QUANTUM MECHANICS B PHY-413 Note Set No. 9

Note: for examination purposes only sections1,3,4,6,7, and 10 need be read

THE QUANTUM MECHANICS OF SPIN.

(1) INTRODUCTION.

We have already seen in our general discussion of angular momentum that besides the orbital angular momentum operators, $\widehat{\mathbf{L}} = \widehat{\mathbf{L}}$, with integral quantum numbers:

$$\ell = 0, 1, 2, \ldots \quad \text{and, for each } \ell, \quad m_\ell = -\ell, -\ell+1, \ldots 0 \ldots, \ell-1, \ell,$$

there emerged the possibility of half-integral angular momentum,

$$j = \frac{1}{2}, \frac{3}{2}, \ldots \qquad \text{and, for each } j, \quad m_j = -j, -j+1, \ldots, j-1, j.$$

The approach taken was abstract, not relying on the detailed properties of the orbital angular momentum operators (such as their representation by differential operators), but rather on the algebraic properties encapsulated in the commutation relations.[1] Does nature take advantage of the existence of this second possibility? The answer is 'yes'; and it is interesting to note that the discovery, as so often in science, preceded the above theoretical analysis in terms of abstract operators, proceeding by analogy with the case of orbital angular momentum. The first evidence came from the existence of doubled spectral lines - doublets - in the spectra of alkali atoms such as Li, Na, K, Rb and Cs. This suggested that *pairs* of quantum states of similar energy were involved in the transitions responsible for the spectra; and by analogy with *orbital* angular momentum where $2\ell+1$ is the number of $m_\ell-$values allowed for a given ℓ, one supposes that the multiplicity $2 = 2\ell + 1$, giving $\ell = 1/2$. In 1925 two graduate students, Goudsmit and Uhlenbeck suggested that this form of angular momentum - the spin - was carried by the atomic electrons and was an *intrinsic*[2] property of electrons. The Stern-Gerlach experiment in 1922,

[1] Although we shall not deal with this topic, there is a very deep fundamental reason why this approach has a wider applicability: the angular momentum operators are the **generators of infinitesimal rotations**, the operators which carry out infinitesimal rotations of the coordinate axes. The commutation relations are the inevitable consequence of the very simple fact that two successive rotations can be represented by a single rotation - a trivial example is that two successive clockwise rotations by 45° about any axis is equivalent to a single rotation by 90° about that axis. These are the defining properties of a mathematical entity - the rotation **group** in 3-dimensions, $O(3)$. Suppose we rotate the coordinate axes clockwise by an angle $\delta\alpha$ about an axis lying in the direction of the unit vector $\mathbf{n}$. In QM the operator describing the effect of this infinitesimal rotation on a wave function is:

$$\widehat{U}_{\mathbf{n}}(\delta\alpha) = 1 - \frac{i}{\hbar}\delta\alpha\, \mathbf{n}.\widehat{\mathbf{J}},$$

and the wave function becomes $U_{\mathbf{n}}(\delta\alpha)\,\psi(\mathbf{r})$. Note the similarity to the much simpler parity operator. For *finite* rotations by α about the same axis the operator becomes

$$\widehat{U}_{\mathbf{n}}(\alpha) = e^{-i\alpha\, \mathbf{n}.\widehat{\mathbf{J}}/\hbar}.$$

You may like to compare this with the tutorial problem TUT.4 in Problems 2, where we saw that the operator carrying out a finite translation by a along the $x-$axis is

$$\widehat{T}_a = e^{ia\widehat{p}_x/\hbar},$$

showing us that $\widehat{p}_x$ is the generator of infinitesimal translations in the $x-$direction.

[2] Intrinsic properties are those possessed by a system which are identifying attributes of the system and which do not change. Thus, an electron can have any orbital quantum number $\ell = 0, 1, 2 \ldots$ depending on its quantum state at the time, but its spin, electric charge, mass and lepton number have fixed values: they are intrinsic properties.

measuring the magnetic moment of Ag atoms, had also shown the surprising occurrence of two states where only an odd number, $2\ell + 1$ was expected. Later we shall discuss this experiment in detail.

Spin cannot be understood classically. Several aspects of spin angular momentum defy our attempts to understand it as the analogue of a top spinning about its axis - despite the fact that one's mental image of spin is precisely this classical system! The key to the difficulty lies in the experimental fact, gleaned from many high energy scattering experiments on electron and positron beams, that the electron is as close to Newton's idealised point particle as one could imagine:

Electron scattering experiments are consistent with: $m_e = 0.51$ MeV and $r_e < 10^{-20}$ m.

Taking the electron as a sphere of mass m_e and radius r_e spinning about its centre with average speed v, and ignoring factors of order 1, the angular momentum is,

$$J \approx r_e m_e v \approx \frac{\hbar}{2}, \quad \text{implying,} \quad r_e \approx \frac{\hbar}{2m_e v} \geq \frac{\hbar}{2m_e c} \approx 2 \times 10^{-13} \text{ m},$$

which is *seven* orders of magnitude greater than the experimental limit. To put this into the context of 1921, compare this with the sizes of atoms (around 10^{-10} m) and of the proton (10^{-15} m); yet at the time it was already known that the electron was very much smaller than either atom or nucleus. We can put this another way: the speed of rotation needed to explain the magnitude of the spin angular momentum for a particle as small as an electron exceeds the speed of light by seven decades:

$$v \approx \frac{\hbar}{2m_e r_e} > 2 \times 10^7 \, c.$$

Thus, like many quantum phenomena, spin has a classical analogue, but the analogue fails to provide a quantitative explanation; it only gives us a convenient mental picture. Electron spin appears to be an entirely quantum mechanical phenomenon. This statement is not, however, quite correct: in 1926 Dirac presented his famous relativistic wave equation for the electron - the Dirac equation - which not only gives the correct spin angular momentum, but also an excellent approximation to its measured magnetic moment *and* a prediction that the electron has an antiparticle, now called the positron. Thus relativity in concert with QM plays an essential role in the explanation of spin.

Having concluded that we cannot picture spin angular momentum simply as matter rotating in some sort of orbit around the electron's centre whose shape we could hope to determine - perhaps by some sort of generalisation of a spherical harmonic - we now face a dilemma: what are the coordinates that play the role of (x, y, z) or (r, θ, φ) in describing orbital angular momentum? Since we do not know we must resort to an abstract approach via the angular momentum operators for spin. This has already been accomplished by our discovery that the commutation rules for the angular momentum operators $\widehat{\mathbf{J}}$ provide for the possibility of spin angular momentum $\hbar/2$. We now continue along that path by following Heisenberg's matrix approach to QM in which no reference need be made to the coordinates.

(2) INTRODUCING HEISENBERG'S MATRIX MECHANICS.

In his matrix version of QM Heisenberg achieved a form of the theory which made reference only to directly measurable physical observables, the expectation values. The formalism is ideally suited to the quantum mechanical study of angular momentum, especially spin. The most appropriate notation to use is Dirac's bra/ket formulation, but since I wish to avoid

introducing yet more new material I shall justify and explain the steps in the argument in terms of orbital angular momentum and simply carry over the results directly to spin angular momentum. However, all results can be proved in a general way without recourse to the integrations I use.

Given the eigenfunctions ψ_{j,m_j} of the angular momentum operators $\widehat{\mathbf{J}}^2$, and $\widehat{J}_z$, the generalised expansion theorem allows us to write, for any wave function Ψ_j corresponding to a given total angular momentum j: [3]

$$\Psi_j = \sum_{m_j=-j}^{j} c_{m_j}\, \psi_{j,m_j}. \tag{1}$$

Note that (a) we are considering a system with definite total angular momentum, Ψ_j being an eigenstate of $\widehat{\mathbf{J}}^2$, so there is no sum over j here; (b) we do not mention any spatial dependence for the wave function because for spin wave functions we don't know whether there is any, and if there is we don't know what the coordinates are.[4] The coefficients in the expansion are normalised,

$$\sum_{m_j=-j}^{j} |c_{m_j}|^2 = 1 \tag{2}$$

This follows from the normalisation of Ψ_j and the orthonormality of the eigenstates ψ_{j,m_j} and shows that the $|c_{m_j}|^2$ are probabilities. For a system in the state Ψ_j a measurement of the $z-$component of angular momentum will yield one of the eigenvalues $\hbar m_j$ with probability $|c_{m_j}|^2$. After a measurement of the $z-$component of angular momentum yielding the result m_j the wave function becomes

$$\Psi_{j\,\text{after}} = \psi_{j,m_j} \tag{3}$$

and if the system is then left undisturbed, a subsequent measurement will yield the same result m_j with probability 1.

Now we consider the case of orbital angular momentum, where the wave function *does* depend on the coordinates, and the expansion eigenfunctions are just the spherical harmonics Y_{j,m_j}. The expectation value of any operator $\widehat{A}$ when the system is in the state Ψ_j is then[5]

$$\langle\, \widehat{A}\, \rangle = \int \Psi_j^* \widehat{A} \Psi_j \, d^3x \tag{4}$$

$$= \sum_{m_j'} \sum_{m_j} c_{m_j'}^* c_{m_j} \int \psi_{j,m_j'}^* \widehat{A} \psi_{j,m_j} \, d^3x \tag{5}$$

$$= \mathbf{C}^\dagger A\, \mathbf{C} \tag{6}$$

(a) The expansion coefficients have been assembled into a $(2j+1)-$dimensional column vector, $\mathbf{C}$ - you can think of the entries as the components of the original wave function Ψ_j in the basis of eigenstates ψ_{j,m_j}:

$$\mathbf{C} = \begin{pmatrix} c_j \\ c_{j-1} \\ \vdots \\ \vdots \\ c_{-j} \end{pmatrix} \qquad \mathbf{C}^\dagger = (\, c_j^* \; c_{j-1}^* \; \ldots\ldots \; c_{-j}^*) \tag{7}$$

[3] Strictly speaking, we should say the total angular momentum is $\hbar\sqrt{j(j+1)}$, but we shall often lapse into this shorthand. In a similar way we refer to spin 1/2 or $\hbar/2$ instead of $\hbar\sqrt{1/2(1/2+1)} = \hbar\sqrt{3}/2$.

[4] This is where the Dirac formulation comes into its own. The states are represented by the kets $|j\rangle$ which are independent of the coordinates, whatever they may be.

[5] We use the notation $\widehat{A}$ for abstract quantum mechanical operators, A for the *matrices* representing them and $(A)_{a,b} \equiv A_{a,b}$ for the a,b element of the matrix. Bold faced symbols without a hat denote column vectors, $\mathbf{C}$. In Dirac notation the right hand side of eq (4) is $\langle j|\widehat{A}|j\rangle$, while the integral in eq. (5) is $\langle j, m_j'|\widehat{A}|j, m_j\rangle$,

The vector $\mathbf{C}$ tells us everything we can know about the wave function Ψ_j and plays the role of the wave function in matrix mechanics, including the normalisation,

$$\mathbf{C}^\dagger \, \mathbf{C} = \sum_{m_j=-j}^{j} |c_{m_j}|^2 = 1. \tag{8}$$

(b) Similarly, the $(2j+1)\times(2j+1)$ matrix A contains all the information about the expectation value of $\widehat{A}$ for all possible situations, ie. for any possible wave function Ψ_j, now represented by the vector $\mathbf{C}$; the **matrix elements** of A are:

$$A_{ab} \equiv (A)_{m'_j m_j} = \int \psi^*_{j,m'_j} \widehat{A} \psi_{j,m_j} \, d^3x \tag{9}$$

We can also show that the rules of matrix multiplication apply, ie. the matrix representation of the operator $\widehat{A}\widehat{B}$ is indeed the product AB of the matrices representing $\widehat{A}$ and $\widehat{B}$ separately. One consequence of this is that if a set of operators, such as the angular momentum operators, satisfy a set of commutation relations then so do the matrices representing them.

An important special case occurs when the eigenfunctions used in the above are eigenstates of the operator $\widehat{A}$ itself, for example with $\widehat{A} = \widehat{J}_z$,

$$\widehat{J}_z \, \psi_{j,m_j} = \hbar m_j \, \psi_{j,m_j} \tag{10}$$

where the eigenvalue is $\hbar m_j$ In that case the matrix representing $\widehat{J}_z$ is **diagonal**:

$$\begin{aligned} (J_z)_{m'_j m_j} &= \int \psi^*_{j,m'_j} \widehat{J}_z \psi_{j,m_j} \, d^3x \\ &= \hbar m_j \int \psi^*_{j,m'_j} \, \psi_{j,m_j} \, d^3x \quad \text{for eigenstates of } \widehat{J}_z, \\ &= \hbar m_j \, \delta_{m'_j,m_j} \end{aligned} \tag{11}$$

where, in the last step, we used the orthonormality of angular momentum eigenstates. Similarly, for the $\widehat{\mathbf{J}}^2$ operator, its matrix representation is not only diagonal, but simply a *multiple of the unit matrix*:

$$\mathbf{J}^2_{m'_j m_j} = \hbar^2 j(j+1) \, \delta_{m'_j,m_j} \tag{12}$$

In its full glory the J_z matrix looks like this:

$$J_z = \begin{pmatrix} (J_z)_{j,j} & (J_z)_{j,j-1} & \dots & (J_z)_{j,-j} \\ (J_z)_{j-1,j} & (J_z)_{j-1,j-1} & \dots & (J_z)_{j-1,-j} \\ \vdots & \vdots & \vdots & \ddots \\ (J_z)_{-j,j} & (J_z)_{-j,j-1} & \dots & (J_z)_{-j,-j} \end{pmatrix} = \hbar \begin{pmatrix} j & 0 & \dots & 0 \\ 0 & j-1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & \dots & -j \end{pmatrix} \tag{13}$$

and the $\mathbf{J}^2$ matrix like this:

$$\mathbf{J}^2 = \hbar^2 j(j+1) \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & \dots & 1 \end{pmatrix} \tag{14}$$

To summarise: All the consequences of quantum mechanics can be obtained from its matrix formulation: all information about the wave function is contained in the column vector $\mathbf{C}$, which is henceforth known as the wave function and which is normalised,

$$\mathbf{C}^\dagger \, \mathbf{C} = 1. \tag{15}$$

All operators are represented by matrices A; measurements yield only expectation values which are calculated from matrix products such as:

$$\langle \widehat{A} \rangle = \mathbf{C}^\dagger \, A \, \mathbf{C} \tag{16}$$

The matrix A is diagonal in a basis of eigenstates of $\widehat{A}$, the diagonal elements being the eigenvalues of the operator. Finally notice that the *general* wave function vector $\mathbf{C}$ is not itself an eigenvector of the matrix A, reflecting the fact that Ψ_j is not an eigenstate but a linear combination of eigenstates ψ_{j,m_j}; however, if we take Ψ_j to be an eigenstate ψ_{j,m_j}, then all the c's vanish except $c_{m_j} = 1$. The wave function would then be represented by the column vector with zeros in every position except the m_j-th,

$$\mathbf{C} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \tag{17}$$

and it would be an eigenvector of the matrix J_z:

$$J_z \, \mathbf{C} = \hbar m_j \, \mathbf{C} \tag{18}$$

The key point for us is that this formalism is *essential* when one deals with purely quantum phenomena where, as in the case of *spin*, it is not known what the appropriate variables, or coordinates actually are.

(3) $j \equiv s = 1/2$: MATRICES FOR THE SPIN-1/2 OPERATORS.

At this stage we have nearly all the necessary equipment to write down the matrices representing the angular momentum operators for *any* value of j. These matrices will obey the commutation relations of the abstract operators they represent, just as do the differential operators $-i\hbar(y\partial/\partial z - z\partial/\partial y)$, etc. used previously to represent the *orbital* angular momentum operators. When we are talking about the special case of orbital angular momentum we know what degrees of freedom are involved - (θ, φ) - but for spin we have no idea, and so the only possibility open to us is to use the matrix representation. This is why the abstract approach to the angular momentum operators is so useful - indeed indispensable. In this section we will obtain these operators for spin-1/2 in as elementary way as possible; for higher values of j this method would become impossibly cumbersome. In Appendices A & B we give a systematic general method for finding the matrices for any j which is quite elementary but involves a multitude of subscripts due to this generality. I recommend that you look through it after reading this section simply to convince yourself that it can be done directly.

Our strategy here will be to construct the 2×2 Hermitian spin-1/2 matrices by requiring that they satisfy the angular momentum commutation relations. We begin with the two diagonal matrices $J_z = S_z$ and $\mathbf{J}^2 = \mathbf{S}^2$, which we have already found in eqs. (13) and (14):

$$S_z = \tfrac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad\qquad \mathbf{S}^2 = \tfrac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{19}$$

where the factor in front of the $\mathbf{S}^2$ matrix is $\hbar^2 j(j+1)$ with $j \equiv s = 1/2$. We now write the as yet unknown complex matrices $J_x = S_x$ and $J_y = S_y$ as

$$S_x = \tfrac{\hbar}{2} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad\qquad S_y = \tfrac{\hbar}{2} \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \tag{20}$$

Step 1: First we show that the two matrices have no diagonal elements, $a = d = a' = d' = 0$ by demanding that they satisfy the commutation relations $[S_y, S_z] = i\hbar\, S_x$ and $[S_z, S_x] = i\hbar\, S_y$:

$$[S_y, S_z] = i\hbar\, S_x \tag{21}$$

$$\text{ie.} \quad \left(\frac{\hbar}{2}\right)^2 \left\{ \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \right\} = i\frac{\hbar^2}{2} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$\text{hence,} \qquad \frac{\hbar^2}{4} \left\{ \begin{pmatrix} a' & -b' \\ c' & -d' \end{pmatrix} - \begin{pmatrix} a' & b' \\ -c' & -d' \end{pmatrix} \right\} = i\frac{\hbar^2}{2} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$\text{or, cancelling } \hbar^2/2, \qquad \begin{pmatrix} 0 & -b' \\ c' & 0 \end{pmatrix} = i \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{22}$$

Hence, equating individual components of the matrices,

$$a = 0, \quad d = 0 \qquad \text{and} \qquad -b' = ib, \quad c' = ic \tag{23}$$

To see the consequences of the commutator

$$[S_z, S_x] = -[S_x, S_z] = i\hbar\, S_y \tag{24}$$

all we need do is notice that the calculation goes precisely as the one above except for the interchange $x \leftrightarrow y$, ie. of primed and unprimed symbols, and an additional minus sign on the left. The result is therefore:

$$a' = 0, \quad d' = 0 \qquad \text{and} \qquad b = ib', \quad -c = ic' \tag{25}$$

Putting all these results together we now have,

$$a = 0, \quad d = 0; \quad a' = 0, \quad d' = 0 \qquad \text{and} \qquad b' = -ib, \quad c' = ic \tag{26}$$

so that the matrices are off-diagonal and can be expresssed in terms of only two unknown numbers b and c,

$$S_x = \tfrac{\hbar}{2} \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \qquad\qquad S_y = \tfrac{\hbar}{2} \begin{pmatrix} 0 & -ib \\ ic & 0 \end{pmatrix} \tag{27}$$

Step 2: The penultimate step is to impose the remaining commutation relation,

$$[S_x, S_y] = i\hbar\, S_z \tag{28}$$

$$\text{ie.} \quad \left(\frac{\hbar}{2}\right)^2 \left\{ \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \begin{pmatrix} 0 & -ib \\ ic & 0 \end{pmatrix} - \begin{pmatrix} 0 & -ib \\ ic & 0 \end{pmatrix} \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \right\} = i\frac{\hbar^2}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\text{hence,} \qquad \frac{\hbar^2}{4} \left\{ \begin{pmatrix} ibc & 0 \\ 0 & -ibc \end{pmatrix} - \begin{pmatrix} -ibc & 0 \\ 0 & ibc \end{pmatrix} \right\} = i\frac{\hbar^2}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\text{or, after cancelling } i\hbar^2/2, \qquad \begin{pmatrix} bc & 0 \\ 0 & -bc \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{29}$$

Hence, equating individual components of the matrices,

$$bc = 1 \tag{30}$$

Step 3: The final step is the phyical requirement that S_x and S_y correspond to physical observables with real expectation values, ie. that they be Hermitian matrices, [6]

$$S_x^\dagger \equiv (S_x^*)^T = S_x \qquad S_y^\dagger = S_y \tag{31}$$

Thus, for S_x:

$$\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}^\dagger = \begin{pmatrix} 0 & c^* \\ b^* & 0 \end{pmatrix} = \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \tag{32}$$

which immediately tells us that the parameters are complex conjugates of each other,

$$c = b^* \tag{33}$$

This condition also ensures that S_y is Hermitian, so there is no further information to be gleaned. Combining our two relations for b and c then yields

$$|b|^2 = |c|^2 = 1 \tag{34}$$

The *general* solution is

$$b = e^{i\alpha}, \quad c = e^{-i\alpha}, \tag{35}$$

where α is a constant real phase angle which has no physical consequences. We are therefore free to choose the simplest solution, $\alpha = 0$; the physics would be the same for any other choice,

$$b = c = 1 \tag{36}$$

We finally have the matrix representation for the spin-1/2 operators:

$$S_x = \tfrac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad S_y = \tfrac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad S_z = \tfrac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \mathbf{S}^2 = \tfrac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{37}$$

[6] Beware: only if operators are being represented by matrices does the adjoint operation (the dagger) correspond to the complex conjugate transpose operation; for representation by differential operators it corresponds to complex conjugation and turning the action from operating on ψ to the right into operating on ψ^* to the left in expectation values.

(4) PAULI MATRICES & SPIN WAVE FUNCTIONS FOR $j \equiv s = 1/2$.

Since spin does not emerge from the quantum mechanics of the Schrödinger equation, it has to be grafted on as an additional postulate: the wave function of a particle with spin is a product of the spatial wave function, $\Psi(\mathbf{r}, t)$ which satisfies the time dependent Schrödinger equation, and a spin wave function, χ_{spin} which is a $(2j+1)-$component vector obtained from the matrix formalism:[7]

$$\Psi_{total} = \Psi(\mathbf{r}, t)\chi_{spin} \tag{38}$$

This is not to say that QM is unable to incorporate spin into the theory in a unified way; the missing ingredient is relativity: the Schrödinger equation is non-relativistic. Shortly after the Schrödinger equation was first introduced Dirac published his relativistic wave equation for the electron - the Dirac equation - which not only gives the correct spin angular momentum, but also an excellent approximation to its measured magnetic moment *and* a prediction that the electron has an antiparticle. Relativistic wave equations also exist for other values of spin. In the Dirac equation the wave function contains all three ingredients in a unified way: the space-time dependence, the spin dependence and the antiparticle wave function.

For spin-1/2, we have seen that the angular momentum operators $\widehat{\mathbf{J}}$, now called $\widehat{\mathbf{S}}$ for 'spin', are represented by 2×2 Hermitian matrices S_i, given by the Pauli matrices σ_i:

$$S_i = \frac{\hbar}{2}\,\sigma_i, \quad i = 1, 2, 3 \quad \text{(or, in an often used equivalent labelling,} \quad i = x, y, z), \tag{39}$$

$$\text{where:} \qquad \sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{40}$$

We emphasise again that the *matrices* satisfy precisely the same commutation relations as the quantum mechanical angular momentum *operators* themselves:

$$\text{for the QM operators:} \quad [\widehat{S}_x, \widehat{S}_y] = i\hbar\,\widehat{S}_z; \qquad \text{for the matrices:} \quad [S_x, S_y] = i\hbar\,S_z \tag{41}$$

In fact, as we saw in the previous section, these matrices can be discovered by simply seeking the Hermitian 2×2 matrices satisfying the commutation relations - the answer is unique. Another useful way to look at the Pauli matrices is that *any* Hermitian 2×2 matrix can be written as a linear combination of Pauli matrices and the unit matrix,

$$I = \sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{42}$$

The eigenvectors (or eigenstates), χ_+ and χ_-, of the diagonal matrix $S_z = \hbar/2\,\sigma_z$ are interpreted as quantum states with definite values of the $z-$component of spin; these values are the eigenvalues $\hbar m_s = +\hbar/2$ and $\hbar m_s = -\hbar/2$. To find these eigenstates we apply a systematic procedure, although the result is fairly obvious:

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{43}$$

We are using the notation χ in the case of spin-1/2 for the column vectors denoted $\mathbf{C}$ in the general discussion of section 2.

[7] Actually the situation is a little more complicated: when a spinning particle is subjected to a time-dependent magnetic interaction the spin part becomes time-dependent. This is achieved through having time-dependent coefficients $a = a(t)$ and $b = b(t)$ in the expansion theorem for χ_{spin} in eqs. (51) & (52). See Appendix D for a worked example.

Proof: The requirement that χ be an eigenstate of S_z is:

$$S_z\,\chi = \frac{\hbar}{2}\lambda\,\chi, \tag{44}$$

where, although we already know that the eigenvalues of S_z correspond to $\lambda = \pm 1$, we shall find it convenient to actually confirm this by calculation. Now the most general form for *any* normalised spin-1/2 wave function is

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \text{with} \qquad |a|^2 + |b|^2 = 1. \tag{45}$$

Our task is to find a, b and λ. Using the known form for the S_z matrix, the requirement that χ be an eigenstate of S_z becomes:

$$\begin{aligned} S_z\,\chi &= \frac{\hbar}{2}\lambda\chi = \frac{\hbar}{2}\lambda\begin{pmatrix} a \\ b \end{pmatrix} \\ &= \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \\ &= \frac{\hbar}{2}\begin{pmatrix} a \\ -b \end{pmatrix} \end{aligned} \tag{46}$$

Equating the elements of column vectors in the first and last lines and cancelling $\hbar/2$, we find:

$$a\lambda = a \tag{47}$$

$$b\lambda = -b \tag{48}$$

These two equations are clearly inconsistent if both a and b are non-zero; but they can be satified by taking one to be zero. This gives two possibilities,

either	$b = 0$	$a \neq 0$	and therefore	$\lambda = +1$
			and normalising,	$a = 1$
or	$a = 0$	$b \neq 0$	and therefore	$\lambda = -1$
			and normalising,	$b = 1$

Q.E.D.

As befits eigenstates of a Hermitian operator with different eigenvalues, they are orthonormal, ie. both normalised and orthogonal:

$$\chi_\pm^\dagger\chi_\pm = 1 \qquad \text{and} \qquad \chi_\pm^\dagger\chi_\mp = 0 \tag{49}$$

Since $\mathbf{S}^2$ commutes with S_z they are also eigenstates of $\mathbf{S}^2$ with common eigenvalue $3\hbar^2/4$:

$$S_z\,\chi_\pm = \pm\frac{\hbar}{2}\,\chi_\pm \qquad \text{and} \qquad \mathbf{S}^2\,\chi_\pm = \frac{3}{4}\hbar^2\,\chi_\pm \tag{50}$$

All these equations can be checked out explicitly by using the matrices for the spin operators, and you should do this for yourself. Moreover you can see that the **Expansion Theorem** holds because *any* 2-dimensional vector, χ, can be written as a linear combination of the two unit vectors $\chi_\pm$ given in eq. (43): [8]

$$\chi = a\,\chi_+ + b\,\chi_- \tag{51}$$

[8] This is a special case of equation (1) for spin-1/2, with spin eigenvalues $j = s = 1/2$ and $m_j = m_s = \pm 1/2$, and the general wave function $\mathbf{C}$ becomes the general two-component vector, now called χ, with $c_{1/2} = a, c_{-1/2} = b$.

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \chi^\dagger = (\, a^* \;\; b^*), \qquad \text{with normalisation: } \sum_{m_s} |c_{m_s}|^2 = |a|^2 + |b|^2 = 1 \tag{52}$$

In general the state χ is not an eigenstate of any of the spin operators. This is just like the case of states obtained from linear combinations of energy eigenstates: they, too, are not eigenstates of the Hamiltonian operator. The case of spin is so much simpler because the space is finite-dimensional (2-D for spin-1/2) and so there are always only a finite number eigenstates and terms in the expansion theorem. However the measurement postulate is the same:
Measurement Postulate: For a spin-1/2 particle in a general state χ given above, the possible outcome of a measurement of the $z-$componenet of spin is one of the eigenvalues $\pm\hbar/2$ (ie. $m_s = \pm 1/2$), with respective probabilities $|a|^2$ and $|b|^2$. If say, a measurement yields a value $m_s = +1/2$, then the wave function 'collapses' to the eigenstate $\chi_{after} = \chi_+$. Since this is an eigenstate of S_z, a subsequent measurement will yield the same value $m_s = +1/2$ with probability 1. We will shortly explore these concepts using the Stern-Gerlach experiment as the method for preparing spin states and measuring spin components.

(5) ROTATED SPIN OPERATORS & THEIR EIGENFUNCTIONS.

Note: The following derivation is a fairly general one leading to the general result, eq. (70). We will actually only use the special case, eq. (72) in discussing the Stern-Gerlach experiment. This is derived separately in Section 6 by a more elementary method, so you may skip the remainder of this section if you wish.

As a preparation for our discussion of the Stern-Gerlach experiment we now ask the following question: Suppose we prepare a spin-1/2 system in an eigenstate of S_z and then subsequently measure the component of spin in some other direction; what are the possible outcomes? An equivalent way to pose this question is: What are the eigenstates of the spin operator corresponding to a direction specified by the unit vector $\mathbf{n}$? Because the spin operator is a vector its component along $\mathbf{n}$ is simply the projection,

$$S_n = \mathbf{n}.\mathbf{S} \qquad \text{and we seek the } \chi'_\pm \text{ for which } \quad S_n\,\chi'_\pm = \pm\frac{\hbar}{2}\,\chi'_\pm \tag{53}$$

For simplicity we only consider the case where the vector $\mathbf{n}$ lies along the direction obtained by rotating the $z-$axis by an angle θ about the $y-$axis. If we call this the $z'-$axis, and the rotated $x-$axis perpendicular to it the $x'-$axis, then we are just asking for the new spin operators in the new coordinate system (x', y', z'):

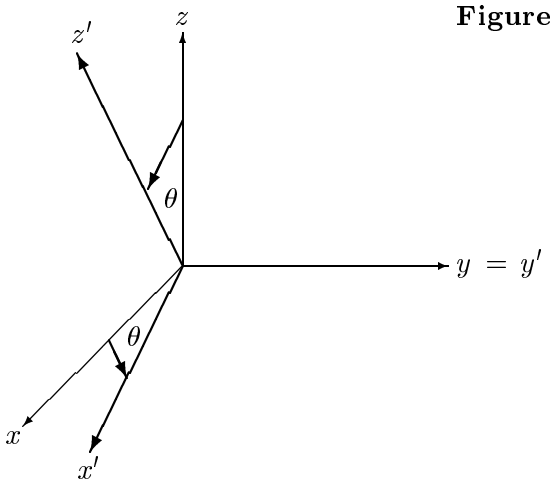


Figure 1: Picturing the rotation by angle θ about the $y-$axis.

Since **S** transforms as a vector the figure shows that the relation between new and old matrices is:

$$S_{x'} = S_x \cos\theta - S_z \sin\theta \tag{54}$$
$$S_{y'} = S_y \tag{55}$$
$$S_{z'} = S_z \cos\theta + S_x \sin\theta \tag{56}$$

so that the new matrices become,

$$S_{x'} = \frac{\hbar}{2}\begin{pmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{pmatrix} \quad S_{y'} = S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad S_{z'} = \frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \tag{57}$$

As a check on our algebra we note that $\theta = 0$ or $\theta = 360^\circ$ gives back our original matrices, while $\theta = 90^\circ$ just exchanges the $x-$ and $z-$matrices, with a sign change on $S_{x'}$, as we would expect since the new $x'-$axis is now playing the role of the original $-z-$axis. This shows rather clearly that it's just a matter of *our* choice onto what axis we decide to measure the projection of the spin.

Now we look for the eigenstates of the new $S_{z'}$ matrix. By the expansion theorem, eq. (51), these eigenstates, χ', must be normalised linear combinations of the original eigenstates of S_z (eq. (43):

$$\chi' = a\chi_+ + b\chi_- = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \text{with} \qquad |a|^2 + |b|^2 = 1. \tag{58}$$

For this to be an eigenstate, with eigenvalue $\lambda\hbar/2$ (where we already know $\lambda = \pm 1$ because the particle has spin-1/2),

$$S_{z'}\,\chi' = \frac{\hbar}{2}\lambda\,\chi'. \tag{59}$$

Using our matrix representation for $S_{z'}$ and the expressions for the original eigenvectors $\chi'_\pm$,

$$S_{z'}\,\chi' = \frac{\hbar}{2}\lambda\,\chi' = \frac{\hbar}{2}\lambda\begin{pmatrix} a \\ b \end{pmatrix} \tag{60}$$
$$= \frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}$$
$$= \frac{\hbar}{2}\begin{pmatrix} a\cos\theta + b\sin\theta \\ a\sin\theta - b\cos\theta \end{pmatrix} \tag{61}$$

Equating the elements of column vectors in the first and last lines, cancelling $\hbar/2$ and collecting coefficients of a and b, we find:

$$a(\lambda - \cos\theta) = b\sin\theta \tag{62}$$
$$a\sin\theta = b(\lambda + \cos\theta) \tag{63}$$

Dividing the equations gives a consistency condition leading to the expected two eigenvalues:

$$(\lambda - \cos\theta)(\lambda + \cos\theta) = \sin^2\theta, \quad \text{ie.} \quad \lambda^2 - \cos^2\theta = \sin^2\theta, \quad \text{or} \quad \lambda^2 = 1, \quad \text{hence} \quad \lambda = \pm 1 \tag{64}$$

Considering each case in turn gives the two eigenstates χ'_+ and χ'_-, where we use the trig. identities $(1 - \cos\theta) = 2\sin^2\theta/2$, $(1 + \cos\theta) = 2\cos^2\theta/2$, $sin\theta = 2\sin\theta/2\cos\theta/2$:

$$\text{For } \lambda = +1: \quad \frac{b}{a} = \frac{(1-\cos\theta)}{\sin\theta} = \frac{2\sin^2\theta/2,}{2\sin\theta/2\cos\theta/2} = \frac{\sin\theta/2,}{\cos\theta/2} \tag{65}$$

$$\text{For } \lambda = -1: \quad \frac{b}{a} = \frac{-(1+\cos\theta)}{\sin\theta} = -\frac{2\cos^2\theta/2,}{2\sin\theta/2\cos\theta/2} = -\frac{\cos\theta/2,}{\sin\theta/2} \tag{66}$$

The magnitude of a in each case is obtained from the condition that $\chi'_\pm$ be normalised:

$$|a|^2 + |b|^2 = |a|^2(1 + |\frac{b}{a}|^2) = 1 \tag{67}$$

$$\text{For } \lambda = +1: \quad a = \cos\theta/2, \quad b = \sin\theta/2 \tag{68}$$

$$\text{For } \lambda = -1: \quad a = \sin\theta/2, \quad b = -\cos\theta/2 \tag{69}$$

Finally we obtain the required eigenstates of $S_{z'}$:

$$\chi'_+ = \begin{pmatrix} \cos\theta/2 \\ \sin\theta/2 \end{pmatrix} \qquad \text{and} \qquad \chi'_- = \begin{pmatrix} \sin\theta/2 \\ -\cos\theta/2 \end{pmatrix} \tag{70}$$

We can check our algebra by noting that we recover the original S_z eigenstates $\chi^\pm$ for $\theta = 0$. But by putting $\theta = 2\pi$ we fail to recover the original eigenstates, even 'though this corresponds to the same point in space as $\theta = 0$. This quite remarkable property is an entirely new phenomenon: **the spin-1/2 wave functions are not single-valued, but double-valued**:

$$\begin{aligned} & \chi'_\pm(\theta + 2\pi) \neq \chi'_\pm(\theta) \\ \text{but,} \qquad & \chi'_\pm(\theta + 4\pi) = \chi'_\pm(\theta) \end{aligned} \tag{71}$$

Finally we note the result we will apply to the double Stern-Gerlach experiment:

$$\text{For} \quad \theta = \frac{\pi}{2}; \qquad \chi'_+ = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \chi'_- = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{72}$$

A glance at Figure 1 will show that $\theta = \pi/2$ corresponds to the new $z-$axis pointing along the old $x-$axis: thus $S_{z'} = S_x$ and so these eigenstates are eigenstates of S_z. In the next section we shall obtain this result directly without quite so much algebra.
If we had repeated the above analysis, but rotated by $\bar{\theta}$ clockwise about the $x-$axis instead, we would have found,

$$S_{x''} = S_x \tag{73}$$

$$S_{y''} = S_y \cos\bar{\theta} - S_z \sin\bar{\theta} \tag{74}$$

$$S_{z''} = S_z \cos\bar{\theta} + S_y \sin\bar{\theta} \tag{75}$$

so that the new matrices become,

$$S_{y''} = \tfrac{\hbar}{2} \begin{pmatrix} -\sin\bar{\theta} & -i\cos\bar{\theta} \\ i\cos\bar{\theta} & \sin\bar{\theta} \end{pmatrix} \quad S_{x''} = S_x = \tfrac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad S_{z''} = \tfrac{\hbar}{2} \begin{pmatrix} \cos\bar{\theta} & -i\sin\bar{\theta} \\ i\sin\bar{\theta} & -\cos\bar{\theta} \end{pmatrix} \tag{76}$$

The eigenstates of $S_{z''}$ would then become:

$$\chi''_+ = \begin{pmatrix} \cos\bar{\theta}/2 \\ i\sin\bar{\theta}/2 \end{pmatrix} \qquad \text{and} \qquad \chi''_- = \begin{pmatrix} \sin\bar{\theta}/2 \\ -i\cos\bar{\theta}/2 \end{pmatrix} \tag{77}$$

Finally we obtain the eigenstates of S_y:

$$\text{For} \quad \bar{\theta} = \frac{\pi}{2}; \qquad \chi''_+ = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \qquad \text{and} \qquad \chi''_- = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \tag{78}$$

We see that $\bar{\theta} = \pi/2$ corresponds to the new $z-$axis pointing along the old $y-$axis: thus $S_{z''} = S_y$ and so these eigenstates are eigenstates of S_y. You should confirm this last result by using the method of the next section to obtain it directly without quite so much algebra.

(6) SPIN EIGENFUNCTIONS FOR QUANTIZATION ALONG THE $x-$AXIS: EIGENSTATES OF S_x.

Given the eigenfunctions, eq. (43), for spin quantised along the $z-$axis, we now wish to find the eigenfunctions for spin quantised along the $x-$axis. This is equivalent to asking for the eigenstates χ' of S_x,

$$S_x\,\chi' = \frac{\hbar}{2}\lambda\,\chi'. \tag{79}$$

where, although we already know that $\lambda = \pm 1$ (because we are describing a spin-1/2 particle), we shall find it convenient to actually confirm this by calculation. Now the information we begin with is that the eigenstates of S_z are given by eq. (43), so we can use the expansion theorem to express our unknown wave functions χ' as a linear combination, eqs. (51), (52):

$$\chi' = a\chi_+ + b\chi_- = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \text{with} \qquad |a|^2 + |b|^2 = 1. \tag{80}$$

and then use the matrix representation S_x given in eq. (37),

$$S_x = \tfrac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{81}$$

Putting these all together, the requirement that χ' be an eigenstate of S_x reads:

$$\begin{aligned} S_x\,\chi' = \frac{\hbar}{2}\lambda\,\chi' &= \frac{\hbar}{2}\lambda \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \frac{\hbar}{2} \begin{pmatrix} b \\ a \end{pmatrix} \end{aligned} \tag{82}$$

Equating the elements of column vectors in the first and last lines and cancelling $\hbar/2$, we find:

$$a\lambda = b \tag{83}$$

$$a = b\lambda \tag{84}$$

Dividing the equations gives a consistency condition leading to the expected two eigenvalues:

$$\lambda^2 = 1 \quad \text{hence} \quad \lambda = \pm 1 \tag{85}$$

Since a and b have the same *magnitude* for both eigenvalues, the condition that $\chi'_\pm$ be normalised requires:

$$|a|^2 + |b|^2 = 2|a|^2 = 1, \qquad \text{or, taking } a \text{ real and positive,} \qquad a = \frac{1}{\sqrt{2}} \tag{86}$$

$$\text{and therefore} \quad a = \frac{1}{\sqrt{2}} \quad b = \pm\frac{1}{\sqrt{2}} \quad \text{for} \quad \lambda = \pm 1 \tag{87}$$

Finally we obtain the two eigenstates χ'_+ and χ'_-,

$$\chi'_\pm = \frac{1}{\sqrt{2}}(\chi_+ \pm \chi_-) \tag{88}$$

or, in explicit form,

$$\chi'_+ = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \chi'_- = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{89}$$

where the $\pm$ denote eigenstates of S_x with eigenvalues $\pm\hbar/2$. It goes without saying that these states are still also eigenstates of $\mathbf{S}^2$ with the same eigenvalue $3\hbar^2/4$ as the original eigenstates $\chi_\pm$ of S_z. We shall see the physical significance of all four of these states when we consider the triple Stern-Gerlach experiment.

(7) THE TRIPLE STERN-GERLACH EXPERIMENT.

Before describing the details of the Stern-Gerlach experiment, we shall first abstract its essence in order to discuss the theory of measurement in QM. The experiment can be thought of as a method for either *measuring* a component of the spin or of *preparing* an ensemble of states having one component of spin determined, ie. of preparing an eigenstate of S_z, S_x or whatever. We represent the apparatus as a 'black box' into which a beam of spin-1/2 particles is injected and from which emerge two spatially separated beams: one with spin projection along the measurement axis $+\hbar/2$; the other with $-\hbar/2$. Figure 2 illustrates two separate measurement experiments in which an unpolarised beam of spin-1/2 particles enters the apparatus. We can think of the incoming beam as having a random mixture of all possible spin orientations: one experiment measures the $z-$component; the other the $x-$component. The two outgoing states are spatially separated and exit in the pure eigenstates $\chi_\pm$ and $\chi'_\pm$ respectively.

Figure 2: Abstract Stern-Gerlach experiments, one measuring the $z-$component of spin, the other the $x-$component. Note that the outgoing beams have equal intensity one-half that of the incoming beam.

These illustrate the angular momentum quantization rule that for a spin-1/2 particle (the incoming particle) the only possible outcome of a measurement of a component of the spin angular momentum is one of the eigenvalues, $+\hbar/2$ or $-\hbar/2$. This is the first stage of the measurement postulate of QM. Note how radically different the outcomes of these experiments are compared with our classical expectation: since the incoming beam is unpolarised we expect the spins to be oriented randomly in *all* possible directions, with projections on the z-axis having a continuum of values lying between $+\hbar/2$ and $-\hbar/2$; instead we only measure two values, and any *single* particle passing through the apparatus will exit either in one beam or the other, not anywhere between the beams. Even more at variance with our classical experience is that the very *same* incoming beam going through the S_x measuring apparatus now seems to have only the maximum possible x-axis projections, $\pm\hbar/2$, a clear impossibility if they already have the maximum z-axis projections as evinced by the S_z measuring experiment! From the quantum mechanical point of view we have to accept that in the incoming beam the spin projection of an *individual* particle is not known, only its probability. Thus it is only *after* a measurement or preparation that we can say for certain what it is, and then only at the moment of measurement, not before. If we are tempted to say, 'Well, it really had the value we determined all along and we just didn't know it until we actually looked', then the fact that a similar measurement of the x-component can only

yield $\pm\hbar/2$ should cast doubt on this interpretation, since now it seems it could 'really' have had a component inconsistent with the value already determined. Thus it seems that a particle's property depends on which experiment we set up. We shall see this in the triple Stern-Gerlach experiment. Thus, just like the two-slit experiment, we have to think of an incoming particle as being in neither one state χ_+ nor the other χ_-, but 'in some sort of equal mixture of the two'; the physical meaning of this statement is that a *measurement* of the z-component of spin for an unpolarised beam will yield *only one* of the two eigenvalues of S_z, but with equal *probability.* Once the measurement is made, the wave function 'collapses' into the appropriate state.[9]

The double Stern-Gerlach experiment enables us to illustrate the next stage of the measurement postulate. First consider *two* Stern-Gerlach experiments, both measuring the $z-$component of spin. We now discard one of the beams emerging from the first stage. This can be viewed as measuring the $z-$component of spin and using this to prepare the system in the eigenstate χ_+, or as the 'collapse' of the wave function into the state χ_+ as a result of the measurement (which is the process of separating the incoming beam *and* discarding the lower one). Since the state entering the second stage has 'collapsed' to the eigenstate χ_+ (or has been *prepared* in that eigenstate) the result of the second measurement is certain: only *one* beam now exits the second, corresponding to spin up.

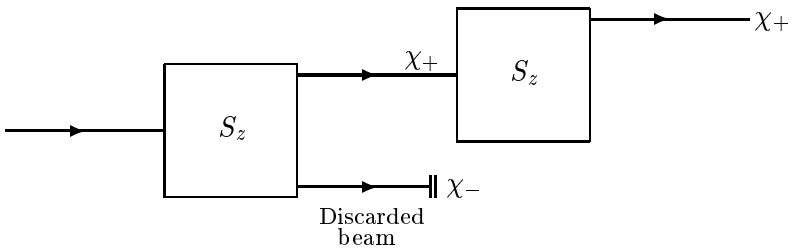
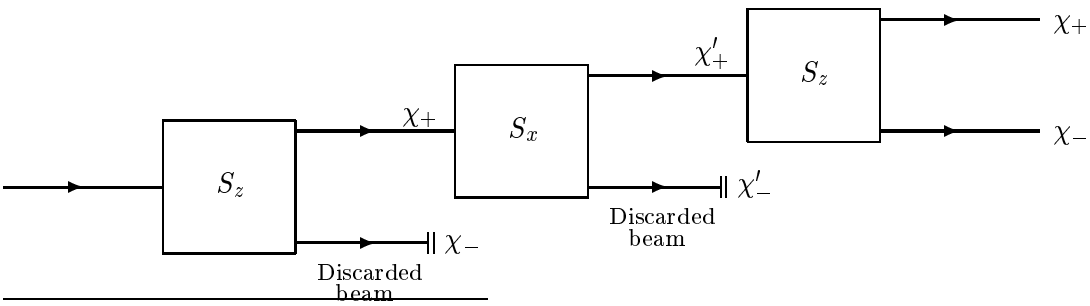


Figure 3: Double Stern-Gerlach experiment illustrating the measurement process in QM. Both measure the $z-$component of spin, but only one beam emerging from the first is admitted into the second.[10]

The triple Stern-Gerlach experiment illustrates several aspects of the quantum measuring process and dramatises the consequences of the expansion theorem in a most remarkable way. We now interpose between the two S_z experiments of the double Stern-Gerlach apparatus a third experiment which measures the $x-$component, S_x:



[9] A subtle point about the wave function representing the unpolarised incoming beam: we have not dealt with such incoherent states in this course, and so I have avoided writing down a wave function for it; but such states have the property that we can still predict the probabilities I am talking about in the text. The formalism used for these so-called mixed states is the density matrix formalism.

[10] Since there is nothing special about the symbol z, we could equally well have illustrated the principle with measurement of the $x-$component. The same picture would serve with the replacements: $S_z \to S_x$ and $\chi_\pm \to \chi'_\pm$

Figure 4: Triple Stern-Gerlach experiment.

The first and last stages measure the $z-$component of spin, but the beam selected from the first is now submitted to an S_x measurement and only one emergent beam is in its turn submitted to the final stage, a second S_z measurement. The intermediate measurement has now influenced the outcome of the last measurement and we see the phenomenon of regeneration: the down spin state discarded in the first stage reappears in the outgoing beam of the whole experiment, as if the intermediate measurement of S_x had 'disturbed' the system and reintroduced some down component. In equations, the sequence is as follows: the incoming state is an unpolarised mixture of spin up and spin down with respect to any chosen axis; the $z-$axis is the one chosen for the first stage measurement, so the spin up and down beams emerging from the S_z measurement are of equal intensity. Choosing to inject only the spin up beam into the S_x measuring apparatus means that the incoming state for *this* stage is, χ_+, which, using the expansion theorem we may write in terms of the S_x eigenstates,

$$\chi_+ = \frac{1}{\sqrt{2}}\chi'_+ + \frac{1}{\sqrt{2}}\chi'_- \tag{90}$$

Again there are only two beams emerging from the S_x measurement with equal intensity, from which we select the one in state χ'_+. This is the initial state for the final S_z measurement, which we again expand, but this time in terms of S_z eigenstates:

$$\chi'_+ = \frac{1}{\sqrt{2}}\chi_+ + \frac{1}{\sqrt{2}}\chi_- \tag{91}$$

Here we have the final outcome: there are two equal intensity beams emerging from the final stage, with *each* S_z eigenstate being present, despite the fact that the first stage of the experiment had discarded the beam in the χ_- spin down state.

Now we come to a crucial point: one may wish to attribute the regeneration observed in the triple Stern-Gerlach experiment to the so-called 'disturbance' introduced by the intermediate S_x measurement; but in fact this is no disturbance at all. We can show this by altering the last experiment by not excluding the S_x spin down beam, but allowing it also to enter the last S_z measuring phase. The result is that *only* the spin up component emerges from the last stage:

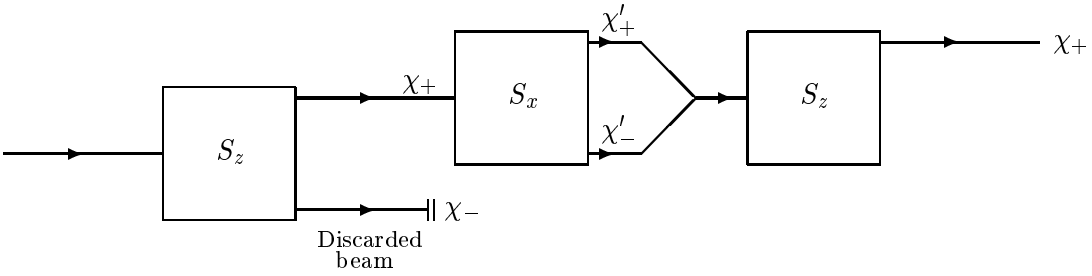


Figure 5: Triple Stern-Gerlach experiment without discarding the down component at the intermediate stage. This shows that the intermediate stage introduces no 'disturbance' when no detection occurs there.

In equations this is simply stated: the two beams exiting the S_z apparatus are recombined when they enter the final stage, maintaining their coherence. Thus the beam entering the final stage is a linear superposition of equal intensity beams, with wave function

$$\begin{aligned} \chi_{S_z out} &= \frac{1}{\sqrt{2}}\chi'_+ + \frac{1}{\sqrt{2}}\chi'_- \\ &= \chi_+ \end{aligned} \tag{92}$$

where, in the last line, we recognise the combination of S_x eigenstates as the spin up S_z eigenstate χ_+. Hence we confirm that the S_x apparatus itself does not disturb the system. Rather than the collapse of the wave function resulting from a disturbance *inside* the S_x apparatus, it is a consequence of a measurement being registered afterwards - in the case of the triple Stern-Gerlach experiment the measurement is the process of completely excluding the lower beam emerging from the S_x apparatus; the upper beam is not touched at all and therefore suffers no disturbance. The analogy with the two-slit experiment should be obvious: if we block the beam from one slit altogether we lose the interference pattern. Thus we see that the act of measurement introduces a degree of decoherence - the triple Stern-Gerlach experiment described is an extreme case where one beam is completely destroyed, but one can imagine introducing a counter in the lower beam instead as one does in attempting to follow the path of the beams in the two-slit experiment. Finally we note that the experiments described here are very closely related to the famous Aspect et al experiments which tested the Bell inequalities and confirmed this apparently strange behaviour of the QM measurement process.

(8) ATOMIC MAGNETIC MOMENTS & ATOMS IN A MAGNETIC FIELD.

The purpose of the Stern-Gerlach experiment is to measure the magnetic moments of *neutral* atoms, so we begin by examining the magnetic properties of a classical atomic electron of charge $-e$ orbiting a nucleus in a circle of radius r with speed v:

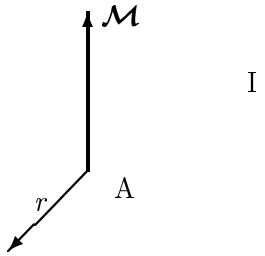


Figure 6: Magnetic moment produced by a current loop.

The orbiting electron produces a current I =charge×(no. of circuits/sec) $= (-e) \times (v/2\pi r)$ circulating around a loop of area $A = \pi r^2$, causing a dipole magnetic field just like that of a tiny bar magnet oriented perpendicular to the orbital plane, with a **magnetic moment**:

$$\mathcal{M} = IA = -\frac{ev}{2\pi r}\,\pi r^2 = -\frac{e}{2m_e}\,m_e v r = -\frac{e}{2m_e}\,L = -\mu_B \frac{L}{\hbar} = -\gamma L \tag{93}$$

or in its full vectorial form,

$$\boldsymbol{\mathcal{M}} = -\mu_B \frac{\mathbf{L}}{\hbar} = -\gamma \mathbf{L} \qquad \text{for an orbiting } e^- \tag{94}$$

where $\gamma \equiv \mu_B/\hbar$ is the **gyromagnetic ratio** and

$$1 \text{ Bohr magneton} \equiv \mu_B = \frac{e\hbar}{2m_e} = 9.27 \times 10^{-24} \text{ J T}^{-1} \quad \text{(Joules per Tesla)} \tag{95}$$

We have established an important result: because the magnetic moment is proportional to the angular momentum, any measurement of the magnetic moment is at the same time a measurement of the angular momentum multiplied by the gyromagnetic ratio. It turns out that this is true whether we are talking about orbital or spin angular momentum or a combination of both: the only difference resides in the numerical values.

If we place our atomic dipole in a magnetic field it will experience a force, which we can express as a potential energy:[11]

$$E_{mag} = -\mathbf{B}.\boldsymbol{\mathcal{M}} \tag{96}$$

$$= \gamma\, \mathbf{B}.\mathbf{L} \qquad \text{for an orbiting } e^- \tag{97}$$

$$= \gamma B_z L_z \qquad \text{for an } e^- \text{ and } z-\text{axis along } \mathbf{B} \tag{98}$$

This term, with $L_z \to \widehat{L}_z$, is added to the Hamiltonian and, acting on an eigenstate of both $\widehat{H}$ and $\widehat{L}_z$, gives an additional energy to the eigenstate,

$$E_{mag} = \gamma \hbar m_\ell B_z \qquad \text{for an } e^- \text{ and } z-\text{axis along } \mathbf{B} \tag{99}$$

$$= \hbar \omega_L\, m_\ell \qquad \text{for an orbiting } e^- \tag{100}$$

[11] See Appendix C.1 and Figure 10 for a simple derivation. The negative gradient of this potential gives the force **F** on the dipole. In general a force acting on the dipole will produce both linear and rotational motion - the former is zero in a uniform field; the latter is caused by a torque, $\mathbf{r} \times \mathbf{F} \neq 0$, which is present even in a uniform field, causing precession of the dipole - see Appendix C.2 and Figures 11, 12, 13.

Thus, we see that the energy corresponds to a characteristic frequency, the Larmour frequency, which will appear as the resonant frequency of radiation needed to induce the angular momentum to change by one unit of $\hbar$ in magnetic resonance experiments. This is also the frequency with which a classical dipole precesses about the direction of the magnetic field - see Appendix C.2 and Figures 11, 12, 13, where we show that both the classical components L_x and L_y precess about the $z-$direction of the magnetic field with the Larmour frequency; and quantum mechanically it is the expectation values $\langle \widehat{L}_x \rangle$ and $\langle \widehat{L}_y \rangle$ which precess at this frequency- see Appendix D.

(9) THE ZEEMAN EFFECT - EVIDENCE FOR AZIMUTHAL QUANTIZATION AND FOR ELECTRONS IN ATOMS.

Since the application of a magnetic field changes the energy of orbiting electrons, it also has an interesting effect on spectra. Historically this was the first evidence for angular momentum quantization as well as for the role of electrons in producing atomic spectra. It also provided early hints of angular momentum 1/2. We have shown that for an atomic electron in a quantum state labelled by the angular momentum quantum numbers ℓ, m_ℓ the application of a magnetic field shifts the energy by

$$\Delta E_{mag} = \gamma \hbar m_\ell B_z \qquad \text{with} \quad m_\ell = -\ell, -\ell+1, \ldots \ell \tag{101}$$

$$= \hbar \omega_L \, m_\ell \tag{102}$$

Now consider two energy levels in an atom, one the ground state with $\ell = 0$, the other with $\ell = 1$. In the absence of a magnetic field the $\ell = 1$ level is 3-fold degenerate: there are three states, all with $\ell = 1$, but with $m_\ell = -1, 0, +1$, and all with the same energy. Thus there is only one spectral line at a frequency ν_0 corresponding to the energy difference $E_1 - E_0 = h\nu_0$ between the two levels. Now with an applied magnetic field the $\ell = 0, m_\ell = 0$ and $\ell = 1, m_\ell = 0$ levels remain unchanged because $m_\ell = 0$ gives $E_{mag} = 0$; but the two $\ell = 1, m_\ell = -1, +1$ levels shift down and up by $\delta E = \hbar\omega_L$ since $m_\ell = \pm 1$. The result is *three separate* levels where one existed before: the magnetic field has **lifted the degeneracy** and created two additional spectral lines on either side of the original line. The three lines illustrated in Figure 7 correspond to the normal Zeeman effect. Spin and other aspects of the magnetic interactions in an atom lead to complicated Zeeman patterns known as the anomalous Zeeman effect. The observation of the Zeeman effect was an important milestone in atomic physics because it provided evidence for the quantization of angular momentum *and* for the electron being in the atom and directly involved in the production of spectral lines: the splitting is both proportional to the applied magnetic field and the charge to mass ratio e/m_e. The measured splitting gave the same value for e/m_e as had been obtained in studies of electron beams (beta rays), providing a direct window into the atom. The Zeeman effect is widely used in astronomy to measure large magnetic fields in stars.

(10) THE STERN-GERLACH EXPERIMENT & ELECTRON SPIN.

It was Otto Stern in 1921 who proposed to use the effect of an **inhomogeneous magnetic field** on a magnetic dipole to measure the magnetic moments of atoms. If $\mathbf{B} = \mathbf{B}(x, y, z)$ then, in addition to a torque the field will exert a net force on the dipole because the little magnet's north and south poles are in slightly different fields. In all our considerations we find it most convenient to choose the $z-$axis to lie along the direction of the magnetic field, $\mathbf{B} = (0, 0, B_z)$. As usual, the force is obtained from the gradient of the potential (see Appendix C.2 and Figure 14 for a derivation):

$$\begin{aligned} \mathbf{F} &= -\boldsymbol{\nabla} E_{mag} && (103)\\ &= \boldsymbol{\nabla}(\boldsymbol{\mathcal{M}}.\mathbf{B}) && (104)\\ &= \boldsymbol{\nabla}(\mathcal{M}_z B_z) \quad \text{with } z-\text{axis along } \mathbf{B} && (105)\\ &= \mathcal{M}_z \left(\frac{\partial B_z}{\partial x}, \frac{\partial B_z}{\partial y}, \frac{\partial B_z}{\partial z}\right) && (106) \end{aligned}$$

In the Stern-Gerlach experiment the magnet poles are shaped to produce a magnetic field that varies rapidly in the $z-$direction only, $B_z = B_z(z)$, [12] so the force also acts in the $z-$direction:

$$\begin{aligned} \mathbf{F} &= (0, 0, \mathcal{M}_z \frac{\partial B_z}{\partial z})\\ &= (0, 0, -\gamma L_z \frac{\partial B_z}{\partial z}) \quad \text{classical orbiting } e^- && (107)\\ &= (0, 0, -\gamma \hbar m_\ell \frac{\partial B_z}{\partial z}) \quad \text{quantum mechanical orbiting } e^- && (108) \end{aligned}$$

where, in the last step, we have introduced the QM quantization rule for the $z-$component of orbital angular momentum. The 1922 Stern-Gerlach experiment was performed with a beam of *neutral* [13] silver atoms, the inhomogeneous magnetic field being produced by one of the poles having a razor-sharp pointed face. The details are shown in the accompanying Figures 8 & 9 which also show the original results. Since the silver atoms' magnetic moments are randomly oriented in the incident beam, the classical expectation was that there be a continuous range of deflections, reflecting the continuous distribution of L_z values in the beam. The QM prediction is that a *discrete* set of deflections should result from the quantization of orbital angular momentum with $2\ell + 1$ possibilities corresponding to $m_\ell = -\ell, -\ell + 1, \ldots 0, \ldots \ell$, including zero deflection for $m_\ell = 0$. In fact no zero deflection was seen and only 2 deflected beams, one above and one below the expected zero deflection, inconsistent with integral values of ℓ but suggesting instead the value $1/2$ to give $2\ell + 1 = 2$.

This experiment and many other hints from atomic spectra and the Zeeman effect led to the Goudschmidt-Uhlenbeck suggestion in 1925 that in fact the electron has an *intrinsic* spin angular momentum $\hbar/2$. The Stern-Gerlach experiment on Ag and later ones on Cu, Au, Na, K, Cs and H finds its explanation in the fact that all these atoms have *closed electron shells plus one extra valence electron.* The closed shell contributes a net zero angular momentum and the valence electron has zero orbital angular momentum, $\ell = 0$, but $1/2$ unit of spin angular momentum. The atom as a whole therefore has a total angular momentum $\hbar/2$ due solely to the

[12] In the experiment depicted in Figure 8 the beam travels along the $y-$direction, with the field uniform along this direction, although there is a small $x-$component to the field which does vary slightly with both x and z. In practice these effects can be seen as distortions in the shape of the outgoing beam seen at the detector as shown in the figure, but do not obscure the main effect seen due to the z variation.

[13] It is essential to use a *neutral* beam in this experiment otherwise the relatively weak deflections due to the magnetic dipole interaction would be completely masked by the large deflections caused by a charged beam passing through a strong magnetic field.

spin of the valence electron, and it is its accompanying spin magnetic moment which is being measured in the Stern-Gerlach experiment. Thus the experiment provided evidence both for the existence of electron spin *and* for the azimuthal quantization of angular momentum. However there was at least one more surprise: the electron also has twice the magnetic moment expected from the analogy with orbital motion. This experimental result is expressed in terms of the g-factor, g_s defined by:

$$\boldsymbol{\mathcal{M}_s} = -g_s \mu_B \frac{\mathbf{S}}{\hbar} = -\gamma \mathbf{S} \qquad \text{for an } e^- \tag{109}$$

The modern value for the electron is $g_s = 2.002381304386 \pm 0.000000000020$ corresponding to a spin magnetic moment of very nearly 1 Bohr magneton: the factor of 1/2 from the spin S is almost exactly cancelled by the 2 from the g-factor. This result has no classical explanation, but the value $g_s = 2$ *is* a consequence of Dirac's relativistic equation for the electron; the small correction is actually in agreement with modifications calculated from Quantum Electrodynamics, the relativistic extension of quantum mechanics to the interaction of photons and electrons. Other particles in nature have been found with spin-1/2, including the proton and the neutron. Because their masses are some 1800 times larger than the electron the nuclear magneton is smaller by that factor:[14]

$$1 \text{ Nuclear magneton} \equiv \mu_N = \frac{e\hbar}{2m_p} = 5.446 \times 10^{-4} \mu_B = 5.05 \times 10^{-27} \text{ J T}^{-1} \tag{110}$$

with $g_p = 5.586$ and $g_n = -3.826$. The reason these particles have g-factors so different from the canonical value of 2 given by the Dirac equation is because the strong interactions play a powerful role in giving these particles a finite size and a complicated and measurable structure - they are surrounded by clouds of so called virtual particles such a pi mesons.

For a spinning particle such as an electron in an atom which also has orbital angular momentum, the magnetic moment is the vector sum:

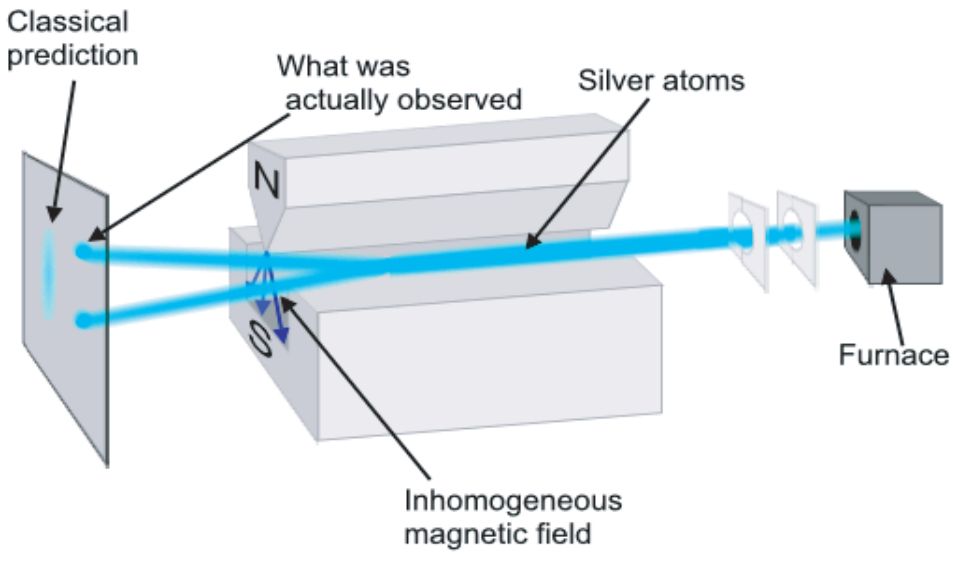
$$\begin{aligned} \boldsymbol{\mathcal{M}_{tot}} &= \boldsymbol{\mathcal{M}_L} + \boldsymbol{\mathcal{M}_s} & (111)\\ &= -\frac{\mu_B}{\hbar}(\mathbf{L} + g_s \mathbf{S}) \qquad \text{for an } e^-. & (112) \end{aligned}$$

Note finally that all our previous expressions for the effects of a magnetic field on a dipole produced by *orbital* motion also hold for the spin magnetic moment with the replacements $L \to S$ and $m_\ell \to m_s$. In particular, for an electron placed in a magnetic field the energy shift due to the interaction with its spin magnetic moment is, since $g_s \approx 2$ to high accuracy:

$$\begin{aligned} E_{mag} &= \gamma \hbar m_s B_z \qquad \text{for an } e^- \text{ and the } z-\text{axis along } \mathbf{B} & (113)\\ &= g_s \mu_B B_z \, m_s & (114)\\ &= \hbar \omega_s \, m_s \qquad \text{where} \quad \omega_s = \gamma B_z = \frac{g_s \mu_B}{\hbar} B_z & (115) \end{aligned}$$

Again, we see that the energy corresponds to a characteristic frequency, but for the electron with $g_s \approx 2$, ω_s is twice the Larmour frequency. This frequency will appear as the resonant frequency of radiation needed to induce the spin angular momentum to change by one unit of $\hbar$ in magnetic resonance experiments. The proton will give a similar expression with the opposite sign, but with the much smaller nuclear magneton giving a much lower energy. It is primarily the spins which are involved in ESR and NMR experiments.

[14] Beware signs: we have used an *explicit* minus sign up to this point for the *electron*'s magnetic moment, so it's g_s is positive. To apply the above equations to the proton we have to remove this minus sign in appropriate places. This doesn't cause any confusion in practice.



APPENDIX A: MATRIX ELEMENTS OF ANGULAR MOMENTUM OPERATORS - GENERAL EXPRESSIONS.

We shall now find expressions for the matrices representing the angular momentum operators.[15] We know that the operator $\widehat{J}_+$ is the raising operator, generating from ψ_{j,m_j} with $\widehat{J}_z$ eigenvalue $\hbar m_j$, the next state up the ladder, $\widehat{J}_+\psi_{j,m_j}$ with eigenvalue $\hbar(m_j+1)$. But the state so generated is not normalised and requires a normalising factor N to become so:

$$\widehat{J}_+\psi_{j,m_j} = N\,\psi_{j,m_j+1} \tag{116}$$

where both ψ_{j,m_j} and ψ_{j,m_j+1} are *normalised* eigenstates. Thus,

$$\begin{aligned}
\int |\widehat{J}_+\psi_{j,m_j}|^2\,d^3x &= N^2\int |\psi_{j,m_j+1}|^2\,d^3x = N^2 \quad \text{since } \psi_{j,m_j+1} \text{ is normalised.} &(117)\\
&= \int (\widehat{J}_+\psi_{j,m_j})^*\,\widehat{J}_+\psi_{j,m_j}\,d^3x \\
&= \int (\psi_{j,m_j})^*\,\widehat{J}_+^\dagger\,\widehat{J}_+\psi_{j,m_j}\,d^3x \quad \text{using definition of adjoint,} \\
&= \int \psi^*_{j,m_j}\widehat{J}_-\widehat{J}_+\psi_{j,m_j}\,d^3x \quad \text{using} \quad \widehat{J}_+^\dagger = \widehat{J}_- \\
&= \int \psi^*_{j,m_j}(\widehat{\mathbf{J}}^2 - \widehat{J}_z^2 - \hbar\widehat{J}_z)\psi_{j,m_j}\,d^3x \\
&\quad \text{using the identity} \quad \widehat{J}_-\widehat{J}_+ = \widehat{\mathbf{J}}^2 - \widehat{J}_z^2 - \hbar\widehat{J}_z \\
&= \hbar^2\left[j(j+1) - m_j(m_j+1)\right]\int \psi^*_{j,m_j}\psi_{j,m_j}\,d^3x \\
&\quad \text{since } \psi_{j,m_j} \text{ is an eigenstate of } \widehat{\mathbf{J}}^2 \text{ and } \widehat{J}_z \\
&= \hbar^2\left[j(j+1) - m_j(m_j+1)\right] \quad \text{since the eigenstate is normalised.} &(118)
\end{aligned}$$

leading to

$$N = \hbar\left[j(j+1) - m_j(m_j+1)\right]^{1/2} \tag{119}$$

A similar argument can be carried out for the state one step down the ladder, $\widehat{J}_-\psi_{j,m_j}$, giving finally for the *normalised* eigenstates,

$$\widehat{J}_\pm\psi_{j,m_j} = \hbar\left[j(j+1) - m_j(m_j\pm1)\right]^{1/2}\,\psi_{j,m_j\pm1} \tag{120}$$

The matrix elements for the raising and lowering operators then follow:[16]

$$\begin{aligned}
(J_\pm)_{m'_j,m_j} &\equiv \int \psi^*_{j,m'_j}\widehat{J}_\pm\psi_{j,m_j}\,d^3x \\
&= \hbar\left[j(j+1) - m_j(m_j\pm1)\right]^{1/2}\int \psi^*_{j,m'_j}\psi_{j,m_j\pm1}\,d^3x \\
&\quad \text{since } \psi_{j,m_j} \text{ is an eigenstate of } \widehat{\mathbf{J}}^2 \text{ and } \widehat{J}_z \\
&= \hbar\left[j(j+1) - m_j(m_j\pm1)\right]^{1/2}\,\delta_{m'_j,m_j\pm1} \\
&\quad \text{since the eigenstates are orthonormalised.} &(121)
\end{aligned}$$

[15] By using the notation $\mathbf{J}$ and j, m_j we are emphasising that the final expressions are generally valid and can be proved so by using the Dirac formalism, although the integrals only have meaning for orbital angular momentum and the eigenstates of $\widehat{\mathbf{L}}^2$ and $\widehat{L}_z$, $\psi_{j,m_j} = Y_{j,m_j}$, with the identifications $j = \ell$ and $m_j = m_\ell$.

[16] The states on the $j-$ladder, ψ_{j,m_j} are orthonormal as we know in general for eigenstates of Hermitian operators and as we have seen explicitly for eigenstates of orbital angular momentum:

$$\int \psi^*_{j,m'_j}\psi_{j,m_j}\,d^3x = \int Y^*_{j,m'_j}Y_{j,m_j}\,d^3x = \delta_{m'_j,m_j}$$

We are now in a position to find the matrix elements of the angular momentum operators $\widehat{J}_x$ and $\widehat{J}_y$ from those of $\widehat{J}_\pm$ using

$$\widehat{J}_\pm \equiv \widehat{J}_x \pm i\widehat{J}_y, \tag{122}$$

$$\text{giving for } \widehat{J}_x \text{ and } \widehat{J}_y, \quad \widehat{J}_x = \frac{1}{2}\left(\widehat{J}_+ + \widehat{J}_-\right), \tag{123}$$

$$\widehat{J}_y = \frac{1}{2i}\left(\widehat{J}_+ - \widehat{J}_-\right). \tag{124}$$

The result, despite its complicated appearance, is very simple when evaluated for individual cases:

$$(J_x)_{m'_j,m_j} = \frac{\hbar}{2}\left\{[j(j+1) - m_j(m_j+1)]^{1/2}\ \delta_{m'_j,m_j+1} + [j(j+1) - m_j(m_j-1)]^{1/2}\ \delta_{m'_j,m_j-1}\right\} \tag{125}$$

$$(J_y)_{m'_j,m_j} = \frac{\hbar}{2i}\left\{[j(j+1) - m_j(m_j+1)]^{1/2}\ \delta_{m'_j,m_j+1} - [j(j+1) - m_j(m_j-1)]^{1/2}\ \delta_{m'_j,m_j-1}\right\} \tag{126}$$

Since the states on the $j-$ladder are eigenstates, their matrix representations are diagonal:

$$(J_z)_{m'_j,m_j} = \hbar\, m_j\, \delta_{m'_j,m_j} \tag{127}$$

$$(\mathbf{J}^2)_{m'_j,m_j} = \hbar^2\, j(j+1)\, \delta_{m'_j,m_j} \tag{128}$$

The hard work is now complete: we have found general expressions for *all* the matrices representing the angular momentum operators in Heisenberg's matrix quantum mechanics.[17] The enormous power of the methods used should be appreciated: without knowing anything about the physical nature of spin, except that it is a form of angular momentum giving the magic multiplicity 2 and suggesting $j = 1/2$, we are now in a position to write down explicit expressions for the spin operators and the wave functions describing the quantum mechanics of the spin degrees of freedom of particles ranging from electrons to quarks.

APPENDIX B: MATRIX ELEMENTS OF ANGULAR MOMENTUM OPERATORS - THE CASES $j = 1/2$ & $j = 1$.

Using the general expressions obtained above in equations (26) to (29), we now list the first two sets of matrix operators representing the abstract quantum mechanical operators $\widehat{\mathbf{J}}$:

$\boxed{j = \frac{1}{2}, \quad m_j = -1/2, +1/2}$ The matrices are $(2j+1) \times (2j+1) = 2 \times 2-$dimensional:

$$J_x = \begin{pmatrix} (J_x)_{1/2,1/2} & (J_x)_{1/2,-1/2} \\ (J_x)_{-1/2,1/2} & (J_x)_{-1/2,-1/2} \end{pmatrix} = \tfrac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{129}$$

Notice how the diagonal elements of the matrix are zero because the kronecker deltas require $m'_j = m_j \pm 1 \neq m_j$. Similar considerations apply to J_y, giving:

$$J_y = \begin{pmatrix} (J_y)_{1/2,1/2} & (J_y)_{1/2,-1/2} \\ (J_y)_{-1/2,1/2} & (J_y)_{-1/2,-1/2} \end{pmatrix} = \tfrac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{130}$$

The diagonal matrices are easily obtained as:

$$J_z = \begin{pmatrix} (J_z)_{1/2,1/2} & (J_z)_{1/2,-1/2} \\ (J_z)_{-1/2,1/2} & (J_z)_{-1/2,-1/2} \end{pmatrix} = \tfrac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{131}$$

[17] In the process we have solved a huge and sophisticated problem in the mathematical theory of groups: we have found all the irreducible representations of the orthogonal group in three dimensions, the rotation group O(3).

$$\mathbf{J}^2 = \begin{pmatrix} (\mathbf{J}^2)_{1/2,1/2} & (\mathbf{J}^2)_{1/2,-1/2} \\ (\mathbf{J}^2)_{-1/2,1/2} & (\mathbf{J}^2)_{-1/2,-1/2} \end{pmatrix} = \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{132}$$

$\boxed{j=1, \quad m_j = -1, 0, +1}$ The matrices are $(2j+1) \times (2j+1) = 3 \times 3-$dimensional:

$$J_x = \begin{pmatrix} (J_x)_{1,1} & (J_x)_{1,0} & (J_x)_{1,-1} \\ (J_x)_{0,1} & (J_x)_{0,0} & (J_x)_{0,-1} \\ (J_x)_{-1,1} & (J_x)_{-1,0} & (J_x)_{-1,-1} \end{pmatrix} = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{133}$$

Again the diagonal elements of the matrix are zero because the kronecker deltas require $m_j' = m_j \pm 1 \neq m_j$; but they also require the elements 2 steps away from the diagonal to vanish. Similar considerations apply to J_y, giving:

$$J_y = \begin{pmatrix} (J_y)_{1,1} & (J_y)_{1,0} & (J_y)_{1,-1} \\ (J_y)_{0,1} & (J_y)_{0,0} & (J_y)_{0,-1} \\ (J_y)_{-1,1} & (J_y)_{-1,0} & (J_y)_{-1,-1} \end{pmatrix} = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \tag{134}$$

The diagonal matrices are easily obtained as:

$$J_z = \begin{pmatrix} (J_x)_{1,1} & (J_z)_{1,0} & (J_z)_{1,-1} \\ (J_z)_{0,1} & (J_z)_{0,0} & (J_z)_{0,-1} \\ (J_z)_{-1,1} & (J_z)_{-1,0} & (J_z)_{-1,-1} \end{pmatrix} = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{135}$$

$$\mathbf{J}^2 = \begin{pmatrix} (\mathbf{J}^2)_{1,1} & (\mathbf{J}^2)_{1,0} & (\mathbf{J}^2)_{1,-1} \\ (\mathbf{J}^2)_{0,1} & (\mathbf{J}^2)_{0,0} & (\mathbf{J}^2)_{0,-1} \\ (\mathbf{J}^2)_{-1,1} & (\mathbf{J}^2)_{-1,0} & (\mathbf{J}^2)_{-1,-1} \end{pmatrix} = 2\hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{136}$$

That these matrices have all the properties required of angular momentum operators can now be checked explicitly. Thus, for the spin-1/2 *matrices* it is easy to check that their commutators are precisely the same as those of the angular momentum *operators.* Thus we find another way to represent operators in QM: instead of using differential operators, Heisenberg's QM uses matrices, and the failure of some operators to commute is here realised through the well known corresponding property of matrices. For angular momentum there is a particular simplification: the matrices are finite dimensional so that there is also a finite number of eigenstates. [18]

[18] Compare with the case of the Hamiltonian operator for the infinite square well or the harmonic oscillator where there is an infinity of eigenstates and therefore the matrices would be infinite-dimensional.

APPENDIX C: CLASSICAL TREATMENT OF SPIN PRECESSION IN A MAGNETIC FIELD. [19]

(C.1) Energy of a magnetic dipole in a magnetic field.

An orbiting charged particle or a spinning particle possesses a magnetic moment $\boldsymbol{\mathcal{M}}$, which can be pictured as a tiny bar magnet with the vector $\boldsymbol{\mathcal{M}}$ pointing from the South to the North pole. In Figure 10 we place this dipole in a uniform magnetic field $\mathbf{B}$, whose direction we choose as the $z-$axis. Even though isolated magnetic monoples do not seem to occur in nature, a magnetic dipole can be thought of as a North pole with a magnetic 'charge' $+m$ separated by a distance $2r$ from a South pole with a magnetic 'charge' $-m$. The dipole moment is then $\mathcal{M} = m\times$(the separation)$= 2mr$ or, taking the vector separation $2\mathbf{r}$ to point from S to N,

$$\boldsymbol{\mathcal{M}} = 2m\,\mathbf{r} \tag{137}$$

If our dipole is at an angle θ to the uniform magnetic field $\mathbf{B} = (0, 0, B_z)$ we can find the force exerted by the field on each monopole from a law exactly like that for electric charges in an electric field: just as an electric charge q in an electric field ϵ_z experiences a force $F_z = q\epsilon_z$, so a monopole of magnetic charge m experiences a force $F_z = mB_z$. These forces, acting in the direction of the fields (chosen as the $z-$direction), can be derived from potentials (ie. potential energies),

$$E_{elec} = -q\epsilon_z z \qquad \text{gives a force} \qquad F_z = -\frac{\partial E_{elec}}{\partial z} = q\epsilon_z \tag{138}$$

$$E_{mag} = -mB_z z \qquad \text{gives a force} \qquad F_z = -\frac{\partial E_{mag}}{\partial z} = mB_z \tag{139}$$

Taking the origin of coordinates at the centre of the dipole, we can now find its total potential energy in the field by summing up the energies of the North pole situated at $z = r\cos\theta$ and the South pole situated at $z = -r\cos\theta$:

$$\begin{aligned} E_{mag} &= E_{+m} + E_{-m} \\ &= -mB_z r\cos\theta - (-m)B_z(-r\cos\theta) \\ &= -(2mr)\cos\theta B_z \\ &= -(\mathcal{M}\cos\theta)B_z \qquad \text{where the dipole moment is} \quad \mathcal{M} = 2mr \\ &= -\boldsymbol{\mathcal{M}}.\mathbf{B} \end{aligned} \tag{140}$$

where, in the last step we recognised θ as the angle between the vectors $\mathbf{B}$ and $\boldsymbol{\mathcal{M}}$. This expression is actually also valid in an inhomogeneous field and is used in section 9 of the text to obtain the forces acting on the dipole in the Stern-Gerlach experiment.

[19] I use the notation L throughout this Appendix because all classical angular momentum is orbital angular momentum. However all the results carry over to quantum mechanics provided due care is taken, so they apply both to orbital, L, and spin, $L \to S$, angular momentum. To convert the equations to spin we need only make the replacements $L \to S$, $\ell \to s$ ($s = 1/2$ for spin 1/2), $m_\ell \to m_s$ ($m_s = \pm 1/2$ for spin 1/2), and the Larmour frequency is different, $\omega_L \to \omega_s$; the gyromagnetic ratio is represented by the same symbol, γ, for both cases, but has different values, $\gamma = \mu_B/\hbar$ for orbiting electrons, $\gamma = g_s\mu_B/\hbar \approx 2\mu_B/\hbar$ for spinning electrons and $\gamma = g_s\mu_N/\hbar$ for neutrons and protons.

(C.2) Dynamics of a dipole in a uniform magnetic field.

We begin with a quick and dirty derivation of the classical law for angular motion:[20] given the angular momentum vector, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, let us find its equation of motion using Newton's second law:

$$\frac{d\mathbf{p}}{dt} = \mathbf{F} \qquad \text{where} \quad \mathbf{p} = m\frac{d\mathbf{r}}{dt} \tag{141}$$

Taking the time derivative of the angular momentum,

$$\begin{aligned} \frac{d\mathbf{L}}{dt} &= \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) \\ &= \frac{d\mathbf{r}}{dt} \times \mathbf{p} + \mathbf{r} \times \frac{d\mathbf{p}}{dt} \\ &= \frac{1}{m}(\mathbf{p} \times \mathbf{p}) + \mathbf{r} \times \mathbf{F} \quad \text{using Newton II} \\ &= \mathbf{r} \times \mathbf{F} \end{aligned} \tag{142}$$

In the penultimate step we recognised that the momentum is mass times velocity, ie. $\mathbf{p} = md\mathbf{r}/dt$; the result is the term with $\mathbf{p} \times \mathbf{p} = 0$ used to obtain the last line. The right hand side is essentially the moment of the force acting on the particle which is known as the **torque $\mathbf{\Gamma}$**,

$$\frac{d\mathbf{L}}{dt} = \mathbf{r} \times \mathbf{F} \equiv \mathbf{\Gamma} \tag{143}$$

In Figure 11 we show this dipole in a uniform magnetic field $\mathbf{B}$, whose direction we choose as the $z-$axis. The magnetic moment experiences a torque $\mathbf{\Gamma}$ given by (see Figure 12):

$$\mathbf{\Gamma} = \boldsymbol{\mathcal{M}} \times \mathbf{B} \tag{144}$$

First let us see how this result is obtained. If we place our dipole at an angle θ to a uniform magnetic field $\mathbf{B}$ we can find the torque exerted by the field from the law exactly like that for electric charges in an electric field: the force on a monopole charge $\pm m$ is $\pm m\mathbf{B}$. Only the components of these forces perpendicular to the dipole contribute to its rotation; the components along the dipole (ie. along $\mathbf{r}$) balance exactly in a uniform field. In our Figure 11 the the forces act anticlockwise on both poles, giving a total torque, in the anticlockwise sense,

$$\Gamma = 2mrB\sin\theta = \mathcal{M}B\sin\theta, \tag{145}$$

where we used $\mathcal{M} = 2mr$. From the figure we notice that this sense of rotation is also the direction of the vector cross product $\boldsymbol{\mathcal{M}} \times \mathbf{B}$; moreover this cross product has the magnitude of the torque since the angle between dipole moment and magnetic field is θ. Hence we arrive at the general result quoted above for the torque, which we now apply by choosing the $z-$axis to lie along the direction of the magnetic field, $\mathbf{B} = (0, 0, B_z)$,

$$\mathbf{\Gamma} = \boldsymbol{\mathcal{M}} \times \mathbf{B} \tag{146}$$

$$= -\gamma\,\mathbf{L} \times \mathbf{B} \qquad \text{for an orbiting } e^- \tag{147}$$

$$= -\gamma\,(L_y, -L_x, 0)B_z \qquad \text{for an orbiting } e^- \tag{148}$$

This torque clearly causes a precession about the $z-$axis. [21] The angular frequency of this precession, the **Larmour frequency,** is

$$\omega_L = \frac{\mu_B}{\hbar}B_z = \gamma B_z \tag{149}$$

[20] I am considering only a single point particle, which is the case of interest to us, but the results hold more generally.

[21] A familiar example of precession is that of a top or gyroscope, where the role of the magnetic field is played by the gravitational field.

Proof 1: This is an elementary proof based on Figure 13 where, in a time dt, the dipole is seen to rotate through an angle $d\varphi = \omega_L dt$ about the $z-$axis. The change in angular momentum dL is then determined by the lever arm, which is the projection of the vector $\mathbf{L}$ pependicular to the $z-$axis, $L\sin\theta$:

$$\begin{aligned} dL &= (\text{ lever arm}) \times (\omega_L dt) \\ &= L\sin\theta\omega_L dt \end{aligned} \tag{150}$$

This gives us an expression for the vectorial rate of change of angular momentum which can be compared directly with that given in equation (148) to give the Larmour frequency in equation (149).

Proof 2: Using Newton's second law for rotational motion,

$$\frac{d\mathbf{L}}{dt} = \mathbf{r}\times\mathbf{F} = \mathbf{\Gamma} \tag{151}$$

we obtain for the $x-$ and $y-$components,

$$\frac{dL_x}{dt} = -\gamma B_z L_y \qquad \frac{d^2L_x}{dt^2} = -\gamma B_z \frac{dL_y}{dt} \tag{152}$$

$$\frac{dL_y}{dt} = +\gamma B_z L_x \qquad \frac{d^2L_y}{dt^2} = +\gamma B_z \frac{dL_x}{dt} \tag{153}$$

where we took the time derivative of each equation to get the equations on the right. Now we can separate the two components of angular momentum by substituting for the right hand sides of these equations from the original equations. We then recognise the result as classical simple harmonic oscillator equations with the Larmour frequency:

$$\frac{d^2L_x}{dt^2} = -(\gamma B_z)^2 L_x = -\omega_L^2 L_x \tag{154}$$

$$\frac{d^2L_y}{dt^2} = -(\gamma B_z)^2 L_y = -\omega_L^2 L_y \tag{155}$$

Thus these two components of angular momentum rotate together about the $z-$axis with angular frequency ω_L, while the $z-$component remains constant. A quantum mechanical treatment given in Appendix D yields the same result, provided we replace the classical angular momentum components L_x and L_y by their corresponding quantum mechanical expectation values $\langle\widehat{L}_x\rangle$ and $\langle\widehat{L}_y\rangle$. These results will be used later when we discuss Electron Spin Resonance (ESR) and Nuclear Magnetic Resonance (NMR).

APPENDIX D: QM TREATMENT OF SPIN PRECESSION IN A MAGNETIC FIELD.

We have seen that a classical dipole in a magnetic field precesses with the Larmour frequency, and that the same applies to the QM case of a magnetic moment produced by orbital motion. It may not be so obvious that the same holds for a magnetic moment produced by spin, especially as electron spin is so difficult to understand classically. Here we show how to study the problem using QM, while at the same time illustrating the role of the Schrödinger equation in spin dynamics and showing that it is the quantum mechanical expectation values that precess. In an applied magnetic field the spin state χ is altered by the fact that the energy changes. This adds a term to the Hamiltonian which acts only on the spin part of the wave function:

$$\widehat{H}_s = -\boldsymbol{\mathcal{M}_s}.\mathbf{B} = \frac{g_s \mu_B B_z}{\hbar}\widehat{S}_z = \omega_s \frac{\hbar}{2}\sigma_z \tag{156}$$

The spin part of the wave function $\chi(t)$ is obtained by demanding it satisfy an appropriate TDSE which takes account of the above interaction, and which induces a time-dependence:

$$\chi(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} \qquad \text{where} \qquad \widehat{H}_s\chi(t) = i\hbar\frac{\partial\chi(t)}{\partial t} \tag{157}$$

$$\text{ie.} \quad \omega_s\frac{\hbar}{2}\sigma_z\chi(t) = \omega_s\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = i\hbar\begin{pmatrix} \dot{c}_1 \\ \dot{c}_2 \end{pmatrix} \tag{158}$$

$$\text{Hence,} \quad \omega_s\frac{\hbar}{2}\begin{pmatrix} c_1 \\ -c_2 \end{pmatrix} = i\hbar\begin{pmatrix} \dot{c}_1 \\ \dot{c}_2 \end{pmatrix} \tag{159}$$

$$\text{ie.} \quad \dot{c}_1 = -i\frac{\omega_s}{2}c_1 \qquad \text{and} \qquad \dot{c}_2 = +i\frac{\omega_s}{2}c_2 \tag{160}$$

$$\text{Hence the solutions,} \qquad c_1(t) = ae^{-i\omega_s t/2} \qquad \text{and,} \qquad c_2(t) = be^{+i\omega_s t/2} \qquad a^2+b^2 = 1 \tag{161}$$

where the constants a and b are chosen to normalise the wave function and to satisfy the initial ($t = 0$) conditions. Two cases illustrate the relation between the classical and quantum pictures:
Case (1) Suppose the wave function is initially in the spin up eigenstate χ_+ of S_z,

$$c_1(t=0) = 1, c_2(t=0) = 0, \quad \text{hence} \quad a = 1,\ b = 0 \quad \text{and} \quad \chi(t) = \begin{pmatrix} e^{-i\omega_s t/2} \\ 0 \end{pmatrix} \tag{162}$$

Now, this remains an eigenstate of S_z, with eigenvalue $+\hbar/2$, for all times; all its expectation values are time-independent and so it is a stationary state with no precession:

$$\begin{aligned} \langle\widehat{S}_z\rangle = \chi^\dagger(t)S_z\chi(t) &= \frac{\hbar}{2}\chi^\dagger\sigma_z\chi \\ &= \frac{\hbar}{2}\begin{pmatrix} e^{+i\omega_s t/2} & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} e^{-i\omega_s t/2} \\ 0 \end{pmatrix} \\ &= \frac{\hbar}{2}\begin{pmatrix} e^{+i\omega_s t/2} & 0 \end{pmatrix}\begin{pmatrix} e^{-i\omega_s t/2} \\ 0 \end{pmatrix} \\ \langle\widehat{S}_z\rangle &= \frac{\hbar}{2} \end{aligned} \tag{163}$$

$$\begin{aligned}
\langle \widehat{S}_x \rangle = \chi^\dagger(t) S_x \chi(t) &= \frac{\hbar}{2} \chi^\dagger \sigma_x \chi \\
&= \frac{\hbar}{2} \begin{pmatrix} e^{+i\omega_s t/2} & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} e^{-i\omega_s t/2} \\ 0 \end{pmatrix} \\
&= \frac{\hbar}{2} \begin{pmatrix} e^{+i\omega_s t/2} & 0 \end{pmatrix} \begin{pmatrix} 0 \\ e^{-i\omega_s t/2} \end{pmatrix} \\
\langle \widehat{S}_x \rangle &= 0
\end{aligned} \tag{164}$$

Similarly $\langle \widehat{S}_y \rangle = 0$. Thus we have the perhaps surprising outcome that a pure spin eigenstate does not precess at all, despite the fact that in the quasi-classical pictures of the quantum rules the spin vector doesn't point exactly in the $z-$direction. However the expectation values correspond precisely to the classical values: if the classical dipole is exactly aligned along the $z-$axis (allowed classically) then not only does the spin have zero projection in the $x-$ and $y-$directions, but there is no torque acting and therefore no precession.

Case (2) Suppose the wave function is not a pure eigenstate but initially an equal mixture of spin up and spin down:

$$c_1(t=0) = c_2(t=0) = \frac{1}{\sqrt{2}}, \quad \text{hence} \quad a = b = \frac{1}{\sqrt{2}}, \quad \text{and} \quad \chi(t) = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\omega_s t/2} \\ e^{+i\omega_s t/2} \end{pmatrix} \tag{165}$$

Since this state is an equal mixture of spin up and spin down it is not surprising to discover that its average projection on the $z-$axis is zero:

$$\begin{aligned}
\langle \widehat{S}_z \rangle = \chi^\dagger(t) S_z \chi(t) &= \frac{\hbar}{2} \chi^\dagger \sigma_z \chi \\
&= \frac{\hbar}{2} \begin{pmatrix} e^{+i\omega_s t/2} & e^{-i\omega_s t/2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} e^{-i\omega_s t/2} \\ e^{+i\omega_s t/2} \end{pmatrix} \\
&= \frac{\hbar}{2} \begin{pmatrix} e^{+i\omega_s t/2} & e^{-i\omega_s t/2} \end{pmatrix} \begin{pmatrix} e^{-i\omega_s t/2} \\ -e^{+i\omega_s t/2} \end{pmatrix} \\
\langle \widehat{S}_z \rangle &= 0
\end{aligned} \tag{166}$$

$$\begin{aligned}
\langle \widehat{S}_x \rangle = \chi^\dagger(t) S_x \chi(t) &= \frac{\hbar}{2} \chi^\dagger \sigma_x \chi \\
&= \frac{\hbar}{4} \begin{pmatrix} e^{+i\omega_s t/2} & e^{-i\omega_s t/2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} e^{-i\omega_s t/2} \\ e^{+i\omega_s t/2} \end{pmatrix} \\
&= \frac{\hbar}{4} \begin{pmatrix} e^{+i\omega_s t/2} & e^{-i\omega_s t/2} \end{pmatrix} \begin{pmatrix} e^{+i\omega_s t/2} \\ e^{-i\omega_s t/2} \end{pmatrix} \\
&= \frac{\hbar}{4} \left(e^{+2i\omega_s t/2} + e^{-2i\omega_s t/2} \right) \\
\langle \widehat{S}_x \rangle &= \frac{\hbar}{2} \cos \omega_s t
\end{aligned} \tag{167}$$

Similarly,

$$\langle \widehat{S}_y \rangle = \frac{\hbar}{2} \sin \omega_s t \tag{168}$$

Here we have the analogue of the classical situation where the dipole lies in the $x-y$ plane with zero $z-$component: the spin vector will then precess around the $z-$axis. The QM example shows that indeed the expectation values have precisely these properties; moreover we now see that the QM precession frequency is $\omega_s \approx 2\omega_L$ for an electron.