Chapter 5

Putting It All Together

In this chapter we acquire the last piece of quantum machinery needed to explain, albeit in general terms, the spin-half and the double slit experiments we considered in the first chapter.

5.1 Tensor Product

There are various ways of multiplying a matrix by another. Up to now, we have considered the Cayley product. Now we need to introduce a new type of product, the *tensor*, or Kronecker, product. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}, \qquad (5.1.1)$$

and B be any other $p \times q$ matrix. Then the tensor product of A and B is

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \ldots & a_{1m}B \\ a_{21}B & a_{22}B & \ldots & a_{2m}B \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1}B & a_{n2}B & \ldots & a_{nm}B \end{pmatrix}. \qquad (5.1.2)$$

For example, let

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \qquad (5.1.3)$$

Then,

$$A \otimes B = \begin{pmatrix} 1\begin{pmatrix} 1 \\ 2 \end{pmatrix} & 2\begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ 3\begin{pmatrix} 1 \\ 2 \end{pmatrix} & 1\begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 1 \\ 6 & 2 \end{pmatrix}. \qquad (5.1.4)$$

If A and B have the same dimensions, then their tensor product is distributive with respect to matrix addition, that is,

$$A \otimes (B + C) = (A \otimes B) + (A \otimes C),$$

and (5.1.5)

$$(B + C) \otimes A = (B \otimes A) + (C \otimes A).$$

Moreover,

$$(A \otimes B)(C \otimes D) = AC \otimes BD. \qquad (5.1.6)$$

Since vectors and operators can be represented by matrices, they can be multiplied tensorially. Hence, given two vectors $|A\rangle$ and $|B\rangle$, we can construct a new vector $|A\rangle \otimes |B\rangle$, at times symbolized by $|AB\rangle$. By extension, vector spaces can be multiplied tensorially as well. So, if V and W are vector spaces, then $Z = V \otimes W$ is the vector space whose vector elements are the tensor products of the basis vectors $|v_1\rangle,...,|v_n\rangle$ of V with the basis vectors $|w_1\rangle,...,|w_m\rangle$ of W and the linear combinations thereof. It turns out that $\{|v_1\rangle \otimes |w_1\rangle, |v_1\rangle \otimes |w_2\rangle ..., |v_n\rangle \otimes |w_m\rangle\}$ is a basis for W. [1]

5.2 Systems of Many Particle and Particles of Many Properties

Consider two particles 1 and 2. Particle 1 is described by $|\Psi_1\rangle$ in state space H_1 and particle 2 by $|\Psi_2\rangle$ in state space H_2. To describe the composite system made up of 1 and 2, we need to construct a new state space H, which is the tensor product of H_1 and H_2 so that $H = H_1 \otimes H_2$. If $\{|e_1\rangle,...,|e_n\rangle\}$ is a basis for H_1, and $\{|d_1\rangle,...,|d_m\rangle\}$ a basis for H_2, then the set of vectors $\{|e_1\rangle \otimes |d_1\rangle, |e_1\rangle \otimes |d_2\rangle ..., |e_n\rangle \otimes |d_m\rangle\}$ is a basis for H. Hence,

[1] In other words, the collection of the tensor products of the basis vectors of the two spaces is a basis of the new space

any vector in H can be expressed as a linear combination of members of that basis. Clearly, the state space H has dimensions $n \times m$.

Suppose now that O_1 is an observable for 1 represented in H_1 by the operator $\hat{O}_1$, and O_2 an observable for 2 represented in H_2 by the operator $\hat{O}_2$. Then, in state space H, O_1 is represented by $\hat{O}_1 \otimes I_2$, the extension of $\hat{O}_1$, where I is the identity operator in H_2; similarly, O_2 is represented by $I_1 \otimes \hat{O}_2$, the extension of $\hat{O}_2$, where I is the identity operator in H_1. Note that since $\hat{O}_1$ and $\hat{O}_2$ belong to different state spaces, they commute. It turns out that $\hat{O}_1 \otimes I_2$ and $I_1 \otimes \hat{O}_2$ are Hermitian and commute as well. In addition, extending an operator does *not* change its spectrum: $\hat{O}_1$ and $\hat{O}_1 \otimes I_2$ have the same eigenvalues, and so do $\hat{O}_2$ and $I_1 \otimes \hat{O}_2$.[2] It is customary, if slightly confusing, to represent an operator and its extension with the same symbol, so that, for example, "$\hat{O}_1$" stands for either $\hat{O}_1 \otimes I_2$ or $\hat{O}_1$.

To fix our ideas, let us look at a system of two spin-half particles. Let $\left(S^1\right)^2$ and S_z^1 be the square of the spin magnitude and the spin z-component of particle (1) in state space H_1. As we know, their operators commute, and the eigenvectors $\left|\uparrow_z^1\right\rangle$ and $\left|\downarrow_z^1\right\rangle$

---

[2] It is also worth noting that given that in state space H_1 the generic eigenvalue equation for $\hat{O}_1$ is $\hat{O}_1|\psi_n\rangle = \alpha_n|\psi_n\rangle$ and that for $\hat{O}_2$ in H_2 is $\hat{O}_2|\chi_m\rangle = \beta_m|\chi_m\rangle$, the generic eigenvalue equation for $\hat{O}_1 + \hat{O}_2$ in H is $\left(\hat{O}_1 + \hat{O}_2\right)|\psi_n\rangle \otimes |\chi_m\rangle = (\alpha_n + \beta_m)|\psi_n\rangle \otimes |\chi_m\rangle$. In other words, the eigenvalues of the sum of the operators are the sum of the eigenvalues of the operators.

can be used as a basis in H_1. Similarly, let $\left(S^2\right)^2$ and S_z^2 be the square of the spin magnitude and the spin z-component of particle (2) in state space H_2. As before, these operators commute, and the eigenvectors $\left|\uparrow_z^2\right\rangle$ and $\left|\downarrow_z^2\right\rangle$ can be used as a basis in H_2. The state space of the system is $H = H_1 \otimes H_2$, and its basis is composed by four vectors:

$$\left|\uparrow_z^1\right\rangle \otimes \left|\uparrow_z^2\right\rangle, \text{ abbreviated as } \left|\uparrow_z^1\uparrow_z^2\right\rangle, \tag{5.2.1}$$

$$\left|\uparrow_z^1\right\rangle \otimes \left|\downarrow_z^2\right\rangle, \text{ abbreviated as } \left|\uparrow_z^1\downarrow_z^2\right\rangle, \tag{5.2.2}$$

$$\left|\downarrow_z^1\right\rangle \otimes \left|\uparrow_z^2\right\rangle, \text{ abbreviated as } \left|\downarrow_z^1\uparrow_z^2\right\rangle, \tag{5.2.3}$$

$$\left|\downarrow_z^1\right\rangle \otimes \left|\downarrow_z^2\right\rangle, \text{ abbreviated as } \left|\downarrow_z^1\downarrow_z^2\right\rangle, \tag{5.2.4}$$

so that every state vector of the system can be expressed in the general form

$$|\Psi\rangle = c_1\left|\uparrow_z^1\uparrow_z^2\right\rangle + c_2\left|\uparrow_z^1\downarrow_z^2\right\rangle + c_3\left|\downarrow_z^1\uparrow_z^2\right\rangle + c_4\left|\downarrow_z^1\downarrow_z^2\right\rangle. \tag{5.2.5}$$

Since S^1 and S^2 belong to different state spaces, they commute, and therefore their extensions to $H = H_1 \otimes H_2$ commute as well. By the same token, every component of S^1 (that is, S_z^1, and so on) commutes with every component of S^2, and the same is true of their extensions.[3] We are now in a position to determine some measurement results, and so let us look at some examples.

EXAMPLE 5.2.1

By simple inspection of $|\Psi\rangle$ in (5.2.5), we can see that if we simultaneously measure S_z^1 and S_z^2 we shall get $S_z^1 = \hbar/2$ and $S_z^2 = \hbar/2$ with probability $|c_1|^2$, $S_z^1 = \hbar/2$ and $S_z^2 = -\hbar/2$ with probability $|c_2|^2$, and so on.

---

[3] Of course, as before, the spin of a particle is incompatible with its components, and the components themselves are incompatible with each other.

EXAMPLE 5.2.2

Now, let us simultaneously measure S_x^1 and S_z^2. What is the probability that we obtain $\hbar/2$ in both cases? The eigenvector in H_1 corresponding to $\hbar/2$ for S_x^1 is $|\uparrow_x^1\rangle$; that in H_2 corresponding to $\hbar/2$ for S_z^2 is $|\uparrow_z^2\rangle$. Hence, the eigenvector in H corresponding to $\hbar/2$ for both S_x^1 and S_z^2 is $|\uparrow_x^1\rangle \otimes |\uparrow_z^2\rangle$, the tensor product of $|\uparrow_x^1\rangle$ and $|\uparrow_z^2\rangle$.

Now,

$$|\uparrow_x^1\langle = \frac{1}{\sqrt{2}}\left[|\uparrow_z^1\langle + |\downarrow_z^1\langle\right], \tag{5.2.6}$$

and therefore

$$|\uparrow_x^1\rangle \otimes |\uparrow_z^2\rangle = \frac{1}{\sqrt{2}}\left[|\uparrow_z^1\rangle + |\downarrow_z^1\rangle\right] \otimes |\uparrow_z^2\rangle. \tag{5.2.7}$$

Distributing, we obtain

$$|\uparrow_x^1\rangle \otimes |\uparrow_z^2\rangle = \frac{1}{\sqrt{2}}\left[|\uparrow_z^1\rangle \otimes |\uparrow_z^2\rangle + |\downarrow_z^1\rangle \otimes |\uparrow_z^2\rangle\right]. \tag{5.2.8}$$

Remembering that $|c_n|^2 = |\langle\psi_n|\Psi\rangle|^2$, we have

$$\Pr\left(S_x^1 = S_z^2 = \frac{\hbar}{2}\right) = \frac{1}{2}\left|\left\langle\left(\langle\uparrow_z^1\uparrow_z^2| + \langle\downarrow_z^1\uparrow_z^2|\right)\middle|\Psi\right\rangle\right|^2. \tag{5.2.9}$$

Distributing, we get

$$\Pr\left(S_x^1 = S_z^2 = \frac{\hbar}{2}\right) = \frac{1}{2}\left[\left|\langle\uparrow_z^1\uparrow_z^2|\Psi\rangle\right| + \left|\langle\downarrow_z^1\uparrow_z^2|\Psi\rangle\right|\right]^2 = \frac{|c_1 + c_3|^2}{2}. \tag{5.2.10}$$

The collapse postulate can be easily extended to many particle systems. To fix our ideas, let us consider the state vector given by (5.2.5). Suppose we measure S_z^1 and obtain $\hbar/2$. To determine the state vector onto which $|\Psi\rangle$ collapses, we throw out all the terms not containing S_z^1's eigenvectors for $\hbar/2$, and then we normalize the resulting vector. This is the vector onto which $|\Psi\rangle$ collapses. So, in our example we obtain

$$|\Psi'\rangle = \alpha|\uparrow_z^1\uparrow_z^2\rangle + \beta|\uparrow_z^1\downarrow_z^2\rangle, \tag{5.2.11}$$

which we suppose normalized.

What was said about many particle systems applies also to single particles with more than one degree of freedom. If *n* independent physical properties are needed fully to describe the state of a system S, then S has *n* degrees of freedom. For example, a particle moving along the *x*-axis has one degree of freedom; one moving on a two-dimensional surface has two; an electron orbiting the nucleus has five: one for each dimension of space, one for orbital angular momentum, and one for spin. Since the properties required to describe S are independent, they are represented in different vector spaces, and therefore commute. Their treatment is analogous to that of multiple particle systems. One constructs the tensor product of the vector spaces, the operator extensions, and so on.

5.3 Accounting for the Spin-half and the Double-Slit Examples

We can now use quantum mechanics to account for the bizarre results of the spin-half and the double-slit experiments discussed in chapter one. Let us start with the spin-half example. An electron emerges from the SGX device in a state of superposition

$$\frac{1}{\sqrt{2}}\left(|\uparrow_x\rangle \otimes |A\rangle + |\downarrow_x\rangle \otimes |B\rangle\right). \tag{5.3.1}$$

Any position or S_x measurement will cause the state vector to collapse either on $|\uparrow_x\rangle \otimes |A\rangle$ or on $|\downarrow_x\rangle \otimes |B\rangle$, each with probability 1/2. Hence, if we place a position detector on path A, for example, there is a 50% chance of detecting an electron, in which case the electron is now in state $|\uparrow_x\rangle \otimes |A\rangle$, and a 50% chance of not detecting the electron, in which case the electron is now in state $|\downarrow_x\rangle \otimes |B\rangle$. Similarly, if we measure S_x on an electron moving along paths A or B, we shall get $\frac{\hbar}{2}$ 50% of the times and $-\frac{\hbar}{2}$ 50% of the times. Upon either measurement, the state vector collapses. By contrast, if we perform no measurement, since $|\uparrow_x\rangle \otimes |A\rangle$ would evolve into $|\uparrow_x\rangle \otimes |C\rangle$ and $|\downarrow_x\rangle \otimes |B\rangle$ into $|\downarrow_x\rangle \otimes |C\rangle$, the linearity of the evolution operator (equation (4.6.2)) tells us that the state vector will eventually evolve into

$$\frac{1}{\sqrt{2}}\left(|\uparrow_x\rangle \otimes |C\rangle + |\downarrow_x\rangle \otimes |C\rangle\right) = |\uparrow_z\rangle \otimes |C\rangle, \qquad (5.3.2)$$

and therefore, upon measuring S_z, we shall get $\frac{\hbar}{2}$ all the times.

The (simplified) account of the double-slit experiment is analogous. Leaving spin aside, let us represent with $|\Psi_A\rangle$ the state of an electron going through slit A and with $|\Psi_B\rangle$ that of an electron going through slit B. Suppose now that slit B is closed, so that the state vector of the electron is $|\Psi_A\rangle$. As $|\Psi_A\rangle$ is a wave function, the position probability density is given by $|\Psi_A|^2$ and its plot by curve (a) in chapter 1, figure 2, which explains the distribution of hits on the screen. If we close slit A, we obtain an analogous situation with $|\Psi_B|^2$ represented by curve (b) in chapter 1, figure 2.

If both slits are open, the state of the electron at the slits is

$$|\Psi\rangle = c_1|\Psi_A\rangle + c_2|\Psi_B\rangle, \qquad (5.3.3)$$

a superposition of the two vectors. If no observation is made, the electron arrives at the screen still in a state of superposition, and the probability density is

$$|c_1\Psi_A + c_2\Psi_B|^2, \qquad (5.3.4)$$

represented by curve (c) in chapter 1, figure 2. If a position observation is carried out at either slit, the state vector collapses onto $|\Psi_A\rangle$ with probability $|c_1|^2$ or onto $|\Psi_B\rangle$ with probability $|c_2|^2$, and the probability density is

$$|\Psi_A|^2 + |\Psi_B|^2, \qquad (5.3.5)$$

represented by curve (c) in chapter 1, figure 1. Crucially, (5.3.4) and (5.3.5) are different in that the former involves interference while the latter does not. [4]

5.4 The Orthodox Interpretation

Up to now we have considered quantum mechanics as an algorithm (a piece of mathematical machinery) that allowed us to predict the statistical frequencies of the returns of measurements performed on ensembles or the probabilities of the returns of measurements performed on individual systems. One might be happy to leave it at that, as indeed many physicists are. The machinery works wonderfully not only at the atomic but also at the nuclear level. However, many of the physicists who contributed directly or indirectly to its birth were prepared to go further, and by the 1930's, on the aftermath of Dirac's, *The Principles of Quantum Mechanics* and von Neumann's *Mathematical Foundations of Quantum Mechanics*, a broad interpretive agreement seems to have formed around a set of principles constituting what is customarily called "the orthodox" or "the standard" interpretation. Many aspects of the tenets of the orthodox interpretation

---

[4] In the example, the plots are correct if we assume that $c_1 = c_2 = \sqrt{1/2}$. If the values are different (for example, if the electron gun is not equidistant from the two slits), the distribution pattern of electron hits on the screen changes.

have already been introduced (somewhat surreptitiously) in this and previous chapters. They are as follows:

1. The quantum state of a system is *completely* represented by its state vector in Hilbert space; there is nothing in the quantum state that is not expressed in the state vector. In this respect, quantum mechanics is complete (*State Completeness Principle*).
2. Every observable is represented by its Hermitian operator in Hilbert space.
3. The physical information contained in the state vector is given by Born's statistical interpretation.
4. A system's observable has a certain value if and only if the system is in the corresponding eigenstate, that is, is represented by the corresponding eigenvector. This is the *eigenstate-eignevalue link*, (EE).[5]
5. In the absence of measurement, the (linear) temporal evolution of a quantum system is governed by TDSE.
6. Upon measurement, the system jumps non-linearly and instantaneously to an eigenstate of the measured observable (*Collapse, or Projection, Principle*).

The orthodox interpretation is appealing because its principles not only fit the mathematical machinery but are also correlated. For example, (1)-(3) warrant (4), and (4) plus the fact that quickly repeated measurements do return the same value provides good

---

[5] Principle (4) is at times modified as saying that upon measurement a system's observable returns a certain value if and only if the system is in the corresponding eigenstate. This provides a more restrictive interpretation because now quantum mechanics is only about measurement returns. However, if we assume that measurement returns give a faithful quantitative representation of the observable, we obtain (4) again.

evidence for (6). Indeed, as the orthodox interpretation can be found, often implicitly, in standard quantum mechanical texts, it has become part and parcel of the quantum mechanics one typically learns. For this reason, often we shall refer to the quantum mathematical machinery plus the orthodox interpretation as "standard quantum mechanics."

Exercises

Exercise 5.1

What is the probability of obtaining $-\hbar/2$ if we measure just S_x^2 on (5.2.5)? [Hint: Notice that there are two eigenvectors in H for S_x^2 corresponding to the eigenvalue $-\hbar/2$. When there are n eigenvectors corresponding to the same eigenvalue, the system is n-fold degenerate. So, the system is two-fold degenerate with respect to $S_x^2 = -\hbar/2$. Then, keep in mind that the overall probability that $S_x^2 = -\hbar/2$ is the sum of the probabilities of all the different ways in which $S_x^2 = -\hbar/2$.]

Answers to the Exercises

Exercise 5.1

First, let us remember that the eigenvector in H_2 corresponding to $-\hbar/2$ for S_x^2 is

$\left|\downarrow_x^2\right\rangle = \frac{1}{\sqrt{2}}\left(\left|\uparrow_z^2\right\rangle - \left|\downarrow_z^2\right\rangle\right)$. With respect to $S_x^2 = -\hbar/2$, the system is two-fold degenerate since there are two eigenvectors in H for S_x^2 corresponding to the eigenvalue $-\hbar/2$, namely $\left|\uparrow_z^1\downarrow_x^2\right\rangle = \frac{1}{\sqrt{2}}\left(\left|\uparrow_z^1\uparrow_z^2\right\rangle - \left|\uparrow_z^1\downarrow_z^2\right\rangle\right)$, and $\left|\downarrow_z^1\downarrow_x^2\right\rangle = \frac{1}{\sqrt{2}}\left(\left|\downarrow_z^1\uparrow_z^2\right\rangle - \left|\downarrow_z^1\downarrow_z^2\right\rangle\right)$. Hence, the probability we are seeking is

$$\frac{1}{2}\left|\left\langle\left(\left\langle\uparrow_z^1\uparrow_z^2\right| - \left\langle\uparrow_z^1\downarrow_z^2\right|\right)\middle|\Psi\right\rangle\right|^2 + \frac{1}{2}\left|\left\langle\left(\left\langle\downarrow_z^1\uparrow_z^2\right| - \left\langle\downarrow_z^1\downarrow_z^2\right|\right)\middle|\Psi\right\rangle\right|^2 = \frac{\left|c_1 - c_2\right|^2}{2} + \frac{\left|c_3 - c_4\right|^2}{2}.$$