

From S. Banerjee and M. Caluay  
Introduction to Spintronics, Chapter 2.

## 2

---

### *The Quantum Mechanics of Spin*

In the 1920s, the old quantum theory was gradually being superseded by the new quantum theory. The cornerstone of the old theory was Bohr's model of the Hydrogen atom which predicted that an electron cannot orbit the proton in the Hydrogen atom in any arbitrary fashion. Orbits are "quantized", meaning that only certain sizes, shapes and magnetic properties are allowed. The principal quantum number  $n$  determined the allowed radii of the orbits, the orbital quantum number  $l$  determined the allowed shapes, and the magnetic quantum number  $m$  determined the magnetic behavior. Additionally, there is a fourth quantum number  $s$  which denotes the fact that the electron has an additional angular momentum, loosely associated with self rotation about its own axis, and that is quantized in units of  $(1/2)\hbar$ . The old quantum theory was useful to infer the existence of discrete energy levels in atoms, calculate energy spacings between these levels, and therefore allowed one to interpret atomic spectra.

The new quantum theory appeared to be more revolutionary and more powerful. It was triggered by Heisenberg's discovery of *matrix mechanics* and Schrödinger's discovery of *wave mechanics*. These two formalisms would not only predict the quantization of energy and provide a prescription to determine the energy difference between the levels (and thus explain the multiplicity of atomic spectra), but also allow one to calculate easily *probabilities of transitions* between different quantized energy states. At first, matrix mechanics and wave mechanics looked entirely different in their mathematical appearance and physical meaning. However, Schrödinger and Eckart [1] independently showed that the two theories are mathematically equivalent. Towards the end of 1926, Dirac unified the two theories using the concept of state vector and thus established the *transformation theory* of quantum mechanics. This ultimately had a profound implication for the quantum mechanical (mathematical) recipe to treat spin, as we will show in this chapter.

The transformation theory is the mathematical recipe to handle modern quantum mechanics. In Heisenberg's matrix mechanics, a physical quantity is expressed by a matrix, whereas in Schrödinger's wave mechanics, a physical quantity is expressed by a linear operator. In the unified transformation theory, physical quantities are represented by abstract linear operators called Dirac's *q-numbers* which are linear operators in an infinite-dimensional linear space. Depending upon which types of orthogonal coordinate systems are used in this linear space, either matrix mechanics or wave mechanics emerges.

In other words, by using coordinate transformation in this linear space, we can derive matrix mechanics from wave mechanics and vice versa. Therefore, this unified theory was named transformation theory. The state of a quantum mechanical object is represented by a so-called *state vector* which is an abstract vector in this linear space (the “wavefunction” in wave mechanics is an example of this) and the linear space is called the *state space*.

Earlier D. Hilbert and J. von Neumann had introduced the notion of a linear space that could absorb the mathematics of matrices and vectors, as well as the mathematics of linear operators and functions. This so-called *Hilbert space* admitted a finite or denumerably (countably) infinite number of co-ordinate axes. Therefore, a state vector in the Hilbert space could have at most a denumerably infinite number of mutually orthogonal components. Dirac extended this concept to a non-denumerably infinite number of coordinate axes in his linear space via the introduction of his famous  $\delta$ -function [3]. The state vector therefore could have non-denumerably infinite number of mutually orthogonal components and could be expressed as

$$\psi(q), \quad q \in [q_1, q_2], \quad (2.1)$$

where the variable  $q$  is a continuous variable in the domain  $[q_1, q_2]$ . On the other hand, if the coordinate axes are countable, then the state vector would be expressed as

$$\psi_n, \quad n = 1, 2, 3, \dots \quad (2.2)$$

where the variable  $n$  is an integer.

According to Dirac’s transformation theory, the state vector (i) evolves in time according to a unitary transformation, and (ii) satisfies a *first order* differential equation with respect to time. This second property is very important as we shall see later. Depending on whether the physical quantity\* represented by the state vector will yield discrete or continuous values upon measurement, the eigenvalues of the linear operator describing this physical quantity will have discrete or continuous values. Accordingly, the coordinate axes in the linear space will be discrete or continuous, and the state vector will be  $\psi_n$  or  $\psi(q)$ . The magnitude squared of the component of the state vector, i.e.,  $|\psi_n|^2$  or  $|\psi(q)|^2$  gives the probability of the physical quantity to take on the  $n$ -th (or  $q$ -th) value when the quantity is measured. This is the physical interpretation (or significance) of the state vector. Therefore, each component of the state vector is called a *probability amplitude*. The familiar “wavefunction” in the Schrödinger formalism of wave mechanics is the probability amplitude where the physical quantities corresponding to the coordinate axes in linear space are the position coordinates and time, i.e.,

$$\psi(\vec{r}) = \psi(x, y, z, t). \quad (2.3)$$

\*A physical quantity, by definition, is anything that can be measured, even if by a gedanken experiment only.

In wave mechanics, the Schrödinger equation for a single particle tells us how the wavefunction evolves in time and space:

$$i\hbar \frac{\partial \psi(\vec{r})}{\partial t} = H_0 \psi(\vec{r}) . \quad (2.4)$$

If we neglect spin, then

$$\begin{aligned} H_0 &= \frac{|\vec{p}|^2}{2m} + V(\vec{r}) \\ \vec{p} &= p_x \hat{x} + p_y \hat{y} + p_z \hat{z} = -i\hbar \frac{\partial}{\partial x} \hat{x} - i\hbar \frac{\partial}{\partial y} \hat{y} - i\hbar \frac{\partial}{\partial z} \hat{z}, \\ \vec{r} &= [x \hat{x} + y \hat{y} + z \hat{z}], t \end{aligned} \quad (2.5)$$

where the quantities with “hats” are unit vectors along the coordinate axes.

Solution of Equation (2.4) yields the wavefunction  $\psi(\vec{r})$ . The quantity  $H_0$  is the so-called Hamiltonian whose first term is the kinetic energy and the second term is the potential energy. The only restriction is that the potential energy term should be a real quantity so that the Hamiltonian remains a Hermitian operator, which guarantees that its eigenvalue (which is its expected value and therefore the expected value of the energy) remains a real quantity.

The question now is how to include “spin” in Equation (2.4)?

This was investigated by Wolfgang Pauli. He derived an equation to replace Equation (2.4) which bears his name and is known as the *Pauli Equation*. But before we discuss this equation, we need to understand an important concept, namely *Pauli spin matrices* since they appear in the Pauli Equation.

## 2.1 Pauli spin matrices

In quantum mechanics, any physical observable is associated with an operator (which would be a linear operator in the Schrödinger formalism, or a matrix in the Heisenberg formalism). The eigenvalues of the linear operator, or the eigenvalues of the matrix, are the expectation values of the physical quantity, i.e., the values we expect to find if we measure the physical quantity in an experiment<sup>†</sup>. Spin is a physical observable since the associated angular momentum can be measured, as was done unwittingly by Stern and Gerlach. Consequently, there must be a quantum mechanical operator associated with spin. Pauli derived the quantum mechanical operators for the spin components along three orthogonal axes –  $S_x$ ,  $S_y$  and  $S_z$ . They are  $2 \times 2$  complex

<sup>†</sup> Repeated measurements of a physical observable will produce a distribution of values whose average will be the expectation value.

matrices that came to be known as the Pauli spin matrices. Pauli's approach was based on the premise that: (1) the measurement of the spin angular momentum component along any coordinate axis for an electron should give the results  $+\frac{\hbar}{2}$  or  $-\frac{\hbar}{2}$ , and (2) the operators for spin components along three mutually orthogonal axes should obey commutation rules similar to those obeyed by the operators associated with components of the orbital angular momentum. This would put spin angular momentum and orbital angular momentum on the same footing.

The operators (matrices) for the orbital angular momentum are known to satisfy the commutation relations

$$\begin{aligned}L_y L_z - L_z L_y &= i\hbar L_x, \\L_z L_x - L_x L_z &= i\hbar L_y, \\L_x L_y - L_y L_x &= i\hbar L_z,\end{aligned}\tag{2.6}$$

which merely reflect the fact that the orbital angular momenta along any two mutually orthogonal axes cannot be simultaneously measured with absolute precision unless the orbital angular momentum along a third axis, perpendicular to both the other two axes, vanishes.

Pauli adopted similar commutation relations for the spin angular momentum operators  $S_x$ ,  $S_y$  and  $S_z$ :

$$\begin{aligned}S_y S_z - S_z S_y &= i\hbar S_x, \\S_z S_x - S_x S_z &= i\hbar S_y, \\S_x S_y - S_y S_x &= i\hbar S_z.\end{aligned}\tag{2.7}$$

Now, in the Stern-Gerlach experiment, assuming that the  $z$ -axis is the axis joining the south to north pole of the magnet, the observation of two traces on the photographic plate was interpreted as being caused by a spin angular momentum  $\vec{S}$  whose  $z$ -component has two values  $\pm\frac{\hbar}{2}$ . Therefore, the matrix operator  $S_z$  must be (i) a  $2\times 2$  matrix (because such a matrix has *two* eigenvalues), and (ii) these eigenvalues must be  $\pm\frac{\hbar}{2}$ .

A  $2\times 2$  matrix that has eigenvalues of  $\pm\frac{\hbar}{2}$  is the matrix

$$M_{2\times 2} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.\tag{2.8}$$

This is not the only  $2\times 2$  matrix with eigenvalues  $\pm\frac{\hbar}{2}$  - there could be many others - but this is the matrix that Pauli chose as a start for the operator  $S_z$ .

Next, he had to find appropriate matrices to serve as operators  $S_x$  and  $S_y$ . Pauli realized that since the choice of the  $z$ -axis as the axis joining the north and south poles of the magnet is completely arbitrary, the result of the Stern-Gerlach measurement should not be affected if he had chosen this axis to be the  $x$ - or  $y$ -axis, instead. This means that the expectation values of  $S_x$  and

$S_y$ , i.e., their eigenvalues, should also be  $\pm\frac{\hbar}{2}$ . Moreover, all three matrices –  $S_x$ ,  $S_y$  and  $S_z$  – must satisfy the commutation relations in Equation (2.7).

Pauli first defined three dimensionless matrices  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$  such that

$$\begin{aligned} S_x &= \frac{\hbar}{2}\sigma_x, \\ S_y &= \frac{\hbar}{2}\sigma_y, \\ S_z &= \frac{\hbar}{2}\sigma_z. \end{aligned} \quad (2.9)$$

Since  $S_x$ ,  $S_y$  and  $S_z$  must have eigenvalues of  $\pm\frac{\hbar}{2}$ , it is obvious that the  $\sigma$ -matrices must have eigenvalues of  $\pm 1$ . Furthermore, Equation (2.7) mandates that

$$\begin{aligned} \sigma_y\sigma_z - \sigma_z\sigma_y &= 2i\sigma_x, \\ \sigma_z\sigma_x - \sigma_x\sigma_z &= 2i\sigma_y, \\ \sigma_x\sigma_y - \sigma_y\sigma_x &= 2i\sigma_z. \end{aligned} \quad (2.10)$$

According to Equations (2.8) and (2.9),

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.11)$$

So now Pauli needed to pick two matrices  $\sigma_x$  and  $\sigma_y$  such that they have eigenvalues of  $\pm 1$  and obey Equation (2.10). Since these matrices will be operators for physical observables (spin components), they must be Hermitian as well. It is easy to verify that  $\sigma_z$  is Hermitian.

We can start our search for  $\sigma_x$  and  $\sigma_y$  with Hermitian matrices that have off-diagonal elements only, i.e.:

$$\sigma_x = \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix}, \quad (2.12)$$

and

$$\sigma_y = \begin{pmatrix} 0 & b \\ b^* & 0 \end{pmatrix}. \quad (2.13)$$

Since the eigenvalues of these matrices are  $\pm 1$ , we must have  $|a|^2 = |b|^2 = 1$ , which leads to the possible choices for  $a$  and  $b = \pm 1$  or  $\pm i$ .

Next, we must satisfy Equation (2.10) and that mandates

$$\text{Im}(ab^*) = 1, \quad (2.14)$$

where  $\text{Im}$  stands for imaginary part.

Therefore, if we select  $a = +1$ , then we must choose  $b = -i$ , and this yields

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (2.15)$$

and

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (2.16)$$

This is how Pauli came up with expressions for  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$ . These matrices are called *Pauli spin matrices* and serve as operators for the spin components according to Equation (2.9).

It is obvious that Pauli's choice was by no means unique. There are other legitimate choices (e.g., we could have chosen  $a = -i$  and  $b = +1$ ), but Pauli's choice is now history and universally adopted.

From the expressions for the Pauli spin matrices, we notice that the square of each of the Pauli matrices is the  $2 \times 2$  unit matrix  $[I]$ . Hence

$$|S|^2 = S_x^2 + S_y^2 + S_z^2 = \frac{3}{4}\hbar^2[I] = \bar{s}(\bar{s} + 1)\hbar^2[I], \quad (2.17)$$

with  $\bar{s} = 1/2$ . This should be compared with the equivalent relation for the orbital angular momentum operator

$$|L|^2 = m(m + 1)\hbar^2[I], \quad m = 1, 2, 3, \dots \quad (2.18)$$

### 2.1.1 Eigenvectors of the Pauli matrices: spinors

The eigenvalues of the Pauli spin matrices are  $\pm 1$ . We now evaluate the corresponding eigenvectors that we denote as  $|\pm\rangle$ .

**Matrix  $\sigma_z$ :** The eigenvectors of  $\sigma_z$  must satisfy

$$\sigma_z|\pm\rangle_z = \pm 1|\pm\rangle_z. \quad (2.19)$$

These eigenvectors (with unit norm) will be

$$|+\rangle_z = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad (2.20)$$

and

$$|-\rangle_z = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.21)$$

It is easy to verify that these two eigenvectors are orthonormal, as they must be since they are eigenvectors of a Hermitian matrix corresponding to distinct (non-degenerate) eigenvalues.

**Matrix  $\sigma_x$ :** The eigenvectors of  $\sigma_x$  must satisfy

$$\sigma_x|\pm\rangle_x = \pm 1|\pm\rangle_x. \quad (2.22)$$

Starting with Equation (2.15), these eigenvectors are found to be

$$|+\rangle_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad (2.23)$$

and

$$|-\rangle_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (2.24)$$

Once again, the two eigenvectors are orthonormal. As can be easily checked, these eigenvectors can also be expressed as

$$|\pm\rangle_x = \frac{1}{\sqrt{2}} [ |+\rangle_z \pm |-\rangle_z ]. \quad (2.25)$$

**Matrix  $\sigma_y$ :** The eigenvectors of  $\sigma_y$  must satisfy

$$\sigma_y |\pm\rangle_y = \pm 1 |\pm\rangle_y. \quad (2.26)$$

Using Equation (2.16), these eigenvectors are found to be

$$|+\rangle_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad (2.27)$$

and

$$|-\rangle_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (2.28)$$

These eigenvectors are orthonormal and can be expressed as

$$|\pm\rangle_y = \frac{1}{\sqrt{2}} [ |+\rangle_z \pm i |-\rangle_z ]. \quad (2.29)$$

The eigenvectors of the Pauli spin matrices are examples of "spinors" which are  $2 \times 1$  column vectors that represent the spin state of an electron. If we know the spinor associated with an electron in a given state, we can deduce the electron's spin orientation, i.e., find the quantities  $\langle S_x \rangle$ ,  $\langle S_y \rangle$  and  $\langle S_z \rangle$ , where the angular brackets  $\langle \dots \rangle$  denote expectation values. We will see this later.

## 2.2 The Pauli Equation and spinors

We can absorb the space and time dependent part of an electron's wavefunction in the spinor, so that the general form of a spinor will be

$$[\psi(\mathbf{x})] = \begin{bmatrix} \phi_1(\mathbf{x}) \\ \phi_2(\mathbf{x}) \end{bmatrix}, \quad (2.30)$$

where  $\mathbf{x} \equiv (x, y, z, t)$ , and  $\phi_1$  and  $\phi_2$  are the two components of the spinor wavefunction (assumed to be properly normalized).

With a 2-component wavefunction, the Schrödinger equation must be recast as

$$\left\{ [H] + \frac{\hbar}{i} \frac{\partial}{\partial t} [I] \right\} [\psi(\mathbf{x})] = [0], \quad (2.31)$$

where the Hamiltonian is a  $2 \times 2$  matrix (since it may contain the  $2 \times 2$  Pauli spin matrices),  $[I]$  is the  $2 \times 2$  identity matrix, and  $[0]$  is the  $2 \times 1$  null vector. Equation (2.31) is a set of two simultaneous differential equations for the two components of the spinor wavefunction –  $\phi_1$  and  $\phi_2$ . Equation (2.31) is referred to as the *Pauli Equation*.

Solution of the Pauli Equation yields the two-component spinor wavefunction  $[\psi(\mathbf{x})]$ . Its practical use is in calculating the expected value of the spin angular momentum of an electron along any coordinate axis. The expected value along the  $n$ -th coordinate axis at location  $(\vec{r}_0 = x_0, y_0, z_0)$  at an instant of time  $t$  will be  $[\psi(\vec{r}_0, t)]^\dagger [S_n] [\psi(\vec{r}_0, t)]$ , where  $S_n = (\hbar/2)\sigma_n$  and the superscript  $\dagger$  (dagger) represents the Hermitian conjugate. Using Equation (2.9), we get

$$\begin{aligned} S_x(\vec{r}_0, t) &= (\hbar/2) [\psi(\vec{r}_0, t)]^\dagger [\sigma_x] [\psi(x_0, y_0, z_0, t)] \\ &= (\hbar/2) [\phi_1^*(\vec{r}_0, t) \ \phi_2^*(\vec{r}_0, t)] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \phi_1(\vec{r}_0, t) \\ \phi_2(\vec{r}_0, t) \end{bmatrix} \\ &= \hbar \text{Re} [\phi_1^*(\vec{r}_0, t) \phi_2(\vec{r}_0, t)], \\ S_y(\vec{r}_0, t) &= (\hbar/2) [\psi(\vec{r}_0, t)]^\dagger [\sigma_y] [\psi(\vec{r}_0, t)] \\ &= (\hbar/2) [\phi_1^*(\vec{r}_0, t) \ \phi_2^*(\vec{r}_0, t)] \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} \phi_1(\vec{r}_0, t) \\ \phi_2(\vec{r}_0, t) \end{bmatrix} \\ &= \hbar \text{Im} [\phi_1^*(\vec{r}_0, t) \phi_2(\vec{r}_0, t)], \\ S_z(\vec{r}_0, t) &= (\hbar/2) [\psi(\vec{r}_0, t)]^\dagger [\sigma_z] [\psi(\vec{r}_0, t)] \\ &= (\hbar/2) [\phi_1^*(\vec{r}_0, t) \ \phi_2^*(\vec{r}_0, t)] \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \phi_1(\vec{r}_0, t) \\ \phi_2(\vec{r}_0, t) \end{bmatrix} \\ &= (\hbar/2) \{ |\phi_1(\vec{r}_0, t)|^2 - |\phi_2(\vec{r}_0, t)|^2 \}, \end{aligned} \quad (2.32)$$

where *Re* stands for the real part, *Im* stands for the imaginary part and the superscript \* (asterisk) represents complex conjugate.

Therefore, if we can find the 2-component wavefunction in Equation (2.30) by solving the Pauli Equation (2.31), then we can find the three components of the expected value of the spin angular momentum at any location at any instant of time. This is why the Pauli Equation and the spinor concept are useful and important. In Chapter 7, we will also show how the Pauli Equation can be used to derive the energy dispersion relations (relation between the energy and the wavevector) of an electron in a solid in the presence of spin-dependent effects.



### 2.7.3 Applications of the Postulates of Quantum Mechanics to a few spin problems

The Postulates of Quantum Mechanics are briefly reviewed in Chapter 15.

**Example 1:** If we measure the  $z$ -component of an electron's spin, apply the postulate of quantum projective measurement (Postulate 3) discussed in Chapter 15 to calculate the probability of the measurement to give the result  $\frac{\hbar}{2}$  or  $-\frac{\hbar}{2}$ , respectively, if prior to the measurement, the state of the system  $|\psi\rangle$  is either

- (1)  $|0\rangle$ ,
- (2)  $|1\rangle$ , or
- (3)  $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ ,

where  $|0\rangle$  is the  $+z$ -polarized state  $|+\rangle_z$  and  $|1\rangle$  is the  $-z$ -polarized state  $|-\rangle_z$ .

#### Solution

(1) The operator  $S_z = \frac{\hbar}{2}\sigma_z$  has the following spectral decomposition

$$S_z = \frac{\hbar}{2}|0\rangle\langle 0| + \left(-\frac{\hbar}{2}\right)|1\rangle\langle 1| = \sum_m m P_m, \quad (2.87)$$

where  $P_m = |m\rangle\langle m|$ , with  $|m\rangle$  being the eigenvectors of  $S_z$ .

Hence, if  $|\psi\rangle = |0\rangle$ , the probability to measure  $+\frac{\hbar}{2}$  is equal to

$$p\left(+\frac{\hbar}{2}\right) = \langle 0|(|0\rangle\langle 0|)|0\rangle = 1, \quad (2.88)$$

and the probability to measure  $-\frac{\hbar}{2}$  is equal to

$$p\left(-\frac{\hbar}{2}\right) = \langle 0|(|1\rangle\langle 1|)|0\rangle = 0 \quad (2.89)$$

and the sum of the probabilities is indeed equal to unity.

(2) If  $|\psi\rangle = |1\rangle$ , we get

$$p\left(\frac{\hbar}{2}\right) = \langle 1|(|0\rangle\langle 0|)|1\rangle = 0, \quad (2.90)$$

and

$$p\left(-\frac{\hbar}{2}\right) = \langle 1|(|1\rangle\langle 1|)|1\rangle = 1. \quad (2.91)$$

(3) Finally, if  $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ , we find

$$p\left(+\frac{\hbar}{2}\right) = \frac{1}{\sqrt{2}}\{|0\rangle + \langle 1|\}\langle 0|\rangle \frac{1}{\sqrt{2}}\{|0\rangle + \langle 1|\} = \frac{1}{2}\{1+0\} \cdot \{0+1\} = \frac{1}{2}, \quad (2.92)$$

and similarly,  $p(-\frac{\hbar}{2}) = \frac{1}{2}$ .

According to *Postulate 3* discussed in Chapter 15, right after the measurement, the spinor collapses into the state

$$|\psi\rangle \rightarrow |\psi^{new}\rangle = \frac{(|0\rangle\langle 0|)}{\sqrt{p(+\frac{\hbar}{2})}} |\psi\rangle = \left(\frac{|0\rangle\langle 0|}{\sqrt{\frac{1}{2}}}\right) \left(\frac{(|0\rangle + |1\rangle)}{\sqrt{2}}\right) = |0\rangle. \quad (2.93)$$

**Exercise:** Repeat the previous exercise if the component  $S_x$  is measured instead.

**Example 2:** Suppose an electron is prepared in the spinor  $|0\rangle$  eigenstate of  $S_z$  with eigenvalue  $+\frac{\hbar}{2}$  and repeated measurements are made of the  $x$ -component of its intrinsic angular momentum, calculate the average value  $\langle S_x \rangle$  and the standard deviation  $\Delta(S_x)$  of these measurements.

**Solution**

The spectral decomposition of the operator  $S_x$  is given by

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \left(+\frac{\hbar}{2}\right) |+\rangle_x \langle +| + \left(-\frac{\hbar}{2}\right) |-\rangle_x \langle -|, \quad (2.94)$$

where  $|+\rangle_x$  is the  $+x$ -polarized state and  $|-\rangle_x$  is the  $-x$ -polarized state.

Hence,

$$\langle S_x \rangle = \frac{\hbar}{2} \left[ \langle 0| S_x |0\rangle = (1 \ 0) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] = 0. \quad (2.95)$$

Furthermore,

$$\begin{aligned} \Delta(S_x) &= \sqrt{\langle S_x^2 \rangle - \langle S_x \rangle^2} = \sqrt{\langle S_x^2 \rangle}, \\ &= \frac{\hbar}{2} \left[ \langle 0| \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} |0\rangle \right]^{1/2}, \\ &= \frac{\hbar}{2} \left[ (1 \ 0) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]^{1/2} = \frac{\hbar}{2}. \end{aligned} \quad (2.96)$$

**Example 3:** Suppose an electron is characterized by the spinor

$$|\psi\rangle = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad (2.97)$$

which is properly normalized, as easily checked. If we measure the  $y$ -component of the spin, what is the probability to find that its value is  $\frac{\hbar}{2}$ ?

**Solution**

The eigenvector of  $\sigma_y$  are  $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$  with eigenvalue +1 and  $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$  with eigenvalue -1

Hence, if we write the spinor  $|\psi\rangle$  as a linear combination of  $|+\rangle_y$  and  $|-\rangle_y$ ,

$$|\psi\rangle = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \alpha \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} + \beta \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (2.98)$$

The probability to find  $+\frac{\hbar}{2}$  when measuring  $S_y$  is given by  $|\alpha|^2$  or

$$p(+\frac{\hbar}{2}) = |\langle + | \psi \rangle|^2 = \left| \frac{1}{\sqrt{2}} (1, -i) \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right|^2 = \frac{1}{20} |(3+i)|^2 = \frac{1}{2}. \quad (2.99)$$

**Example 4:** Suppose an electron is characterized by the spinor

$$|\psi\rangle = \frac{4}{5}|0\rangle + \frac{3}{5}|1\rangle. \quad (2.100)$$

(1) What is the probability that a measurement of z-component of the spin will be  $+\frac{\hbar}{2}$ ? and  $-\frac{\hbar}{2}$ ?

**Solution**

$$p(+\frac{\hbar}{2}) = \left| \langle 0 | \left( \frac{4}{5}|0\rangle + \frac{3}{5}|1\rangle \right) \right|^2 = \frac{16}{25}, \quad (2.101)$$

and

$$p(-\frac{\hbar}{2}) = \left| \langle 1 | \left( \frac{4}{5}|0\rangle + \frac{3}{5}|1\rangle \right) \right|^2 = \frac{9}{25}. \quad (2.102)$$

(2) What is the expectation value of  $S_z = \frac{\hbar}{2}\sigma_z$ ?

**Solution**

$$\begin{aligned} \langle \psi | S_z | \psi \rangle &= \left( \frac{4}{5} \langle 0 | + \frac{3}{5} \langle 1 | \right) \frac{\hbar}{2} \sigma_z \left( \frac{4}{5} |0\rangle + \frac{3}{5} |1\rangle \right), \\ &= \frac{16}{25} \langle 0 | \sigma_z | 0 \rangle \frac{\hbar}{2} + \left( \frac{3}{5} \right)^2 \frac{\hbar}{2} \langle 1 | \sigma_z | 1 \rangle = \frac{16}{25} \left( \frac{\hbar}{2} \right) - \frac{9}{25} \left( \frac{\hbar}{2} \right), \\ &= \frac{7}{25} \left( \frac{\hbar}{2} \right). \end{aligned} \quad (2.103)$$

(3) What is the standard deviation of  $S_z$  if measurements are made on many electrons prepared in the state  $|\psi\rangle$  above?

**Solution**

By definition

$$\Delta(S_z) = \sqrt{\langle \psi | S_z^2 | \psi \rangle - \langle S_z \rangle^2}. \quad (2.104)$$

Since

$$\langle S_z^2 \rangle = \langle \psi | S_z^2 | \psi \rangle = \langle \psi | \frac{\hbar^2}{4} \sigma_z^2 | \psi \rangle = \frac{\hbar^2}{4}, \quad (2.105)$$

and from the previous problem

$$\langle S_z \rangle = \frac{7}{25} \left( \frac{\hbar}{2} \right), \quad (2.106)$$

we get

$$\Delta(S_z) = \sqrt{\frac{\hbar^2}{4} \left( 1 - \frac{49}{625} \right)} = \frac{12}{25} \hbar. \quad (2.107)$$

**Exercise:** Proceeding as above, calculate the standard deviations  $\Delta(S_x)$  and  $\Delta(S_y)$ .

**Answer:**

$$\Delta(S_x) = \frac{7}{25} \frac{\hbar}{2}, \quad (2.108)$$

and

$$\Delta(S_y) = \frac{\hbar}{2}. \quad (2.109)$$

**2.7.4 The Heisenberg Principle for spin components**

In Chapter 15, we prove the following general Heisenberg inequality

$$\Delta(C)\Delta(D) \geq \frac{|\langle \psi | [C, D] | \psi \rangle|}{2}, \quad (2.110)$$

where  $\Delta(C)$  and  $\Delta(D)$  are the standard deviations associated with measurements of the observables  $C$  and  $D$  on an electron prepared many times in the same state  $|\psi\rangle$ .

If we take as observables the two components of the intrinsic angular momentum of the electron  $\sigma_x$  and  $\sigma_y$ , the inequality above becomes

$$\Delta(S_x)\Delta(S_y) \geq \frac{\hbar}{2} |\langle \psi | S_z | \psi \rangle| = \frac{\hbar^2}{4} |\langle \psi | \sigma_z | \psi \rangle|. \quad (2.111)$$

**Exercise:** Prove that, in the last inequality, the equality signs holds when the state of the spinor is given by

$$|\psi\rangle = \frac{3}{\sqrt{10}} |0\rangle + \frac{-1}{\sqrt{10}} |1\rangle, \quad (2.112)$$

considered in Example 3.

### Solution

For an electron prepared in the state  $|\psi\rangle$  above, we use the results of Example 3 and find right-hand-side of Equation (2.111) to be equal to

$$\frac{\hbar^2}{4} |\langle \psi | \sigma_z | \psi \rangle| = \frac{\hbar^2}{5}. \quad (2.113)$$

The left hand side of the inequality is

$$\Delta(S_x)\Delta(S_y) = \left(\frac{2}{5}\hbar\right)\left(\frac{\hbar}{2}\right) = \frac{\hbar^2}{5}. \quad (2.114)$$

Therefore, the inequality (2.111) reduces to an equality in this case.

**Exercise:** The following inequalities also hold

$$\Delta(S_y)\Delta(S_z) \geq \frac{\hbar}{2} |\langle \psi | S_x | \psi \rangle| = \frac{\hbar^2}{4} |\langle \psi | \sigma_x | \psi \rangle|, \quad (2.115)$$

and

$$\Delta(S_x)\Delta(S_z) \geq \frac{\hbar}{2} |\langle \psi | S_y | \psi \rangle| = \frac{\hbar^2}{4} |\langle \psi | \sigma_y | \psi \rangle|. \quad (2.116)$$

Prove that these inequalities are indeed satisfied for a spinor in the state (2.112) given above. The solution of this exercise is left for the reader.

## 2.8 References

- [1] C. Eckart, "Operator Calculus and the Solution of the Equations of Dynamics", *Phys. Rev.*, **28**, 711 (1926).
- [2] R. P. Feynman, R. B. Leighton and M. Sands, *The Feynman Lectures on Physics*, Volume II, chapter 34, Addison-Wesley, Reading, Mass., (1965).
- [3] P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford University Press (1958).
- [4] J. J. Sakurai, *Advanced Quantum Mechanics*, Addison-Wesley, Reading, Mass. (1967).
- [5] E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra*, Cambridge University Press, Cambridge (1935).

### 2.7.2 Two useful theorems

A trivial decomposition of any  $2 \times 2$  matrix  $M$  is obviously

$$M = m_{11} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + m_{12} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + m_{21} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + m_{22} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad (2.74)$$

since the 4 matrices on the right hand side form a complete basis for all  $2 \times 2$  matrices. Now that we have introduced the Pauli spin matrices, a more subtle decomposition of any  $2 \times 2$  complex matrix can be found, as discussed next.

**Theorem I:** Any  $2 \times 2$  matrix  $M$

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \quad (2.75)$$

can be decomposed as follows

$$M = \frac{m_{11} + m_{22}}{2} I + \frac{m_{11} - m_{22}}{2} \sigma_z + \frac{m_{12} + m_{21}}{2} \sigma_x + \frac{m_{12} - m_{21}}{2} \sigma_y. \quad (2.76)$$

The proof of which is left as an exercise. In other words, the 4 matrices  $(I, \sigma_x, \sigma_y, \sigma_z)$  form a complete set of bases in the space of  $2 \times 2$  complex matrices.

The last equation can be written in the more condensed form

$$M = a_0 I + \vec{a} \cdot \vec{\sigma}, \quad (2.77)$$

where

$$a_0 = \frac{1}{2} Tr(M), \quad (2.78)$$

and

$$\vec{a} = \frac{1}{2} Tr(M \vec{\sigma}), \quad (2.79)$$

where  $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ , and  $Tr$  stands for the trace of the matrix.

A comparison of Equations (2.76) and (2.77) shows that  $M$  is Hermitian if  $a_0$  and the three components of the vector  $\vec{a}$  are real.

---

**Exercise:** Calculate the values of  $(a_0, \bar{a})$  for the matrix  $M$  given by

$$M = \begin{pmatrix} 2 & \frac{i\sqrt{3}}{3} \\ \frac{-i\sqrt{3}}{3} & 4 \end{pmatrix}. \quad (2.80)$$


---

Next, we prove an identity which will be used in the next chapter to interpret geometrically the Pauli matrices after the introduction of the Bloch sphere concept.

**Theorem II:** If  $\theta$  is real and if the matrix  $A$  is such that  $A^2 = I$ , the following identity holds:

$$e^{i\theta A} = \cos\theta I + i\sin\theta A. \quad (2.81)$$

This is the generalization to operators of the well-known Euler relation for complex numbers, i.e.,  $e^{iz} = \cos z + i\sin z$ .

From the Taylor series expansion

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad (2.82)$$

and the definition of the function of an operator, we get

$$e^{i\theta A} = I + (i\theta)A + \frac{(i\theta)^2 A^2}{2!} + \frac{(i\theta)^3 A^3}{3!} + \frac{(i\theta)^4 A^4}{4!} + \dots \quad (2.83)$$

or

$$\begin{aligned} e^{i\theta A} &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots + (-1)^k \frac{\theta^{2k}}{(2k)!}\right) I \\ &+ i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots + (-1)^k \frac{i\theta^{2k+1}}{(2k+1)!}\right) A, \end{aligned} \quad (2.84)$$

which is indeed Equation (2.81) if we use the the Taylor expansions

$$\sin x = \sum_{k=0}^{\infty} (-1)^k x^{2k+1} / (2k+1)!, \quad (2.85)$$

and

$$\cos x = \sum_{k=0}^{\infty} (-1)^k x^{2k} / (2k)!. \quad (2.86)$$