1 Very brief background

In 1922, a famous experiment conducted by Otto Stern and Walther Gerlach, involving particles subject to a nonuniform magnetic field, identified two surprising properties of the angular momentum: 1) the components of the angular momentum are quantized, i.e. they can only take certain discrete values; 2) there exists a type of angular momentum that is intrinsic to a particle, even a point particle, and that cannot be put in the form $\mathbf{r} \times \mathbf{p}$: this is the so-called "spin angular momentum". The total angular momentum $\mathbf{J}$ is the sum of the orbital angular momentum $\mathbf{L}$ and the spin angular momentum $\mathbf{S}$: $\mathbf{J} = \mathbf{L} + \mathbf{S}$. In this lecture, we will start from standard postulates for the angular momenta to derive the key characteristics highlighted by the Stern-Gerlach experiment.

2 General properties of angular momentum operators

2.1 Commutation relations between angular momentum operators

Let us first consider the orbital angular momentum $\mathbf{L}$ of a particle with position $\mathbf{r}$ and momentum $\mathbf{p}$. In classical mechanics, $\mathbf{L}$ is given by

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

so by the correspondence principle, the associated operator is

$$\widehat{L} = \frac{\hbar}{i}\mathbf{r} \times \nabla$$

The operator for each components of the orbital angular momentum thus are

$$\begin{cases} \widehat{L}_x & = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y = \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \\ \widehat{L}_y & = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z = \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \\ \widehat{L}_z & = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \end{cases}$$

We also define the operator

$$\widehat{L}^2 = \widehat{L}_x^2 + \widehat{L}_y^2 + \widehat{L}_z^2$$

Let us start with the commutation relation between $\widehat{L}_x$ and $\widehat{L}_y$:

$$\begin{aligned} [\widehat{L}_x, \widehat{L}_y] &= [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}p_z] = (\hat{y}\hat{p}_z - \hat{z}\hat{p}_y)(\hat{z}\hat{p}_x - \hat{x}p_z) - (\hat{z}\hat{p}_x - \hat{x}p_z)(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y) \\ &= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] - [\hat{z}\hat{p}_y, \hat{z}\hat{p}_x] - [\hat{y}\hat{p}_z, \hat{x}\hat{p}_z] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] \\ &= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] = \hat{y}\hat{p}_z[\hat{p}_z, \hat{z}] + \hat{x}\hat{p}_y[\hat{z}, \hat{p}_z] \end{aligned}$$

Now, we know that $[\hat{z}, \hat{p}_z] = i\hbar$, so we conclude that

$$[\widehat{L}_x, \widehat{L}_y] = i\hbar(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) = i\hbar\widehat{L}_z$$

By cyclical permutations one easily obtain the other relations:

$$[\widehat{L}_y, \widehat{L}_z] = i\hbar\widehat{L}_x \qquad [\widehat{L}_z, \widehat{L}_x] = i\hbar\widehat{L}_y$$

We note a very important results: the three components of the orbital angular momentum are not compatible with one another, and have associated uncertainty relations.

The operator $\widehat{L}^2$, on the other hand, commutes with $\widehat{L}_x$, $\widehat{L}_y$ and $\widehat{L}_z$. Indeed,

$$[\widehat{L}^2, \widehat{L}_z] = [\widehat{L}_x^2, \widehat{L}_z] + [\widehat{L}_y^2, \widehat{L}_z]$$

and

$$[\widehat{L}_x^2, \widehat{L}_z] = \widehat{L}_x\widehat{L}_x\widehat{L}_z - \widehat{L}_z\widehat{L}_x\widehat{L}_x = \widehat{L}_x\widehat{L}_z\widehat{L}_x - i\hbar\widehat{L}_x\widehat{L}_y - \widehat{L}_x\widehat{L}_z\widehat{L}_x - i\hbar\widehat{L}_y\widehat{L}_x = -i\hbar(\widehat{L}_x\widehat{L}_y + \widehat{L}_y\widehat{L}_x)$$

Repeating the same calculation for $[\widehat{L}_y^2, \widehat{L}_z]$, we would find

$$[\widehat{L}_y^2, \widehat{L}_z] = i\hbar(\widehat{L}_x\widehat{L}_y + \widehat{L}_y\widehat{L}_x)$$

so that finally

$$[\widehat{L}^2, \widehat{L}_z] = 0$$

and in exactly the same way, we would prove

$$[\widehat{L}^2, \widehat{L}_x] = 0 \qquad [\widehat{L}^2, \widehat{L}_y] = 0$$

A postulate of quantum mechanics is that all types of angular momentum operator $\widehat{J}$, orbital or spin, satisfy the following commutation relations:

$$[\widehat{J}^2, \widehat{J}_x] = 0 \qquad [\widehat{J}^2, \widehat{J}_y] = 0 \qquad [\widehat{J}^2, \widehat{J}_z] = 0 \tag{1}$$

$$[\widehat{J}_x, \widehat{J}_y] = i\hbar\widehat{J}_z \qquad [\widehat{J}_y, \widehat{J}_z] = i\hbar\widehat{J}_x \qquad [\widehat{J}_z, \widehat{J}_x] = i\hbar\widehat{J}_y \tag{2}$$

We will now take this relations as a starting point, and derive general properties of any angular momentum operator $\widehat{J}$ that satisfies these properties.

2.2 Eigenvalues of the operators $\widehat{J}^2$ and $\widehat{J}_z$

We take the commutation relations given by Eq. (1) and Eq. (2) as our postulate, and show that alone they allow us to prove that the eigenvalues or $\widehat{J}^2$ and $\widehat{J}_z$ are quantized. Since $\widehat{J}^2$ and $\widehat{J}_z$ commute, there exists a basis of eigenvectors that are common to these two operators. Let us call $|a, b\rangle$ an eigenstate of both $\widehat{J}^2$, with eigenvalue $\hbar^2 a$, and of $\widehat{J}_z$, with eigenvalue $\hbar b$. The factors $\hbar^2$ and $\hbar$ appear because we have normalized the eigenvalues so that a and b are dimensionless numbers. We thus have

$$\widehat{J}^2|a, b\rangle = \hbar^2 a|a, b\rangle \qquad \widehat{J}_z|a, b\rangle = \hbar b|a, b\rangle$$

We also have the additional normalization condition

$$\langle a, b|a, b\rangle = 1$$

Let us now construct the operators

$$\widehat{J}^+ = \widehat{J}_x + i\widehat{J}_y \qquad \widehat{J}^- = \widehat{J}_x - i\widehat{J}_y$$

Note that $\widehat{J}^+$ and $\widehat{J}^-$ are not Hermitian, but Hermitian conjugates of one another: $(\widehat{J}^+)^\dagger = \widehat{J}^-$. We will now see what happens when one applies $\widehat{J}+$ and $\widehat{J}-$ to the state $|a, b\rangle$.

Since $\widehat{J}^2$ commutes with $\widehat{J}_x$ and $\widehat{J}_y$, we can write

$$\widehat{J}^2\left(\widehat{J}^\pm|a, b\rangle\right) = \widehat{J}^\pm\left(\widehat{J}^2|a, b\rangle\right) = \widehat{J}\pm\left(\hbar^2 a|a, b\rangle\right) = \hbar^2 a\widehat{J}^\pm|a, b\rangle$$

where $\widehat{J}^\pm$ stands for either $\widehat{J}^+$ or $\widehat{J}^-$. We see that $\widehat{J}^\pm|a, b\rangle$ is an eigenvector of $\widehat{J}^2$ with eigenvalue $\hbar^2 a$.

Likewise, using the commutation relations

$$[\widehat{J}_z, \widehat{J}^+] = [\widehat{J}_z, \widehat{J}_x + i\widehat{J}_y] = i\hbar\widehat{J}_y + i(-i\hbar)\widehat{J}_x = \hbar\widehat{J}^+$$

we find

$$\widehat{J}_z\widehat{J}^+|a, b\rangle = \widehat{J}^+\widehat{J}_z|a, b\rangle + \hbar\widehat{J}^+|a, b\rangle = \hbar(b+1)|a, b\rangle$$

Following the same procedure, we could also show that

$$\widehat{J}_z\widehat{J}^-|a, b\rangle = \hbar(b-1)|a, b\rangle$$

We see that $\widehat{J}^+|a, b\rangle$ is an eigenvector of $\widehat{J}_z$ with eigenvalue $\hbar(b+1)$, and $\widehat{J}^-|a, b\rangle$ is an eigenvector of $\widehat{J}_z$ with eigenvalue $\hbar(b-1)$. For these reasons, $\widehat{J}^+$ and $\widehat{J}^-$ are sometimes called *ladder operators*.

We just showed that $\widehat{J}^+|a,b\rangle$ is colinear with the normalized eigenstate $|a,b+1\rangle$, and that $\widehat{J}^-|a,b\rangle$ is colinear with the normalized eigenstate $|a,b-1\rangle$. There exist complex numbers c_+ and c_- such that

$$\widehat{J}^+|a,b\rangle = c_+|a,b+1\rangle$$
$$\widehat{J}^-|a,b\rangle = c_-|a,b-1\rangle$$

Since $\widehat{J}^-$ is the Hermitian conjugate of $\widehat{J}^+$, the square of the norm of the ket of $\widehat{J}^+|a,b\rangle$ is

$$|c_+|^2 = \left(\langle a,b|\widehat{J}^-\right)\left(\widehat{J}^+|a,b\rangle\right) = \langle a,b|\widehat{J}^-\widehat{J}^+|a,b\rangle$$

We can easily compute

$$\widehat{J}^-\widehat{J}^+ = \widehat{J}_x^2 + i\left[\widehat{J}_x,\widehat{J}_y\right] + \widehat{J}_y^2 = \widehat{J}^2 - \widehat{J}_z^2 - \hbar\widehat{J}_z$$

Thus

$$|c_+|^2 = \hbar^2\left[a - b(b+1)\right]$$

In the same way, you can convince yourself that

$$|c_-|^2 = \hbar^2\left[a - b(b-1)\right]$$

Since $|c_+|^2$ and $|c_-|^2$ are positive quantities, we have restrictions on the allowable values for the eigenvalues a and b. Indeed, if for a given value of a, any value of b would be allowed, then by applying $\widehat{J}^+$ or $\widehat{J}^-$ to $|a,b\rangle$ multiple times, we would keep increasing the products $b(b+1)$ until $|c_+|^2$ becomes negative. The only way to avoid this contradiction is by saying that b must be restricted to a finite interval. Indeed, the only way of stopping the iterative process is to find $|c_+| = 0$ and $c_-| = 0$ at some point. There must therefore exist:

1. A maximum value b_{max} of b such that $a = b_{max}(b_{max}+1)$
2. A minimum value b_{min} of b such that $a = b_{min}(b_{min}-1)$

The ket $\widehat{J}^+|a,b_{max}\rangle$ then is 0, as is the ket $\widehat{J}^-|a,b_{min}\rangle$. We can write

$$b_{max}(b_{max}+1) = b_{min}(b_{min}-1)$$

The only solution of this equation with $b_{max} > b_{min}$ is $b_{max} = -b_{min} > 0$. And since one can go from the eigenvalue b_{min} to b_{max} by steps of size 1 through the iterative application of $\widehat{J}^+$, we have the additional condition

$$b_{max} = b_{min} + k$$

where $k \in \mathbb{N}$, we conclude:

$$b_{max} = \frac{k}{2} = -b_{min} \qquad k \in \mathbb{N}$$

Let us define $j \equiv k/2$, where j is either an integer or a half-integer. We then have $a = j(j+1)$, and b can vary from $-j$ to j in steps of size 1. We can summarize this as the following important result:

- The eigenvalues of $\widehat{J}^2$ are of the form $\hbar^2 j(j+1)$ with j positive integer or half integer
- For a fixed j, the eigenvalues of $\widehat{J}_z$ can be written in the form $\hbar m$, where m can take the following $(2j+1)$ values:
$$-j, -j+1, -j+2, \ldots, j-2, j-1, j$$
- If j is an integer, there is an odd number of eigenvalues of $\widehat{J}_z$ for that j
- If j is a half-integer, there is an even number of eigenvalues of $\widehat{J}_z$ for that j

Note that the eigenstates of $\widehat{J}^2$ with eigenvalue $\hbar^2 j(j+1)$ belong to a subspace of dimension at least $2j+1$. Indeed, we just found $2j+1$ of them that were also eigenstates of $\widehat{J}_z$ with distinct eigenvalues; they are therefore orthogonal.

In order to motivate the next section, let us talk some more about the Stern-Gerlach experiment. It found that angular momentum was quantized, which can be seen as a consequence of the commutation relations (1) and (2). At this point, it could still be the case that angular momentum only consists of $\mathbf{L}$, the orbital angular momentum, since $\widehat{\mathbf{L}}$ does satisfy (1) and (2). However, the Stern-Gerlach found a second important property of angular momentum: it can be such that j is half-integer. We will now show that this impossible if the angular momentum is only made of $\mathbf{L}$, for reasons that have not yet been mentioned.

3 Orbital angular momentum

3.1 Orbital angular momentum in spherical coordinates

We use here the usual spherical coordinate system (ρ, θ, φ), with associated basis vectors $\mathbf{e}_\rho$, $\mathbf{e}_\theta$, and $\mathbf{e}_\varphi$, where θ is the colatitude and φ the azimuth. In these coordinates, the orbital angular momentum operator is

$$\widehat{\mathbf{L}} = \rho \mathbf{e}_\rho \times \frac{\hbar}{i}\nabla = \frac{\hbar}{i}\left(\mathbf{e}_\varphi \frac{\partial}{\partial\theta} - \frac{\mathbf{e}_\theta}{\sin\theta}\frac{\partial}{\partial\varphi}\right)$$

The Cartesian components of $\widehat{\mathbf{L}}$ are

$$\widehat{L}_x = \frac{\hbar}{i}\left(-\sin\varphi\frac{\partial}{\partial\theta} - \cot\theta\cos\varphi\frac{\partial}{\partial\varphi}\right) \qquad \widehat{L}_y = \frac{\hbar}{i}\left(\cos\varphi\frac{\partial}{\partial\theta} - \cot\theta\sin\varphi\frac{\partial}{\partial\varphi}\right) \qquad \widehat{L}_z = \frac{\hbar}{i}\frac{\partial}{\partial\varphi}$$

For the sake of completeness, one may also calculate the expression for $\widehat{L}^2$. This is most conveniently done by applying $\widehat{L}^2$ to an arbitrary function $\Psi(\mathbf{r})$:

$$\begin{aligned}\widehat{L}^2\Psi &= -\hbar^2\left(\mathbf{e}_\varphi\frac{\partial}{\partial\theta} - \frac{\mathbf{e}_\theta}{\sin\theta}\frac{\partial}{\partial\varphi}\right)\left(\mathbf{e}_\varphi\frac{\partial\Psi}{\partial\theta} - \frac{\mathbf{e}_\theta}{\sin\theta}\frac{\partial\Psi}{\partial\varphi}\right)\\ &= -\hbar^2\left[\mathbf{e}_\varphi\frac{\partial}{\partial\theta}\left(\mathbf{e}_\varphi\frac{\partial\Psi}{\partial\theta}\right) - \mathbf{e}_\varphi\frac{\partial}{\partial\theta}\left(\frac{\mathbf{e}_\theta}{\sin\theta}\frac{\partial\Psi}{\partial\varphi}\right) - \frac{\mathbf{e}_\theta}{\sin\theta}\frac{\partial}{\partial\varphi}\left(\mathbf{e}_\varphi\frac{\partial\Psi}{\partial\theta}\right) + \frac{\mathbf{e}_\theta}{\sin\theta}\frac{\partial}{\partial\varphi}\left(\frac{\mathbf{e}_\theta}{\sin\theta}\frac{\partial\Psi}{\partial\varphi}\right)\right]\end{aligned}$$

Now, since $\partial\mathbf{e}_\theta/\partial\theta = -\mathbf{e}_\rho$, $\partial\mathbf{e}_\theta/\partial\varphi = \mathbf{e}_\varphi\cos\theta$, $\partial\mathbf{e}_\varphi/\partial\theta = 0$, $\partial\mathbf{e}_\varphi/\partial\varphi = -(\mathbf{e}_\rho\sin\theta + \mathbf{e}_\theta\cos\theta)$, this becomes

$$\widehat{L}^2\Psi = -\hbar^2\left(\frac{\partial^2\Psi}{\partial\theta^2} + \cot\theta\frac{\partial\Psi}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2\Psi}{\partial\varphi^2}\right)$$

In other words,

$$\widehat{L}^2 = -\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right) \tag{3}$$

Lastly, it can be convenient to use the variable $u \equiv \cos\theta$ instead of θ. It is easy to show that in the u, φ variables, $\widehat{L}^2$ is

$$\widehat{L}^2 = -\hbar^2\left[(1-u^2)\frac{\partial^2}{\partial u^2} - 2u\frac{\partial}{\partial u} + \frac{1}{1-u^2}\frac{\partial^2}{\partial\varphi^2}\right] \tag{4}$$

3.2 Eigenvalues and eigenfunctions of the orbital angular momenta

Eigenfunctions of $\widehat{L}_z$

Let us solve for the eigenfunction of the operator $\widehat{L}_z$ associated with the eigenvalue $\hbar b$, which we call $Y_b(\rho, \theta, \varphi)$. By definition, we have

$$\begin{aligned} \widehat{L}_z Y_b &= \hbar b Y_b\\ \Leftrightarrow\ \frac{\hbar}{i}\frac{\partial Y_b}{\partial\varphi} &= \hbar b Y_b\\ \Leftrightarrow\ \frac{1}{Y_b}\frac{\partial Y_b}{\partial\varphi} &= ib\end{aligned}$$

We solve this equation by separation of the variables: we write $Y_b(\rho, \theta, \varphi) = P(\rho, \theta)e^{ib\varphi}$, where P is an arbitrary function of ρ and θ (at this point). Now, here is a crucial point that we did not highlight enough in the previous lectures: wavefunctions in $\mathbf{r}$ representation must be single-valued function of the space coordinates, so that we can define their Fourier transforms. For Y_b to be a single-valued function, we must have

$$Y_b(\rho, \theta, \varphi) = Y_b(\rho, \theta, \varphi + 2\pi)$$

which means that b is an integer. We write $b = m \in \mathbb{N}$, and we can say that the eigenvalues of $\widehat{L}_z$ are $\hbar m$ with $m \in \mathbb{N}$. This result does not contradict the general results of Section 2.2, but adds an additional constraint regarding orbital angular momenta: half-integer values of m and therefore j are not allowed.

In conclusion, the eigenvalues of $\widehat{L}^2$ are of the form $\hbar^2 l(l+1)$ with $l \in \mathbb{N}$, and the eigenvalues $\widehat{L}_z$ for a given l take all the integer values from $-l$ to $+l$.

Common eigenfunctions of $\widehat{L}^2$ and $\widehat{L}_z$

We call Y_l^m the eigenfunction that is common to $\widehat{L}^2$ and $\widehat{L}_z$ with respective eigenvalues $\hbar^2 l(l+1)$ and $\hbar m$. We already know that $l \in \mathbb{N}$ and that m can take integer values between $-l$ and $+l$. We also saw that since Y_l^m is an eigenfunction of $\widehat{L}_z$, we can write

$$Y_l^m = P_l^m(\rho, \cos\theta) e^{im\varphi}$$

Note that in principle P_l^m is a function of ρ and θ, but it is more convenient to use $\cos\theta$ as a variable instead of θ, as we will now see.

Since the variable ρ does not appear in $\widehat{L}^2$, let us ignore it in the notation. Following Eq.(4), the eigenvalue equation for $\widehat{L}^2$ is

$$(1-u^2)\frac{d^2 P_l^m(u)}{du^2} - 2u\frac{dP_l^m(u)}{du} - \left[l(l+1) - \frac{m^2}{1-u^2}\right] P_l^m(u) = 0$$

All we have to do is solve the equation above for $m = 0$, because we can then use the ladder operators $\widehat{L}^{\pm} = \widehat{L}_x \pm i\widehat{L}_y$ to go from Y_l^0 to Y_l^m for all desired m values. Writing $P_l(u) = P_l^0(u)$, $P_l(u)$ satisfies

$$(1-u^2)P_l''(u) - 2uP_l'(u) + l(l+1)P_l(u) = 0 \tag{5}$$

Eq. (5) is the *Legendre differential equation.* The only solutions that do not have singularities at $u = \pm 1$ are polynomials called *Legendre polynomials*, which we already encountered in this course.

Note that the eigenfunctions Y_l^m presented here, to within a normalization that we will not discuss here, are called *spherical harmonics.*

4 Spin angular momentum

The Stern-Gerlach experiments showed that there are atoms and particles for which the angular momentum J_z can only take two values. We therefore are in the case for which $j = 1/2$, which is the smallest allowable angular momentum. We saw that this angular momentum could not be put in the usual form $\mathbf{r} \times \mathbf{p}$, and does not have an equivalent in classical mechanics. It is an intrinsic angular momentum that physicists have decided to call spin. At this point, we know plenty of its mathematical properties to investigate it in detail: it satisfies the commutation relations (1) and (2), and is such that $j = 1/2$.

4.1 Representation of the states of a particle with spin 1/2

Electrons, protons, and neutrons, and quarks inside the protons and neutrons, are called *fermions*, meaning that their spin angular momentum is a half-integer. More precisely, let $\mathbf{S}$ be the spin angular momentum; fermions are such that they are always in an eigenstate of $\widehat{S}^2$ with eigenvalue $\hbar^2 \times 1/2 \times (1/2+1) = 3/4\hbar^2$. We saw that the state of a particle without spin was given by the wavefunction $\Psi(\mathbf{r})$. For a particle with spin, one adds $\widehat{S}^2$ and $\widehat{S}_z$ to $\widehat{x}$, $\widehat{y}$, and $\widehat{z}$ to form to complete set of observables. In fact, we know that for fermions the measurement of S^2 always leads to $3/4\hbar^2$, so only $\widehat{S}_z$ provides new information. A complete measurement leads to simultaneous knowledge of the coordinates x, y, and z, as well as S_z, which can only take the two values $s_z = \pm\frac{\hbar}{2}$. We thus define a new probability amplitude $\Psi(\mathbf{r}, s_z)$ such that the probability of finding the particle in the volume $d\mathbf{r}$ around $\mathbf{r}_0$ with the value s_z for the component S_z is given by $|\Psi(\mathbf{r}_0, s_z)|^2$.

While acceptable, this notation treats continuous variables (the coordinates) on an equal footing with the discrete variable s_z. For this reason, one will preferably adopt the following notation:

$$\Psi_+(\mathbf{r}) \equiv \Psi(\mathbf{r}, +\frac{\hbar}{2}) \qquad \Psi_-(\mathbf{r}) \equiv \Psi(\mathbf{r}, -\frac{\hbar}{2})$$

A general wavefunction thus has two components, which we write in the form of a column vector:

$$\Psi(\mathbf{r}) = \begin{pmatrix} \Psi_+(\mathbf{r}) \\ \Psi_-(\mathbf{r}) \end{pmatrix} = \Psi_+(\mathbf{r}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \Psi_-(\mathbf{r}) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6}$$

In the sum above, each term is the product of an ordinary wavefunction ($\Psi_+(\mathbf{r})$ or $\Psi_-(\mathbf{r})$), and a column vector ($\chi_+ = (1,\ 0)^T$ or $\chi_- = (0,\ 1)^T$) called a *spinor*. Spin operators such as $\widehat{S}_x$, $\widehat{S}_y$ or $\widehat{S}_z$ only act on spinors and in this representation can be seen as 2×2 matrices. Since our spinors are eigenstates of $\widehat{S}_z$, we write

$$\widehat{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7}$$

Note that spinors have played an important role in geometry, and were discovered and studied by the French mathematician Elie Cartan much before their usefulness in quantum mechanics was understood.

4.2 Pauli's spin matrices

We now work in the subspace $\mathcal{S}$ of the eigenstates of $\widehat{S}^2$. The dimension of $\mathcal{S}$ is 2. We choose a representation that uses the basis of eigenstates of $\widehat{S}_z$:

$$|z_+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |z_-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We already know the expression for $\widehat{S}_z$ in this representation, given by Eq. (7). What is the expression for $\widehat{S}_x$ and $\widehat{S}_y$? This is what we derive now.

Note first that with a rotation, one can make the z-axis coincide with the x-axis or with the y-axis. This implies that the matrices $\widehat{S}_x$, $\widehat{S}_y$ and $\widehat{S}_z$ have the same trace and the same determinant. From $\widehat{S}_z$, we thus know that the trace of the three matrices is 0 and the determinant is $\hbar^2/4$.

For the simplicity of the notation, let us define the matrices $\widehat{\sigma}_x$, $\widehat{\sigma}_y$ and $\widehat{\sigma}_z$ such that

$$\widehat{S}_x \equiv \frac{\hbar}{2}\widehat{\sigma}_x \qquad \widehat{S}_y \equiv \frac{\hbar}{2}\widehat{\sigma}_y \qquad \widehat{S}_z \equiv \frac{\hbar}{2}\widehat{\sigma}_z$$

$\widehat{\sigma}_x$, $\widehat{\sigma}_y$ and $\widehat{\sigma}_z$ are called *Pauli's matrices*, and are subject, as we have seen, to the following 3 constraints:

1. They are Hermitian matrices. This is because of the components of the spin angular momentum are physical observables. The most general form for Pauli's matrices must therefore be

$$\begin{pmatrix} a & c - id \\ c + id & b \end{pmatrix} \tag{8}$$

   where a, b, c, and d are real numbers.

2. The trace of Pauli's matrices is zero, so that $a = -b$. Any Pauli matrix can therefore be written as the sum

$$a\widehat{\sigma}_3 + c\widehat{\sigma}_1 + d\widehat{\sigma}_2$$

   with

$$\widehat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \widehat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \widehat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

   Note that $\widehat{\sigma}_3 = \widehat{\sigma}_z$

3. The determinant of Pauli's matrices is -1, so that $a^2 + c^2 + d^2 = 1$

We need one more ingredient to fully characterize $\widehat{S}_x$ and $\widehat{S}_y$, namely the commutation relations

$$[\widehat{\sigma}_1, \widehat{\sigma}_2] = 2i\widehat{\sigma}_3 \qquad [\widehat{\sigma}_2, \widehat{\sigma}_3] = 2i\widehat{\sigma}_1 \qquad [\widehat{\sigma}_3, \widehat{\sigma}_1] = 2i\widehat{\sigma}_2$$

Let us now write the most general form for $\widehat{S}_x$:

$$\widehat{S}_x = \frac{\hbar}{2}(c\widehat{\sigma}_1 + d\widehat{\sigma}_2 + a\widehat{\sigma}_3) \qquad \text{with } a^2 + c^2 + d^2 = 1$$

We know from Eq. (2) that $[\widehat{S}_x, \widehat{S}_z] = -i\hbar\widehat{S}_y$. Plugging our expressions for $\widehat{S}_x$ and $\widehat{S}_z$ in terms of $\widehat{\sigma}_1$, $\widehat{\sigma}_2$, and $\widehat{\sigma}_3$ in this commutation relation, we find

$$\widehat{S}_y = \frac{\hbar}{2}(c\widehat{\sigma}_2 - d\widehat{\sigma}_1)$$

We then use this form for $\widehat{S}_y$ into $[\widehat{S}_y, \widehat{S}_z] = i\hbar \widehat{S}_x$ to find

$$\widehat{S}_x = \frac{\hbar}{2}(c\widehat{\sigma}_1 + d\widehat{\sigma}_2)$$

Hence, $a = 0$, so $c^2 + d^2 = 1$ and we can write

$$\widehat{S}_x = \frac{\hbar}{2}(\cos\alpha\widehat{\sigma}_1 - \sin\alpha\widehat{\sigma}_2)$$

for some real number α. We then also have

$$\widehat{S}_y = \frac{\hbar}{2}(\sin\alpha\widehat{\sigma}_1 + \cos\alpha\widehat{\sigma}_2)$$

For any angle α, the matrices above satisfy the appropriate commutation relation. At this point, we have fixed the z-axis, but are still free to rotate the x and y axes by an angle $-\alpha$ so that $\widehat{\sigma}_x = \widehat{\sigma}_1$ and $\widehat{\sigma}_y = \widehat{\sigma}_2$. With this choice of axes, we have

$$\widehat{S}_x = \frac{\hbar}{2}\widehat{\sigma}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \widehat{S}_y = \frac{\hbar}{2}\widehat{\sigma}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \widehat{S}_z = \frac{\hbar}{2}\widehat{\sigma}_3 = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The eigenvectors of $\widehat{S}_x$ and $\widehat{S}_y$ in this representation are

$$|x^+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$|x^-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
$$|y^+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$$
$$|y^-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}$$

and they can be written as follows in the basis of eigenvectors of $\widehat{S}_z$:

$$|x^+\rangle = \frac{1}{\sqrt{2}}\left(|z^+\rangle + |z^-\rangle\right)$$
$$|x^-\rangle = \frac{1}{\sqrt{2}}\left(|z^+\rangle - |z^-\rangle\right)$$
$$|y^+\rangle = \frac{1}{\sqrt{2}}\left(|z^+\rangle + i|z^-\rangle\right)$$
$$|y^-\rangle = \frac{1}{\sqrt{2}}\left(|z^+\rangle - i|z^-\rangle\right)$$