studied isolated spin-1/2 nuclei under the influence of external magnetic fields. In reality, however, nuclei interact. There is a **long range interaction** between spins of nuclei in different molecules, that we ignore, but there are also **short range interactions** between spins within the same molecule.

These interactions come in various flavors. To simplify our treatment, we now specialize to the case of a **liquid probe**. As the rotational and spatial degrees of freedom are then completely randomized, many interaction terms average out. The only interactions terms that are relevant in this situation (see chapter 9 of [7]) are the **chemical shift** and the **J-coupling**.

The chemical shift arises from the interaction of the external magnetic field with the electrons surrounding the nuclei. Essentially, the same nuclei at different positions in the molecule will not experience the same external magnetic field. Instead, the shielding of the magnetic field by the electron cloud will lead to different effective magnetic fields and therefore different Larmor frequencies.

Thus the Larmor frequency is shifted by a term characteristic for the position of the nucleus inside the molecule. This shift is called chemical shift and is extremely useful in practical applications as it allows us to stimulate the spins of different nuclei in the same molecule separately and because its value reveals information on the valence and bonding structures.

The second type of coupling that is relevant is called **J-coupling**. This coupling is an indirect coupling - the magnetic moment created by a spin interacts with the electric field of an electron, which in turn interacts with the magnetic moment of a different spin.

Let us now look at an ensemble of quantum systems, where each system describes two spin-1/2 nuclei. We make the assumption that we are in the **weak coupling regime**, i.e. that the strength of the J-coupling is much smaller than the difference between the Larmor frequencies of the two nuclei. Then, the Hamiltonian is - in good approximation - given by (see [7], section 14.5 or [8], section 2.3)

$$H = \omega_1 I_z^1 + \omega_2 I_z^2 + \frac{2\pi}{\hbar} J_{12} I_z^1 I_z^2$$

Here ω_i is the Larmor frequency of nucleus i , the upper index at an operator indicates the spin on which the operator is acting and J_{12} is the strength of the J-coupling, measured in Hz. The weak coupling condition is then the condition that

$$|J_{12}| \ll |\omega_1 - \omega_2|$$

The system consisting of the two nuclei is described by the basis states $|0\rangle|0\rangle$, $|0\rangle|1\rangle$, $|1\rangle|0\rangle$ and $|1\rangle|1\rangle$. These states are all eigenstates of the I_z^i and therefore the Hamiltonian is already diagonal in this basis. The energy levels are

$$\begin{aligned} & \frac{\hbar}{2}(\omega_1 + \omega_2 + 2\pi J_{12}) \\ & \frac{\hbar}{2}(\omega_1 - \omega_2 - 2\pi J_{12}) \\ & \frac{\hbar}{2}(-\omega_1 + \omega_2 - 2\pi J_{12}) \\ & \frac{\hbar}{2}(-\omega_1 - \omega_2 + 2\pi J_{12}) \end{aligned}$$

Due to the weak coupling condition, these are four different energies and the system is fully non-degenerate. The transformation into the rotating frame system is given by the unitary transformation

$$T = e^{i\omega_{ref}t(I_z^1 + I_z^2)} = e^{i\omega_{ref}tI_z^1} e^{-i\omega_{ref}tI_z^2}$$

As all parts of the Hamiltonian commute with this transformation, the Hamiltonian in the rotating frame is easily seen to be

$$\tilde{H} = \Omega_1^0 I_z^1 + \Omega_2^0 I_z^2 + \frac{2\pi}{\hbar} J_{12} I_z^1 I_z^2$$

where again

$$\Omega_i^0 = \omega_i - \omega_{ref}$$

is the difference between the Larmor frequency of spin i and the rotation frequency. This Hamiltonian has three parts:

- The first two terms $\Omega_i^0 I_z^i$ which are called the **resonance offset** Hamiltonian and result in precession around the z-axis of the respective spin
- The third term which is the **J-coupling term**

Note that these three terms all commute. Therefore we can write the time evolution as a product of the individual time evolution operators, which are nothing but rotations around the z-axis, and these operators will commute as well. These operators act on the density matrix by conjugation. As the density is hermitian, we can always expand it into a sum of products of Pauli operators (or angular momentum operators, which is the same up to a factor) and use the relations in appendix A to calculate how the time evolution operators act on these terms. This is sometimes called the **operator product formalism** and very useful in our situation.

As in the case of an ensemble of a single spin-1/2 system, let us now try to determine the state at thermal equilibrium. Again, we assume that, at thermal equilibrium, our ensemble can be described as a canonical ensemble. The density matrix is then given as

$$\rho = \frac{1}{Z} e^{-\frac{H}{kT}}$$

Again, we use the high temperature approximation, i.e. we use the fact that all energies are much smaller than the thermal energy at room temperature. Then we can make the approximation

$$\rho = \frac{1}{Z} (1 - \frac{1}{kT} H)$$

If we set

$$\beta_i = \frac{\gamma_i B \hbar}{kT}$$

and

$$\beta_{12} = \frac{2\pi J_{12} \hbar}{kT}$$

the density matrix at equilibrium becomes

$$\rho = \frac{1}{4} [1 + \frac{\beta_1}{\hbar} I_z^1 + \frac{\beta_2}{\hbar} I_z^2 - \frac{\beta_{12}}{\hbar^2} I_z^1 I_z^2]$$

Following [7], chapter 15, we can make a second approximation in a few special cases. If the two nuclei that we consider are the same isotope and the chemical shift is small compared to the Larmor frequency, the values of the β_i are very close to

$$\beta = \frac{-\omega \hbar}{kT}$$

In addition, the J-coupling is usually some orders of magnitude smaller, so that it is a good approximation to use

$$\rho = \frac{1}{4} + \frac{1}{4} \frac{\beta}{\hbar} (I_z^1 + I_z^2)$$

for the density matrix in the thermal equilibrium state. This is often written as

$$\rho = \frac{1}{4} + \frac{1}{4} \frac{\beta}{\hbar} I_z$$

with the total spin $I = I^1 + I^2$. Systems of this type are called **homonuclear systems**. An early quantum algorithm carried out using NMR, for instance, was using two protons in a cytosin molecule (see [9]), which would fall into this category.

In a **heteronuclear system**, the two nuclei can be different isotopes, so the values of β_i can be significantly different. If the J-coupling is small, however, we can still make the approximation

$$\rho = \frac{1}{4} + \frac{1}{4} \frac{\beta_1}{\hbar} I_z^1 + \frac{1}{4} \frac{\beta_2}{\hbar} I_z^2$$

Let us now again try to understand how this state evolves over time. Clearly, ρ commutes with the Hamiltonian, so this state is stable. Further, ρ commutes with the transformation T describing the rotating frame, so we find that the thermal equilibrium density matrix in the rotating frame is given by

$$\tilde{\rho} = \frac{1}{4} + \frac{1}{4} \frac{\beta_1}{\hbar} I_z^1 + \frac{1}{4} \frac{\beta_2}{\hbar} I_z^2$$

Clearly, the second and third term can again be interpreted as the magnetization, i.e. we find that there is a net magnetization in the direction of the z-axis.

Next, let us again consider the impact of a rotating RF field with frequency ω_{ref} . Similar to the case of a single spin system, the additional term in the Hamiltonian in the rotating frame is given by

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

with the difference that now, the operator I is the total spin $I^1 + I^2$. For simplicity, let us assume that Φ_p is zero and that the pulse is exactly on resonance for the first spin. Then, during the presence of the pulse, the Hamiltonian in the rotating frame is

$$[\Omega_2^0 I_z^2 + \omega_{nut} I_x^2] + \omega_{nut} I_x^1 + \frac{2\pi}{\hbar} J_{12} I_z^1 I_z^2$$

Note that the resonance offset part for the first spin vanishes. Now, in most situations, the coupling is much smaller than the contribution of the other terms. In [8], for example, the molecule in question contains two ^{13}C nuclei with a Larmor frequency of 125 MHz and a chemical shift between 600-900 Hz. The J-coupling corresponds to a frequency of 100 Hz. Assuming that the RF field is weaker than the static field by a factor 10^3 , the nutation frequency will be in the order of 100 kHz.

The first term in this Hamiltonian acts on the second spin only. According to our discussion of off-resonance effects, this will induce a rotation around an axis almost parallel to the z-axis with a frequency almost equal to the difference between the two Larmor frequencies, i.e. several hundred MHz. If we compare this to the coupling term - the only other term acting on the second spin - we see that it is around six orders of magnitude weaker and can therefore safely be ignored. Thus the impact of the pulse on the second spin will be that of an RF pulse very far off resonance. This will have almost no impact if the system is already parallel to the z-axis. In general, it will contribute a very rapidly

oscillating term that averages out over the duration of the RF pulse. Thus we are left with the contribution of the term

$$\omega_{nut} I_x^1$$

The upshot of the discussion is that essentially, the pulses can be tuned such that the pulse only acts on one of the spins and has the same impact on this spin as in the case of a single spin system. Thus we can build pulses that rotate each of the spins individually. If, however, we apply so called hard pulses with a nutation frequency much higher than the energy shift between the spins, the pulse will rotate both spins simultaneously (see also the discussion in section 2.2. of [8]).

Let us now study the role of the coupling term during a free evolution, i.e. when no pulse is active. To illustrate this, let us consider two thought experiments. First, suppose we were able to prepare an ensemble of the form (we will get back to the question whether this can be done in a later section)

$$\rho = \frac{1}{4}(1 - \epsilon) + \epsilon|00\rangle\langle 00|$$

where we work with an eigenbasis for the operators I_z^i with eigenvalues $\pm \frac{\hbar}{2}$. To be able to easily carry out calculations, let us first rewrite this in terms of the Pauli matrices. A short calculation shows that

$$|00\rangle\langle 00| = \frac{1}{4}(1 + \sigma_z^1 + \sigma_z^2 + \sigma_z^1 \sigma_z^2)$$

so that our state becomes

$$\frac{1}{4} + \epsilon[\sigma_z^1 + \sigma_z^2 + \sigma_z^1 \sigma_z^2]$$

After absorbing some constants into the definition of ϵ , we can write this as

$$\rho = \frac{1}{4} + \epsilon[\frac{\hbar}{2}I_z + I_z^1 I_z^2]$$

This is a good point in time to understand the physical meaning of the individual terms, i.e. their impact on the expectation value of macroscopically observable quantities. First, let us see how the expectation value of the magnetization I can be calculated. When we take the trace of ρI , only the term proportional to I_z will make a contribution, and therefore we find that we observe a net magnetization in the direction of the z-axis, as in the case of the uncoupled system.

Now assume that we wanted to observe macroscopic correlations, corresponding to the product $I_z^1 I_z^2$ of the spins in z-direction which is of course again a hermitian operator. If we multiply ρ by this matrix and take the trace, all terms in ρ that have no or only one spin operator produce a tracefree term. Thus, in the final trace, only the term proportional to $I_z^1 I_z^2$ survives. Thus we need to interpret this term as the observable correlation between the spins of the two nuclei.

Let us now study how this term behaves under the impact of the Hamiltonian in the absence of a pulse. We have already seen that we can study the impact of the different terms - the resonance offset term and the J-coupling term -

individually. The resonance offset terms commute with the density matrix and therefore do not contribute to the time evolution. The J-coupling

$$\frac{2\pi}{\hbar} J_{12} I_z^1 I_z^2$$

does contribute and creates a time evolution operator

$$U(t) = \exp \left[-i \frac{2\pi J_{12}}{\hbar^2} I_z^1 I_z^2 t \right] = \exp \left[-i \frac{\pi J_{12}}{2} \sigma_z^1 \sigma_z^2 t \right]$$

Now the operator $\sigma_z^1 \sigma_z^2$ squares to one. Therefore, we can again apply the exponential series to see that

$$U(t) = \cos\left(\frac{\pi J_{12}}{2} t\right) - i \sigma_z^1 \sigma_z^2 \sin\left(\frac{\pi J_{12}}{2} t\right) \quad (5)$$

This formula allows us to easily compute the action of conjugation with $U(t)$ on each of the Pauli matrices. In particular, the time evolution commutes with all product of σ_z -matrices alone. Therefore it leaves the density matrix ρ considered above invariant.

Now let us try to rotate one of the spins in the direction onto the x-axis, say the first spin. We know that the time evolution operator induced by a pulse at resonance acts (in good approximation) only on this spin and is given by a rotation

$$R_x(\omega_{nut}\tau)$$

where τ is the duration of the pulse. Let us see how this process of rotation is expressed in terms of the density matrix, i.e. let us calculate the result of conjugating ρ with this matrix. Conjugation is a group homomorphism, so the conjugate of a product is the product of the conjugates. We also know that the rotation only acts on the first system. So the only calculation that we actually have to make is

$$R_x(\omega_{nut}\tau) I_z^1 R_x(-\omega_{nut}\tau)$$

which yields, using the formulas from the appendix, that

$$R_x(\omega_{nut}\tau) I_z^1 R_x(-\omega_{nut}\tau) = \cos(\omega_{nut}\tau) I_z^1 - \sin(\omega_{nut}\tau) I_y^1$$

Let us now apply what is called a $\frac{\pi}{2}$ -**pulse**, i.e. a pulse for which

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

Then

$$R_x(-\omega_{nut}\tau) I_z^1 R_x(\omega_{nut}\tau) = -I_y^1$$

Consequently, we find that the transformed state is described by the density matrix

$$\rho' = \frac{1}{4} + \epsilon \left[\frac{\hbar}{2} I_z^2 - \frac{\hbar}{2} I_y^1 - I_y^1 I_z^2 \right]$$

in which we have simply replaced all occurrences of I_z^1 by $-I_y^1$. This is a nice example that demonstrates the operator product formalism - we simply have to calculate the time evolution for the angular momentum operators and then can reduce our calculation to simple replacements.

When we look at the result, we see that two things have changed - we now have magnetizations in the z-axis and the y-axis, and we have a correlation between the y-axis of spin 1 and the z-axis of spin 2. How does this state evolve under the J-coupling?

To see this, we need to calculate the action of the conjugation by the time evolution operator (5) on the operators that appear in ρ' . We already know that I_z^2 is mapped to itself. What is the action on I_y^1 ? To see this, let us first calculate

$$\begin{aligned} U(t)\sigma_y^1U(-t) &= [\cos a - i\sigma_z^1\sigma_z^2\sin a]\sigma_y^1[\cos a + i\sigma_z^1\sigma_z^2\sin a] \\ &= \cos^2 a\sigma_y^1 + i(\cos a\sin a)\sigma_y^1\sigma_z^1\sigma_z^2 - i(\cos a\sin a)\sigma_z^1\sigma_z^2\sigma_y^1 - \sin^2 a\sigma_y^1 \\ &= (\cos^2 a - \sin^2 a)\sigma_y^1 - (2\cos a\sin a)\sigma_x^1\sigma_z^2 \\ &= \sigma_y^1\cos 2a - \sigma_x^1\sigma_z^2\sin 2a \end{aligned}$$

where we have used the abbreviation

$$a = \frac{\pi J_{12}}{2}t$$

Now let us apply this to the individual terms of the density matrix ρ' . We know that the first two terms are left invariant. The action on the third term is

$$\begin{aligned} U(t)\frac{\hbar}{2}I_y^1U(t^{-1}) &= \frac{\hbar^2}{4}U(t)\sigma_y^1U(t^{-1}) \\ &= \frac{\hbar^2}{4}[\sigma_y^1\cos 2a - \sigma_x^1\sigma_z^2\sin 2a] \\ &= \frac{\hbar}{2}I_y^1\cos 2a - I_x^1I_z^2\sin 2a \end{aligned}$$

and the action on the last term is

$$\begin{aligned} U(t)I_y^1I_z^2U(t^{-1}) &= I_z^2U(t)I_y^1U(t^{-1}) \\ &= I_z^2I_y^1\cos 2a - \frac{2}{\hbar}I_z^2I_x^1I_z^2\sin 2a \\ &= I_z^2I_y^1\cos 2a - \frac{\hbar}{2}I_x^1\sin 2a \end{aligned}$$

This gives a rather complicated expression for the density matrix which is an oscillation with frequency given by J_{12} . However, if we look at the I_x and I_y terms, we see that these two terms describe a rotation of the spin polarization vector in the x-y-plane. To get an intuition for what is going on, let us consider the case that $2a = \frac{\pi}{2}$. Then we obtain the following substitution rules

$$\begin{aligned} \frac{\hbar}{2}I_y^1 &\mapsto -I_x^1I_z^2 \\ I_y^1I_z^2 &\mapsto -\frac{\hbar}{2}I_x^1 \end{aligned}$$

Thus we exchange correlation terms for magnetization terms and vice versa. In our case, the new density matrix is

$$\rho'' = \frac{1}{4} + \epsilon\left[\frac{\hbar}{2}I_z^2 + \frac{\hbar}{2}I_x^1 + I_x^1I_z^2\right]$$

We can think of the coupling as rotating the original net magnetization of spin 1 along the y-axis into a net magnetization along the x-axis, thereby turning the correlation between y-axis and z-axis into a correlation between x-axis and z-axis. If we had flipped the second spin as well, we would see a similar oscillation here.

Again, it is instructive to numerically simulate the impact of the J-coupling term on the NMR spectrum. Diagram shows the result of such a simulation, again using the two carbon nuclei of a TCE molecule.

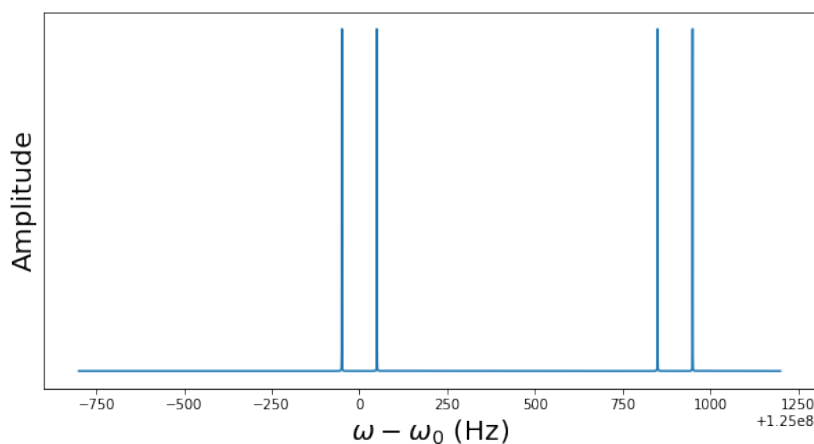


Figure 6: Impact of J-coupling on the NMR spectrum

In this simulation, we have first placed the system in a state given by the density matrix

$$\rho = \frac{1}{4}(1 - \epsilon) + \epsilon|00\rangle\langle 00|$$

as studied before. We then apply two pulses to turn both spins into the direction of the x-axis. Then we let the system evolve under the influence of the coupling term and take a sample of the resulting magnetization over time, to which we again apply a Fourier transform. We see that the two peaks belonging to the two carbon nuclei have split into two subpeaks, separated by a distance of roughly 100 Hz which is the strength of the coupling in this case. This splitting corresponds to the rotation of the spin polarization vectors in the x-y-plane caused by the J-coupling that we have derived above.

7 Realizing quantum gates

In the previous sections, we have studied the basic principles of NMR. We will now turn our attention to the use of NMR devices as quantum computers, using the individual nuclei in a molecule as qubits.

We have seen that we can use RF pulses to manipulate the state of individual qubits, and that the J-coupling is causing a correlation between different qubits, which will allow us to implement multi-qubit gates. But there is an issue with this approach. In quantum computing, we are typically manipulating individual qubits which are initially in a pure state. In an NMR probe, we usually have

a highly mixed state. It appears almost impossible to prepare the system in a pure state, i.e. to prepare all molecules consistently in the same quantum state. So we first need to understand how the usual description of a quantum computer in terms of qubits and unitary operations relates to the description of an NMR experiment in terms of density matrices and mixed states.

The formalism that we will apply is called the formalism of **pseudo-pure states**. A pseudo-pure state is a density matrix of the form

$$\rho = \frac{1}{2^N}(1 - \epsilon) + \epsilon|\psi\rangle\langle\psi|$$

where N is the number of involved qubits, i.e. nuclei. This is a convex combination of two density matrices and therefore again a density matrix. It describes an ensemble for which almost all molecules are in purely statistically distributed states, but a small part - measured by ϵ - of them are in a pure state.

Given such a state, we can reconstruct the number ϵ as the difference of the two eigenvalues and therefore - up to phase - the state $|\psi\rangle$. The part

$$\rho_\Delta = \epsilon|\psi\rangle\langle\psi|$$

is often called the **deviation** of the density matrix. The reason why pseudo-pure states are useful is that their time evolution and their expectation values are directly related to the time evolution and expectation values of the corresponding pure states.

Let us start with the time evolution. In general, the time evolution of a density matrix is given by conjugation with a unitary operator $U(t)$. Now

$$\begin{aligned} U(t)\rho U(t)^{-1} &= U(t)\left[\frac{1}{2^N}(1 - \epsilon) + \epsilon|\psi\rangle\langle\psi|\right]U(t)^{-1} \\ &= \frac{1}{2^N}(1 - \epsilon) + \epsilon U(t)|\psi\rangle\langle\psi|U(t)^{-1} \\ &= \frac{1}{2^N}(1 - \epsilon) + \epsilon|U(t)\psi\rangle\langle U(t)\psi| \end{aligned}$$

This is the pseudo-pure state that we would have obtained if we had started from the vector

$$U(t)|\psi\rangle$$

In other words, under the correspondence of pseudo-pure states and pure states, the time evolution of the density matrix transforms into the time evolution of the corresponding pure state with the same time evolution operator. In particular, once we have placed the system in a pseudo-pure state, it will be at (a different) pseudo-pure state at any later point in time.

A similar relation holds for observables. Suppose that A is an observable, i.e. a hermitian operator, with vanishing trace (this holds for most of the observables interesting in this context, especially for the Pauli operators). Then we can use the usual relation between traces and expectation values to write

$$\begin{aligned} \langle\psi|A|\psi\rangle &= \text{tr}(|\psi\rangle\langle\psi|A) \\ &= (2^N\epsilon)^{-1}\text{tr}(\rho_\Delta A) = (2^N\epsilon)^{-1}\text{tr}(\rho A) \end{aligned}$$

where we have used that the difference between ρ and ρ_Δ is a multiple of the identity so that its product with A is a multiple of A and therefore traceless.

This is remarkable, because it relates a macroscopically observable quantity, i.e. the average of the expectation values of A across the ensemble, to the quantum mechanical expectation value of A . Of course the value of ϵ might not be known, but is the same for all observables and therefore cancels out if we measure only ratios, for instance the value of A before and after a computation.

Using these relations, we can now translate the typical process of a quantum computation into NMR terminology as follows.

1. The process of initializing a quantum computer into an initial state $|\psi\rangle$ corresponds to putting the NMR probe into the corresponding pseudo-pure state with $\rho_\Delta \sim |\psi\rangle\langle\psi|$
2. If the quantum algorithm is described as a unitary operator U (typically presented as a sequence of gates U_i), we apply the same unitary operator to the density matrix, i.e. we realize the gates as an NMR experiment
3. We then measure the macroscopic quantity A using the procedure described above

In this section, we will discuss how quantum gates are realized in NMR, and leave the subject of state preparation and measurement to separate sections. To discuss gates, it is useful to again use a rotating frame of reference. This time, however, we are dealing with several qubits, each realized by a nucleus with a specific Larmor frequency ω_i . We therefore use a rotating frame in which the spin polarization vector of nucleus i is rotating with frequency ω_i , i.e. the transformation T between states in the laboratory frame and the rotating frame is given by

$$T = \prod_i R_{z^i}(-\omega_i t) = \prod_i e^{i\frac{\omega_i}{\hbar} t I_z^i} \quad (6)$$

A short calculation shows that the transformed state

$$|\widetilde{\psi}\rangle = T|\psi\rangle$$

evolves according to the equation

$$i\hbar \frac{d}{dt} |\widetilde{\psi}\rangle = [THT^{-1} - \sum_i \omega_i I_z^i] |\widetilde{\psi}\rangle$$

Thus the Hamiltonian in the rotating frame is given by

$$\widetilde{H} = THT^{-1} - \sum_i \omega_i I_z^i$$

Now, as before, the Hamiltonian in the absence of an RF pulse H is the sum of the chemical shift terms and the coupling term. We see that the chemical shift terms cancel with the terms $\omega_i I_z^i$. The J-coupling term

$$\frac{2\pi}{\hbar} J_{12} I_z^1 I_z^2$$

is invariant under rotations around the z-axis. Therefore, only the coupling term survives and the Hamiltonian in the rotating frame is

$$\widetilde{H}_C = \frac{2\pi}{\hbar} J_{12} I_z^1 I_z^2$$

Thus, in this rotating frame, the Hamiltonian during free evolution, i.e without an RF field, takes a very convenient form. For later reference, let us call this frame the **logical frame of reference**.

Let us now again see how the Hamiltonian changes in the presence of an RF field. In the laboratory frame, the RF field will add a term

$$- \sum_i \gamma_i B_{RF} [\cos(\omega_{ref}t + \phi_p) I_x^i + \sin(\omega_{ref}t + \phi_p) I_y^i]$$

which we can again write as

$$\sum \omega_{nut}^i R_{z^i}(\omega_{ref}t + \Phi_p) I_x^i R_{z^i}(-\omega_{ref}t - \Phi_p)$$

with the nutation frequency

$$\omega_{nut}^i = -\gamma_i B_{RF}$$

Thus, in the rotating frame, we obtain an additional RF term in the Hamiltonian which is given by

$$\widetilde{H_{RF}} = \sum \omega_{nut}^i R_{z^i}((\omega_{ref} - \omega_i)t + \Phi_p) I_x^i R_{z^i}(-(\omega_{ref} - \omega_i)t - \Phi_p)$$

Let us now again consider the case of a sharp x-pulse (i.e. a pulse with $\Phi_p = 0$) that is on-resonance with qubit i . Then, the RF part of the Hamiltonian will be

$$\omega_{nut}^i I_x^i + \sum_{j \neq i} \omega_{nut}^j R_{z^j}((\omega_i - \omega_j)t) I_x^j R_{z^j}(-(\omega_i - \omega_j)t)$$

Similar to our previous discussion for the case of two spins, we find that with a proper choice of chemical shifts, field strengths and pulse length, this can safely be approximated by

$$\widetilde{H_{RF}} = \omega_{nut}^i I_x^i$$

What do these calculations imply for our ability to manipulate our qubits, i.e. for the set of gates that we can apply? First, let us consider the most fundamental of all gates - the "no operation" gate, i.e. the gate that represents an unchanged qubit. To implement this gate should be trivial, but in NMR, there is a complication. In fact, even in the absence of RF pulses, we have non-trivial time evolution given by the coupling Hamiltonian. So it appears that we cannot even create a stable qubit.

Fortunately, the arsenal of NMR techniques contains a method to turn off the coupling that is called **refocussing**. To understand the idea, suppose that we let the system evolve freely over some time t , so that the evolution is described by the unitary transformation (see equation (5))

$$U_1 = \cos\left(\frac{\pi J_{12}}{2}t\right) - i\sigma_z^1 \sigma_z^2 \sin\left(\frac{\pi J_{12}}{2}t\right)$$

Next, we use an RF pulse with length $\tau = \pi \omega_{nut}^{-1}$ to apply a rotation

$$U_2 = R_x(\omega_{nut}\tau) = R_x(\pi) = -i\sigma_x^1$$

around the x-axis of spin one, assuming that the pulse is sufficiently short compared to the coupling frequency such that we can neglect the coupling Hamiltonian for the duration of the pulse. We know that this will change the sign

of the z-component of spin one. Thus, intuitively, we expect that the sign of the Hamilton coupling Hamiltonian is reverted. When now the free evolution continues, the reverted sign of the Hamiltonian has the same impact on the time evolution operator as reverting the sign of the time t . Therefore, we will now evolve "backwards in time" and undo the evolution during the first time interval t . We can then finally apply the inverse of the pulse again to also undo the change of spin one and should now be in the same state as before.

Formally, the entire time evolution during this sequence of events is given by

$$U_1 U_2 U_1 U_2^{-1}$$

Let us try to calculate this to see that our intuitive picture is correct. First, we calculate the product

$$\begin{aligned} U_1 U_2 &= -i \left[\cos\left(\frac{\pi J_{12}}{2} t\right) - i \sigma_z^1 \sigma_z^2 \sin\left(\frac{\pi J_{12}}{2} t\right) \right] \sigma_x^1 \\ &= -i \cos\left(\frac{\pi J_{12}}{2} t\right) \sigma_x^1 - i \sin\left(\frac{\pi J_{12}}{2} t\right) \sigma_y^1 \sigma_z^2 \end{aligned}$$

Multiplying this from the right with U_1 , we obtain, after applying the commutation relations for the Pauli matrices to combine terms, that

$$U_1 U_2 U_1 = -i \sigma_x^1$$

This is already what we did expect - the second free evolution reverts the change caused by the first free evolution. Finally, we can multiply by U_2 again to find that

$$U_1 U_2 U_1 U_2^{-1} = (-i \sigma_x^1)(i \sigma_x^1) = 1$$

so that we obtain the initial state back.

What does this imply for the set of possible one-qubit gates? First, we can realize the trivial gate that does not change the state by applying the refocusing pulse sequence. Second, we know that an RF pulse will describe a rotation around an axis in the x-y-plane, determined by the phase Φ_p . Now it is not difficult to see that any rotation of the Bloch sphere can be written as a product of rotations around the x-axis and the y-axis (see also appendix A). Therefore, RF pulses can be used to realize any rotation of the Bloch sphere, i.e. any one-qubit quantum gate.

To develop an intuition, let us write down a few examples. We will use the formula

$$R_a(\phi) = \cos \frac{\phi}{2} - i \sigma_a \sin \frac{\phi}{2}$$

for the rotation around an axis a . If we choose the angle, i.e. the duration of the pulse, such that $\phi = \pi$, we obtain multiples of the Pauli matrices. For $\phi = \frac{\pi}{2}$, we get linear combinations of the Pauli matrices. For instance, we find that

$$R_z\left(\frac{\pi}{2}\right) = \frac{1}{\sqrt{2}}(1 - i \sigma_z) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 - i & 0 \\ 0 & 1 + i \end{pmatrix} = \begin{pmatrix} e^{-i\frac{\pi}{4}} & \\ & e^{i\frac{\pi}{4}} \end{pmatrix}$$

which changes the relative phase of the two qubits. Similarly, a rotation by that angle around the y-axis gives

$$R_y\left(\frac{\pi}{2}\right) = \frac{1}{\sqrt{2}}(1 - i \sigma_y) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = H \sigma_z$$

where H is the Hadamard gate. We have already seen above that $-i\sigma_z$ can be realized by a single pulse. Consequently, we can realize the Hadamard gate, up to a phase, as a sequence of two pulses:

$$H = iR_y\left(\frac{\pi}{2}\right)R_z(\pi)$$

To prove that NMR techniques can be used to realize a universal set of quantum gates, we still have to find a realization of the CNOT gate. It turns out that the J-coupling can be used to do this. In fact, if we apply a free evolution without refocussing over the time t , we have seen that the J-coupling leads to the time evolution

$$\cos\left(\frac{\pi J_{12}}{2}t\right) - i\sigma_z^1\sigma_z^2 \sin\left(\frac{\pi J_{12}}{2}t\right)$$

If we choose t such that

$$\frac{\pi J_{12}}{2}t = \frac{\pi}{4}$$

then we obtain the transformation

$$U = \frac{1}{\sqrt{2}}(1 - i\sigma_z^1\sigma_z^2)$$

It is instructive to write this as a matrix. Clearly, the matrix is already diagonal on the standard basis, and the second term is $-i$ if both qubits have the same value and i otherwise. Thus

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1-i & & & \\ & 1+i & & \\ & & 1+i & \\ & & & 1-i \end{pmatrix}$$

Let us now compute the product

$$\begin{aligned} R_{z^2}\left(-\frac{\pi}{2}\right)R_{z^1}\left(-\frac{\pi}{2}\right)U &= \frac{1}{2\sqrt{2}}(1 + i\sigma_z^1)(1 + i\sigma_z^2)(1 - i\sigma_z^1\sigma_z^2) \\ &= \frac{1+i}{2\sqrt{2}}[1 + \sigma_z^1 + \sigma_z^2 - \sigma_z^1\sigma_z^2] \end{aligned}$$

Again, this matrix is diagonal. If both qubits have the same value, the first and the last term cancel, and the term in the parentheses is either $+2$ or -2 . If both qubits have a different value, the two terms in the middle cancel and the parentheses gives a $+2$. Thus the matrix is

$$R_{z^2}\left(-\frac{\pi}{2}\right)R_{z^1}\left(-\frac{\pi}{2}\right)U = \frac{1+i}{\sqrt{2}} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}$$

The matrix in this equation is known as the **controlled phase gate** C_{PHASE} . It is well known (see for instance [6] section 7.4.2) that this gate is equivalent to the CNOT gate, up to single qubit gates. Specifically,

$$C_{NOT} = (I \otimes H)C_{PHASE}(I \otimes H)$$

which can easily be seen by applying the sequence on the basis vectors.

It is interesting to compare the switching times for the gates considered so far. Single qubit gates are realized by applying an RF field for a time that is in the order of the inverse of the nutation frequency. For instance, the Larmor frequency for a proton at 11.74 Tesla is 500 MHz. If the RF field is weaker by a factor 10^3 than the static field, the nutation frequency will be 500 kHz. This results in switching times in the order of a few microseconds. The switching time of the CNOT gate, in contrast, is determined by the coupling term J_{12} . This is usually in the order of a few hundred Hz, so that several milliseconds will be required to apply a CNOT gate. In addition, in molecules with more than two qubits, the coupling decreases quickly with the spatial distance between the qubits. To implement operations involving distant qubits, we first need to apply swap gates (which can be implemented as a combination of three CNOT gates) to propagate the states of the qubits that we want to combine across the molecule and then back, consuming additional time. Thus quantum gates implemented in NMR computing are rather slow.

8 Measurements

Let us now take a closer look at the process of measurement and understand how the actual measurement process in an NMR experiment works. Our starting point will be the density matrix $\tilde{\rho}$ in the logical frame that we have introduced in the previous section. This density matrix is related to the density matrix in the laboratory frame by

$$\tilde{\rho} = T\rho T^{-1}$$

where T is the transformation described by equation (6). At any point in time, the density matrix, being a hermitian matrix, can be written as a sum of products of Pauli matrices. Thus, we can decompose

$$\tilde{\rho} = \frac{1}{2^N} + \sum_i \tilde{m}^i(t) \cdot I^i + \tilde{c}(t)$$

where $\tilde{c}(t)$ collects all terms that contain Pauli matrices acting on at least two different nuclei and I^i is the vector whose components are the spin angular momentum operators acting on qubit i .

If we calculate the average of a spin, say I_x^i for some i , i.e. if we calculate $\text{tr}(\rho I_x^i)$, the only term that will contribute the trace is the term $m_x^i I_x^i$. All other terms contain at least one Pauli matrix and therefore have trace zero. Thus we can interpret the time dependent and vector valued function \tilde{m} as the magnetic moment of the spin ensemble.

Let us now make the transformation back into the laboratory frame. We know that conjugation with T will map a term that contains at least two Pauli matrices into a similar term. A term containing no Pauli matrices at all is invariant, and a term that contains one Pauli matrix will again yield a term containing one Pauli matrix. Thus we can write

$$\rho = T^{-1}\tilde{\rho}T = \frac{1}{2^N} + \sum_i T^{-1}\tilde{m}^i \cdot I^i T + c(t)$$

Again, the second term will represent the magnetic moment in the laboratory frame. Let us take a closer look at this term. First, the transformation T is a product of terms that each act on one qubit only. So in a product

$$T^{-1} \tilde{m}^i \cdot I^i T$$

almost all factors in T commute through the term, and we find that we can write the magnetization in the laboratory frame as

$$m(t) = \sum_i R^{z^i}(\omega_i t) \tilde{m}^i \cdot I^i R^{z^i}(-\omega_i t)$$

Using again the relations in the appendix and sorting terms, we find that

$$\begin{aligned} m_x^i(t) &= \tilde{m}_x(t) \cos \omega_i t - \tilde{m}_y(t) \sin \omega_i t \\ m_y^i(t) &= \tilde{m}_y(t) \cos \omega_i t + \tilde{m}_x(t) \sin \omega_i t \\ m_z^i(t) &= \tilde{m}_z(t) \end{aligned}$$

This is not very surprising and simply says that the magnetization in the laboratory frame is related to the magnetization in the logical frame by a rotation around the angle $\omega_i t$.

Let us now again suppose that we have a receiver coil placed at a point on the x-axis that is also oriented along the x-axis. This coil is able to detect the change of the magnetic field in the x-direction. This change is proportional to the change of the magnetic moment in the x-direction. Therefore the induced voltage is proportional to

$$v(t) \approx \sum_i a_i [\tilde{m}_x(t) \sin \omega_i t + \tilde{m}_y(t) \cos \omega_i t]$$

where we assume that the magnetization in the logical frame changes so slowly that the time derivative of the magnetization \tilde{m} can be ignored and the a_i are constants depending on the geometry of the coil, the gyromagnetic ratio of the nucleus, the number of nuclei in the probe and the strength of the magnetic field.

Now that voltage is sent into a circuit that looks as displayed in diagram 7 and is known as a **quadrature receiver** (see also A.5 in [7]).

In this circuit, the signal from the coil is first split and then follows two paths through the circuit. Each path contains a mixer, that essentially multiplies its inputs, and a low pass filter, that filters out high frequencies. The second input of each mixer is reference signal.

Let us look at the upper path first. Here, the reference signal is $\sin \omega_{ref} t$ with a reference frequency that can be adjusted. Thus, the output of the mixer is a sum of terms of the form

$$\sin(\omega_{ref} t) a_i [\tilde{m}_x(t) \sin \omega_i t + \tilde{m}_y(t) \cos \omega_i t]$$

We can calculate this, using the well known formulas for the products of trigonometric functions, and obtain

$$\begin{aligned} &a_i \left[\frac{1}{2} \tilde{m}_x(t) (\cos(\omega_{ref} - \omega_i)t) - \cos((\omega_{ref} + \omega_i)t) \right] + \\ &a_i \left[\frac{1}{2} \tilde{m}_y(t) (\sin(\omega_{ref} - \omega_i)t) + \sin((\omega_{ref} + \omega_i)t) \right] \end{aligned}$$

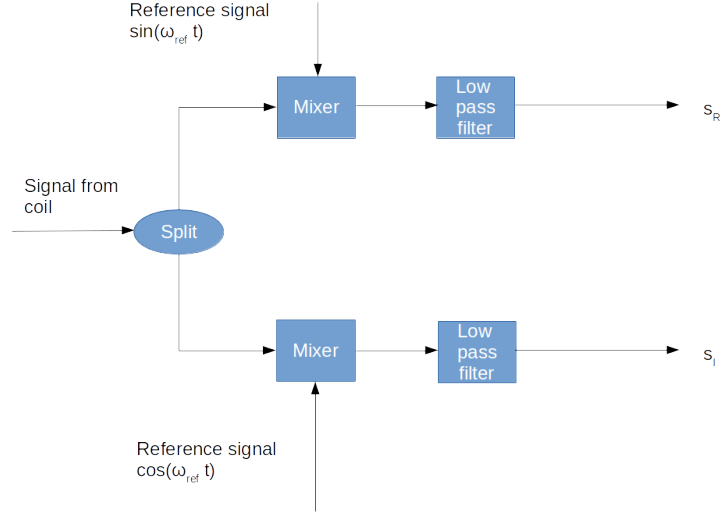


Figure 7: Quadrature receiver

Now let us assume that the reference frequency ω_{ref} is close to one of the ω_i . The low pass filter will then let the oscillations with frequency $\omega_{ref} - \omega_i$ pass, but will suppress the oscillations with frequency $\omega_{ref} + \omega_i$. Thus the output of the upper path of the circuit will be proportional to

$$s_R(t) = \sum_i a_i [\tilde{m}_x(t) \cos \Omega^i t + \tilde{m}_y(t) \sin \Omega^i t]$$

where

$$\Omega^i = \omega_{ref} - \omega_i$$

Similarly, we can calculate the output of the lower path. Each term contributes a term of the form

$$\cos(\omega_{ref} t) \omega_i [\tilde{m}_x(t) \sin \omega_i t + \tilde{m}_y(t) \cos \omega_i t]$$

which can be written as

$$\begin{aligned} & a_i \left[\frac{1}{2} \tilde{m}_x(t) (\sin(\omega_i - \omega_{ref})t + \sin((\omega_{ref} + \omega_i)t)) \right] + \\ & a_i \left[\frac{1}{2} \tilde{m}_y(t) (\cos(\omega_{ref} - \omega_i)t + \cos((\omega_{ref} - \omega_i)t)) \right] \end{aligned}$$

Again, the low pass filter will suppress the terms that oscillate with frequency $\omega_i + \omega_{ref}$, and we obtain that the output of the lower path is proportional to

$$s_I(t) = \sum_i a_i [-\tilde{m}_x(t) \sin \Omega^i t + \tilde{m}_y(t) \cos \Omega^i t]$$

Thus we find that the two terms differ by a phase shift of $\frac{\pi}{2}$. In fact, the term s_I is proportional to the induced voltage that a coil placed along the y-axis would

detect. It is common practice to combine these two signals into one complex signal given by

$$s(t) = s_R(t) + is_I(t)$$

which can then be written as

$$s(t) = \sum_i a_i (\tilde{m}_x(t) + i\tilde{m}_y(t)) e^{i\Omega^i t}$$

Thus the signal that we receive is a sum of oscillations with frequencies Ω^i , overlayed by the change of the magnetization, so that we can recover the magnetization in the x- and y-direction in the local frame from the Fourier transform of $s(t)$. Essentially, we are therefore able to measure both, the magnetization in the x-direction and the magnetization in the y-direction, at the same time, i.e. the expectation values of the operators I_x^i and I_y^i for arbitrary i (assuming that the chemical shifts are larger than the resolution of the measurement so that we can distinguish the contributions from the individual Ω^i). Equivalently, we are able to measure the expectation values of the (non-hermitian) operators

$$I_x^i + iI_y^i$$

for every spin i in the logical frame.

It is worth noting that there is a fundamental difference between the measurement in an NMR quantum computer and the usual measurement in quantum computing. To a very good approximation, the measurement is a macroscopic measurement, i.e. it is not projecting the state of the system onto an eigenstate. Instead, we measure averages, i.e. expectation values. This is very useful, as it allows us - as we have seen - to measure the values of the non-commuting operators I_x and I_y . On the other hand, some quantum algorithms use the projective measurement as a means of preparing a certain state. Shor's algorithm, for instance, ends with a measurement and has a success probability different from one. In an NMR setup, the algorithm would not fail or be successful, but we would measure a - useless - average of failed and successful execution. To deal with this, we either have to find a way to extract useful information from the average - which works in some cases - or move the classical post processing into the quantum part, i.e. execute the entire post processing as a quantum algorithm until we obtain a definite result, see the discussions in [10] and [12].

9 State preparation

In the previous sections, we have made use of pseudo-pure states of the form

$$\rho = \frac{1}{2^N} (1 - \epsilon) + \epsilon |\psi\rangle\langle\psi|$$

and demonstrated that for the purposes of quantum computing, these states can be employed in a way very similar to pure states of an individual spin. In this section, we will briefly touch on methods to prepare a system in such a state. The main references for this section are [11], [12], [13] and the material in [6].

First, let us start with an approach that is called **temporal averaging**. To explain this, let us take a closer look at the density matrix of a state in thermal

equilibrium

$$\rho = \frac{1}{Z} \exp \left[-\frac{H}{kT} \right] \approx 1 - \sum_i \frac{1}{2} \frac{\omega_i \hbar}{kT} \sigma_z^i$$

We can write this in matrix notation as

$$\rho = \begin{pmatrix} a & & \dots \\ & b_1 & \dots \\ & & \ddots \\ & & & b_n \end{pmatrix}$$

where a is determined by the energy of the ground state and the b_i are related to the energies of the excited states. Assuming that the gyromagnetic ratios are all positive, $a > b_i$ for all i .

Now suppose for a moment that all the b_i are equal, i.e. $b_i = b$ for some b and all i . Then, this matrix could be written as

$$\rho = \begin{pmatrix} b & & \dots \\ & b & \dots \\ & & \ddots \\ & & & b \end{pmatrix} + \begin{pmatrix} a-b & & \dots \\ & 0 & \dots \\ & & \ddots \\ & & & 0 \end{pmatrix} = b + (a-b)|0\rangle\langle 0|$$

which is a pseudo-pure state. Of course, in general, the b_i are not all equal. The idea of temporal averaging is to take the average over several experiments to obtain the same results as if all the b_i were equal.

To achieve this, let U_i be a unitary transformation that cyclically permutes the basis vectors $|1\rangle, \dots, |2^N - 1\rangle$ by i positions, i.e.

$$\begin{aligned} U_1|1\rangle &= |2\rangle \\ U_1|2\rangle &= |3\rangle \\ &\dots \\ U_1|2^N - 1\rangle &= |1\rangle \\ U_2|1\rangle &= |3\rangle \end{aligned}$$

and so forth. Conjugation of the density matrix with one of the U_i then simply cyclically permutes the entries b_i along the diagonal, for instance

$$U_1 \rho U_1^{-1} = \begin{pmatrix} b & & \dots \\ & b_2 & \dots \\ & & \ddots \\ & & & b_1 \end{pmatrix}$$

There are $n = 2^N - 1$ cyclic permutations, and each permutation i gives us a new density matrix, obtained by acting with U_i . As the U_i are unitary, they can be realized as quantum gates, so starting from a state ρ in thermal equilibrium, we can prepare each of the states $U_i \rho U_i^{-1}$. Let us now sum up all these states and divide by n to normalize again. We find that

$$\bar{\rho} = \sum_i U_i \rho U_i^{-1} = \begin{pmatrix} a & & \dots \\ & \frac{1}{n} \sum b_i & \dots \\ & & \ddots \\ & & & \frac{1}{n} \sum_i b_i \end{pmatrix}$$

This matrix is now a pseudo-pure state, as discussed before. Even though we cannot prepare this state, we can prepare any of the $U_i \rho U_i^{-1}$. To execute a quantum algorithm, we can now prepare each of these states, run the algorithm, measure the outcome, and eventually average over all these executions. The result will be the same as if we had started with the state $\bar{\rho}$, i.e. with a pseudo-pure state. As we average over several executions, this method is known as **temporal averaging**.

To implement this idea, we need to execute $2^N - 1$ experiments instead of just one. Obviously, this simple approach does not scale well and only works for a small number of qubits. However, in practice, there are additional relations between the b_i . If, for instance, some of our qubits are realized by the same isotope, some of the b_i that differ only by the spin polarizations of these qubits will be very close to each other. Therefore, one can usually simplify this approach considerably and work with a much smaller number of repetitions, so that this method becomes feasible - see for instance [13] for a more detailed analysis.

Alternatively, modern NMR spectrometers usually have coils that can create a magnetic field depending on the location of the nucleus in space, so called **gradient coils**. Using this, it is theoretically possible to prepare different initial states in different, spatially separated regions of the probe. We can then run the algorithm only once and average the resulting signal over the entire probe (in fact, an ordinary receiver coil will do that anyway). This technique is called **spatial averaging**.

As an alternative to method using averaging, that utilize all spins for computations, there is a method known as **labeling** that restricts the computations to a suitable subspace on which the density matrix appears to be pure state. To illustrate this idea, let us consider the case of a homonuclear system with three spins. The traceless part of the thermal equilibrium density matrix will be proportional to

$$3|0\rangle\langle 0| + 1|1\rangle\langle 1| + 1|2\rangle\langle 2| - 1|3\rangle\langle 3| + 1|4\rangle\langle 4| - 1|5\rangle\langle 5| - 1|6\rangle\langle 6| - 3|7\rangle\langle 7|$$

As all spins have roughly the same Larmor frequency, the diagonal entries at position (k, k) only depend on the number of bits set in the binary representation of k . By rearranging the order of the basis vectors, which is a unitary transformation, we can bring this into a form where all the +1 and -1 coefficients are grouped, i.e.

$$3|0\rangle\langle 0| + 1|1\rangle\langle 1| + 1|2\rangle\langle 2| + 1|3\rangle\langle 3| - 1|4\rangle\langle 4| - 1|5\rangle\langle 5| - 1|6\rangle\langle 6| - 3|7\rangle\langle 7|$$

On the subspace spanned by the first four basis vectors, i.e. those having the third bit equal to zero, this now looks like a pseudo-pure state. In fact, we can write this as

$$\begin{aligned} &|0\rangle\langle 0| \otimes [3|0\rangle\langle 0| + 1|1\rangle\langle 1| + 1|2\rangle\langle 2| + 1|3\rangle\langle 3|] \\ &- \\ &|1\rangle\langle 1| \otimes [1|0\rangle\langle 0| + 1|1\rangle\langle 1| + 1|2\rangle\langle 2| + 3|3\rangle\langle 3|] \end{aligned}$$

where we have adjusted our notation by writing the most significant bit as a separate register, so that for instance

$$|1\rangle \otimes |2\rangle = |1\rangle \otimes |10\rangle = |110\rangle = |5\rangle$$

Thus, we can write the thermal equilibrium state as

$$\rho = |0\rangle\langle 0| \otimes \rho_0 + |1\rangle\langle 1| \otimes \rho_1$$

where ρ_0 is a pseudo-pure state. When we now apply a unitary transformation U only on the space on which ρ_0 is acting, the result of the calculation will again be of this type:

$$U\rho U^{-1} = |0\rangle\langle 0| \otimes U\rho_0 U^{-1} + |1\rangle\langle 1| \otimes U\rho_1 U^{-1}$$

We can think of the state of the first qubit as a label that carves out the subspace on which the pseudo-pure density matrix acts. This technique is known as **logical labeling**. More generally, we can prepare states of the form

$$\rho = \sum_k |k\rangle\langle k| \otimes \rho_k$$

where ρ_0 is a pseudo-pure state and the other ρ_k are garbage states that we ignore. We have shown explicitly how such a state can be prepared for three qubits, but it is not difficult to see that this works for eneral values of N , see also [14].

How do we measure the result of a calculation in that approach? Typically, we eventually want to measure σ_z^i for some $i > 0$. Let us assume that we have a method to do this in place (namely by rotating the spin polarization axis around the y-axis and measuring the resulting magnetization in the x-y-plane as discussed in the previous sections). We perform a first read-out operation, i.e. we determine

$$\begin{aligned} \text{tr}(\rho \sigma_z^i) &= \sum_k \text{tr}(|k\rangle\langle k| \otimes \rho_k \sigma_z^i) \\ &= \sum_k \text{tr}(\rho_k \sigma_z^i) = \text{tr}(\rho_0 \sigma_z^i) + \sum_{k>0} \text{tr}(\rho_k \sigma_z^i) \end{aligned}$$

The first term is what we actually want to measure, the second term is garbage that we somehow have to filter out. Following [14], this can be achieved as follows. We first apply the unitary transformation

$$R = |0\rangle\langle 0| \otimes I + \sum_{k>0} |k\rangle\langle k| \otimes \sigma_x^i$$

In other words, we apply σ_x on qubit i conditioned on the value of k , which can be realized using CNOT gates. Intuitively, we flip all spins except the one labeled by $k = 0$. The resulting state will now be

$$\rho' = |0\rangle\langle 0| \otimes \rho_0 + \sum_{k>0} |k\rangle\langle k| \otimes \sigma_x^i \rho_k \sigma_x^i$$

We now conduct another measurement of σ_z^i , i.e. we measure

$$\begin{aligned} \text{tr}(\rho' \sigma_z^i) &= \text{tr}(\rho_0 \sigma_z^i) + \sum_{k>1} \text{tr}(\sigma_x^i \rho_k \sigma_x^i \sigma_z^i) \\ &= \text{tr}(\rho_0 \sigma_z^i) - \sum_{k>1} \text{tr}(\sigma_x^i \rho_k \sigma_z^i \sigma_x^i) \\ &= \text{tr}(\rho_0 \sigma_z^i) - \sum_{k>1} \text{tr}(\rho_k \sigma_z^i \sigma_x^i \sigma_x^i) = \text{tr}(\rho_0 \sigma_z^i) - \sum_{k>1} \text{tr}(\rho_k \sigma_z^i) \end{aligned}$$

If we add the result of this measurement with the previous measurement, the second terms cancel out and we are left with the result that the pseudo-pure state alone would give us.

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

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

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

$$R_x(\phi)I_y R_x(-\phi) = \cos \phi I_y + \sin \phi I_z \quad (7)$$

This obviously implies that

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

for all integers k and therefore

$$R_x(\phi)R_y(\Theta)R_x(-\phi) = \exp \left[-i \frac{\Theta}{\hbar} (\cos \phi I_y + \sin \phi I_z) \right] \quad (8)$$

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

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

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 ,

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

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

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

Another fact that we will need is the fact that any rotation of the Bloch sphere can be written as a product of rotations around the x-axis and y-axis. To see that this is true, assume that are given a vector n and a rotation around n . First, let us consider the special case that n is in the y-z-plane, i.e.

$$n = ae_y + be_z$$

Let v be a vector in the y-z-plane perpendicular to n , such that (e_x, n, v) is a positively oriented triple, i.e. $v = e_x \times n$. Then the rotation R around n is given by the mapping rules

$$\begin{aligned} e_x &\mapsto e_x \cos \varphi - v \sin \varphi \\ v &\mapsto e_x \sin \varphi + v \cos \varphi \\ n &\mapsto n \end{aligned}$$

for some rotation angle φ . Let now R_1 be the rotation around the x-axis that rotates n onto e_y . Of course $R_1(e_x) = e_x$ and by our choice of v , $R_1(v) = e_z$. Thus R_1 maps the coordinate axes

$$(e_x, n, v)$$

to

$$(e_x, e_y, e_z)$$

Let R_2 be the rotation around e_y by the angle φ , i.e. R_2 is given by

$$\begin{aligned} e_x &\mapsto e_x \cos \varphi - e_z \sin \varphi \\ e_z &\mapsto e_x \sin \varphi + e_z \cos \varphi \\ e_y &\mapsto e_y \end{aligned}$$

We now claim that

$$R = R_1^{-1} R_2 R_1$$

To see that this is true, it suffices to verify this equation on the axis. First, we have

$$\begin{aligned} (R_1^{-1} R_2 R_1) e_x &= (R_1^{-1} R_2) e_x \\ &= R_1^{-1} [e_x \cos \varphi - e_z \sin \varphi] \\ &= e_x \cos \varphi - v \sin \varphi = R e_x \end{aligned}$$

Next, let us evaluate this on n . We find that

$$(R_1^{-1}R_2R_1)n = R_1^{-1}R_2e_y = R_1^{-1}e_y = n$$

as expected. Finally,

$$\begin{aligned} (R_1^{-1}R_2R_1)v &= (R_1^{-1}R_2)e_z \\ &= R_1^{-1}[e_x \sin \varphi + e_z \cos \varphi] \\ &= e_x \sin \varphi + v \cos \varphi = Rv \end{aligned}$$

This proves our claim for the special case of a rotation around an axis in the y-z-plane. Now suppose we are given a rotation R around an arbitrary axis n . Let R_1 be a rotation around the z-axis that rotates n onto a vector m in the y-z-plane. Then $R' = R_1RR_1^{-1}$ is a rotation that maps m to itself, i.e. a rotation around m . Thus we can decompose R' as a product of rotations around the x-axis and the y-axis. The same is true for R_1 , being a rotation around an axis in the y-z-plane, and therefore for R .

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

$$\langle A \rangle = \sum_a a \langle \psi | a \rangle \langle a | \psi \rangle = \sum_a a |\langle \psi | a \rangle|^2 \quad (9)$$

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 (9) as matrix elements of an operator. In fact,

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

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 (9) 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**. In case of a time-independent Hamiltonian, this is solved by

$$\rho(t) = U(t)\rho(0)U(t)^{-1}$$

where

$$U(t) = e^{-\frac{i}{\hbar}Ht}$$

is the usual time evolution operator. In fact, using the product rule, we can

differentiate this expression and find that

$$\begin{aligned}
 \frac{d}{dt}U(t)\rho(0)U(t)^{-1} &= \left[\frac{d}{dt}U(t)\right]\rho(0)U(t)^{-1} + U(t)\rho(0)\frac{d}{dt}U(t)^{-1} \\
 &= -\frac{i}{\hbar}HU(t)\rho(0)U(t)^{-1} + \frac{i}{\hbar}U(t)\rho(0)U(t)^{-1}H \\
 &= -\frac{i}{\hbar}[H\rho(t) - \rho(t)H] \\
 &= -\frac{i}{\hbar}[H, \rho]
 \end{aligned}$$

where we have used that $U(t)$ and H commute.

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_x . 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
- [8] E. Knill et.al., Introduction to NMR Quantum Information Processing, arXiv:quant-ph/0207172
- [9] J.A. Jones, M. Mosca, Implementation of a Quantum Algorithm to Solve Deutsch's Problem on a Nuclear Magnetic Resonance Quantum Computer, *J.Chem.Phys.* 109 (1998) 1648-1653 or arXiv:quant-ph/9801027
- [10] L.M.K. Vandersypen et. al., Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance, *Nature* Vol. 414, Dec. 2001
- [11] D.G. Cory, A.F.Fahmy, T.F. Havel, Ensemble Quantum Computing by NMR Spectroscopy, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 94, No. 5 (Mar. 4, 1997), pp. 1634-1639
- [12] N.A. Gershenfeld, I.L Chuang, Bulk Spin-Resonance Quantum Computation, *Science* Vol. 275, Jan. 1997
- [13] E. Knill, I.L. Chuang, R. Laflamme, Effective pure states for bulk quantum computation, *Phys. Rev. A* Vol. 57 No. 5
- [14] I.L. Chuang, N. Gershenfeld, M.G. Kubinec, D.W. Leung, Bulk Quantum Computation with Nuclear Magnetic Resonance: Theory and Experiment, *Proceedings: Mathematical, Physical and Engineering Sciences*, Vol. 454, No. 1969, Quantum Coherence and Decoherence (Jan. 8, 1998), pp. 447-467
- [15] D.P. DiVincenzo, The Physical Implementation of Quantum Computation, arXiv:quant-ph/0002077