Chapter 3

Principles and Spin

3.1 Four Principles of Quantum Mechanics

There are many axiomatic presentations of quantum mechanics that aim at conceptual economy and rigor. However, it would be unwise for us to get into that, as our present aim is merely to learn how to use the mathematical machinery of quantum mechanics. For this reason, we might as well start with a run of the mill account that aims neither at intellectual economy nor at great rigor.[1] With this in mind, we are now in a position to state four tenets of quantum mechanics that will allow us to study the simplest non-trivial quantum mechanical case, namely spin-half systems.

1. The state of a system (every physical object or collection of such objects) is described by the state vector $|\Psi\rangle$.
2. An observable (a measurable property) O is represented by a Hermitian operator $\hat{O}$.
3. Given an observable O represented by the Hermitian operator $\hat{O}$, and a system in a state represented by $|\Psi\rangle = \sum_n c_n |\psi_n\rangle$, where $|\psi_n\rangle$ are the orthonormal eigenvectors of $\hat{O}$ (that is, $\hat{O}|\psi_n\rangle = \lambda_n |\psi_n\rangle$, where λ_n is the eigenvalue) spanning the space, upon

[1] A run of the mill account is, of course, infected (if that is the word) by the standard interpretation, which we shall explicitly introduce at the end of chapter 5. However, the standard interpretations is, well, standard and therefore we might as well start, if somewhat surreptitiously, with something like it.

measurement of *O*, one always obtains one of the λ_n, and the probability of obtaining λ_n is $|c_n|^2$, the square of the modulus of the expansion coefficient of $|\psi_n\rangle$.[2]

4. Once the measurement of *O* returns λ_n, the state vector is instantaneously transformed from $|\Psi\rangle$ into $|\psi_n\rangle$, the eigenvector corresponding to the eigenvalue λ_n.

Let us briefly comment on these principles.

In order to be successful, a physical theory must give us predictions (which may be statistical) on what returns we shall obtain if we measure a property *O* (an observable, in quantum mechanical jargon) on physical systems in some quantum state or other. Consequently, quantum mechanics must contain the representation of the quantum state of any physical system on which we intend to perform a measurement and that of any observable we intend to measure. Principles (1) and (2) do just that. Principle (1) tells us that all the information that quantum mechanics has with respect to a system's physical state is encapsulated in a vector $|\Psi\rangle$, exactly the mathematical object we just studied. For reasons that will become clear shortly $|\Psi\rangle$ must be normalized. Principle (2) states that the theoretical counterpart of an observable is a Hermitian operator.[3]

[2] For any expression *x*, $|x|^2 = x^* x$. That is, to obtain the squared modulus of an expression, one multiplies the expression times its complex conjugate.

[3] It is worth noticing that although to every observable there corresponds a Hermitian operator, whether the converse is true is a different matter. For example, D'Espagnat has argued that it not the case that to every Hermitian operator there corresponds an observable. See D'Espagnat, B., (1995): 98-9.

Principle (3) tells us what to expect if we perform a measurement of an observable *O* on a system in a state represented by $|\Psi\rangle$. Since an observable is represented by a Hermitian operator, we can take the orthonormal eigenvectors $|\psi_n\rangle$ of $\hat{O}$ as basis vectors of the vector space. Consequently, we can express $|\Psi\rangle$ as a linear combination of $\hat{O}$'s eigenvectors so that $|\Psi\rangle = \sum_n c_n|\psi_n\rangle$. The measurement return is always an eigenvalue λ_n of $\hat{O}$ with respect to a $|\psi_n\rangle$; the probability of obtaining λ_n is $|c_n|^2$, the square of the modulus of the expansion coefficient of $|\psi_n\rangle$.[4] For example, suppose that $\hat{O}$ has only two eigenvectors, so that $\hat{O}\psi_1 = \lambda_1\psi_1$ and $\hat{O}\psi_2 = \lambda_2\psi_2$. Then $|\Psi\rangle = c_1\psi_1 + c_2\psi_2$, and upon measuring *O*, we shall obtain λ_1 with probability $|c_1|^2$ and λ_2 with probability $|c_2|^2$.[5]

Since upon measuring *O* we must obtain one of the λ_n, it must be the case that

---

[4] In reality, because of a quantum mechanical result (the energy-time uncertainty, so called), that measurement returns are always eigenvalues is not quite true. However, statistically they oscillate in predictable ways around eigenvalues; see appendix 6. Note that it follows from (2.9.6) that $|c_n|^2 = |\langle\psi_n|\Psi\rangle|^2$, where $\langle\psi_n|$ is the bra of the ket $|\psi_n\rangle$. That is, the probability of obtaining λ_n is the square of the magnitude of the modulus of the inner product between $|\psi_n\rangle$, the eigenvector for λ_n with respect to $\hat{O}$, and the pre-measurement state vector $|\Psi\rangle$.

[5] At the cost of being pedantic, we should note that principle (3) does not ask us to apply $\hat{O}$ to $|\Psi\rangle$ but to decompose $|\Psi\rangle$ into $\hat{O}$'s eigenvectors, something that can always be done because $\hat{O}$ is Hermitian.

$$\sum_n |c_n|^2 = 1. \tag{3.1.1}$$

In other words, the sum of the probabilities of all the possible measurement returns must be equal to one. Keeping in mind that the norm of a vector $|\Psi\rangle$ is $\sqrt{\langle\Psi|\Psi\rangle}$, where $\langle\Psi|\Psi\rangle = \sum_n |c_n|^2$, one can easily see that (3.1.1) is satisfied if and only if $|\Psi\rangle$ is normalized.

Principle (4) embodies what is called "the collapse" or "the reduction" of the state vector. As we shall see, it is a most mysterious feature of the theory. Note that the combination of principles (3) and (4) guarantees that if a measurement of *O* returns λ_n, a new measurement of *O*, if performed sufficiently quickly (before the system has started evolving again), will return the same eigenvalue λ_n. For, principle (4) tells us that the first measurement results in the collapse of $|\Psi\rangle$ onto $|\psi_n\rangle$, so that $|\Psi\rangle$ is now an eigenvector of $\hat{O}$, which entails that $c_n = 1$, and consequently principle (3) tells us that the probability of obtaining λ_n upon re-measuring is $|c_n|^2 = 1$.[6]

3.2 Spin

We can decompose the total spin *S* of a quantum particle into its three components S_x, S_y, and S_z, one for each dimension. As we noted, measuring one of them randomizes the measurement returns of the others, and for this reason S_x, S_y, and S_z, are called "incompatible observables". In quantum mechanics, this is expressed by the fact

[6] Suppose we leave the system alone instead of performing a measurement on it. What will its temporal evolution be? The answer to this central question will concern us in the next chapter. For now, however, let us put it on the back burner.

that their corresponding operators $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ do not commute. In fact, the following relations hold:

$$\left[\hat{S}_x,\hat{S}_y\right]=i\hbar\hat{S}_z;\ \left[\hat{S}_y,\hat{S}_z\right]=i\hbar\hat{S}_x;\ \left[\hat{S}_z,\hat{S}_x\right]=i\hbar\hat{S}_y.^7 \tag{3.2.1}$$

By contrast, S^2, the square of the total spin, is compatible with each of the spin components, and consequently $\hat{S}^2$ and $\hat{S}_z$ commute. Since they are commuting Hermitian operators, their common set of eigenvectors spans the space, and therefore can be used as an orthonormal basis. When studying spin, it is customary to use exactly such basis, which is called the "standard basis". (We shall see later how to apply the operators $\hat{S}_x$ and $\hat{S}_y$ when the state vector is expressed in the standard basis). If we measure a spin component, S_z, for example, the possible measurement returns depend on a number *s*, the spin quantum number. It turns out that every species of elementary particle has an unchangeable spin quantum number, confusingly called "the spin" of that species. For example, protons, neutrons, and electrons, the particles constituting ordinary matter, have spin 1/2 (spin-half), while photons have spin 1.

Spin-half is the simplest case of spin: $\hat{S}^2$ and $\hat{S}_z$ have just two eigenvectors, namely, $|\downarrow_z\rangle$, called "*z*-spin down", and $|\uparrow_z\rangle$, called "*z*-spin up". The eigenvector equations are:

$$\hat{S}^2|\uparrow_z\rangle=\frac{3}{4}\hbar^2|\uparrow_z\rangle,\hat{S}^2|\downarrow_z\rangle=\frac{3}{4}\hbar^2|\downarrow_z\rangle \tag{3.2.2}$$

and

$$\hat{S}_z|\uparrow_z\rangle=\frac{1}{2}\hbar|\uparrow_z\rangle,\ \hat{S}_z|\downarrow_z\rangle=-\frac{1}{2}\hbar|\downarrow_z\rangle. \tag{3.2.3}$$

---

[7] To obtain the other commutators, note that $[\hat{A},\hat{B}]=-[\hat{B},\hat{A}]$.

Since $\hat{S}^2$ and $\hat{S}_z$ are Hermitian and commute, we can use $\{|\uparrow_z\rangle, |\downarrow_z\rangle\}$ as a basis, and in this basis

$$|\uparrow_z\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ and } |\downarrow_z\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.2.4}$$

Hence, the general state $|\Psi\rangle$ of a spin-half particle is a linear combination of the basis vectors:

$$|\Psi\rangle = c_1|\uparrow_z\rangle + c_2|\downarrow_z\rangle = c_1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}. \tag{3.2.5}$$

Now let us find the matrix forms of $\hat{S}^2$ and $\hat{S}_z$ in the standard basis. As they must be conformable with their eigenvectors, they are 2×2 matrices. Moreover, they are diagonalized because expressed with respect to the basis $\{|\uparrow_z\rangle, |\downarrow_z\rangle\}$ made up of their eigenvectors; in other words, all their elements will be zeros, with the exception of the ones making up the main diagonal, which will be the eigenvalues of the operator with respect to the basis vectors. Hence,

$$\hat{S}^2 = \begin{pmatrix} \frac{3}{4}\hbar & 0 \\ 0 & \frac{3}{4}\hbar \end{pmatrix} = \frac{3}{4}\hbar\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.2.6}$$

Similarly,

$$\hat{S}_z = \begin{pmatrix} \frac{\hbar}{2} & 0 \\ 0 & -\frac{\hbar}{2} \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.2.7}$$

Discovering the matrices of $\hat{S}_x$ and $\hat{S}_y$ in the standard basis is a bit more complicated; suffice it to say that they are

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.2.8}$$

and

$$\hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{3.2.9}$$

Since we know the matrices for $\hat{S}_x, \hat{S}_y, \hat{S}_z$, and $\hat{S}^2$, we can now perform a few calculations.

3.3 Measuring Spin

Consider a spin-half particle in state $|\Psi\rangle = c_1|\uparrow_z\rangle + c_2|\downarrow_z\rangle$.[8] Suppose now that we want to predict the results of a measurement of S_z on the particle. As we know from principle (3), there are only two possible returns, $\hbar/2$ (the eigenvalue associated with $|\uparrow_z\rangle$) and $-\hbar/2$ (the eigenvalue associated with $|\downarrow_z\rangle$). Moreover, the probability of getting $\hbar/2$ is $|c_1|^2$ and that of getting $-\hbar/2$ is $|c_2|^2$.

EXAMPLE 3.3.1

A spin-half particle is in state $|\Psi\rangle = \frac{1}{\sqrt{3}}\begin{pmatrix} 1-i \\ 1 \end{pmatrix}$. We can easily infer that upon measuring S_z the probability of obtaining $\hbar/2$ is $\left|\frac{1-i}{\sqrt{3}}\right|^2 = \frac{2}{3}$, while the probability of getting $-\hbar/2$ is $\left|\frac{1}{\sqrt{3}}\right|^2 = \frac{1}{3}$.[9] Suppose we have obtained $\hbar/2$. What shall we get if we

---

[8] Every time we say that particle is in some state, we assume that the corresponding vector is normalized.

[9] Remember that $|a+b|^2 = (a+b)^*(a+b)$. Note also that the sum of the two probabilities is one, as it should be.

immediately measure S_z again? As we know from the combination of principles (3) and (4), we shall surely get $\hbar/2$, as $|\Psi\rangle$ has collapsed onto $|\uparrow_z\rangle$.

Suppose now that the particle is in the generic state $|\Psi\rangle = c_1|\uparrow_z\rangle + c_2|\downarrow_z\rangle = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$, and we measure S_x. What shall we get? Finding the answer is relatively simple, albeit somewhat laborious. Basically, we need to obtain the information necessary to apply principle (3). Hence, first we must determine the (normalized) eigenvectors $|\uparrow_x\rangle$, $|\downarrow_x\rangle$, and the eigenvalues λ_1, λ_2 of S_x. Since S_x is Hermitian, its eigenvectors span the space. Hence, we can expand $|\Psi\rangle$ into a linear combination of the normalized eigenvectors of S_x, so that $|\Psi\rangle = \alpha|\uparrow_x\rangle + \beta|\downarrow_x\rangle$. Consequently, upon measuring S_x, we shall get λ_1 with probability $|\alpha|^2$, and λ_2 with probability $|\beta|^2$. So, let us proceed.

Since $\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, if we take $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ as the generic eigenvector and λ as the generic eigenvalue, we have

$$\begin{pmatrix} 0 & \frac{\hbar}{2} \\ \frac{\hbar}{2} & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{3.3.1}$$

that is, a system of two simultaneous equations

$$\begin{cases} \frac{\hbar}{2}x_2 - \lambda x_1 = 0 \\ \frac{\hbar}{2}x_1 - \lambda x_2 = 0 \end{cases} \tag{3.3.2}$$

The characteristic equation is

$$0 = \begin{vmatrix} -\lambda & \frac{\hbar}{2} \\ \frac{\hbar}{2} & -\lambda \end{vmatrix} = \lambda^2 - \frac{\hbar^2}{4} , \tag{3.3.3}$$

which yields two eigenvalues, $\lambda_1 = -\hbar/2$, and $\lambda_2 = \hbar/2$. By plugging the former into either of (3.3.2), we get $x_1 = -x_2$. By taking $x_1 = 1$, we obtain the eigenvector for S_x with eigenvalue $-\hbar/2$, namely $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Similarly, by plugging $\lambda_2 = \hbar/2$ into either (3.3.2) we obtain $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

However, the two eigenvectors must be normalized in order for them to represent real spin states. As we know, to normalize a vector $|A\rangle$, we must divide it by its norm $\sqrt{\langle A|A\rangle}$, and in an orthonormal basis $\langle A|A\rangle = |a_1|^2 + |a_2|^2 + ... + |a_n|^2$, where $a_1 ..., a_n$ are the components of the vector. Applying this to vector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, we obtain

$$|\downarrow_x\rangle = \frac{1}{\sqrt{|1|^2 + |-1|^2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{3.3.4}$$

Similarly,

$$|\uparrow_x\rangle = \frac{1}{\sqrt{|1|^2 + |1|^2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{3.3.5}$$

Now let us apply the equation $|c_n|^2 = |\langle \psi_n | \Psi \rangle|^2$, introduced in note 4, p. 59. As $|\uparrow_x\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, the corresponding bra is $\frac{1}{\sqrt{2}} (1 \quad 1)$. Hence, the probability of obtaining $S_x = \hbar/2$ is

$$\Pr(S_x = \hbar/2) = \left| \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \right|^2 = \frac{|c_1 + c_2|^2}{2}, \tag{3.3.6}$$

and that of obtaining $S_x = -\hbar/2$ is

$$\Pr(S_x = -\hbar/2) = \left| \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \right|^2 = \frac{|c_1 - c_2|^2}{2}. \tag{3.3.7}$$

The eigenvalues, eigenvectors, and related probabilities for $\hat{S}_y$ can be found in a similar fashion. The following table gives the operators, eigenvalues, eigenvectors, and the related probability formulas for $\hat{S}_x, \hat{S}_y$, and $\hat{S}_z$ in the standard basis.[10]

Operator	**Eigenvalue**	**Eigenvector**	**Pr(eigenvalue)**
$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{\hbar}{2}\sigma_x$	$\frac{\hbar}{2}$	$\vert\uparrow_x\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\frac{\vert c_1 + c_2\vert^2}{2}$
$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{\hbar}{2}\sigma_x$	$-\frac{\hbar}{2}$	$\vert\downarrow_x\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$	$\frac{\vert c_1 - c_2\vert^2}{2}$
$\hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{\hbar}{2}\sigma_y$	$\frac{\hbar}{2}$	$\vert\uparrow_y\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$	$\frac{\vert c_1 - ic_2\vert^2}{2}$
$\hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{\hbar}{2}\sigma_y$	$-\frac{\hbar}{2}$	$\vert\downarrow_y\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}$	$\frac{\vert c_1 + ic_2\vert^2}{2}$
$\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\hbar}{2}\sigma_z$	$\frac{\hbar}{2}$	$\vert\uparrow_z\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\vert c_1\vert^2$
$\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\hbar}{2}\sigma_z$	$-\frac{\hbar}{2}$	$\vert\downarrow_z\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\vert c_2\vert^2$

Note that every time we measure one of the components of spin of a particle of spin-half, we obtain $\hbar/2$ or $-\hbar/2$.

---

[10] $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, and $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ are the famous Pauli operators. Note that if we use $\hbar/2$ as unit of measure when measuring spin, then the spin operators reduce to their Pauli operators.

EXAMPLE 3.3.2

Let a particle be in state $|\Psi\rangle = \frac{1}{\sqrt{9}}\begin{pmatrix} 2-i \\ 2 \end{pmatrix}$, and suppose we make just one measurement. The table above contains all the information needed to determine what returns we may expect and their associated probabilities. If we measure S_z, we shall get $\hbar/2$ with probability

$$\Pr(S_z = \hbar/2) = \frac{|2-i|^2}{9} = \frac{5}{9}, \tag{3.3.8}$$

and $-\hbar/2$ with probability

$$\Pr\left(S_z = -\frac{\hbar}{2}\right) = \frac{|2|^2}{9} = \frac{4}{9}. \tag{3.3.9}$$

If we measure S_x, we shall obtain $\hbar/2$ with probability

$$\Pr(S_x = \hbar/2) = \frac{1}{2}\left|\frac{2-i}{\sqrt{9}} + \frac{2}{\sqrt{9}}\right|^2 = \frac{1}{2}\frac{|4-i|^2}{9} = \frac{17}{18}, \tag{3.3.10}$$

and $-\hbar/2$ with probability

$$\Pr\left(S_x = -\frac{\hbar}{2}\right) = \frac{1}{2}\left|\frac{2-i}{\sqrt{9}} - \frac{2}{\sqrt{9}}\right|^2 = \frac{1}{2}\frac{|-i|^2}{9} = \frac{1}{18}. \tag{3.3.11}$$

If we measure S_y, we shall get $\hbar/2$ with probability

$$P(S_y = \hbar/2) = \frac{1}{2}\left|\frac{2-i}{\sqrt{9}} - \frac{2i}{\sqrt{9}}\right|^2 = \frac{1}{2}\frac{|2-3i|^2}{9} = \frac{13}{18}, \tag{3.3.12}$$

and $-\hbar/2$ with probability

$$\Pr\left(S_y = -\frac{\hbar}{2}\right) = \frac{1}{2}\left|\frac{2-i}{\sqrt{9}} + \frac{2i}{\sqrt{9}}\right|^2 = \frac{1}{2}\frac{|2-i|^2}{9} = \frac{5}{18}. \tag{3.3.13}$$

What happens if we perform two quick successive measurements of the same observable? As we know, once we have carried out a measurement the state vector collapses, and consequently if the system did not have time to evolve, a quick new measurement will return the same value. What happens if we measure first S_x and quickly after S_z? Upon measuring S_x, we shall get $\hbar/2$ with probability 17/18, and $-\hbar/2$ with probability 1/18. Suppose we get $\hbar/2$. Then the state vector will collapse onto $|\uparrow_x\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$. Hence, a measurement of S_z will yield $\hbar/2$ or $-\hbar/2$, both with probability 1/2.[11]

EXAMPLE 3.3.3

We can now start to make sense of some of the peculiar results of the spin-half experiment discussed in chapter one. Since $|\uparrow_z\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_x\rangle + |\downarrow_x\rangle\right)$, half the particles will exit the SGX device (chapter 1, figure 4) in state $|\uparrow_x\rangle$ and half in state $|\downarrow_x\rangle$. In addition, when we measure S_x on either path, $|\uparrow_z\rangle$ collapses onto $|\uparrow_x\rangle$ or $|\downarrow_x\rangle$; however, $|\uparrow_x\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle + |\downarrow_z\rangle\right)$ and $|\downarrow_x\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle - |\downarrow_z\rangle\right)$, and therefore in either case when we measure S_z on electrons on path C we obtain $\hbar/2$ or $-\hbar/2$, both with probability 1/2. Similarly, if we block path A, only $|\downarrow_x\rangle$ electrons will reach path C, and upon measuring S_z we obtain the same result.[12]

3.4 Series of Stern-Gerlach Devices

[11] As we noted before, measuring S_x randomizes the return values for S_z. As we shall see, this is a manifestation of the Generalized Uncertainty Principle.

[12] We still do not know why we obtain the same results if we measure position instead of spin before the electrons reach C. We shall address this problem later.

Let us look at some examples in which Stern-Gerlach devices are in series.

EXAMPLE 3.4.1

Now let us consider the following arrangement of Stern-Gerlach devices, and let us shoot a beam of spin-half particles in state $|\uparrow_x\rangle$ through it (Fig. 1).

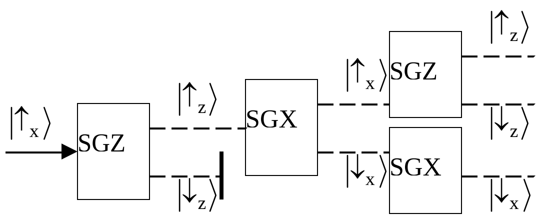

Figure 1

Half the particles will exit the first SGZ in state $|\uparrow_z\rangle$ and half in state $|\downarrow_z\rangle$. So, only a probabilistic prediction can be given for the behavior of any given particle. Let us now block the spin down particles and send the spin up particles into a Stern-Gerlach apparatus oriented to measure the x-component of spin (SGX). A quick look at the table shows that half of the particles will exit in state $|\uparrow_x\rangle$ and half in state $|\downarrow_x\rangle$. Again, only a probabilistic prediction can be given for the behavior of any given particle. If we send the $|\downarrow_x\rangle$ particles through another SGX apparatus, all of them will exit with x-spin down. Here, we can actually predict the behavior of any particle because each is in the eigenstate $|\downarrow_x\rangle$. If we send the $|\uparrow_x\rangle$ particles through a SGZ apparatus, half of them will come out with z-spin up and half with z-spin down; again, only a probabilistic prediction can be given for the behavior of any given particle.

EXAMPLE 3.4.2

Consider now the following setting with a stream of spin-half particles in state $|\uparrow_x\rangle$ (Fig. 2).

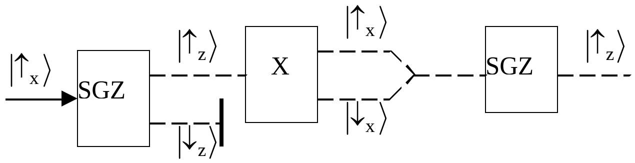

Figure 2

Here the device marked as X is not quite a SGX device because it lacks the detector. In other words, as the particles go though it, *no* measurement takes place. If we want, we can expand the state of the particles entering X as

$$|\uparrow_z\rangle = \frac{1}{\sqrt{2}}\left(|\downarrow_x\rangle + |\uparrow_x\rangle\right). \tag{3.4.1}$$

However, as there is no measurement, there is no collapse of the state function that, therefore, remains unaltered as the particles enter the second SGZ device. To determine what we are going to observe, we need to express the state vector of the particles in the standard basis, so that we obtain

$$\frac{1}{\sqrt{2}}\left(|\downarrow_x\rangle + |\uparrow_x\rangle\right) = |\uparrow_z\rangle. \tag{3.4.2}$$

Hence, all the particles will exit the second SGZ device in state $|\uparrow_z\rangle$ and therefore upon measurement, we obtain $S_z = \hbar/2$.[13]

[13] Although we can manipulate the relevant mathematical apparatus successfully and make correct predictions, one can fairly say that nobody really knows much about the nature of spin. To be sure, the existence of spin can be inferred from experiment and from relativistic quantum mechanical considerations based on the assumption that angular momentum is conserved. However, we have no concrete picture of it since a classically rotating electron would have a surface linear velocity far exceeding the speed of light. There is general agreement that spin is an intrinsic property of quantum particles, but its

3.5 Expectation Values

How do we test that quantum mechanics handles spin-half correctly? Obviously, not by experimenting on one particle. The reason is not merely that in general it is a good idea to perform many experiments to make sure that our measurements are accurate, but that in order to test a probabilistic prediction we need a very large number of identical particles all in the very same state $|\Psi\rangle$ on which we perform the exact same measurement. Such a collection of particles is called an "ensemble." For example, if an electron is in state

$$|\Psi\rangle = \frac{1}{\sqrt{9}}\begin{pmatrix} 2-i \\ 2 \end{pmatrix} \tag{3.5.1}$$

and we measure S_z, quantum mechanics predicts that we shall get $\hbar/2$ with probability 5/9 and $-\hbar/2$ with probability 4/9. Hence, if we perform the measurement on an ensemble of 900,000 electrons in that state and we obtain $\hbar/2$ 500,000 times and $-\hbar/2$ 400,000 times, then we may consider the prediction verified. A way of finding out whether these proportions are in fact borne out experimentally is by checking out the average of the measurement returns for S_z, and for that we need a brief statistical detour.

Suppose we have a cart with seven boxes in it. Two boxes weigh 5*kg* each; three weigh 10*kg* each; one weighs 40*kg*, and one 60*kg*. Let us denote the number of boxes with *N* and the number of boxes of weight *w* with $N(w)$. Obviously,

---

cause is unknown. Perhaps spin originates from the internal structure of quantum particles, although electrons seem to have no components, or perhaps it is a fundamental property not amenable to any explanation; its nature remains mysterious.

$$N = \sum_w N(w), \tag{3.5.2}$$

that is, $7 = N(5kg) + N(10kg) + N(40kg) + N(60kg)$. Suppose now that we pick a box randomly out of the cart. Let us denote the probability of picking a box of weight w with $\Pr(w)$. Then,

$$\Pr(w) = \frac{N(w)}{N}. \tag{3.5.3}$$

For example, since there are 3 boxes weighing 10*kg* and 7 boxes in all, $\Pr(10kg) = 3/7$. Notice now that

$$\Pr(5kg) + \Pr(10kg) + \Pr(40kg) + \Pr(60kg) = 1, \tag{3.5.4}$$

and in general,

$$\sum_w \Pr(w) = 1, \tag{3.5.5}$$

which is another way of saying that the probability of picking a box of any weight out of the cart is one.

Suppose now that we want to determine $<w>$, the average or *mean* weight of the boxes in the cart. Obviously, we add the weights of all the boxes, and divide by the number of boxes. So,

$$< w >= \frac{\sum_w wN(w)}{N}. \tag{3.5.6}$$

For example,

$$< w >= \frac{1}{N}(5 \cdot 2 + 10 \cdot 3 + 40 \cdot 1 + 60 \cdot 1)kg = 20kg, \tag{3.5.7}$$

and therefore the average weight of a box is 20*kg*. [14] Let us now notice that

[14] Notice, by the way, that no box weighs 20*kg*!

$$\frac{\sum_w wN(w)}{N} = \sum_w w \frac{N(w)}{N}. \tag{3.5.8}$$

Consequently, by plugging (3.5.8) into (3.5.6) and using (3.5.3), we obtain

$$< w >= \sum_w w \Pr(w).^{15} \tag{3.5.9}$$

If we apply (3.5.6) to the previous case, we obtain:

$$< w >= \left(5\frac{2}{7} + 10\frac{3}{7} + 40\frac{1}{7} + 60\frac{1}{7}\right)kg = 20kg. \tag{3.5.10}$$

In quantum mechanics, the average is generally the quantity of interest, and it is called "the expectation value". It is a very misleading, but ingrained, terminology because it strongly suggests that this quantity is the most likely outcome of a single measurement. But that is the most probable value, *not* the average value. For example, the most probable outcome of weighing (measuring) a box randomly picked from the cart is 10*kg*, since $P(10kg)$ is the highest. By contrast, the expectation value is the average value, that is, 20*kg*.

We can now go back to our quantum mechanical problem. Recall that we wanted to determine the predicted expectation value for S_z for a particle in the state given by (3.5.1). By applying (3.5.8) we obtain

$$< S_z >= \frac{\hbar}{2}\frac{5}{9} - \frac{\hbar}{2}\frac{4}{9} = \frac{\hbar}{18}. \tag{3.5.11}$$

[15] This is a special case of the important general law $< f(w) >= \sum_{w=0}^{\infty} f(w)\Pr(w)$, where $f(w)$ is any function of *w*.

Hence, if the actual measurement returns satisfy (3.5.11), the prediction will be borne out.

There is a nice formula for determining the expectation value of an observable *O*, namely,

$$< O >= \left\langle \Psi \middle| \hat{O}\Psi \right\rangle; \tag{3.5.12}$$

in other words, we sandwich the relevant operator in an inner product between the vector states of the system. We leave the proof of (3.5.12) as an exercise.

EXAMPLE 3.5.1

Consider again an ensemble of particles in state $\left|\Psi\right\rangle = \frac{1}{\sqrt{9}}\begin{pmatrix} 2-i \\ 2 \end{pmatrix}$. Let us calculate the expectation value for S_z. By sandwiching, we obtain

$$< S_z >= \left\langle \Psi \middle| \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \Psi \right\rangle = \begin{pmatrix} \frac{2+i}{3} & \frac{2}{3} \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{2-i}{3} \\ \frac{2}{3} \end{pmatrix}, \tag{3.5.13}$$

once we remember that a bra is the complex conjugate transpose of its ket. Then,

$$< S_z >= \frac{\hbar}{2} \begin{pmatrix} \frac{2+i}{3} & -\frac{2}{3} \end{pmatrix} \begin{pmatrix} \frac{2-i}{3} \\ \frac{2}{3} \end{pmatrix} = \frac{\hbar}{2}\left(\frac{5}{9} - \frac{4}{9}\right) = \frac{\hbar}{18}. \tag{3.5.14}$$

We can easily verify that this result jibes with the results obtained in the previous example.

Exercises

Exercise 3.1

1. Let $|\Psi\rangle = \frac{1}{\sqrt{11}}\begin{pmatrix} i \\ 3i+1 \end{pmatrix}$ be the state of a spin-half particle. Determine: a) the probability of obtaining $\hbar/2$ upon measuring S_x; b) the probability of obtaining $\hbar/2$ upon measuring S_z.
2. Let $|\Psi\rangle = \frac{1}{\sqrt{15}}\begin{pmatrix} 3+i \\ 2-i \end{pmatrix}$ be the state of a spin-half particle. Determine: a) the probability of obtaining $-\hbar/2$ upon measuring S_y; b) the probability of obtaining $\hbar/2$ upon measuring S_z.
3. Determine the eigenvalues, the normalized eigenvectors, and the related probabilities for $\hat{S}_y$.

Exercise 3.2

1. Consider an ensemble of particles in state $|\Psi\rangle = \frac{1}{\sqrt{15}}\begin{pmatrix} 3+i \\ 2-i \end{pmatrix}$. Determine the expectation value for S_z.
2. Consider an ensemble of particles in state $|\Psi\rangle = \frac{1}{\sqrt{13}}\begin{pmatrix} 3i \\ 2 \end{pmatrix}$. Determine the expectation value for S_x

Exercise 3.3

1. A stream of spin-half particles in state $|\Psi\rangle = \frac{1}{\sqrt{15}}\begin{pmatrix} 3+i \\ 2-i \end{pmatrix}$ enters a SGX device. Provide the probabilities and states of the out-coming particle streams.

2. A stream of spin-half particles in state $\left|\Psi\right\rangle = \frac{1}{\sqrt{13}}\begin{pmatrix} 3i \\ 2 \end{pmatrix}$ enters a SGY device. Provide the probabilities and states of the out-coming particle streams.

3. Prove that $<O> = \left\langle \Psi \middle| \hat{O}\Psi \right\rangle$ for a two-dimensional vector space. [Hint: $\hat{O}$ is Hermitian; therefore $\left\langle \Psi \middle| \hat{O}\Psi \right\rangle = \left\langle \hat{O}\Psi \middle| \Psi \right\rangle$ and $\left|\Psi\right\rangle$ can be expressed in terms of $\hat{O}$'s eigenvectors.]

Answers to the Exercises

Exercise 3.1

1a: $\Pr\left(S_x=\frac{\hbar}{2}\right)=\frac{1}{11}\frac{1}{2}\left|i+3i+1\right|^2=\frac{1}{22}(4i+1)(-4i+1)=\frac{17}{22}$.

1b: $\Pr\left(S_z=\frac{\hbar}{2}\right)=\frac{1}{11}(i)(-i)=\frac{1}{11}$.

2a: $\Pr\left(S_y=-\frac{\hbar}{2}\right)=\frac{1}{15}\frac{1}{2}\left|4+3i\right|^2=\frac{1}{30}(4+3i)(4-3i)=\frac{5}{6}$.

2b: $\Pr\left(S_z=\frac{\hbar}{2}\right)=\frac{1}{15}(3+i)(3-i)=\frac{2}{3}$.

3. We start by determining eigenvalues and eigenvectors of $\hat{S}_y=\frac{\hbar}{2}\begin{pmatrix}0 & -i\\ -i & 0\end{pmatrix}$.

$$\frac{\hbar}{2}\begin{pmatrix}0 & -i\\ -i & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\lambda\begin{pmatrix}x_1\\ x_2\end{pmatrix}\Rightarrow\begin{cases}\lambda x_1+\frac{\hbar i}{2}x_2=0\\ -\frac{\hbar i}{2}x_1-\lambda x_2=0\end{cases}\Rightarrow 0=\begin{vmatrix}\lambda & \frac{\hbar i}{2}\\ -\frac{\hbar i}{2} & -\lambda\end{vmatrix}=-\lambda^2+\left(\frac{\hbar}{2}\right)^2.$$ Hence,

$\lambda_1=-\hbar/2$, and $\lambda_2=\hbar/2$. For λ_1, we obtain $x_2=-ix_1$, and therefore the eigenvector $\left|X_{\lambda_1}\right\rangle=x_1\begin{pmatrix}1\\ -i\end{pmatrix}$; since $\left\langle X_{\lambda_1}\middle|X_{\lambda_1}\right\rangle=1+1=2$, the normalized eigenvector is $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix}$. For λ_2 we obtain $x_2=ix_1$, and therefore the eigenvector $\left|X_{\lambda_2}\right\rangle=x_1\begin{pmatrix}1\\ i\end{pmatrix}$; since $\left\langle X_{\lambda_2}\middle|X_{\lambda_2}\right\rangle=1+1=2$, the normalized eigenvector is $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix}$. Since $\hat{S}_y$ is Hermitian, its eigenvectors span the space, and therefore the standard $\left|sm\right\rangle$ vector can be expressed as a linear combination of them. Hence, $\begin{pmatrix}a\\ b\end{pmatrix}=c_1\left|\downarrow_y\right\rangle+c_2\left|\uparrow_y\right\rangle=\frac{1}{\sqrt{2}}\begin{pmatrix}c_1\\ -ic_1\end{pmatrix}+\frac{1}{\sqrt{2}}\begin{pmatrix}c2\\ ic_2\end{pmatrix}$. This is equivalent to the following two simultaneous equations: $\begin{cases} a = \dfrac{c_1 + c_2}{\sqrt{2}} \\ b = \dfrac{i(c_2 - c_1)}{\sqrt{2}} \end{cases}$. The top equation yields

$c_1 = a\sqrt{2} - c_2$, which we plug into the bottom equation to obtain $c_2 = \dfrac{b + ai}{i\sqrt{2}} = \dfrac{a - ib}{\sqrt{2}}$.

Hence, $c_1 = a\sqrt{2} - \dfrac{a}{\sqrt{2}} + \dfrac{ib}{\sqrt{2}} = \dfrac{a + ib}{\sqrt{2}}$. Consequently, we have

$|\Psi\rangle = \dfrac{a+ib}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} + \dfrac{a-ib}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$, the state vector expressed in terms of the eigenvectors of $\hat{S}_y$.

So, the probability of obtaining $\lambda_1 = -\hbar/2$ is $\dfrac{|a+ib|^2}{2}$, and that of obtaining $\lambda_2 = \hbar/2$ is

$\dfrac{|a-ib|^2}{2}$.

Exercise 3.2

1. $< S_z >= \dfrac{\hbar}{2}\begin{pmatrix} \dfrac{3-i}{\sqrt{15}} & \dfrac{2+i}{\sqrt{15}} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \dfrac{3+i}{\sqrt{15}} \\ \dfrac{2-i}{\sqrt{15}} \end{pmatrix} = \dfrac{\hbar}{2}\begin{pmatrix} \dfrac{3-i}{\sqrt{15}} & -\dfrac{2+i}{\sqrt{15}} \end{pmatrix}\begin{pmatrix} \dfrac{3+i}{\sqrt{15}} \\ \dfrac{2-i}{\sqrt{15}} \end{pmatrix} = \dfrac{\hbar}{6}$.

2. $< S_x >= \dfrac{\hbar}{2}\begin{pmatrix} \dfrac{-3i}{\sqrt{13}} & \dfrac{2}{\sqrt{13}} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \dfrac{3i}{\sqrt{13}} \\ \dfrac{2}{\sqrt{13}} \end{pmatrix} = \dfrac{\hbar}{2}\begin{pmatrix} \dfrac{2}{\sqrt{13}} & \dfrac{-3i}{\sqrt{13}} \end{pmatrix}\begin{pmatrix} \dfrac{3i}{\sqrt{13}} \\ \dfrac{2}{\sqrt{13}} \end{pmatrix} = 0$.

Exercise 3.3

1. Two particle streams come out. The first is composed of particles in state $|\uparrow_x\rangle$ with spin $S_x = \hbar/2$. By using the table, we find that the probability that a particle belongs to this stream is $\dfrac{1}{15}\dfrac{|3+i+2-i|^2}{2} = \dfrac{5}{6}$. The second particle stream is composed of particles

in state $|\downarrow_x\rangle$ with spin $S_x = -\hbar/2$. By using the table, we find that the probability that a particle belongs to this stream is $\frac{1}{15}\frac{|3+i-2+i|^2}{2} = \frac{1}{6}$. Note that, of course, $\frac{5}{6}+\frac{1}{6}=1$.

2. Two particle streams come out. The first is composed of particles in state $|\uparrow_y\rangle$ with spin $S_y = \hbar/2$. By using the table, we find that the probability that a particle belongs to this stream is $\frac{1}{13}\frac{|3i-2i|^2}{2} = \frac{1}{26}$. The second particle stream is composed of particles in state $|\downarrow_y\rangle$ with spin $S_y = -\hbar/2$. By using the table, we find that the probability that a particle belongs to this stream is $\frac{1}{13}\frac{|3i-2i|^2}{2} = \frac{25}{26}$.

3. Let $|\Psi\rangle = c_1|e_1\rangle + c_2|e_2\rangle$, where $\{|e_1\rangle,|e_2\rangle\}$ is the basis made up of the eigenvectors of $\hat{O}$. Since $\hat{O}$ is Hermitian,

$$\left\langle \Psi \middle| \hat{O}\Psi \right\rangle = \left\langle \hat{O}\Psi \middle| \Psi \right\rangle = \left\langle \hat{O}\left(c_1\langle e_1| + c_2\langle e_2|\right) \middle| \Psi \right\rangle = \left\langle \left(c_1\lambda_1\langle e_1| + c_2\lambda_2\langle e_2|\right) \middle| \Psi \right\rangle,$$

where λ_1 and λ_2 are $\hat{O}$'s eigenvalues. But

$$\left\langle \left(c_1\lambda_1\langle e_1| + c_2\lambda_2\langle e_2|\right) \middle| \Psi \right\rangle = c_1^*\lambda_1\langle e_1|\Psi\rangle + c_2^*\lambda_2\langle e_2|\Psi\rangle = c_1^*\lambda_1 c_1 + c_2^*\lambda_2 c_2 = \lambda_1|c_1|^2 + \lambda_2|c_2|^2.$$

However, $\lambda_1|c_1|^2 + \lambda_2|c_2|^2 = \lambda_1 \Pr(\lambda_1) + \lambda_2 \Pr(\lambda_2) = \langle O \rangle$.

The extension to higher dimensional spaces is immediate.