Chapter 2

States of spin one-half

The Stern-Gerlach experiment demonstrated that the magnetic moment of the electron takes quantized values. By inference one is led to the concept of spin angular momentum and the idea that the electron is a spin one-half particle. We introduce the appropriate two-dimensional complex vector space of states and operators $\hat{S}_x, \hat{S}_y$, and $\hat{S}_z$ that satisfy the algebra of angular momentum. We explicitly construct spin operators pointing in arbitrary directions and the associated spin states.

2.1 The Stern-Gerlach Experiment

In 1922, at the University of Frankfurt in Germany, Otto Stern and Walther Gerlach conducted fundamental experiments to measure the deflection of beams of silver atoms as they were sent through inhomogeneous magnetic fields. Rather than finding the expected continuous range of deflections, the incoming beam was split into two beams as deduced from the two separate spots on the target screen. These experiments demonstrated that these silver atoms have quantized magnetic moments that can have one of two values.

Although consistent with the idea that the electron had spin, this suggestion took a few more years to develop. Pauli introduced a "two-valued" degree of freedom for electrons, without suggesting a physical interpretation. Kronig suggested in 1925 that this degree of freedom originated from the self-rotation of the electron. This idea was severely criticized by Pauli, and Kronig did not publish it. In the same year Uhlenbeck and Goudsmit had a similar idea, and Ehrenfest encouraged them to publish it. They are now credited with the discovery that the electron has an intrinsic spin with value "one-half". Much of the mathematics of spin one-half was developed by Pauli himself in 1927. In fact, it took until 1927 for anyone to realize that the Stern-Gerlach experiment measured the magnetic moment of the electron.

A current on a closed loop induces a magnetic dipole moment. The magnetic moment

vector $\vec{\mu}$ is proportional to the current I on the loop and the vector area $\mathbf{A}$ of the loop:

$$\boldsymbol{\mu} \;=\; I\mathbf{A}\,. \tag{2.1.1}$$

For a planar loop the vector area is a vector normal to the loop with length equal to the value of the area. The direction of the normal is determined from the direction of the current and the right-hand rule. The product μB of the magnitude μ of the magnetic moment times the magnetic field has units of energy, so the units of μ are

$$[\mu] \;=\; \frac{\text{erg}}{\text{gauss}} \quad \text{or} \quad \frac{\text{Joule}}{\text{Tesla}} \tag{2.1.2}$$

A rotating charge distribution results in a magnetic moment and, if the distribution has mass, an angular momentum. The magnetic moment and the angular momentum are proportional to each other, and the constant of proportionality is universal. To see this consider a ring of charge with radius R that has a uniform charge distribution and total charge Q. Assume that the ring is rotating about an axis perpendicular to the plane of the ring and going through its center. Let the tangential velocity at the ring be v. The current at the loop is equal to the linear charge density λ times the velocity:

$$I \;=\; \lambda\, v \;=\; \frac{Q}{2\pi R} v\,. \tag{2.1.3}$$

It follows that the magnitude μ of the dipole moment of the loop is

$$\mu \;=\; IA \;=\; \frac{Q}{2\pi R}\, v\, \pi R^2 \;=\; \frac{Q}{2}\, Rv\,. \tag{2.1.4}$$

Let the mass of the ring be M. The magnitude L of the angular momentum of the ring is then $L = R(Mv)$. As a result

$$\mu \;=\; \frac{Q}{2M}\, RMv \;=\; \frac{Q}{2M}\, L\,, \tag{2.1.5}$$

leading to the notable ratio

$$\boxed{\frac{\mu}{L} \;=\; \frac{Q}{2M}\,.} \tag{2.1.6}$$

Note that the ratio does not depend on the radius of the ring, nor on its velocity. By superposition, any rotating distribution with uniform mass and charge density will have a ratio μ/L as above, with Q the total charge and M the total mass. The above is also written as

$$\mu \;=\; \frac{Q}{2M}\, L\,. \tag{2.1.7}$$

This formula would certainly apply to a uniformly charged sphere that is rotating.

We can ask if the electron can have an intrinsic magnetic moment $\boldsymbol{\mu}$, as if it were, a tiny spinning ball. Well, it has an intrinsic $\boldsymbol{\mu}$ but it cannot really be viewed as a rotating little

ball of charge (this was part of Pauli's objection to the original idea of spin). Moreover, we currently view the electron as an elementary particle with zero size, so the idea that it rotates is just not sensible. The classical relation, however, points to the correct result. Even if it has no size, the electron has an intrinsic spin $\mathbf{S}$ –intrinsic angular momentum. One could guess that

$$\boldsymbol{\mu} \;=\; -\frac{e}{2m_e}\,\mathbf{S}\;? \tag{2.1.8}$$

Note that we included a minus sign because the charge of the electron is $-e$, with $e > 0$. Since angular momentum and spin have the same units we can write this as

$$\boldsymbol{\mu} \;=\; -\frac{e\hbar}{2m_e}\,\frac{\mathbf{S}}{\hbar}\;? \tag{2.1.9}$$

This is not exactly right, however. For electrons the magnetic moment is actually twice as large. One uses a constant "g-factor" to describe this effect

$$\boldsymbol{\mu} \;=\; -g\,\frac{e\hbar}{2m_e}\,\frac{\mathbf{S}}{\hbar}\,, \quad g = 2 \;\text{ for an electron.} \tag{2.1.10}$$

This factor of two is in fact predicted by the Dirac equation for the electron, and has been verified experimentally. To describe the above more briefly, one introduces a canonical value μ_B of the dipole moment called the Bohr-magneton:

$$\mu_B \;\equiv\; \frac{e\hbar}{2m_e} \;=\; 9.27\times 10^{-24}\frac{\mathrm{J}}{\mathrm{Tesla}}\,. \tag{2.1.11}$$

With this formula we get, for an electron

$$\boxed{\boldsymbol{\mu} \;=\; -\,g\,\mu_B\,\frac{\mathbf{S}}{\hbar}\,, \quad g = 2\,.} \tag{2.1.12}$$

Another feature of magnetic dipoles is needed for our discussion: a dipole placed in a non-uniform magnetic field experiences a force. An illustration is given in Figure 2.1, where to the left we show a current ring whose associated dipole moment $\boldsymbol{\mu}$ points upward. The magnetic field lines diverge as we move up, so the magnetic field is stronger as we move down. This dipole will experience a force pointing down, as can be deduced as follows. On a small piece of wire the force $d\mathbf{F}$ is proportional to $\mathbf{I}\times\mathbf{B}$. The vectors $d\mathbf{F}$ are sketched in the right part of the figure. Their horizontal components cancel out, but the result is a net force downwards.

In general the equation for the force on a dipole $\boldsymbol{\mu}$ in a magnetic field $\mathbf{B}$ is given by

$$\mathbf{F} \;=\; \nabla(\boldsymbol{\mu}\cdot\mathbf{B})\,. \tag{2.1.13}$$

Note that the force points in the direction for which $\boldsymbol{\mu}\cdot\mathbf{B}$ increases the fastest. Given that in our situation $\boldsymbol{\mu}$ and $\mathbf{B}$ are parallel, this direction is the direction in which the magnitude of $\mathbf{B}$ increases the fastest.

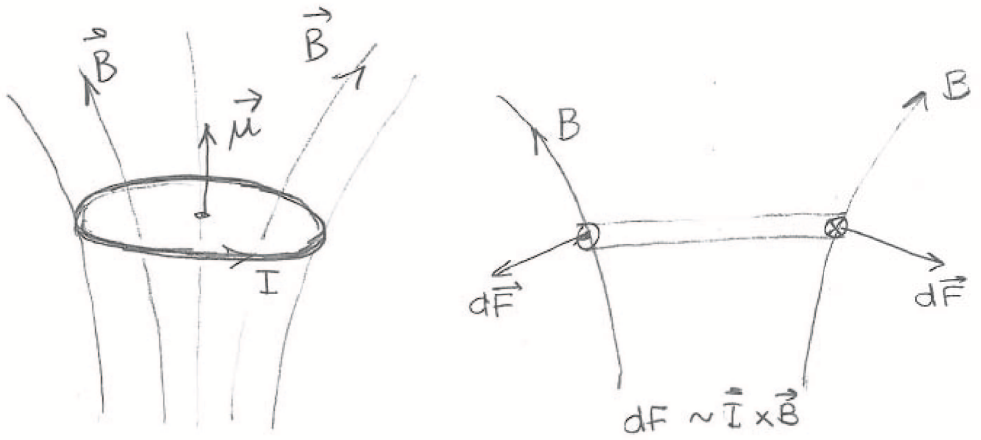

Figure 2.1: A magnetic dipole in a non-uniform magnetic field will experience a force. The force points in the direction for which $\boldsymbol{\mu} \cdot \mathbf{B}$ grows the fastest. In this case the force is downward.

The Stern-Gerlach experiment uses silver atoms, which have 47 electrons. Forty-six of them fill completely the $n = 1, 2, 3$, and 4 energy levels. The last electron is an $n = 5$ electron with zero orbital angular momentum –a $5s$ state. The only possible angular momentum is the intrinsic angular momentum of the last electron. Thus the magnetic dipole moment is also that of the last electron. The nucleus has a much smaller dipole moment and can be ignored. The silver is vaporized in an oven and, with a help of a collimating slit, a narrow beam of silver atoms is sent down to a magnet configuration. In the situation described by Figure 2.2 the magnetic field points mostly in the positive z direction, and the gradient is also in the positive z-direction. As a result, the above equation gives

$$\mathbf{F} \simeq \nabla(\mu_z B_z) = \mu_z \nabla B_z \simeq \mu_z \frac{\partial B_z}{\partial z} \mathbf{e}_z \,, \tag{2.1.14}$$

and thus, classically, the atoms experience a force and deflection in the z-direction proportional to the z-component of their magnetic moment. Undeflected atoms would hit the detector screen at the point P. Atoms with positive μ_z should be deflected upwards and atoms with negative μ_z should be deflected downwards.

The oven produces atoms with magnetic moments pointing in random directions and thus the expectation was that the z-component of the magnetic moment would have a smooth probability distribution leading to a detection that would be roughly like the one indicated on the left side of Figure 2.3. Surprisingly, the observed result was two separate peaks as if all atoms had either a fixed positive μ_z or a fixed negative μ_z. This is shown on the right side of the figure. The fact that the peaks are spatially separated led to the original unfitting name of "space quantization." The Stern-Gerlach experiment demonstrates the quantization of the dipole moment, and by theoretical inference from (2.1.12), the quantization of the spin (or intrinsic) angular momentum.

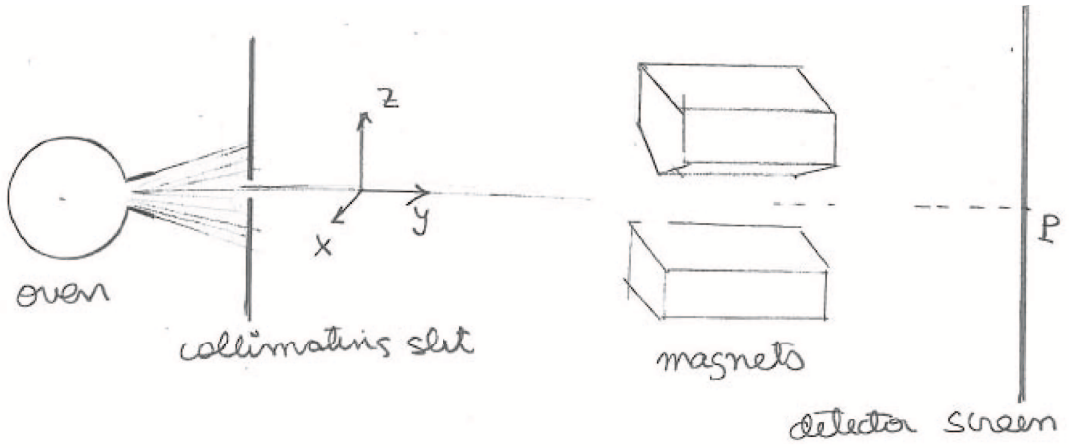

Figure 2.2: A sketch of the Stern-Gerlach apparatus. An oven and a collimating slit produces a narrow beam of silver atoms. The beam goes through a region with a strong magnetic field and a strong gradient, both in the z-direction. The screen, to the right, acts as a detector.

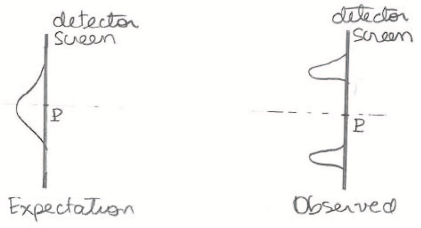

Figure 2.3: Left: the pattern on the detector screen that would be expected from classical physics. Right: the observed pattern, showing two separated peaks corresponding to positive and negative magnetic moment μ_z.

Given the known magnetic field gradient, the measured deflections lead to a determination of the two values of μ_z. Since (2.1.12) relates magnetic moments to spin angular momentum as

$$\mu_z \;=\; -\,2\,\mu_B\,\frac{S_z}{\hbar}\,, \tag{2.1.15}$$

this leads to a determination of two values for S_z:

$$S_z \;=\; \pm\frac{\hbar}{2}\,, \quad \text{or} \quad \frac{S_z}{\hbar} \;=\; \pm\frac{1}{2}\,. \tag{2.1.16}$$

A particle with such possible values of $S_z/\hbar$ is called a spin one-half particle. The values of the magnetic moments of the electron are plus-minus one Bohr magneton.

With the magnetic field and its gradient along the z-direction, the Stern-Gerlach apparatus measures the component of the spin $\mathbf{S}$ in the z direction. To streamline our pictures we will denote such an apparatus as a box with a $\hat{z}$ label, as in Figure 2.4. The input beam comes in from the left and the box lets out two beams from the right. If we placed a detector to the right, the top beam would be identified as having atoms with $S_z = \hbar/2$ and the bottom beam would be identified as having atoms with $S_z = -\hbar/2$.[1]

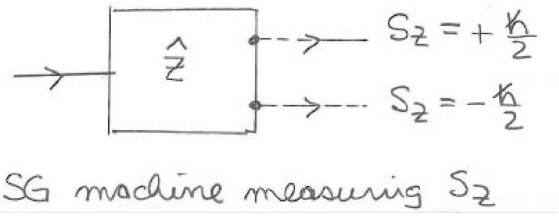

Figure 2.4: A schematic representation of the SG apparatus, minus the screen.

Let us now consider thought experiments in which we put a few SG apparatus in series. In the first configuration, shown at the top of Figure 2.5, the first box is a $\hat{z}$-type SG machine, where we block the $S_z = -\hbar/2$ output beam and let only the $S_z = \hbar/2$ beam go into the next machine. This machine acts as a filter. The second SG apparatus is also a $\hat{z}$-type machine. Since all ingoing particles have $S_z = \hbar/2$ the second machine lets those out the top output, and nothing comes out the bottom output. The quantum mechanical lesson here is that $S_z = \hbar/2$ states have no component or amplitude along $S_z = -\hbar/2$ states. $S_z = \hbar/2$ and $S_z = -\hbar/2$ are thus said to be orthogonal states.

The second configuration in the figure shows the outgoing $S_z = \hbar/2$ beam from the first machine entering an $\hat{x}$-type machine. The outputs of this machine are –in analogy to the $\hat{z}$-type machine– $S_x = \hbar/2$ and $S_x = -\hbar/2$. Classically an object with angular momentum along the z axis has no component of angular momentum along the x axis; these are orthogonal directions. But the result of the experiment indicates that this is not true for quantum spins. About half of the $S_z = \hbar/2$ atoms exit through the top $S_x = \hbar/2$ output, and the other half exit through the bottom $S_x = -\hbar/2$ output. Quantum mechanically, a state with a definite value of S_z has an amplitude along the state $S_x = \hbar/2$ as well as an amplitude along the state $S_x = -\hbar/2$.

In the third and bottom configuration the $S_z = \hbar/2$ beam from the first machine goes into the $\hat{x}$-type machine and the top output is blocked so that we only have an $S_x = -\hbar/2$ output. That beam is fed into a $\hat{z}$-type machine. One could speculate that the beam entering the third machine has both $S_x = -\hbar/2$ *and* $S_z = \hbar/2$, as it is composed of silver atoms that made it through both machines. If that were the case the third machine would

[1] In the quantum mechanical view of the experiment, a single atom can be in both beams, with different amplitudes. Only the act of measurement, which corresponds to the act of placing the detector screen, forces the atom to decide in which beam it is.

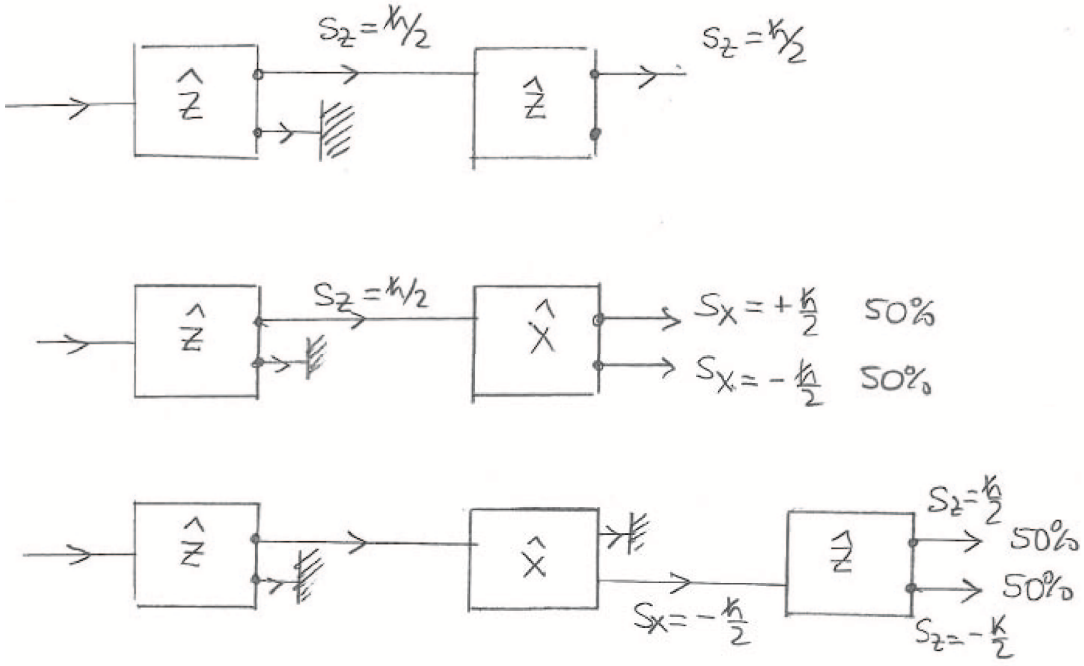

Figure 2.5: Three configurations of SG boxes.

let all atoms out the top output. This speculation is falsified by the result. There is no memory of the first filter: the particles out of the second machine do not have $S_z = \hbar/2$ anymore. We find that half of the particles make it out of the third machine with $S_z = \hbar/2$ and the other half with $S_z = -\hbar/2$. In the following section we discuss a mathematical framework consistent with the results of the above thought experiments.

2.2 Spin one-half states and operators

The Stern-Gerlach experiment suggests that the spin states of the electron can be described using two basis vectors (or kets):

$$|z;+\rangle \quad \text{and} \quad |z;-\rangle\,. \tag{2.2.1}$$

The first corresponds to an electron with $S_z = \frac{\hbar}{2}$. The label z indicates that the z component of spin is being specified, and the $+$ indicates that this component is positive. This state is also called 'spin up' along z. The second state corresponds to an electron with $S_z = -\frac{\hbar}{2}$, that is a 'spin down' along z. Mathematically, we have a spin operator $\hat{S}_z$ for which the

above states are eigenstates with opposite eigenvalues:

$$\begin{aligned} \hat{S}_z|z;+\rangle &= +\frac{\hbar}{2}|z;+\rangle \\ \hat{S}_z|z;-\rangle &= -\frac{\hbar}{2}|z;-\rangle\,. \end{aligned} \tag{2.2.2}$$

With two basis states, the the state space of electron spin is a two-dimensional *complex vector space*. Each vector in this vector space represents a possible state of the electron spin. Note that we are not discussing other degrees of freedom of the electron, such as its position, momentum, or energy, just the spin. The general vector in the two-dimensional space is an arbitrary linear combination of the basis states (or basis vectors) and thus takes the form

$$|\Psi\rangle = c_1|z;+\rangle + c_2|z;-\rangle\,, \quad \text{with} \quad c_1, c_2 \in \mathbb{C}\,. \tag{2.2.3}$$

It is customary to call the state $|z;+\rangle$ the *first* basis state and it denote by $|1\rangle$. The state $|z;-\rangle$ is called the *second* basis state and is denoted by $|2\rangle$. In a two-dimensional vector space a vector is explicitly *represented* as a column vector with two components. The first basis vector is represented as $\begin{pmatrix}1\\0\end{pmatrix}$ and the second basis vector is represented as $\begin{pmatrix}0\\1\end{pmatrix}$. Thus we have the following names for states and their concrete representation as column vectors

$$\begin{aligned} |z:+\rangle &= |1\rangle = \begin{pmatrix}1\\0\end{pmatrix}\,, \\ |z:-\rangle &= |2\rangle = \begin{pmatrix}0\\1\end{pmatrix}\,. \end{aligned} \tag{2.2.4}$$

Using these options the state in (2.2.3) takes the possible forms

$$|\Psi\rangle = c_1|z;+\rangle + c_2|z;-\rangle = c_1|1\rangle + c_2|2\rangle = c_1\begin{pmatrix}1\\0\end{pmatrix} + c_2\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}c_1\\c_2\end{pmatrix}\,. \tag{2.2.5}$$

Recall that the top experiment in Figure 2.5 suggests that we have an orthonormal basis: the state $|z;+\rangle$ entering the second machine must have zero overlap with $|z,-\rangle$ since no such down spins emerge. Moreover the overlap of $|z;+\rangle$ with itself must be one, as all states emerge from the second machine top output. We will thus write

$$\langle z;-|z;+\rangle = 0\,, \quad \langle z;+|z;+\rangle = 1\,, \tag{2.2.6}$$

using bra-ket notation to be clarified shortly. Similarly, we expect

$$\langle z;+|z;-\rangle = 0\,, \quad \langle z;-|z;-\rangle = 1\,. \tag{2.2.7}$$

Labeling the basis states as $|1\rangle$ and $|2\rangle$ we have a simple form that summarizes the four equations above:

$$\langle i|j\rangle = \delta_{ij}\,, \quad i,j = 1,2. \tag{2.2.8}$$

We define the basis 'bras' as the *row vectors* obtained by transposition and complex conjugation:

$$\langle 1| \;=\; (1,0)\,, \quad \langle 2| \;=\; (0,1)\,. \tag{2.2.9}$$

Given states $|\alpha\rangle$ and $|\beta\rangle$

$$\begin{aligned} |\alpha\rangle \;&=\; \alpha_1|1\rangle + \alpha_2|2\rangle \;=\; \begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} \\ |\beta\rangle \;&=\; \beta_1|1\rangle + \beta_2|2\rangle \;=\; \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} \end{aligned} \tag{2.2.10}$$

we define the bra $\langle\alpha|$ associated with the ket $|\alpha\rangle$ as follows:

$$\langle\alpha| \;\equiv\; \alpha_1^*\langle 1| + \alpha_2^*\langle 2| \;=\; (\alpha_1^*\,, \alpha_2^*)\,. \tag{2.2.11}$$

The 'bra-ket' inner product is a number defined by the matrix product of the corresponding row and column vector representatives:

$$\langle\alpha|\beta\rangle \;\equiv\; (\alpha_1^*\,, \alpha_2^*)\cdot\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} \;=\; \alpha_1^*\beta_1 + \alpha_2^*\beta_2\,. \tag{2.2.12}$$

The matrix product is, of course, a number. You should check that this definition is consistent with (2.2.8).

When we represent the states as two-component column vectors the operators that act on the states to give new states can be *represented* as two-by-two matrices. We can thus represent the operator $\hat{S}_z$ as a 2×2 matrix which we claim takes the form

$$\boxed{\hat{S}_z \;=\; \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\,.} \tag{2.2.13}$$

To test this, it suffices to verify that the matrix $\hat{S}_z$ acts on the column vectors that represent the basis states as expected from (2.2.2). Indeed,

$$\begin{aligned} \hat{S}_z|z;+\rangle \;&=\; \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} \;=\; \frac{\hbar}{2}\begin{pmatrix}1\\ 0\end{pmatrix} \;=\; \frac{\hbar}{2}\,|z;+\rangle \\ \hat{S}_z|z;-\rangle \;&=\; \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix} \;=\; -\frac{\hbar}{2}\begin{pmatrix}0\\ 1\end{pmatrix} \;=\; -\frac{\hbar}{2}\,|z;-\rangle\,. \end{aligned} \tag{2.2.14}$$

The states $|1\rangle$ and $|2\rangle$, viewed as column vectors, are the eigenvectors of the matrix $\hat{S}_z$.

There is nothing particular about the z axis. We could have started with a SG apparatus that measures spin along the x axis and we would have been led to an operator $\hat{S}_x$. Had we used the y axis we would have been led to the operator $\hat{S}_y$. Since spin represents angular momentum (albeit of intrinsic type) it is expected to have three components, just like orbital angular momentum has three components $\hat{L}_x, \hat{L}_y$, and $\hat{L}_z$. Each component is a hermitian

operator and can be written as a sum of products of coordinates and momenta in three-dimensional space. Writing $\hat{L}_x = \hat{L}_1, \hat{L}_y = \hat{L}_2$, and $\hat{L}_z = \hat{L}_3$, their commutation relations can be concisely stated as

$$\left[\hat{L}_i\,,\hat{L}_j\right] \;=\; i\hbar\,\epsilon_{ijk}\,\hat{L}_k \tag{2.2.15}$$

This is the famous algebra of angular momentum; repeated indices are summed over the values 1,2,3, and ϵ_{ijk} is the totally antisymmetric symbol with $\epsilon_{123} = +1$. Make sure that you understand this notation clearly, and can use it to see that it implies the relations

$$\begin{aligned}
[\hat{L}_x\,,\hat{L}_y] &= i\hbar\hat{L}_z\,,\\
[\hat{L}_y\,,\hat{L}_z] &= i\hbar\hat{L}_x\,,\\
[\hat{L}_z\,,\hat{L}_x] &= i\hbar\hat{L}_y\,.
\end{aligned} \tag{2.2.16}$$

While $\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$, for example, is a hermitian operator written in terms of coordinates and momenta, we have no such construction for $\hat{S}_z$. The latter is a more abstract operator: it does not act on wavefunctions $\psi(\vec{x})$ but rather on the 2-component column vectors introduced above. The operator $\hat{S}_z$ is just a two-by-two *hermitian*[2] matrix with constant entries! If spin is a quantum mechanical angular momentum, we must have a triplet of operators $\hat{S}_x, \hat{S}_y$, and $\hat{S}_z$ that satisfy the algebra of angular momentum:

$$\begin{aligned}
[\hat{S}_x\,,\hat{S}_y] &= i\hbar\hat{S}_z\,,\\
[\hat{S}_y\,,\hat{S}_z] &= i\hbar\hat{S}_x\,,\\
[\hat{S}_z\,,\hat{S}_x] &= i\hbar\hat{S}_y\,.
\end{aligned} \tag{2.2.17}$$

Again, using numerical subscripts for the components ($\hat{S}_1 = \hat{S}_x, \hat{S}_2 = S_y, \hat{S}_3 = S_z$) we must have

$$\left[\hat{S}_i\,,\hat{S}_j\right] \;=\; i\hbar\epsilon_{ijk}\,\hat{S}_k\,. \tag{2.2.18}$$

We can now try to figure out how the matrices for $\hat{S}_x$ and $\hat{S}_y$ must look, given that we know the matrix for $\hat{S}_z$. We have a few constraints. First the matrices must be hermitian, just like the angular momentum operators are. The most general two-by-two hermitian matrix takes the form

$$\begin{pmatrix} a+d & b-ic \\ b+ic & a-d \end{pmatrix}\;,\;\text{with}\;\; a,b,c,d\in\mathbb{R} \tag{2.2.19}$$

Indeed, you can easily see that transposing and complex conjugating results in exactly the same matrix. This matrix is parameterized by four real constants and can be written as

$$a\begin{pmatrix}1&0\\0&1\end{pmatrix} + b\begin{pmatrix}0&1\\1&0\end{pmatrix} + c\begin{pmatrix}0&-i\\i&0\end{pmatrix} + d\begin{pmatrix}1&0\\0&-1\end{pmatrix} \tag{2.2.20}$$

We see here four simple hermitian two-by-two matrices: the ones multiplying each of the real constants. These matrices are independent in the sense that all of them are needed to

[2] A Hermitian matrix remains unchanged by the successive operations of transposition and complex conjugation.

construct the most general hermitian matrix by combination of linear sums. We also note that the last matrix is $\hat{S}_z$ up to the factor of $\hbar/2$. So we may try to identify $\hat{S}_x$ and $\hat{S}_y$ among the other three. Since the first matrix is the identity, and commutes with any other matrix, it is natural to guess that $\hat{S}_x$ and $\hat{S}_y$ correspond to the second and third matrices

$$\hat{S}_x \sim \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{S}_y \sim \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{2.2.21}$$

We will discuss later if there is a unique guess that works (the answer is no!). The overall scale of the matrices can be fixed by the constraint that their eigenvalues be $\pm\hbar/2$, just like they are for $\hat{S}_z$. Let us give the eigenvalues (denoted by λ) and the associated normalized eigenvectors for these two matrices. Short computations (can you do them?) give

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}: \qquad \lambda = 1, \text{ for } \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \lambda = -1, \text{ for } \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \tag{2.2.22}$$

for the first matrix and

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}: \qquad \lambda = 1, \text{ for } \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}, \quad \lambda = -1, \text{ for } \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}, \tag{2.2.23}$$

for the second matrix. In case you are puzzled by the normalizations, note that a vector $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ is normalized if $|c_1|^2 + |c_2|^2 = 1$. Since the eigenvalues of both matrices are ± 1, we simply need to multiply the matrices by $\hbar/2$ to get matrices with eigenvalues $\pm\hbar/2$. Thus, we tentatively identify

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{2.2.24}$$

which have, at least, the correct eigenvalues. But in fact, these also satisfy the commutation relations! Indeed, we check that, as desired,

$$\begin{aligned} [\hat{S}_x, \hat{S}_y] &= \frac{\hbar^2}{4}\Big(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\Big) \\ &= \frac{\hbar^2}{4}\Big(\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}\Big) \\ &= \frac{\hbar^2}{4}\begin{pmatrix} 2i & 0 \\ 0 & -2i \end{pmatrix} = i\hbar\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = i\hbar\hat{S}_z\,. \end{aligned} \tag{2.2.25}$$

All in all we have

$$\boxed{\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.} \tag{2.2.26}$$

Exercise. Verify that these matrices satisfy the other two commutation relations in (2.2.17).

Is ours the unique solution for $\hat{S}_x$ and $\hat{S}_y$ given the choice of $\hat{S}_z$? The answer is no, as illustrated by the following exercise:

Exercise. Check that the set of commutation relations of the spin operators are in fact preserved when we replace $\hat{S}_x \to -\hat{S}_y$ and $\hat{S}_y \to \hat{S}_x$.

The solution we gave is the one that, happily, is used by all physicists. Any other solution is physically equivalent to the one we gave (as will be explained in more detail after we develop more results). The solution defines the **Pauli matrices** σ_i by writing

$$\hat{S}_i \;=\; \frac{\hbar}{2}\,\sigma_i\,. \tag{2.2.27}$$

We then have that the Pauli matrices are

$$\boxed{\sigma_1 \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \quad \sigma_2 \;=\; \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\,, \quad \sigma_3 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,.} \tag{2.2.28}$$

We are now in a position to describe spin states that point along x or $-x$ as superposition of our basis states that point along z and $-z$. The spin states that point along $\pm x$ are the eigenstates of $\hat{S}_x$, are are called $|x;\pm\rangle$. By definition, they satisfy

$$\hat{S}_x\,|x;\pm\rangle \;=\; \pm\frac{\hbar}{2}\,|x;\pm\rangle\,. \tag{2.2.29}$$

Given (2.2.24) their form can be read from (2.2.22):

$$\boxed{\begin{aligned} |x;+\rangle \;&=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \;=\; \frac{1}{\sqrt{2}}|z;+\rangle + \frac{1}{\sqrt{2}}|z;-\rangle\,, \\ |x;-\rangle \;&=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \;=\; \frac{1}{\sqrt{2}}|z;+\rangle - \frac{1}{\sqrt{2}}|z;-\rangle\;\;, \end{aligned}} \tag{2.2.30}$$

Note that these states are orthogonal to each other. The above equations can be inverted to find

$$\begin{aligned} |z;+\rangle \;&=\; \frac{1}{\sqrt{2}}|x;+\rangle + \frac{1}{\sqrt{2}}|x;-\rangle \\ |z;-\rangle \;&=\; \frac{1}{\sqrt{2}}|x;+\rangle - \frac{1}{\sqrt{2}}|x;-\rangle \end{aligned} \tag{2.2.31}$$

These relations are consistent with the second experiment shown in Figure 2.5. The state $|z;+\rangle$ entering the second, $\hat{x}$-type SG apparatus, has equal probability to be found in $|x;+\rangle$ as it has probability to be found in $|x;-\rangle$. This is reflected in the first of the above relations, since we have the amplitudes

$$\langle x;+|z;+\rangle \;=\; \frac{1}{\sqrt{2}}\,, \qquad \langle x;-|z;+\rangle \;=\; \frac{1}{\sqrt{2}}\,. \tag{2.2.32}$$

The probabilities, being equal to the norm squared of the amplitudes, are 1/2 in both cases. The relative minus sign on the second equation above is needed to make it orthogonal to the state on the first equation.

We can finally consider the eigenstates of $\hat{S}_y$, namely the spin states that point along $\pm y$. We have

$$\hat{S}_y\,|y;\pm\rangle \;=\; \pm\frac{\hbar}{2}\,|y;\pm\rangle\,. \tag{2.2.33}$$

and using (2.2.23) we read

$$\boxed{\begin{aligned} |y;+\rangle \;&=\; \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix} \;=\; \frac{1}{\sqrt{2}}|z;+\rangle + \frac{i}{\sqrt{2}}|z;-\rangle\;\;, \\ |y;-\rangle \;&=\; \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix} \;=\; \frac{1}{\sqrt{2}}|z;+\rangle - \frac{i}{\sqrt{2}}|z;-\rangle\,. \end{aligned}} \tag{2.2.34}$$

Note that this time the superposition of $|z;\pm\rangle$ states involves complex numbers!

2.3 Properties of Pauli matrices and index notation

Since we know the commutation relations for the spin operators

$$\big[\hat{S}_i\,,\,\hat{S}_j\big] = i\hbar\,\epsilon_{ijk}\hat{S}_k\,, \tag{2.3.35}$$

and we have $S_i = \frac{\hbar}{2}\sigma_i$, it follows that

$$\frac{\hbar}{2}\frac{\hbar}{2}\big[\sigma_i\,,\,\sigma_j\big] = i\hbar\,\epsilon_{ijk}\frac{\hbar}{2}\sigma_k\,. \tag{2.3.36}$$

Cancelling the $\hbar$'s and some factors of two, we find

$$\boxed{[\sigma_i,\sigma_j] \;=\; 2i\,\epsilon_{ijk}\sigma_k\,.} \tag{2.3.37}$$

Another important property of the Pauli matrices is that they square to the identity matrix. This is best checked explicitly (do it!):

$$\boxed{(\sigma_1)^2 \;=\; (\sigma_2)^2 \;=\; (\sigma_3)^2 \;=\; \mathbf{1}\,.} \tag{2.3.38}$$

This property "explains" that the eigenvalues of each of the Pauli matrices can only be plus or minus one. Indeed, the eigenvalues of a matrix satisfy the algebraic equation that the matrix satisfies. Take for example a matrix M that satisfies the matrix equation

$$M^2 + \alpha M + \beta\mathbf{1} \;=\; 0 \tag{2.3.39}$$

Let v be an eigenvector of M with eigenvalue λ: $Mv = \lambda v$. Let the above equation act on v

$$M^2 v + \alpha M v + \beta \mathbf{1} v \;=\; 0 \quad \rightarrow \quad \lambda^2 v + \alpha\lambda v + \beta v = 0 \quad \rightarrow \quad (\lambda^2 + \alpha\lambda + \beta)v = 0\,, \tag{2.3.40}$$

and since $v \neq 0$ (by definition an eigenvector cannot be zero!) we conclude that $\lambda^2+\alpha\lambda+\beta = 0$, as claimed. For the case of the Pauli matrices we have $(\sigma_i)^2 = 1$ and therefore the eigenvalues must satisfy $\lambda^2 = 1$. As a result, $\lambda = \pm 1$ are the only options.

We also note, by inspection, that the Pauli matrices have zero trace, namely, the sum of entries on the diagonal is zero:

$$\boxed{\mathrm{tr}(\sigma_i) \;=\; 0\,, \qquad i = 1,2,3.} \tag{2.3.41}$$

A fact from linear algebra is that the trace of a matrix is equal to the sum of its eigenvalues. So each Pauli matrix must have two eigenvalues that add up to zero. Since the eigenvalues can only be plus or minus one, we must have one of each. This shows that each of the Pauli matrices has a plus one and a minus one eigenvalue.

If you compute a commutator of Pauli matrices by hand you might notice a curious property. Take the commutator of σ_1 and σ_2:

$$[\sigma_1\,, \sigma_2\,] \;=\; \sigma_1\sigma_2 - \sigma_2\sigma_1\,. \tag{2.3.42}$$

The two contributions on the right hand side give

$$\begin{aligned} \sigma_1\sigma_2 \;&=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \;=\; \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\,, \\ \sigma_2\sigma_1 \;&=\; \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \;=\; \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}\,. \end{aligned} \tag{2.3.43}$$

The second contribution is minus the first, so that both terms contribute equally to the commutator! In other words,

$$\sigma_1\sigma_2 \;=\; -\,\sigma_2\sigma_1\,. \tag{2.3.44}$$

This equation is taken to mean that σ_1 and σ_2 *anticommute*. Just like we define the commutator of two operators X, Y by $[X, Y] \equiv XY - YX$, we define the **anticommutator**, denoted by curly brackets, by

$$\boxed{\text{Anticommutator:}\;\; \{X\,, Y\} \;\equiv\; XY + YX\,.} \tag{2.3.45}$$

In this language we have checked that

$$\{\sigma_1\,, \sigma_2\} \;=\; 0\,, \tag{2.3.46}$$

and the property $\sigma_1^2 = \mathbf{1}$, for example, can be rewritten as

$$\{\sigma_1\,, \sigma_1\} \;=\; 2\cdot\mathbf{1}\,. \tag{2.3.47}$$

In fact, you can check (by examining the two remaining cases) that any two different Pauli matrices anticommute:

$$\{\sigma_i\,,\sigma_j\} \;=\; 0\,, \quad \text{for} \;\; i \neq j\,. \tag{2.3.48}$$

We can improve this equation to make it work also when i is equal to j. We claim that

$$\boxed{\{\sigma_i\,,\sigma_j\} \;=\; 2\delta_{ij}\,\mathbf{1}\,.} \tag{2.3.49}$$

Indeed, when $i \neq j$ the right-hand side vanishes, as needed, and when i is equal to j, the right-hand side gives $2 \cdot \mathbf{1}$, also as needed in view of (2.3.47) and its analogs for the other Pauli matrices.

Both the commutator and anti-commutator identities for the Pauli matrices can be summarized in a single equation. This is possible because, for any two operators X, Y we have

$$X\,Y \;=\; \tfrac{1}{2}\,\{X,Y\} \;+\; \tfrac{1}{2}\,[X,Y]\,, \tag{2.3.50}$$

as you should confirm by expansion. Applied to the product of two Pauli matrices and using our expressions for the commutator and anticommutator we get

$$\boxed{\sigma_i\sigma_j \;=\; \delta_{ij}\,\mathbf{1} \;+\; i\,\epsilon_{ijk}\,\sigma_k\,.} \tag{2.3.51}$$

This equation can be recast in vector notation. Denote three-component vectors by bold symbols: $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$. Then the dot product

$$\mathbf{a}\cdot\mathbf{b} \;=\; a_1b_1 + a_2b_2 + a_3b_3 = a_ib_i\,. \tag{2.3.52}$$

We use of the sum convention in which repeated indices are summed over: $a_ib_i \equiv \sum_{i=1}^{3} a_ib_i$. Using the sum convention you can also see that, for example, $b_j\delta_{ij} = b_i$. We also have that

$$\mathbf{a}\cdot\mathbf{a} \;=\; |\mathbf{a}|^2\,. \tag{2.3.53}$$

Cross products use the epsilon symbol. Make sure you understand why

$$(\mathbf{a}\times\mathbf{b})_k \;=\; a_ib_j\,\epsilon_{ijk}\,. \tag{2.3.54}$$

We can also have triplets of matrices! They are denoted by the same bold symbol we use for vectors. For the Pauli matrices we denote

$$\boldsymbol{\sigma} \;\equiv\; (\sigma_1, \sigma_2, \sigma_3)\,. \tag{2.3.55}$$

We can construct a matrix by dot product of a vector $\mathbf{a}$ with the 'vector' $\boldsymbol{\sigma}$. We define

$$\mathbf{a}\cdot\boldsymbol{\sigma} \;\equiv\; a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 \;=\; a_i\sigma_i\,. \tag{2.3.56}$$

Note that $\mathbf{a}\cdot\boldsymbol{\sigma}$ is just a single two-by-two matrix. Since the components of $\mathbf{a}$ are numbers, and numbers commute with matrices, this dot product is commutative: $\mathbf{a}\cdot\boldsymbol{\sigma} = \boldsymbol{\sigma}\cdot\mathbf{a}$. We are now ready to rewrite (2.3.51). Multiply this equation by $a_i b_j$ to get

$$\begin{aligned} a_i\sigma_i\; b_j\sigma_j &= a_i b_j\delta_{ij}\,\mathbf{1} \;+\; i\,(a_i b_j\epsilon_{ijk})\,\sigma_k \\ &= (\mathbf{a}\cdot\mathbf{b})\,\mathbf{1} \;+\; i\,(\mathbf{a}\times\mathbf{b})_k\,\sigma_k\,, \end{aligned} \tag{2.3.57}$$

so that, finally, we get the matrix equation

$$\boxed{(\mathbf{a}\cdot\boldsymbol{\sigma})(\mathbf{b}\cdot\boldsymbol{\sigma}) = (\mathbf{a}\cdot\mathbf{b})\,\mathbf{1} + i\,(\mathbf{a}\times\mathbf{b})\cdot\boldsymbol{\sigma}\,.} \tag{2.3.58}$$

As a simple application we take $\mathbf{b} = \mathbf{a}$. We then have $\mathbf{a}\cdot\mathbf{a} = |\mathbf{a}|^2$ and $\mathbf{a}\times\mathbf{a} = 0$, so that the above equation gives

$$(\mathbf{a}\cdot\boldsymbol{\sigma})^2 = |\mathbf{a}|^2\,\mathbf{1}\,. \tag{2.3.59}$$

When $\mathbf{a}$ is a unit vector this becomes

$$(\mathbf{n}\cdot\boldsymbol{\sigma})^2 = \mathbf{1}\,, \quad \mathbf{n}\text{ a unit vector.} \tag{2.3.60}$$

The epsilon symbol satisfies useful identities. One can show that the product of two epsilons with one index contracted is a sum of products of Kronecker deltas:

$$\epsilon_{ijk}\,\epsilon_{ipq} \;=\; \delta_{jp}\delta_{kq} \;-\; \delta_{jq}\delta_{kp}\,. \tag{2.3.61}$$

The further contraction in which we set $p = j$ is also useful:

$$\epsilon_{ijk}\,\epsilon_{ijq} \;=\; 2\delta_{kq}\,. \tag{2.3.62}$$

The first of these two allows one to prove the familiar vector identity

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) \;=\; \mathbf{b}\,(\mathbf{a}\cdot\mathbf{c}) \;-\; (\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}\,. \tag{2.3.63}$$

It will be useful later on to consider the dot and cross products of *operator* triplets. Given the operators $\mathbf{X} = (\hat{X}_1, \hat{X}_2, \hat{X}_3)$ and $\mathbf{Y} = (\hat{Y}_1, \hat{Y}_2, \hat{Y}_3)$ we define

$$\begin{aligned} \mathbf{X}\cdot\mathbf{Y} &\equiv \hat{X}_i\,\hat{Y}_i\,, \\ (\mathbf{X}\times\mathbf{Y})_i &\equiv \epsilon_{ijk}\,\hat{X}_j\,\hat{Y}_k\,. \end{aligned} \tag{2.3.64}$$

In these definitions the order of the operators on the right hand side is as in the left-hand side. This matters, since the $\hat{X}_i$ and $\hat{Y}_j$ operators may not commute. The dot product of two operator triplets is not necessarily commutative, nor is the cross product necessarily antisymmetric.

2.4 Spin states in arbitrary direction

We consider here the description and analysis of spin states that point in arbitrary directions, as specified by a unit vector $\mathbf{n}$:

$$\mathbf{n} \;=\; (n_x, n_y, n_z) \;=\; (\sin\theta\cos\phi,\, \sin\theta\sin\phi,\, \cos\theta)\;. \tag{2.4.65}$$

Here θ and ϕ are the familiar polar and azimuthal angles used in spherical coordinates. We view the spatial vector $\mathbf{n}$ as a triplet of numbers. Just like we did for $\boldsymbol{\sigma}$, we can define $\mathbf{S}$ as the triplet of operators

$$\mathbf{S} \;\equiv\; (\hat{S}_x\,, \hat{S}_y\,, \hat{S}_z)\,. \tag{2.4.66}$$

Note that, in fact,

$$\mathbf{S} \;=\; \frac{\hbar}{2}\boldsymbol{\sigma}\,. \tag{2.4.67}$$

We take the *dot* product of $\mathbf{S}$ with $\mathbf{n}$ to build a spin operator $\hat{S}_\mathbf{n}$ that has a simple interpretation:

$$\boxed{\hat{S}_\mathbf{n} \;\equiv\; \mathbf{n}\cdot\mathbf{S} \;\text{ is the spin operator in the direction of the unit vector }\mathbf{n}\,.} \tag{2.4.68}$$

More explicitly

$$\hat{S}_\mathbf{n} \;\equiv\; \mathbf{n}\cdot\mathbf{S} \;\equiv\; n_x\hat{S}_x + n_y\hat{S}_y + n_z\hat{S}_z \;=\; \frac{\hbar}{2}\,\mathbf{n}\cdot\boldsymbol{\sigma}\,. \tag{2.4.69}$$

Note that $\hat{S}_\mathbf{n}$ is an operator, a hermitian two-by-two matrix, in fact. To convince you that the interpretation above makes sense note that, for example, when $\mathbf{n}$ points along z, we have $(n_x, n_y, n_z) = (0, 0, 1)$ and $\hat{S}_\mathbf{n}$ becomes $\hat{S}_z$. The same holds, of course, for the x and y directions. Moreover, just like all the $\hat{S}_i$, the eigenvalues of $\hat{S}_\mathbf{n}$ are $\pm\hbar/2$. This is needed physically, since all directions are physically equivalent and those two values for spin must be the only ones allowed for *any* direction. To see that this is true we first compute the square of the matrix $\hat{S}_\mathbf{n}$:

$$(\hat{S}_\mathbf{n})^2 \;=\; \Big(\frac{\hbar}{2}\Big)^2(\mathbf{n}\cdot\boldsymbol{\sigma})^2 \;=\; \Big(\frac{\hbar}{2}\Big)^2, \tag{2.4.70}$$

using (2.3.60). Moreover, since the Pauli matrices are traceless so is $\hat{S}_\mathbf{n}$:

$$\mathrm{tr}(\hat{S}_\mathbf{n}) \;= n_i\,\mathrm{tr}(\hat{S}_i) \;=\; n_i\,\frac{\hbar}{2}\,\mathrm{tr}(\sigma_i) \;=\; 0\,. \tag{2.4.71}$$

By the same argument we used for Pauli matrices, we conclude that the eigenvalues of $\hat{S}_\mathbf{n}$ are indeed $\pm\hbar/2$. In classical physics $\mathbf{n}\cdot\mathbf{V}$ denotes the component of the vector $\mathbf{V}$ along $\mathbf{n}$. This component is just a number. In our case $\mathbf{n}\cdot\mathbf{S}$ gives a component of $\mathbf{S}$, but now this component is an operator. What we have explained above is that this component can be thought as the operator "along" the direction of $\mathbf{n}$.

Since the eigenvalues of $\hat{S}_{\mathbf{n}}$ are $\pm\hbar/2$ the associated spin eigenstates, denoted as $|\mathbf{n};\pm\rangle$, satisfy

$$\hat{S}_{\mathbf{n}}\,|\mathbf{n};\pm\rangle \;=\; \pm\frac{\hbar}{2}\,|\mathbf{n};\pm\rangle\,. \tag{2.4.72}$$

$$\boxed{|\mathbf{n};+\rangle \text{ and } |\mathbf{n};-\rangle \text{ are spin states that point along } \mathbf{n} \text{ and } -\mathbf{n} \text{ respectively.}} \tag{2.4.73}$$

It is useful to calculate the explicit form of those spin states as a function of $\mathbf{n}$. For this we first write the matrix $\hat{S}_{\mathbf{n}}$ for an arbitrary direction

$$\hat{S}_{\mathbf{n}} \;=\; \frac{\hbar}{2}\Big[n_x\begin{pmatrix}0&1\\1&0\end{pmatrix}+n_y\begin{pmatrix}0&-i\\i&0\end{pmatrix}+n_z\begin{pmatrix}1&0\\0&-1\end{pmatrix}\Big] \;=\; \frac{\hbar}{2}\begin{pmatrix}n_z & n_x-in_y\\ n_x+in_y & -n_z\end{pmatrix}. \tag{2.4.74}$$

Using the values of the components of $\mathbf{n}$ we find:

$$\hat{S}_{\mathbf{n}} \;=\; \frac{\hbar}{2}\begin{pmatrix}\cos\theta & \sin\theta e^{-i\phi}\\ \sin\theta e^{i\phi} & -\cos\theta\end{pmatrix}. \tag{2.4.75}$$

Exercise. Confirm by direct computation that the eigenvalues of $S_{\mathbf{n}}$ are $\pm\hbar/2$.

To find the eigenvector v associated with the eigenvalue λ we must solve the linear equation $(\hat{S}_{\mathbf{n}}-\lambda\mathbf{1})v=0$. For the eigenvector $|\mathbf{n};+\rangle$ we write the ansatz

$$|\mathbf{n};+\rangle = c_1|+\rangle + c_2|-\rangle \;=\; \begin{pmatrix}c_1\\c_2\end{pmatrix}, \qquad |\pm\rangle \equiv |z;\pm\rangle\,, \tag{2.4.76}$$

where for notational simplicity we have introduced the abbreviations $|\pm\rangle$ for $|z;\pm\rangle$. The eigenvector equation $(\hat{S}_{\mathbf{n}}-\frac{\hbar}{2}\mathbf{1})|\mathbf{n};+\rangle=0$ reads

$$\frac{\hbar}{2}\begin{pmatrix}\cos\theta-1 & \sin\theta e^{-i\phi}\\ \sin\theta e^{i\phi} & -\cos\theta-1\end{pmatrix}\begin{pmatrix}c_1\\c_2\end{pmatrix}=0\,. \tag{2.4.77}$$

Either equation gives the same relation between c_1 and c_2. The top equation, for example gives

$$c_2 = e^{i\phi}\,\frac{1-\cos\theta}{\sin\theta}\,c_1 = e^{i\phi}\,\frac{\sin\frac{\theta}{2}}{\cos\frac{\theta}{2}}\,c_1\,. \tag{2.4.78}$$

Exercise. Check that the second equation gives the same relation.

We want normalized states, and therefore

$$|c_1|^2+|c_2|^2=1 \quad\to\quad |c_1|^2\left[1+\frac{\sin^2\frac{\theta}{2}}{\cos^2\frac{\theta}{2}}\right]=1 \quad\to\quad |c_1|^2=\cos^2\frac{\theta}{2}\,. \tag{2.4.79}$$

Since the overall phase of the state is not observable we take the simplest option for c_1:

$$c_1 \;=\; \cos\tfrac{\theta}{2}\,, \qquad c_2=\sin\tfrac{\theta}{2}\exp(i\phi)\;, \tag{2.4.80}$$

leading to

$$|\mathbf{n};+\rangle \;=\; \cos\tfrac{\theta}{2}|+\rangle \;+\; \sin\tfrac{\theta}{2}e^{i\phi}|-\rangle\,. \tag{2.4.81}$$

As a quick check we see that for $\theta = 0$, which corresponds to a unit vector $\mathbf{n} = \mathbf{e}_3$ along the plus z direction we get $|\mathbf{e}_3;+\rangle = |+\rangle$. Note that even though ϕ is ambiguous when $\theta = 0$, this does not affect our answer, since the term with ϕ dependence vanishes. In the same way one can obtain the normalized eigenstate corresponding to $-\hbar/2$. A simple phase choice gives

$$|\mathbf{n};-\rangle \;=\; \sin\tfrac{\theta}{2}|+\rangle \;-\; \cos\tfrac{\theta}{2}e^{i\phi}|-\rangle\,. \tag{2.4.82}$$

If we again consider the $\theta = 0$ direction, this time the ambiguity of ϕ remains in the term that contains the $|z;-\rangle$ state. It is convenient to multiply this state by the phase $-e^{-i\phi}$. Doing this, the pair of eigenstates read[3]

$$\boxed{\begin{aligned} |\mathbf{n};+\rangle \;&=\; \cos\tfrac{\theta}{2}\,|+\rangle \;+\; \sin\tfrac{\theta}{2}e^{i\phi}|-\rangle\,,\\ |\mathbf{n};-\rangle \;&=\; -\sin\tfrac{\theta}{2}e^{-i\phi}|+\rangle \;+\; \cos\tfrac{\theta}{2}|-\rangle\,. \end{aligned}} \tag{2.4.83}$$

The vectors are normalized. Furthermore, they are orthogonal

$$\langle\mathbf{n};-|\mathbf{n};+\rangle = -\sin\tfrac{\theta}{2}e^{i\phi}\cos\tfrac{\theta}{2} + \cos\tfrac{\theta}{2}\sin\tfrac{\theta}{2}e^{i\phi} = 0\,. \tag{2.4.84}$$

Therefore, $|\mathbf{n};+\rangle$ and $|\mathbf{n};-\rangle$ are an orthonormal pair of states.

Let us verify that the $|\mathbf{n};\pm\rangle$ reduce to the known results as $\mathbf{n}$ points along the z, x, and y axes. Again, if $\mathbf{n} = (0,0,1) = \mathbf{e}_3$, we have $\theta = 0$, and hence

$$|\mathbf{e}_3;+\rangle = |+\rangle\,, \quad |\mathbf{e}_3;-\rangle = |-\rangle\,, \tag{2.4.85}$$

which are, as expected, the familiar eigenstates of $\hat{S}_z$. If we point along the x axis, $\mathbf{n} = (1,0,0) = \mathbf{e}_1$ which corresponds to $\theta = \pi/2$, $\phi = 0$. Hence

$$\begin{aligned} |\mathbf{e}_1;+\rangle \;&=\; \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) \;=\; |x;+\rangle\,,\\ |\mathbf{e}_1;-\rangle \;&=\; \frac{1}{\sqrt{2}}(-|+\rangle + |-\rangle) \;=\; -|x;-\rangle\,, \end{aligned} \tag{2.4.86}$$

where we compared with (2.2.30). The overall minus sign in the second state is physically irrelevant, so we recovered the eigenvectors of $\hat{S}_x$. Finally, if $\mathbf{n} = (0,1,0) = \mathbf{e}_2$, we have $\theta = \pi/2$, $\phi = \pi/2$ and hence, with $e^{\pm i\phi} = \pm i$, we have

$$\begin{aligned} |\mathbf{e}_2;+\rangle \;&=\; \frac{1}{\sqrt{2}}(|+\rangle + i|-\rangle) = |y;+\rangle\,,\\ |\mathbf{e}_2;-\rangle \;&=\; \frac{1}{\sqrt{2}}(i|+\rangle + |-\rangle) = i\frac{1}{\sqrt{2}}(|+\rangle - i|-\rangle) = i|y;-\rangle\,, \end{aligned} \tag{2.4.87}$$

which are, up to an immaterial phase for the second one, the eigenvectors of $\hat{S}_y$.

[3]The formula (2.4.83) works nicely at the north pole ($\theta = 0$), but at the south pole ($\theta = \pi$) the ϕ ambiguity shows up again. If one works near the south pole multiplying the results in (2.4.83) by suitable phases will do the job. The fact that no formula works well unambiguously through the full the sphere is a topological property of spin states!