3 Spin $\frac{1}{2}$

3.1 The Stern-Gerlach experiment

In the Stern-Gerlach experiment, a beam of silver atoms with a relatively uniform velocity in the y direction is prepared.[9] This beam then is taken through a region with an inhomogeneous magnetic field, approximately in the z-direction. Of course, one cannot have an inhomogeneous magnetic field exactly in a single direction as this would contradict Maxwell's equation $\mathrm{div}\vec{B} = 0$. Instead the magnetic field slightly spreads out in the x direction in the regions where it is weaker. However, one may arrange that the beam of atoms goes through the region around $x = 0$ which is the symmetry plane of the magnetic field and where it is only in the z-direction. After having travelled through the region with the inhomogeneous magnetic field the silver atoms are detected on a screen.

Let us begin by giving the classically expected description of this experiment. The silver atoms are known to possess a magnetic moment $\vec{m}$. The energy of such a magnetic moment in the magnetic field is $E = -\vec{m}\cdot\vec{B} = -m_z B_z - m_x B_x - m_y B_y$. Since $\vec{B}$ is inhomogeneous, this energy also varies with the position and taking its gradient results in a force :

$$\vec{F} = -\vec{\nabla}E = m_z\vec{\nabla}B_z + m_x\vec{\nabla}B_x + m_y\vec{\nabla}B_y \ . \tag{3.1}$$

One must keep in mind that, although B_z is much larger than the other components (in the region of our beam), at least two of the derivatives $\frac{\partial B_x}{\partial x}$, $\frac{\partial B_y}{\partial y}$ and $\frac{\partial B_z}{\partial z}$ must be of the same order to ensure $\mathrm{div}\vec{B} = 0$.

Another effect of the magnetic field on the magnetic moment is to produce a torque $\vec{\Gamma} = \vec{m}\times\vec{B}$. The magnetic moment of the atoms is expected to be due to some angular momentum of the atom. One can make a model where this magnetic moment is due to the supposedly circular orbit of an electron which results in the relation[10]

$$\vec{m} = \gamma\vec{L} \ , \tag{3.2}$$

where $\vec{L}$ is the resulting angular momentum, $\gamma = \frac{q_e}{2m_e}$ is called the gyromagnetic factor, and q_e and m_e are the electric charge and mass of the electron. The torque implies that

$$\frac{\mathrm{d}\vec{L}}{\mathrm{d}t} = \vec{\Gamma} = \vec{m}\times\vec{B} \quad \Rightarrow \quad \frac{\mathrm{d}\vec{m}}{\mathrm{d}t} = \frac{q_e}{2m_e}\,\vec{m}\times\vec{B} \ . \tag{3.3}$$

[9] This can be done e.g. by evaporating the atoms in a hot cavity. They can escape from the cavity through a hole, with their thermal velocities. A first screen with a single hole selects only the atoms with velocities in the y direction. As the atoms continue they also "fall" in the z direction under the effect of gravity. By placing a second screen with a single hole further in the y-direction but at a slightly lower z coordinate than the first hole, one selects those atoms with a precise value of the velocity in the y-direction.

[10] If an electric charge q of mass m moves on a circular orbit of radius a with angular frequency ω, its angular momentum is $L = ma^2\omega$. On the other hand, the period is $T = 2\pi/\omega$ and this is equivalent to a loop of current $I = q/T = q\omega/(2\pi)$. The magnetic moment of such a loop is known to be $\pi a^2\, I = a^2 q\omega/2 = \frac{q}{2m}L$, so that the gyromagnetic ratio is $\gamma = \frac{q}{2m}$.

The second equation implies that the magnetic moment precesses around the direction of the magnetic field (the z-direction) with the so-called Larmor frequency $\omega_0 = \frac{q_e B}{2m_e}$. Since the electron mass is small, this frequency is large enough, so that the time spent in the magnetic field corresponds to several periods. This means that m_x and m_y average to zero, while m_z remains constant. Then, (3.1) reduces to $\vec{F} = m_z \vec{\nabla} B_z \simeq m_z \frac{\partial B_z}{\partial z} \vec{e}_z$, and the effect of the inhomogeneous magnetic field on the beam is to produce a deviation by an angle α in the z direction. Obviously, this deviation depends on the value of m_z for a given atom. This angle can be shown to be[11]

$$\alpha = \frac{\partial B_z}{\partial z} \frac{a}{m_{\text{atom}} v_y^2} m_z = \gamma \frac{\partial B_z}{\partial z} \frac{a}{m_{\text{atom}} v_y^2} L_z \ , \tag{3.4}$$

where a is the spatial extension of the magnetic field in the y-direction. The important feature of this formula is that measuring the deviation angle immediately gives us the measurement of the z-component of the magnetic moment and, if we can trust the classical value of the gyromagnetic ratio γ, also of the z-component of the angular momentum. Classically, we expect that the magnetic moment for the different atoms is randomly oriented, so that m_z takes any value between $-|\vec{m}|$ and $+|\vec{m}|$. Accordingly, we expect that the impacts on the detection screen form a continuous vertical bar.

What did Stern and Gerlach observe ? Instead of a continuous vertical bar, they just observed *two* spots, symmetrically distributed above and below the central point (where the impacts would occur without magnetic field). The separation between the two spots is consistent with

$$m_z = \pm \frac{q_e}{2m_e} \hbar = \frac{q_e}{m_e} \left(\pm \frac{\hbar}{2} \right) \equiv \pm \mu_0 \ . \tag{3.5}$$

If the experiment is repeated with helium atoms, no deviation is observed and there is just a single central spot on the detection screen. Indeed, helium atoms have no magnetic moment. This confirms that the only relevant "degree of freedom" for a quantum mechanical description of the Stern-Gerlach experiment is the magnetic moment, and that we can totally disregard the "spatial degrees of freedom" which are simply described classically (as we indeed did above).

3.2 Quantum-mechanical description of the Stern-Gerlach experiment

If we apply the postulates of quantum mechanics as enumerated above, the z-component of the magnetic moment should be an observable that has 2 eigenvalues and that the corresponding Hilbert space should be 2-dimensional, i.e. our $\mathcal{H} \simeq \mathbf{C}^2$ studied in detail above. Actually we can immediately give the complete description as follows. Let us denote the quantum mechanical

[11] If v_y is the well-defined velocity of the atoms in the beam, and a the spatial extension of the magnetic field in the y direction, a given atom will experience the force for a time interval $\Delta t = \frac{a}{v_y}$ and hence from $m_{\text{atom}} \frac{\mathrm{d}v_z}{\mathrm{d}t} = F_z$ we get $m_{\text{atom}} \Delta v_z = F_z \Delta t = m_z \frac{\partial B_z}{\partial z} \frac{a}{v_y}$ But initially, $v_z = 0$ so that Δv_z is just the z-component of the velocity as the atom exists the region of the magnetic field. Now, $\frac{v_z}{v_y} = \tan \alpha \simeq \alpha$ is the angle of deviation of the beam and we get $\alpha \simeq \frac{\partial B_z}{\partial z} \frac{a}{m_{\text{atom}} v_y^2} m_z$.

observable (linear operator) corresponding to m_z by $\widehat{m}_z$. Then we can choose any orthonormal basis in $\mathcal{H}$, say $|0\rangle$ and $|1\rangle$ and declare that they are the eigenstates corresponding to the two measured eigenvalues (3.5). Then

$$\widehat{m}_z = \mu_0\big(\,|0\rangle\,\langle 0| - |1\rangle\,\langle 1|\,\big) \quad , \quad \widehat{m}_z\,|0\rangle = \mu_0\,|0\rangle \quad , \quad \widehat{m}_z\,|1\rangle = -\mu_0\,|1\rangle \ . \tag{3.6}$$

Since the $|0\rangle$ and $|1\rangle$ correspond to the eigenvalues $\pm\mu_0$ of $\widehat{m}_z$ it will actually be useful to rename these states as

$$|0\rangle \equiv |+\rangle_z \quad , \quad |1\rangle \equiv |-\rangle_z \ , \tag{3.7}$$

so that (3.6) is rewritten as

$$\widehat{m}_z = \mu_0\big(\,|+\rangle_z\,\langle +|_z - |-\rangle_z\,\langle -|_z\,\big) \quad , \quad \widehat{m}_z\,|+\rangle_z = \mu_0\,|+\rangle_z \quad , \quad \widehat{m}_z\,|-\rangle_z = -\mu_0\,|-\rangle_z \ . \tag{3.8}$$

In any case, the matrix corresponding to $\widehat{m}_z$ is

$$\begin{pmatrix} \langle +|_z\,\widehat{m}_z\,|+\rangle_z & \langle +|_z\,\widehat{m}_z\,|-\rangle_z \\ \langle -|_z\,\widehat{m}_z\,|+\rangle_z & \langle -|_z\,\widehat{m}_z\,|-\rangle_z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \equiv \sigma_z \ . \tag{3.9}$$

For every single atom, if we measure the z-component of the magnetic moment one finds either $+\mu_0$ or $-\mu_0$. For any given atom, the relative probabilities depend on what exactly was the state ψ before the measurement. Once we measure $+\mu_0$ we know that the atom is in the state $|+\rangle_z$: the "wave packet" has been reduced, i.e. the initial state has been projected onto the eigenstate $|+\rangle_z$. Similarly if we measure $-\mu_0$. To try to get information on the initial state $|\psi\rangle$, we now repeat the experiment with two Stern-Gerlach setups. The first setup is used to select only the atoms with z-component $+\mu_0$ by using only the atoms with the corresponding deviation. Then we know that the state is $|+\rangle_z$ for all the atoms in this beam. We then use this beam as the incoming beam for the second Stern-Gerlach setup. If the magnetic field (and its gradient) is again along the z-axis there is no surprise, one just gets a single spot on the detection screen corresponding to a further deviation in the z-direction. This is consistent with the postulates : Since the initial state now is an eigenstate of $\widehat{m}_z$ we get $+\mu_0$ with probability one.

Let us now turn the second Stern-Gerlach by 90^o so that the magnetic field and its gradient now are in the x-direction, so that one measures $\widehat{m}_x$. One then finds that there are again two spots consistent with $m_x = \pm\mu_0$, and with equal intensity. This means that the initial state $|+\rangle_z$ must decompose on the eigenstates of $\widehat{m}_x$ (which we call $|\pm\rangle_x$) with amplitudes of equal modulus squared $\frac{1}{2}$, i.e. $|+\rangle_z = \alpha\,|+\rangle_x + \beta\,|-\rangle_x$ with $|\alpha|^2 = |\beta|^2 = \frac{1}{2}$. A similar discussion holds if we had prepared $|-\rangle_z$ by the first Sterrn-Gerlach. But the required relation is exactly what the change of basis (2.18) expresses. We conclude that we can identify $|\pm\rangle_x$ with the $|\pm\rangle$ of (2.18). Of course, it would have been equivalent to identify the $|\pm\rangle_x$ with the $|R\rangle$ and $|L\rangle$, but let's make the *choice* to keep with the first option. We can also repeat two Stern-Gerlach experiments with orientations along the z-axis and then along the y-axis with a similar conclusion. Finally we may prepare a state $|+\rangle_x$ or $|-\rangle_x$ with a first Stern-Gerlach and then measure m_y with the second Stern-Gerlach,

again with the conclusion that $|+\rangle_x$ and $|-\rangle_x$ must both be superpositions of $|+\rangle_y$ and $|-\rangle_y$ with amplitudes of modulus squared $\frac{1}{2}$. In conclusion, the eigenstates of $\widehat{m}_x$, $\widehat{m}_y$ and $\widehat{m}_z$ must consist of three different orthonormal basis, such that any vector of one basis can be expressed as a superposition of the two vectors of any other basis with amplitudes of modulus squared equal to $\frac{1}{2}$. This is exactly the case for the 3 orthonormal basis introduced in (2.13) with the relations (2.18). Hence, we identify :

$$|+\rangle_z = |0\rangle \ , \ |-\rangle_z = |1\rangle \quad , \quad |+\rangle_x = |+\rangle \ , \ |-\rangle_x = |-\rangle \quad , \quad |+\rangle_y = |R\rangle \ , \ |-\rangle_y = |L\rangle \ . \tag{3.10}$$

For historical reasons, the z-basis $|\pm\rangle_z$ is the "preferred" basis. We can then express the $|\pm\rangle_x$ and $|\pm\rangle_y$ in the z-basis by rewriting (2.18) as

$$\begin{aligned} |+\rangle_x &= \frac{1}{\sqrt{2}}(|+\rangle_z + |-\rangle_z) \quad , \quad |-\rangle_x = \frac{1}{\sqrt{2}}(|+\rangle_z - |-\rangle_z) \\ |+\rangle_y &= \frac{1}{\sqrt{2}}(|+\rangle_z + i\,|-\rangle_z) \quad , \quad |-\rangle_y = \frac{1}{\sqrt{2}}(|+\rangle_z - i\,|-\rangle_z) \end{aligned} \tag{3.11}$$

We have already given the observable $\widehat{m}_z$ in the z-basis. The observables $\widehat{m}_x$ and $\widehat{m}_y$ must then be given in total analogy by

$$\widehat{m}_x = \mu_0\big(\,|+\rangle_x\,\langle+|_x - |-\rangle_x\,\langle-|_x\,\big) \quad , \quad \widehat{m}_y = \mu_0\big(\,|+\rangle_y\,\langle+|_y - |-\rangle_y\,\langle-|_y\,\big) \ , \tag{3.12}$$

so that, of course, $\widehat{m}_x\,|\pm\rangle_x = \pm\mu_0\,|\pm\rangle_x$ and $\widehat{m}_y\,|\pm\rangle_y = \pm\mu_0\,|\pm\rangle_y$.

3.3 Spin $\frac{1}{2}$

In the classical picture, we had associated the magnetic moment $\vec{m}$ with an angular momentum $\vec{L}$ according to $\vec{m} = \gamma\vec{L}$ with $\gamma = \frac{q_e}{2m_e}$. From the measured value (3.5) one would then conclude that the angular momentum L_z can only take the values $\pm\hbar$. However, as we will see later-on, when discussing angular momenta in quantum mechanics, if L_z can take values $\pm\hbar$ then it can also take a third value 0, so that there should also be a third eigenstate with $\mu_z = 0$. Now Stern and Gerlach did not observe any impact for zero deviation. They observed two spots and not three. This is resolved by assuming that the atoms carry an angular momentum of a different kind, called spin, and which is not due to the orbital movement of the electrons, but is an intrinsic property of the electron. This spin angular momentum is such that when measured one only finds the values $\pm\frac{\hbar}{2}$ (one says the particle has a spin $\frac{1}{2}$) and instead the gyromagnetic factor for this spin is twice as large, i.e.

$$\gamma_e = g_e \times \frac{q_e}{2m_e} \quad , \quad g_e = 2 \ . \tag{3.13}$$

The value $g_e = 2$ can easily be obtained by studying the relativistic Dirac equation for an electron, as we will briefly explain in the second part of these lectures. Actually this prefactor g_e is not exactly 2 but the difference $g_e - 2$ was too small to be observed in the Stern-Gerlach experiment.

However, modern experiments have been able to measure this $g_e - 2$ with incredible precision (up to 8 relevant digits), and this same value – with the same precision – has been confirmed by computing sophisticated correction terms ("four-loop diagrams") within relativistic quantum field theory ! Here, we simply continue to use the approximate value $g_e = 2$. The angular momentum associated with the spin is usually denoted by $\vec{S}$ or S_x, S_y and S_z, and we will use the same notation (without a hat) to denote the quantum mechanical observables (linear operators). We then arrive at

$$\widehat{m}_j = \frac{2\mu_0}{\hbar} S_j = \frac{q_e}{m_e} S_j \ , \quad j = x,\, y,\, z \quad \Leftrightarrow \quad \vec{\widehat{m}} = \frac{q_e}{m_e} \vec{S} \ . \tag{3.14}$$

We can then rewrite our previous relations for $\widehat{m}_j$ in terms of S_j (with the same bras and kets as before) as

$$S_x = \frac{\hbar}{2}\big(|+\rangle_x \langle +|_x - |-\rangle_x \langle -|_x \big) \ , \ S_y = \frac{\hbar}{2}\big(|+\rangle_y \langle +|_y - |-\rangle_y \langle -|_y \big) \ , \ S_z = \frac{\hbar}{2}\big(|+\rangle_z \langle +|_z - |-\rangle_z \langle -|_z \big) \ . \tag{3.15}$$

It is instructive to express all three operators entirely with the bras and kets of the z-basis. Let's do this very explicitly for S_y. First, from (3.11) we also get

$$\langle +|_y = \frac{1}{\sqrt{2}}(\langle +|_z - i \langle -|_z) \quad , \quad \langle -|_y = \frac{1}{\sqrt{2}}(\langle +|_z + i \langle -|_z) \ . \tag{3.16}$$

Note that upon changing the kets into bras (taking the adjoint) one needs to complex conjugate the coefficients ($i \to -i$). Then

$$\begin{aligned} |+\rangle_y \langle +|_y - |-\rangle_y \langle -|_y \ &= \ \frac{1}{2}\Big(|+\rangle_z \langle +|_z + i\, |-\rangle_z \langle +|_z - i\, |+\rangle_z \langle -|_z + |-\rangle_z \langle -|_z \Big) \\ &- \ \frac{1}{2}\Big(|+\rangle_z \langle +|_z - i\, |-\rangle_z \langle +|_z + i\, |+\rangle_z \langle -|_z + |-\rangle_z \langle -|_z \Big) \\ &= \ +i\, |-\rangle_z \langle +|_z - i\, |+\rangle_z \langle -|_z \ . \end{aligned} \tag{3.17}$$

Similarly for S_x we get

$$|+\rangle_x \langle +|_x - |-\rangle_x \langle -|_x = |-\rangle_z \langle +|_z + |+\rangle_z \langle -|_z \ . \tag{3.18}$$

We can then rewrite (3.15) entirely in the z-basis as

$$S_x = \frac{\hbar}{2}\big(|-\rangle_z \langle +|_z + |+\rangle_z \langle -|_z \big) \ , \ S_y = \frac{\hbar}{2}\big(i\, |-\rangle_z \langle +|_z - i\, |+\rangle_z \langle -|_z \big) \ , \ S_z = \frac{\hbar}{2}\big(|+\rangle_z \langle +|_z - |-\rangle_z \langle -|_z \big) \ . \tag{3.19}$$

As in (3.9) we can now determine the matrix elements of these 3 spin operators in the z-basis. Arranging the matrix elements in the same order as in (3.9) one finds

$$\begin{pmatrix} \langle +|_z\, S_a\, |+\rangle_z & \langle +|_z\, S_a\, |-\rangle_z \\ \langle -|_z\, S_a\, |+\rangle_z & \langle -|_z\, S_a\, |-\rangle_z \end{pmatrix} = \frac{\hbar}{2}\, \sigma_a \quad , \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \ \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{3.20}$$

The three matrices σ_a are called the Paul matrices.

Exercise 3.1 : *Explicitly check that in* (3.20) *one indeed gets these 3 Pauli matrices. Explicitly verify that all Pauli matrices square to the identity and that the following relations* (3.21) *hold.*

The Pauli matrices satisfy the following relations

$$\sigma_x\sigma_y = i\sigma_z \ , \ \sigma_y\sigma_z = i\sigma_x \ , \ \sigma_z\sigma_x = i\sigma_y \quad , \quad \sigma_y\sigma_x = -i\sigma_z \ , \ \sigma_z\sigma_y = -i\sigma_x \ , \ \sigma_x\sigma_z = -i\sigma_y \ . \tag{3.21}$$

Using the antisymmetric symbol ϵ_{abc} which equals $+1$ if abc is an even permutation of x, y, z (or of $1, 2, 3$) and equals -1 for an odd permutation, and vanishes otherwise, all relations can be summarised as

$$\sigma_a\sigma_b = i\epsilon_{abc}\sigma_c + \delta_{ab}\,\mathbf{1}_{2\times 2} \ , \tag{3.22}$$

where we used the convenient convention that one has to sum over the repeated index c. It then also follows that the commutator of two Pauli matrices gives the third Pauli matrix as

$$[\,\sigma_a, \sigma_b\,] = 2i\epsilon_{abc}\sigma_c \ . \tag{3.23}$$

Also note that two different Pauli matrices anti-commute. One usually denotes the anti-commutator of two matrices or operators by $\{\ ,\ \}$. This, together with the fact that each Pauli matrix squares to the identity, is summarised as[12]

$$\{\sigma_a, \sigma_b\} = 2\delta_{ab}\,\mathbf{1} \ . \tag{3.24}$$

All these relations of the Pauli matrices reflect the corresponding relations of the spin operators, in particular

$$S_aS_b = i\,\frac{\hbar}{2}\,\epsilon_{abc}\,S_c + \frac{\hbar^2}{4}\,\delta_{ab}\,\mathbf{1} \quad \Rightarrow \quad [\,S_a, S_b\,] = i\,\hbar\,\epsilon_{abc}\,S_c \ . \tag{3.25}$$

These relations may be checked directly from the form (3.19) of the spin operators in the z-basis, e.g. $S_xS_z = \frac{\hbar^2}{4}\big(\,|-\rangle_z\langle+|_z + |+\rangle_z\langle-|_z\big)\big(\,|+\rangle_z\langle+|_z - |-\rangle_z\langle-|_z\big) = \frac{\hbar^2}{4}\big(\,|-\rangle_z\langle+|_z - |+\rangle_z\langle-|_z\big) = -i\frac{\hbar}{2}S_y$.

Exercise 3.2 : *To verify that the relations* (3.25) *necessarily are equivalent to the relations* (3.22) *of the corresponding matrices show that taking the matrix elements of* $S_aS_b \equiv S_a\mathbf{1}S_b = S_a\big(\,|+\rangle_z\langle+|_z + |-\rangle_z\langle-|_z\big)S_b$ *directly translates into the matrix product* $\frac{\hbar}{2}\sigma_a\,\frac{\hbar}{2}\sigma_b$.

From our discussion of the Stern-Gerlach experiment and the eigenvectors of the different spin operators it should be clear that there is no way to determine simultaneously two different components of the spin of a given particle. If the measurement with the first Stern-Gerlach of S_z gives $+\frac{\hbar}{2}$ the subsequent measurement of S_x or S_y will give with equal probabilities $+\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$, and this latter measurement destroys the initial information about S_z. This is reminiscent of, but different from the uncertainty relation (2.65) which now reads

$$(\Delta S_z)_\psi(\Delta S_x)_\psi \geq \frac{\hbar}{2}\,|\,\langle\,S_y\,\rangle_\psi\,| \ . \tag{3.26}$$

[12]Viewed as algebraic relations between objects γ_a, the algebra $\{\gamma_a, \gamma_b\} = 2\delta_{ab}\,\mathbf{1}$ is known as defining a Clifford algebra. Clifford algebras play an important role when studying the Dirac equation which occurs in the relativistic description of spin-$\frac{1}{2}$ particles like the electron.

Let us check that this relation is indeed satisfied. If e.g. ψ is an eigenstate of S_y then the right-hand side equals $\frac{\hbar^2}{4}$ and indeed $(\Delta S_x)_\psi = (\Delta S_z)_\psi = \frac{\hbar}{2}$. On the other hand, if ψ is an eigenstate of say S_z then both $\langle\, S_y \,\rangle_\psi$ and $(\Delta S_z)_\psi$ vanish, while $(\Delta S_x)_\psi \neq 0$ which is of course also compatible with (3.26).

Exercise 3.3 : *Check these last statements explicitly.*

The reader may have realised that our discussion of preparing and measuring the spin in different directions with the Stern-Gerlach setups is somewhat similar to preparing and measuring the polarisation of a photon using polarisers with the polarisation say in the x-direction corresponding to $|+\rangle_z$ and the polarisation in the y direction to $|-\rangle_z$. One could also measure the polarisations in an arbitrary direction in the $x-y$ plane., At present one can measure the spin along an arbitrary direction $\vec{u}$ in space. In this case the operator is simply

$$S_{\vec{u}} \equiv \vec{u}\cdot\vec{S} \equiv u_x S_x + u_y S_y + u_z S_z \ . \tag{3.27}$$

This should be clear from thinking about measuring the corresponding magnetic moment. The vector $\vec{u}$ has unit norm and can be written in spherical coordinates as

$$u_z = \cos\theta \quad , \quad u_x = \sin\theta\cos\varphi \quad , \quad u_y = \sin\theta\sin\varphi \ . \tag{3.28}$$

Inserting the expressions for the S_a in the z-basis we get

$$S_{\vec{u}} = \frac{\hbar}{2}\Big[\cos\theta\big(\,|+\rangle_z\langle+|_z - |-\rangle_z\langle-|_z\,\big) + \sin\theta\big(e^{i\varphi}\,|-\rangle_z\langle+|_z + e^{-i\varphi}\,|+\rangle_z\langle-|_z\,\big)\Big] \ . \tag{3.29}$$

Maybe somewhat simpler, the corresponding matrix (without the factor $\frac{\hbar}{2}$) is

$$\sigma_{\vec{u}} \equiv \vec{u}\cdot\vec{\sigma} = \begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\varphi} \\ \sin\theta\, e^{i\varphi} & -\cos\theta \end{pmatrix} \ . \tag{3.30}$$

Recall the fact that the eigenvalues of an operator and of its matrix are the same and that the eigenvectors of the matrix are the components of the eigenvector of the operator in the corresponding orthonormal basis. So let us then determine the eigenvalues of (3.30). Now for any matrix, its trace (sum off the diagonal elements) equals the sum of the eigenvalues and its determinant is the product of all eigenvalues. For the matrix (3.30) the trace is obviously 0 and the determinant is $-\cos^2\theta - \sin^2\theta = -1$. Hence $a_1 + a_2 = 0$ and $a_1 a_2 = -1$ so that the eigenvalues are ± 1, just as for any of the three Pauli matrices corresponding to the three axis' x, y or z. Let us also determine the corresponding eigenvectors $v_\pm = \begin{pmatrix} v_\pm^1 \\ v_\pm^2 \end{pmatrix}$. We write $\sigma_u v_\pm = \pm v_\pm$ as $(\sigma_u \mp \mathbf{1})v_\pm = 0$ which yields two equations. Such a system of homogeneous equations in general has as its only solution $v_\pm = 0$ which is not what we want, except if the matrix $\sigma_u \mp \mathbf{1}$ is singular, i.e. has a vanishing determinant. But this condition is precisely the condition that determines the eigenvalues. The matrix being singular also means that one only needs to solve one of the two linear equations, e.g. the first

one which reads $(\cos\theta \mp 1)v_\pm^1 + \sin\theta e^{-i\varphi}v_\pm^2 = 0$. Using $1+\cos\theta = 2\cos^2\frac{\theta}{2}$, $1-\cos\theta = 2\sin^2\frac{\theta}{2}$ and $\sin\theta = 2\cos\frac{\theta}{2}\sin\frac{\theta}{2}$ this yields $\sin\frac{\theta}{2}\,v_+^1 = \cos\frac{\theta}{2}e^{-i\varphi}\,v_+^2$ and $\cos\frac{\theta}{2}\,v_-^1 = -\sin\frac{\theta}{2}e^{-i\varphi}\,v_+^2$. Imposing normalisation and doing a convenient choice of phase, this gives

$$v_+ \equiv v_+(\vec{u}) = \begin{pmatrix} \cos\frac{\theta}{2}\,e^{-i\varphi/2} \\ \sin\frac{\theta}{2}\,e^{i\varphi/2} \end{pmatrix} \quad , \quad v_- \equiv v_-(\vec{u}) = \begin{pmatrix} -\sin\frac{\theta}{2}\,e^{-i\varphi/2} \\ +\cos\frac{\theta}{2}\,e^{i\varphi/2} \end{pmatrix} \ . \tag{3.31}$$

The $v_\pm(\vec{u})$ are the eigenvectors of the spin operator in the $\vec{u}$ direction, i.e. of $S_{\vec{u}} = \vec{u}\cdot\vec{S}$ with eigenvalues $\pm\frac{\hbar}{2}$. The corresponding kets are written in the z-basis as

$$|+\rangle_{\vec{u}} = \cos\frac{\theta}{2}\,e^{-i\varphi/2}\,|+\rangle_z + \sin\frac{\theta}{2}\,e^{i\varphi/2}\,|-\rangle_z \quad , \quad |-\rangle_{\vec{u}} = -\sin\frac{\theta}{2}\,e^{-i\varphi/2}\,|+\rangle_z + \cos\frac{\theta}{2}\,e^{i\varphi/2}\,|-\rangle_z \ . \tag{3.32}$$

Let us come back to the coupling of the magnetic field and the magnetic moment. Classically there was an energy $E = -\vec{m}\cdot\vec{B}$ which can now be written as $E = -\gamma\vec{S}\cdot\vec{B}$. If we take the magnetic field in the $\vec{u}$-direction, $\vec{B} = B\vec{u}$, the energy becomes ("classically") $E = -\gamma B S_{\vec{u}}$. Since the Hamiltonian operator is associated with the energy, we can suppose that the appropriate Hamiltonian is

$$H = -\gamma B\, S_{\vec{u}} \ . \tag{3.33}$$

The eigenvalues of H are the energies and we find $E_\pm = \mp\gamma B\frac{\hbar}{2}$. For an electron γ is negative and the lower energy is indeed $E_- = -|\gamma|B\frac{\hbar}{2}$. If B is very large we expect that an electron in the eigenstate of energy E_- will remain there even if subject to small perturbations.

There is a very curious fact about this. Suppose that we slowly change the direction $\vec{u}$ of the magnetic field, so that $S_{\vec{u}}$ changes accordingly, and suppose we ensure that the electron stays in this eigenstate corresponding to E_-, i.e. in $v_-(\vec{u})$ which then will also change with $\vec{u}$. To make the argument clearest, suppose $\vec{u}$ lies in the $x-y$ plane, i.e. $\theta = \frac{\pi}{2}$ so that $\cos\frac{\theta}{2} = \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}$ and also $\sin\frac{\theta}{2} = \sin\frac{\pi}{4} = \frac{1}{\sqrt{2}}$. Then $v_+(\vec{u}) = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-i\varphi/2} \\ e^{i\varphi/2}\end{pmatrix}$ and $v_-(\vec{u}) = \frac{1}{\sqrt{2}}\begin{pmatrix} -e^{-i\varphi/2} \\ e^{i\varphi/2}\end{pmatrix}$. If $\vec{u}$ undergoes a full rotation in the $x-y$ plane, then $\varphi \to \varphi + 2\pi$ and the eigenvectors change each by an overall factor $e^{i\pi} = -1$. Of course, this overall phase is not observable and the eigenvectors rotated by 2π are equivalent to (are in the same ray as) the original eigenvectors.

3.4 The group of space rotations

3.4.1 Spin operators and rotation generators

We have seen that the spin operators S_a satisfy the commutation relations $[S_a, S_b] = i\hbar\epsilon_{abc}S_c$ or equivalently $[S_x, S_y] = i\hbar S_z$ and cyclical permutations of (xyz). Correspondingly, the Pauli matrices satisfy

$$[\frac{\sigma_x}{2}, \frac{\sigma_y}{2}] = i\frac{\sigma_z}{2} \quad \text{and cyclic permutations} \quad \Leftrightarrow \quad [\frac{\sigma_a}{2}, \frac{\sigma_b}{2}] = i\,\epsilon_{abc}\,\frac{\sigma_c}{2} \ . \tag{3.34}$$

Actually this commutator algebra is the algebra of infinitesimal rotations around the different axis'. To see this, consider how a rotation by an angle α around an arbitrary axis $\vec{u}$ acts. One has

to distinguish active and passive transformations. Active rotations act on some physical system, while passive transformations act "on the space", i.e. on the orthonormal frame $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$ used to define the coordinates of a given point in space. To be very precise, consider first the example of a passive rotation by an angle α around the z-axis acting as $\vec{e}_x \to \vec{e}_x{}' = \cos\alpha\, \vec{e}_x + \sin\alpha\, \vec{e}_y$ and $\vec{e}_y \to \vec{e}_y{}' = \cos\alpha\, \vec{e}_y - \sin\alpha\, \vec{e}_x$, leaving $\vec{e}_z$ unchanged. A given point in space can be expressed in the old and new orthonormal frame with coordinates (x, y, z) and (x', y', z') as $\vec{OP} = x\vec{e}_x + y\vec{e}_y + z\vec{e}_z = x'\vec{e}_x{}' + y'\vec{e}_y{}' + z'\vec{e}_z{}'$. One easily sees that the primed coordinates are related to the unprimed by the *same* relations[13] $x' = \cos\alpha\, x + \sin\alpha\, y$ and $y' = \cos\alpha\, y - \sin\alpha\, x$. In general, we then write for a passive rotation by an angle α around an axis of direction $\vec{u}$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathcal{R}(\vec{u}, \alpha) \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad , \quad \begin{pmatrix} \vec{e}_x \\ \vec{e}_y \\ \vec{e}_z \end{pmatrix} \to \begin{pmatrix} \vec{e}_x{}' \\ \vec{e}_y{}' \\ \vec{e}_z{}' \end{pmatrix} = \mathcal{R}(\vec{u}, \alpha) \begin{pmatrix} \vec{e}_x \\ \vec{e}_y \\ \vec{e}_z \end{pmatrix} \ . \tag{3.35}$$

Specifying to the three axis' and writing $\mathcal{R}_x(\alpha) \equiv \mathcal{R}(\vec{e}_x, \alpha)$ etc., we have

$$\mathcal{R}_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{pmatrix} \ , \ \mathcal{R}_y(\alpha) = \begin{pmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{pmatrix} \ , \ \mathcal{R}_z(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \ . \tag{3.36}$$

If we let α become infinitesimal, $\alpha \to \epsilon$ the rotation matrices can all be written as

$$\mathcal{R}_j(\epsilon) = \mathbf{1}_{3\times 3} + i\epsilon J_j \ , \tag{3.37}$$

with the 3 matrices J_j, called the (infinitesimal) rotation generators, given by

$$J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \ , \ J_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \ , \ J_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ . \tag{3.38}$$

They satisfy the commutator algebra

$$[J_x, J_y] = iJ_z \quad \text{and cyclic permutations} \quad \Leftrightarrow \quad [J_a, J_b] = i\,\epsilon_{abc}\, J_c \tag{3.39}$$

This is indeed exactly the same commutator algebra as for the $\sigma_a/2$ or the $\frac{1}{\hbar}S_a$.

Mathematically, the space rotations form a group, given by the group of orthogonal 3×3 matrices of unit determinant, called SO(3) :

$$\mathcal{R} \in \mathrm{SO}(3) \quad \Leftrightarrow \quad \mathcal{R} \text{ is a } 3 \times 3 \text{ matrix} \ , \quad \mathcal{R}\mathcal{R}^T = \mathcal{R}^T\mathcal{R} = \mathbf{1}_{3\times 3} \ , \quad \det\mathcal{R} = 1 \ . \tag{3.40}$$

Of course, one immediately verifies that all 3 matrices (3.36) are orthogonal of unit determinant.

What about active rotations ? Active rotations rotate a physical system, while keeping the orthonormal frame fixed. It should be intuitively clear that rotating the system by an angle α

[13] In general, if the basis vectors are transformed with a matrix M, then the components are transformed with $(M^T)^{-1}$. For rotations, the matrices are orthogonal and $(M^T)^{-1} = M$.

while keeping the frame fixed is equivalent to keeping the system in place and rotating the frame by the opposite angle $-\alpha$. One may look at the example of a particle with coordinates $x = r\cos\varphi$, $y = r\sin\varphi$. If we actively rotate this system by an angle β around the z-axis, the new position of the particle will have coordinates $x' = r\cos(\varphi+\alpha),\ y' = r\sin(\varphi+\alpha)$. Expanding the trigonometric functions one gets $x' = \cos\alpha\ x - \sin\alpha\ y$ and $y' = \sin\alpha\ x + \cos\alpha\ y$, which are the same as for the passive transformations but with $\alpha \to -\alpha$.

3.4.2 Some remarks about Lie groups and Lie algebras

At this point it is probably useful to recall - or introduce - some basic facts about groups and in particular Lie groups and their Lie algebras. This is by no means meant to be rigorous or general, but just to provide some basic notions and vocabulary. If the reader feels that most of the material of this subsection is too advanced, he or she should not worry, as it is not strictly necessary to understand the rest of these lectures. In any case, we will say more about this subject in sect. 15, in the second part of these lectures.

A group G is a set for which an associative binary composition law is specified, called the group multiplication, $G\times G \to G,\ (g_1, g_2) \mapsto g_1 \circ g_2 \equiv g_1 g_2$, such that there is an identity element $e \in G,\ g\circ e = e\circ g = g$, and every group element has an inverse g^{-1} such that $g^{-1}\circ g = g\circ g^{-1} = e$. An example is the set of integers $\mathbf{Z}$ with the composition being the addition, the identity being 0 and the inverse of k being $-k$. Another example is the group S_n of permutations of n elements. While S_n is a finite group (it contains finitely many elements), $\mathbf{Z}$ is not, but it contains countably many elements. An example of a group containing uncountably many elements is U(1), the group of phases of the form $g(\alpha) = e^{i\alpha},\ \alpha \in \mathbf{R}$, with the group multiplication being simply the multiplication of these phases. In this last example, the group elements depend in a continuous (actually smooth = C^∞) way on the real parameter α. This is an example of a Lie group.

A Lie group is a group where the elements are parametrised by a certain number n of real parameters such that the dependence is smooth.[14] More precisely, this means that the set of group elements must be endowed with a topology so that one can define neighbourhoods for any group element and continuous, bijective maps (charts) between these neighbourhoods and open sets in $\mathbf{R}^n$. These mappings provide the real parameters for the group elements. Of course one needs to make sure that the parameters for different overlapping neighbourhoods are "correctly" related by some transition functions. We require that these transition functions are C^∞. If this is done, this provides an atlas of charts on the group and turns it into a (smooth) differentiable manifold. In the example of U(1) one needs (at least) two charts, one for $g = e^{i\alpha}$ with $\alpha \in (-\frac{3}{4}\pi, \frac{3}{4}\pi)$ and one for $g = e^{i\beta}$ with $\beta \in (\frac{1}{4}\pi, \frac{7}{4}\pi)$ with transition functions $\alpha = \beta$ for the first overlap $(\frac{1}{4}\pi, \frac{3}{4}\pi)$ and $\alpha = \beta - 2\pi$ for the second overlap. In general one can define a Lie group by requiring it to be

[14] It has been conjectured by Hilbert and proven much later that it is actually enough to require a continuous dependence on the parameters and that the group structure then automatically implies a smooth dependence. Here we will suppose from the outset that the dependence is smooth.

at the same time a group and a differentiable manifold. The dimension n of this manifold then is also referred to as the dimension of the Lie group. In the example of U(1) the manifold is simply the circle and the dimension is one.

For the rotation group SO(3) there are 3 parameters, the rotation angle and two other angles specifying the direction of the rotation axis. The overall manifold structure is somewhat more complicated, but locally, the neighbourhood of any fixed rotation can be mapped to an open set of $\mathbf{R}^3$ with coordinates $\theta^1, \theta^2, \theta^3$, (which parametrise the rotation axis and rotation angle) which we conveniently package into $\vec{\theta}$. We then write $\mathcal{R}(\vec{\theta})$. In the chart containing the identity $\mathbf{1}$, we can always choose this parametrisation such that $\mathcal{R}(0) = \mathbf{1}$. On any manifold of dimension n there are n linearly independent tangent vectors at each point. As the point changes, these tangent vectors change, so they actually constitute vector fields. One can go on in this geometric setting to define the Lie algebra in terms of these vector fields, but let us proceed in a somewhat less general but simpler way : For all groups that are realised as a group of matrices - like the rotation group SO(3) - there is a natural notion of adding matrices and multiplying them with real (or complex) numbers. In particular, we can differentiate the matrices $\mathcal{R}(\vec{\theta})$ with respect to the different parameters θ^j. If we do this in the vicinity of the identity we get the different "tangent vectors" which we call J_j, namely[15]

$$\frac{\partial}{\partial\theta^j}\mathcal{R}(\vec{\theta})\Big|_{\vec{\theta}=0} = i\,J_j \ , \tag{3.41}$$

where the additional factor i is customary in the physics conventions and such that this definition agrees with (3.37). We will mostly refer to the J_j as the infinitesimal generators rather than as basis vectors of the tangent space. Obviously here, there are three different J_j and they form indeed a basis for the tangent space. The important thing is that this tangent space can be given the structure of a Lie algebra as follows.

Lie algebra: In general one considers some one-parameter families of group elements $g(t)$ and $h(t)$ with $\frac{\partial}{\partial t}g(t)|_{t=0} = iA$ and $\frac{\partial}{\partial t}h(t)|_{t=0} = iB$. We let $k(t) = g(t)h(t)g^{-1}(t)h^{-1}(t)$ and then $\frac{\partial}{\partial t}k(\sqrt{t})|_{t=0} = iC$. One can then see that

$$C = i[A,B] \equiv i(AB - BA) \ , \tag{3.42}$$

is the commutator of A and B. In general, $[.,.]$ is called the Lie bracket and corresponds in the more geometrical formulation to the Lie bracket of the tangent vector fields at the identity. It is clear from (3.42) that this commutator is again a tangent vector, i.e. an element of the Lie algebra. It is important to realise that if g_* is any fixed group element, then $\frac{\partial}{\partial t}\big(g_* g(t)\big)|_{t=0}$ gives a tangent vector at g_* in terms of the tangent vector at the identity, and in this way the tangent spaces at all points are isomorphic. For this reason it is enough to consider the tangent space at the identity, and this is the Lie algebra we just described. The structure of the Lie algebra is

[15] Obviously, the J_j look more like matrices than vectors. But one should remember that every point of our manifold is a group element, i.e. a rotation matrix, and then the tangent vectors are also matrices.

completely captured by the commutators of its basis elements. If $\{T_j,\ j = 1, \dots n\}$ is a basis of the Lie algebra, then

$$[T_j, T_k] = iC^l{}_{jk}\, T_l \ , \tag{3.43}$$

where it is again implicit that the repeated index l is summed from 1 to n. The n^3 numbers $C^l{}_{jk}$ are called the structure constants of the Lie algebra. They are not all independent. First of all, they are obviously antisymmetric in k and l, and, since $[A,[B,C]] + \text{cyclic perm} = 0$, they must also obey the so-called Jacobi identity

$$3\, C^l{}_{j[k} C^j{}_{pq]} = C^l{}_{jk} C^j{}_{pq} + (\text{cyclic permutations of } k,p,q) = 0 \ . \tag{3.44}$$

Any set of structure constants satisfying antisymmetry and the Jacobi identity actually defines a Lie algebra, and for every Lie algebra there is a (generally non unique) corresponding Lie group.

Exercise 3.4 : *By writing out the commutators, check that* $[A,[B,C]]+\text{cyclic perm} = 0$ *is trivially true, and show that it implies indeed* (3.44).

Representations: Finally note that one defines an r-dimensional representation of a Lie group and correspondingly of its Lie algebra as any "realisation" in terms of $r \times r$-matrices that satisfy the same algebraic relations. One says one has a homomorphism from the group into $\mathrm{GL}(r, \mathbf{C})$, the group of invertible $r \times r$-matrices. This means that if $M(g)$ is the matrix that represents g, we must have $M(\mathbf{1}) = \mathbf{1}_{r\times r}$, $M(g^{-1}) = (M(g))^{-1}$, $M(gh) = M(g)M(h)$. In particular, for any group that is already defined from matrices, the definition itself provides a particular, "defining" representation. For the rotation group this defining representation is 3-dimensional, but we will see later-on that there are representations of any dimension, and so-called irreducible representations of all odd dimensions.

3.4.3 Back to the rotations

As we have just discussed, the J_a of (3.38) are a basis of the Lie algebra of the rotation group, and the ϵ_{abc} are the corresponding structure constants. Any other realisation of this same commutator algebra by linear operators (matrices) on an r-dimensional vector space then is an r-dimensional representation of this Lie algebra. Hence the $\sigma_a/2$ or the $\frac{1}{\hbar}S_a$ constitute a 2-dimensional representation of this Lie algebra.

An important question is to which extent this Lie algebra determines the full group structure. We leave this question to the appropriate mathematics course and only make a few remarks. First of all, it is easy to see that the group SU(2) of unitary 2×2 matrices also has the $\sigma_a/2$ as generators. Indeed, the Pauli matrices are hermitian and hence $e^{i\vec{\alpha}\cdot\vec{\sigma}/2}$ is a unitary matrix and one can show that any unitary 2×2 matrix is of this form. Taking derivatives with respect to the α_j at $\vec{\alpha} = 0$ shows that the $\sigma_j/2$ are indeed the infinitesimal generators, i.e. form a basis of the corresponding Lie algebra of SU(2). We see that the Lie algebra of SU(2) is identical to the Lie

algebra of SO(3) although these two groups are different. (Of course, they could be isomorphic, but as we will see, they are not.)

It is not a coincidence that the exponential $e^{i\vec{\alpha}\cdot\vec{\sigma}/2}$ of the generators gives the finite group elements. This is actually a general feature. In particular, a rotation by a finite angle α around a given axis can be obtained by composing N rotations around this same axis by angles $\frac{\alpha}{N}$ and then from (3.37) it follows that

$$\mathcal{R}_j(\alpha) = \lim_{N\to\infty}\Big(\mathbf{1}_{3\times 3} + i\frac{\alpha}{N}J_j\Big)^N = e^{i\alpha J_j} \ . \tag{3.45}$$

Here, since we are dealing with the spin $\frac{1}{2}$, we are mostly interested in the 2-dimensional representation in terms of the Pauli matrices. Let us then explicitly compute $e^{i\vec{\alpha}\cdot\vec{\sigma}/2}$. The exponential of a matrix (or of any linear operator) is defined by the corresponding power series. First note that

$$\begin{aligned}\big(\vec{\alpha}\cdot\vec{\sigma}\big)^2 &= \sum_{a,b}\alpha_a\sigma_a\,\alpha_b\sigma_b = \frac{1}{2}\sum_{a,b}\big(\alpha_a\sigma_a\,\alpha_b\sigma_b + \alpha_b\sigma_b\,\alpha_a\sigma_a\big) = \frac{1}{2}\sum_{a,b}\alpha_a\alpha_b\big(\sigma_a\sigma_b + \sigma_b\sigma_a\big)\\ &= \frac{1}{2}\sum_{a,b}\alpha_a\alpha_b\,2\delta_{ab}\mathbf{1}_{2\times 2} = \vec{\alpha}\cdot\vec{\alpha}\ \mathbf{1}_{2\times 2} \equiv \alpha^2\ \mathbf{1}_{2\times 2}\ ,\end{aligned} \tag{3.46}$$

where α is the length of the vector $\vec{\alpha}$, i.e. $\vec{\alpha} = \alpha\,\widehat{\vec{\alpha}}$, and we will interpret α as the rotation angle and the unit vector $\widehat{\vec{\alpha}}$ as the direction of the rotation axis. In particular, we see that $\big(\vec{\alpha}\cdot\vec{\sigma}\big)^{2k} = \alpha^{2k}\ \mathbf{1}_{2\times 2}$ and $\big(\vec{\alpha}\cdot\vec{\sigma}\big)^{2k+1} = \alpha^{2k+1}\,\widehat{\vec{\alpha}}\cdot\vec{\sigma}$. Then, separating the even and odd powers,

$$\begin{aligned}e^{i\vec{\alpha}\cdot\vec{\sigma}/2} &= \sum_{n=0}^{\infty}\frac{1}{n!}\Big(\frac{i}{2}\Big)^n\big(\vec{\alpha}\cdot\vec{\sigma}\big)^n = \sum_{k=0}^{\infty}\frac{1}{(2k)!}\Big(\frac{i}{2}\Big)^{2k}\big(\vec{\alpha}\cdot\vec{\sigma}\big)^{2k} + \sum_{k=0}^{\infty}\frac{1}{(2k+1)!}\Big(\frac{i}{2}\Big)^{2k+1}\big(\vec{\alpha}\cdot\vec{\sigma}\big)^{2k+1}\\ &= \sum_{k=0}^{\infty}\frac{1}{(2k)!}\Big(\frac{i}{2}\Big)^{2k}\alpha^{2k}\,\mathbf{1}_{2\times 2} + \sum_{k=0}^{\infty}\frac{1}{(2k+1)!}\Big(\frac{i}{2}\Big)^{2k+1}\alpha^{2k+1}\,\widehat{\vec{\alpha}}\cdot\vec{\sigma}\ .\end{aligned} \tag{3.47}$$

We recognise the series expansions of $\cos\frac{\alpha}{2}$ and $\sin\frac{\alpha}{2}$, so that

$$e^{i\vec{\alpha}\cdot\vec{\sigma}/2} = \cos\frac{\alpha}{2}\ \mathbf{1}_{2\times 2} + i\sin\frac{\alpha}{2}\ \widehat{\vec{\alpha}}\cdot\vec{\sigma}\ . \tag{3.48}$$

From our preceding discussion that the $\frac{\sigma_j}{2}$ are a representation of the Lie algebra of the rotation group, and in view of (3.45), we expect that these $e^{i\vec{\alpha}\cdot\vec{\sigma}/2}$ correspond to a two-dimensional representation of the finite elements of the rotation group. As such they should correspond to the way a rotation acts on a spin-$\frac{1}{2}$ state. If one looks more precisely how we have defined the signs of the generators and rotation parameters, the $e^{i\vec{\alpha}\cdot\vec{\sigma}/2}$ correspond to passive rotations by an angle α. However, the action on a spin-$\frac{1}{2}$ is an active transformation, and from our previous discussion about active and passive transformations we know that we should flip the sign of the angle α. Thus, an active rotation of some physical spin-$\frac{1}{2}$ should be implemented by the operator $e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}$. In particular, for any of the 3 Pauli matrices we define

$$U(\alpha, j) = e^{-i\alpha\sigma_j/2} = \cos\frac{\alpha}{2}\ \mathbf{1}_{2\times 2} - i\sin\frac{\alpha}{2}\ \sigma_j\ , \tag{3.49}$$

which should implement rotations by an angle α around the j-axis on a spin-$\frac{1}{2}$ state

Exercise 3.5 : *Show that*

$$U(\frac{\pi}{2},x)\,U(\alpha,y)\,U(-\frac{\pi}{2},x)=U(\alpha,z)\quad,\quad U(\frac{\pi}{2},z)\,U(\alpha,x)\,U(-\frac{\pi}{2},z)=U(\alpha,y)\ . \tag{3.50}$$

Let us check that our interpretation is correct. Before, in (3.31) we had determined the eigenstates of the spin operator $\vec{u}\cdot\vec{S}$. We can now act with our $e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}$ on these states and check whether the resulting state is indeed the correctly rotated spin, i.e. the eigenstate of $\vec{v}\cdot\vec{S}$, where the vector $\vec{v}$ should be the vector that results from rotating $\vec{u}$ by an angle α around the axis $\widehat{\vec{\alpha}}$. In order not to deal with complicated trigonometric relations let us just check this in some simple case. Take (3.31) with $\theta=\pi/2$ so that this is a spin in the $x-y$-plane. Acting with $e^{-i\alpha\sigma_z/2}$ on it gives

$$\begin{aligned} e^{-i\alpha\sigma_z/2}\ \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\varphi/2}\\ e^{i\varphi/2}\end{pmatrix} &= \big(\cos\frac{\alpha}{2}\,\mathbf{1}_{2\times2}-i\sin\frac{\alpha}{2}\,\sigma_z\big)\ \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\varphi/2}\\ e^{i\varphi/2}\end{pmatrix}\\ &= \begin{pmatrix}e^{-i\alpha/2}&0\\0&e^{i\alpha/2}\end{pmatrix}\ \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\varphi/2}\\ e^{i\varphi/2}\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i(\varphi+\alpha)/2}\\ e^{i(\varphi+\alpha)/2}\end{pmatrix}\ , \end{aligned} \tag{3.51}$$

and we see that this is indeed the spin state rotated by an angle α around the z-axis. One can similarly work out other examples, and with more energy also the general case, confirming that the $e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}$ indeed are the operators that rotate the spin-$\frac{1}{2}$.

One can also show that when two operators (matrices) $e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}$ and $e^{-i\vec{\beta}\cdot\vec{\sigma}/2}$ are multiplied, this gives the matrix that corresponds to the resulting composed rotation, although with one subtlety to be discussed shortly. Note that the relation (3.50) translates the fact that the composition of a rotation by $-\frac{\pi}{2}$ around the x-axis followed by a rotation by α around the y axis and another rotation by $\frac{\pi}{2}$ around the x axis is equivalent to a rotation by α around the z-axis ! Now about the subtlety : We had already observed that a spin-$\frac{1}{2}$ changes sign after a full 2π rotation, and this is now also obvious for the rotation operators. Taking $\alpha=2\pi$ in (3.48) we see that

$$e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}\Big|_{\alpha=2\pi}=-\mathbf{1}_{2\times2}\ . \tag{3.52}$$

This means that on states of spin $\frac{1}{2}$ a rotation by 2π - which should be the identity - is represented by -1. We could also similarly write this as

$$e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}\Big|_{\alpha=\pi}\ e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}\Big|_{\alpha=\pi}=-\mathbf{1}_{2\times2}=-e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}\Big|_{\alpha=0}\ , \tag{3.53}$$

and read this as "while the product of two rotations by π around the same axis should be the identity, for the corresponding product of the $e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}\big|_{\alpha=\pi}$ there is an extra sign." This means that the present matrices reproduce the group multiplication law of the rotation group, except possibly up to a sign. But in quantum mechanics the overall sign (or phase) of a state is irrelevant, and this is why the $e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}$ still are "good quantum mechanical representations of the rotation group". Mathematically they are called projective representations of the group of rotations of space. These

projective representations of SO(3) actually are "true" representations of the group SU(2) which is known to be the so-called covering group [16] of SO(3).

3.5 Time evolution of a spin-$\frac{1}{2}$

Let us now return to a much simpler, but equally important question, which is the time evolution of a spin-$\frac{1}{2}$ state, As always, to obtain the time evolution one has to solve the Schrödinger equation with the appropriate Hamiltonian and for a given initial condition. We have already mentioned this Hamiltonian which results from the coupling of the magnetic moment to an external magnetic field. Let us then assume that the magnetic field is constant in time and - without loss of generality - is directed in the z-direction, $\vec{B} = B_0 \vec{u}_z$ at the (supposedly fixed) position of the spin. Then

$$H = -\gamma B_0 \, S_z \equiv \omega_0 \, S_z \quad , \quad \omega_0 = -\gamma B_0 \ . \tag{3.54}$$

Here we have defined ω_0 as $-\gamma B_0$ as appropriate for negatively charged particles like electrons or muons where γ is negative, so that ω_0 is positive. The Hamiltonian is proportional to S_z and then its eigenvectors are the eigenvectors of S_z :

$$H \, |\pm\rangle_z = E_\pm \, |\pm\rangle_z \quad , \quad E_\pm = \pm \frac{\hbar\omega_0}{2} \ . \tag{3.55}$$

Suppose one has prepared some state at the initial time t_0 as

$$|\psi(t_0)\rangle = \alpha_+(t_0) \, |+\rangle_z + \alpha_-(t_0) \, |-\rangle_z \quad , \quad |\alpha_+(t_0)|^2 + |\alpha_-(t_0)|^2 = 1 \ . \tag{3.56}$$

This could e.g. be done by a Stern-Gerlach along a direction $\vec{u}$ singling out the state $|+\rangle_{\vec{u}}$ as given in (3.32). In this case $\alpha_+(t_0) = \cos\frac{\theta}{2} \, e^{-i\varphi/2}$ and $\alpha_-(t_0) = \sin\frac{\theta}{2} \, e^{i\varphi/2}$. But let us continue with general $\alpha_\pm(t_0)$. At any time the state $|\psi(t)\rangle$ can again be decomposed on this same basis as

$$|\psi(t)\rangle = \alpha_+(t) \, |+\rangle_z + \alpha_-(t) \, |-\rangle_z \ . \tag{3.57}$$

In subsection 2.9 we have explained in general how the solution is obtained. But let us do it once more for this present case. We apply the Schrödinger equation,

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \, |\psi(t)\rangle = H \, |\psi(t)\rangle \quad \Rightarrow \quad i\hbar \big(\dot{\alpha}_+(t) \, |+\rangle_z + \dot{\alpha}_-(t) \, |-\rangle_z \big) = E_+ \alpha_+(t) \, |+\rangle_z + E_- \alpha_-(t) \, |-\rangle_z \ , \tag{3.58}$$

where a dot indicates, as usual, a time derivative. The result is simple since we have written $|\psi(t)\rangle$ in the basis of eigenstates of the Hamiltonian. We now identify the coefficients of $|+\rangle_z$, as well as the coefficients of $|-\rangle_z$ to get two equations that are immediately solved :

$$i\hbar \dot{\alpha}_\pm(t) = E_\pm \alpha_\pm(t) \quad \Rightarrow \quad \alpha_\pm(t) = e^{-iE_\pm(t-t_0)/\hbar} \, \alpha_\pm(t_0) \ , \tag{3.59}$$

[16] The group SO(3) is not simply connected but its first homotopy group is $\mathbf{Z}_2 \simeq \{1, -1\}$. Its covering group is SU(2) which is simply connected and such that $\mathrm{SU}(2)/\mathbf{Z}_2 \simeq \mathrm{SO}(3)$. There is a 2 to 1 projection from SU(2) to SO(3). Locally the structure of SU(2) and SO(3) is identical, but globally they are not the same. A much simpler example is the group SO(2) which are just rotations by an angle α around a single axis. The group multiplication law is just addition of the rotation angles, and locally this is the same as addition on the real numbers. But for the rotation angles α and $\alpha + 2\pi n$ must be identified, corresponding to a projection from $\mathbf{R}$ to the circle S^1.

where we have also taken into account the initial condition (3.56). Note that we automatically have $|\alpha_+(t)|^2 + |\alpha_-(t)|^2 = 1$. This follows from the fact that the time evolution with the Schrödinger equation preserves the norm, as shown in the previous section. We can then write

$$\begin{aligned} |\psi(t)\rangle &= e^{-iE_+(t-t_0)/\hbar}\,\alpha_+(t_0)\,|+\rangle_z + e^{-iE_-(t-t_0)/\hbar}\,\alpha_-(t_0)\,|-\rangle_z \\ &= e^{-i\omega_0(t-t_0)/2}\,\alpha_+(t_0)\,|+\rangle_z + e^{i\omega_0(t-t_0)/2}\,\alpha_-(t_0)\,|-\rangle_z \ , \end{aligned} \tag{3.60}$$

where the first line is the generic answer, while the second line takes into account the values of $E_\pm$ for our spin-$\frac{1}{2}$ in the magnetic field.

Let us now assume that indeed at $t = t_0$ the initial state is $|+\rangle_{\vec{u}}$ with the $\alpha_\pm(t_0)$ as given above. Then

$$\begin{aligned} |\psi(t)\rangle &= e^{-i\omega_0(t-t_0)/2}\,\cos\frac{\theta}{2}\,e^{-i\varphi/2}\,|+\rangle_z + e^{i\omega_0(t-t_0)/2}\,\sin\frac{\theta}{2}\,e^{i\varphi/2}\,|-\rangle_z \\ &= \cos\frac{\theta}{2}\,e^{-\frac{i}{2}(\varphi+\omega_0(t-t_0))}\,|+\rangle_z + \sin\frac{\theta}{2}\,e^{\frac{i}{2}(\varphi+\omega_0(t-t_0))}\,|-\rangle_z \ . \end{aligned} \tag{3.61}$$

We see that at t this is the eigenstate $|+\rangle_{\vec{u}(t)}$ of $\vec{u}(t)\cdot\vec{S}$ where $\vec{u}(t)$ is the vector that points in the direction with spherical coordinates θ and $\varphi + \omega_0(t-t_0)$. This means that the direction of $\vec{u}(t)$ precesses around the z-axis with an angular frequency ω_0. We have obtained the quantum mechanical analogue of the classical Larmor precession !

We may ask what is the probability that a measurement of $\vec{u}(t_0)\cdot\vec{S}$ gives $-\frac{\hbar}{2}$. This is usually phrased as "what is the probability that the spin has flipped from the initial direction $\vec{u}(t_0) \equiv \vec{u}$ to $-\vec{u}$?". We know from the postulates that this is[17]

$$\begin{aligned} \mathcal{P}_{\vec{u}\to-\vec{u}} &= \big|\,{}_{\vec{u}}\langle -|\psi(t)\rangle\,\big|^2 \\ &= \Big|\,\Big(-\sin\frac{\theta}{2}\,e^{i\varphi/2}\,\langle+|_z + \cos\frac{\theta}{2}\,e^{-i\varphi/2}\,\langle-|_z\Big)\Big(\cos\frac{\theta}{2}\,e^{-\frac{i}{2}(\varphi+\omega_0(t-t_0))}\,|+\rangle_z + \sin\frac{\theta}{2}\,e^{\frac{i}{2}(\varphi+\omega_0(t-t_0))}\,|-\rangle_z\Big)\,\Big|^2 \\ &= \Big|\,-\sin\frac{\theta}{2}\cos\frac{\theta}{2}\,e^{-\frac{i}{2}\omega_0(t-t_0)} + \cos\frac{\theta}{2}\sin\frac{\theta}{2}\,e^{\frac{i}{2}\omega_0(t-t_0)}\,\Big|^2 = \Big|\,i\,\sin\theta\,\sin\frac{\omega_0(t-t_0)}{2}\,\Big|^2 \\ &= \sin^2\theta\,\sin^2\frac{\omega_0(t-t_0)}{2} = \sin^2\theta\,\frac{1}{2}\big(1-\cos\omega_0(t-t_0)\big) \ . \end{aligned} \tag{3.62}$$

We see that this probability (that the initial spin has flipped) starts at 0 at the initial time and then oscillates with period ω_0 around the mean value $\frac{1}{2}\sin^2\theta$. If we want to maximise this amplitude we must choose $\theta = \frac{\pi}{2}$ which means that the initial $\vec{u}$ and the magnetic field $\vec{B}$ are orthogonal ($\vec{u}$ is in the $x-y$ plane), in agreement with the above discussion of the quantum mechanical Larmor precession. Measuring the period of oscillation of this probability amounts to measuring ω_0 and hence γ. In the next subsection we will discuss a refinement of this idea which allows to determine γ to a high precision. One can also compute the probability $\mathcal{P}_{\vec{u}\to\vec{u}}$ that the spin does not flip. Since flipping or not flipping are the only options, both probabilities should add to 1 and one should find

$$\mathcal{P}_{\vec{u}\to\vec{u}} = 1 - \mathcal{P}_{\vec{u}\to-\vec{u}} = 1 - \sin^2\theta\,\sin^2\frac{\omega_0(t-t_0)}{2} \ . \tag{3.63}$$

[17]Note that $-\vec{u}$ corresponds to $\pi-\theta$ and $\varphi+\pi$. One then easily sees that $|-\rangle_{\vec{u}}$ is not $|+\rangle_{-\vec{u}}$ but $i\,|+\rangle_{-\vec{u}}$.

Exercise 3.6 : *Explicitly compute* $\mathcal{P}_{\vec{u}\to\vec{u}} = \big|\,\langle +|_{\vec{u}}\,|\psi(t)\rangle\,\big|^2$ *and show that one indeed gets* (3.63).

3.6 Magnetic resonance

The technique of magnetic resonance goes back to Rabi in 1939 who devised a method to measure ω_0 very precisely due to a resonance phenomenon. The setup is very simple. We put the spin whose magnetic moment we want to determine in a magnetic field that is the superposition of a fixed (uniform) magnetic field in the z-direction and a small magnetic field that rotates in the $x-y$ plane at an angular frequency ω :

$$\vec{B}(t) = B_0\,\vec{u}_z + B_1\cos\omega t\;\vec{u}_x + B_1\sin\omega t\;\vec{u}_y\ . \tag{3.64}$$

As before, we define the angular frequencies as

$$\omega_0 = -\gamma B_0 \quad , \quad \omega_1 = -\gamma B_1\ , \tag{3.65}$$

so that the Hamiltonian reads

$$H(t) = -\gamma\vec{B}(t)\cdot\vec{S} = \omega_0\,S_z + \omega_1\big(\cos\omega t\,S_x + \sin\omega t\,S_y\big)\ . \tag{3.66}$$

A general state $|\psi\rangle$ can again be written as (3.57), and inserting this ansatz into the Schrödinger equation with the Hamltonian (3.66) yields the two coupled first-order differential equations for the coefficients $\alpha_\pm(t)$ as

$$i\dot\alpha_+ = \frac{\omega_0}{2}\,\alpha_+ + \frac{\omega_1}{2}\,e^{-i\omega t}\,\alpha_- \quad , \quad i\dot\alpha_- = -\frac{\omega_0}{2}\,\alpha_- + \frac{\omega_1}{2}\,e^{i\omega t}\,\alpha_+\ . \tag{3.67}$$

To solve these, one realises that multiplying the first equation with $e^{i\omega t/2}$ and the second with $e^{-i\omega t/2}$, one arrives at a system that looks more "symmetric" :

$$ie^{i\omega t/2}\dot\alpha_+ = \frac{\omega_0}{2}\,e^{i\omega t/2}\alpha_+ + \frac{\omega_1}{2}\,e^{-i\omega t/2}\,\alpha_- \quad , \quad ie^{-i\omega t/2}\dot\alpha_- = -\frac{\omega_0}{2}\,e^{-i\omega t/2}\alpha_- + \frac{\omega_1}{2}\,e^{i\omega t/2}\,\alpha_+\ . \tag{3.68}$$

This suggests to define two new functions as $\beta_\pm(t) = e^{\pm i\omega t/2}\alpha_\pm(t)$, so that $ie^{\pm i\omega t/2}\dot\alpha_\pm = i\dot\beta_\pm \pm \frac{\omega}{2}\beta_\pm$. Then

$$i\dot\beta_+ = \frac{\omega_0-\omega}{2}\beta_+ + \frac{\omega_1}{2}\beta_- \quad , \quad i\dot\beta_- = -\frac{\omega_0-\omega}{2}\beta_- + \frac{\omega_1}{2}\beta_+\ . \tag{3.69}$$

The change from the $\alpha_\pm$ to the $\beta_\pm$ has resulted in a system of two coupled differential equations with *constant* coefficients. We may rewrite it in matrix form as[18]

$$i\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\beta_+\\ \beta_-\end{pmatrix} = \widetilde{H}\begin{pmatrix}\beta_+\\ \beta_-\end{pmatrix} \quad , \quad \widetilde{H} = \begin{pmatrix}\frac{\omega_0-\omega}{2} & \frac{\omega_1}{2}\\ \frac{\omega_1}{2} & -\frac{\omega_0-\omega}{2}\end{pmatrix} = \frac{(\omega_0-\omega)}{2}\,\sigma_z + \frac{\omega_1}{2}\,\sigma_x\ . \tag{3.70}$$

[18]One can interpret $\hbar\widetilde{H}$ as the matrix corresponding to a "new" Hamiltonian. Comparing with the initial Hamiltonian (3.66), the time-dependence has now disappeared. One could say that the change of functions from $\alpha_\pm$ to $\beta_\pm$ is the analogue of going to a rotating reference frame in which the magnetic field no longer turns in the $x-y$ plane but is fixed in the x-direction. But one should always be careful not to take too seriously such analogies with classical reasonings.

To solve (3.70) one uses the following trick. First note that, since $\widetilde{H}$ does not depend on time, one can easily iterate this equation as

$$\Big(i\frac{\mathrm{d}}{\mathrm{d}t}\Big)^2\begin{pmatrix}\beta_+\\ \beta_-\end{pmatrix} = i\frac{\mathrm{d}}{\mathrm{d}t}\Big(\widetilde{H}\begin{pmatrix}\beta_+\\ \beta_-\end{pmatrix}\Big) = \widetilde{H}i\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\beta_+\\ \beta_-\end{pmatrix} = \widetilde{H}\widetilde{H}\begin{pmatrix}\beta_+\\ \beta_-\end{pmatrix}\ , \tag{3.71}$$

with

$$\widetilde{H}\widetilde{H} = \frac{1}{4}\Big((\omega_0-\omega)^2\sigma_z^2+\omega_1^2\sigma_x^2+(\omega_0-\omega)\omega_1(\sigma_z\sigma_x+\sigma_x\sigma_z)\Big)\ . \tag{3.72}$$

Note that when computing the square, the cross term is not $2\sigma_z\sigma_x$ but $\sigma_z\sigma_x+\sigma_x\sigma_z$ which vanishes by virtue of the anti-commutativity of the Pauli matrices, see (3.24). Also $\sigma_z^2=\sigma_x^2=\mathbf{1}$, so that

$$\widetilde{H}^2 \equiv \widetilde{H}\widetilde{H} = \frac{1}{4}\big((\omega_0-\omega)^2+\omega_1^2\big)\,\mathbf{1} \equiv \frac{1}{4}\,\Omega^2\,\mathbf{1}\ . \tag{3.73}$$

Thus both functions satisfy $\ddot{\beta}_\pm+\frac{\Omega^2}{4}\beta_\pm=0$ and are superpositions of $e^{i\Omega t/2}$ and $e^{-i\Omega t/2}$, or equivalently of $\sin\frac{\Omega t}{2}$ and $\cos\frac{\Omega t}{2}$. One could continue and solve (3.70) in general and then impose the initial condition, but let's simplify things a bit by supposing that initially, at $t=t_0\equiv 0$, the state is $|\Psi(0)\rangle = |+\rangle_z$. This means that $\alpha_-(0)=0$ and hence $\beta_-(0)=0$. Hence $\beta_-(t)=\beta_0\sin\frac{\Omega t}{2}$. We can then use the second equation (3.69) to compute β_+ with the result $\beta_+(t)=i\beta_0\big(\frac{\Omega}{\omega_1}\cos\frac{\Omega t}{2}-i\frac{\omega_0-\omega}{\omega_1}\sin\frac{\Omega t}{2}\big)$. Again the initial condition gives $\alpha_+(0)=\beta_+(0)=1$ which implies $\beta_0=-i\frac{\omega_1}{\Omega}$ so that finally

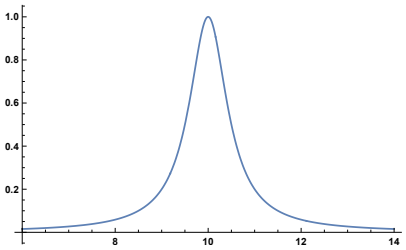
$$\begin{aligned}\beta_+(t) &= \cos\frac{\Omega t}{2}-i\frac{\omega_0-\omega}{\Omega}\sin\frac{\Omega t}{2}\ ,\\ \beta_-(t) &= -i\frac{\omega_1}{\Omega}\sin\frac{\Omega t}{2}\quad , \qquad\qquad \Omega=\sqrt{(\omega_0-\omega)^2+\omega_1^2}\ .\end{aligned} \tag{3.74}$$

The probability to find the spin flipped at time t, i.e. that measuring S_z gives $-\frac{\hbar}{2}$ is

$$\begin{aligned}\mathcal{P}_{+\to-}(t) &= \big|\,\langle-|\psi(t)\rangle\,\big|^2 = \big|\alpha_-(t)\,\big|^2 = \big|\beta_-(t)\,\big|^2 = \frac{\omega_1^2}{\Omega^2}\ \sin^2\frac{\Omega t}{2}\\ &= \frac{\omega_1^2}{(\omega_0-\omega)^2+\omega_1^2}\ \sin^2\frac{\Omega t}{2}\ .\end{aligned} \tag{3.75}$$

Much as in the previous example of the Larmor precession, we find that the flip probability varies periodically in time, but the interesting feature now is the prefactor $\frac{\omega_1^2}{(\omega_0-\omega)^2+\omega_1^2}$ characteristic of a resonance, as shown in the figure below (for $\omega_0=10$ and $\omega_1=0.5$ in some units, say GHz). Since ω_1 is much smaller than ω_0, this factor presents a pronounced and narrow maximum equal to 1 if $\omega=\omega_0$. The width of this curve at half height is $|\omega-\omega_0|=\omega_1$. By changing the frequency ω of the turning magnetic field and tuning it until the flip probability is maximum allows to determine ω_0 to within a precision of the order ω_1. If we know the magnitudes of the magnetic field this allows to determine γ with a relative precision of $\frac{\omega_1}{\omega_0}$ which can be made very small !

Rabi's original experimental setup to measure this probability consists of a source for the beam of the atoms "carrying" the spin-$\frac{1}{2}$, with a first Stern-Gerlach preparing them in the state $|+\rangle_z$.



Then comes the region with the fixed and turning magnetic field. The turning magnetic field can be realised by two electromagnets in the x and y direction and alternating currents with a $\frac{\pi}{2}$ phase difference. The z-component of the spin in the outgoing beam is then measured by a second Stern-Gerlach. Only 5 years later, in 1944, Rabi won the Nobel prize for "having established a radio contact with the ultimate constituents of matter".

A few years later, in 1945, Bloch and Purcell made another breakthrough. The spectacular development of radars during the war had made available more sophisticated techniques to manipulate and measure radio-frequency waves. In particular the turning magnetic field is directly produced by a radio-frequency wave with a circular polarisation.[19] This allowed Bloch and Purcell to measure the spin-flip no longer on individual atoms in a beam but directly in condensed matter samples. These samples are made from a macroscopic number of spins and the spin flips are detected directly by the absorption of the radio-frequency wave. This allows for an even more precise determination of the factor γ. While for isolated electrons we have seen that in a good approximation $\gamma_e \simeq 2\frac{q_e}{2m_e}$, for protons and neutrons, both also spin-$\frac{1}{2}$ particles, one finds $\gamma_p \simeq 5.58\frac{q_p}{2m_p}$ and $\gamma_n \simeq 3.82\frac{q_p}{2m_p}$. These values are much smaller than γ_e since the proton or neutron mass is about 2000 times larger than the electron mass. Similarly, in a more complex atomic nucleus, the effective magnetic moment is of this same order of magnitude. By choosing the radio-frequency appropriately one can be sensitive to the electron or the nuclear magnetic moment. Moreover, in a molecule, the neighbouring nuclei slightly modify the magnetic moment of a given nucleus and, hence, the precise measurement of the resonance frequency gives valuable information on the molecular configuration. This is at the basis of the modern imaging techniques by nuclear resonance.

[19] For circular polarisation of an electromagnetic wave of frequency ω_1 one has in complex notation $\vec{B}_{\text{complex}} \sim (\vec{u}_x + i\vec{u}_y)e^{i\kappa_1 z - i\omega_1 t}$ so that the real part is $\vec{u}_x \cos(k_1 z - \omega_1 t) - \vec{u}_y \sin(k_1 z - \omega_1 t)$, and for $z = 0$ e.g. this is indeed $\vec{B} \sim \vec{u}_x \cos(\omega_1 t) + \vec{u}_y \sin(\omega_1 t)$.