

Cambridge Books Online

<http://ebooks.cambridge.org/>



Quantum Computing for Computer Scientists

Noson S. Yanofsky, Mirco A. Mannucci

Book DOI: <http://dx.doi.org/10.1017/CBO9780511813887>

Online ISBN: 9780511813887

Hardback ISBN: 9780521879965

Chapter

4 - Basic Quantum Theory pp. 103-137

Chapter DOI: <http://dx.doi.org/10.1017/CBO9780511813887.006>

Cambridge University Press

Basic Quantum Theory

Reality is that which, when you stop believing in it, does not go away.

Philip K. Dick¹

In Chapters 1 and 2 we developed the necessary mathematical apparatus and terminology that will be used throughout this book. Chapter 3 has provided some heuristics and gently led us to the threshold of quantum mechanics. It is now time to open the door, introduce the basic concepts and tools of the trade, and continue our journey to quantum computing.²

In Section 4.1 we spend a few words on the motivations behind quantum mechanics. We then introduce quantum states and how they are distinguishable from one another through observations. Section 4.2 describes observable physical quantities within the quantum framework. How observable quantities are measured is the topic of Section 4.3. The dynamics of quantum systems, i.e., their evolution in time, is the focus of Section 4.4. Finally, in Section 4.5, we revisit the tensor product and show how it describes the way in which larger quantum systems are assembled from smaller ones. In the process, we meet the crucial notion of entanglement, a feature of the quantum world that pops up again in the chapters ahead.

4.1 QUANTUM STATES

Why quantum mechanics? To answer this question, we have to hearken back in time to the dawn of the twentieth century. Classical mechanics still dominated the scene, with its double-pronged approach: **particles** and **waves**. Matter was considered to be

¹ The quotation is taken from Dick's 1978 lecture *How to build a Universe that does not fall apart two days later*, freely available on the Web at <http://deoxy.org/pkd.how2build.htm>.

² No attempt will be made to present the material in an exhaustive historical manner. The curious reader can refer to the references at the end of this chapter for a plethora of good, comprehensive introductions to quantum mechanics.

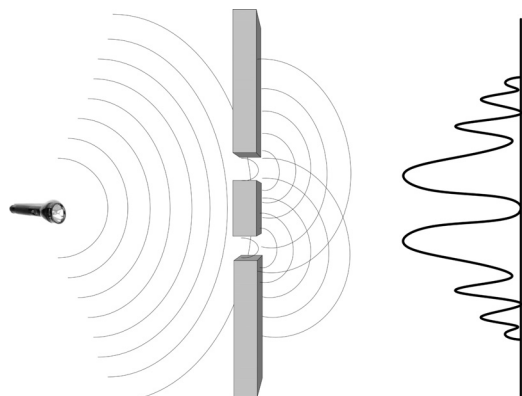


Figure 4.1. Young's double-slit experiment.

ultimately composed of microscopic particles, and light was thought of as continuous electromagnetic waves propagating in space.

The dichotomy – particles versus waves – was proven false by several groundbreaking experiments. For instance, the **diffraction experiment** shows that a beam of subatomic particles hitting a crystal diffract following a wave-like pattern, entirely similar to the diffraction pattern of light itself. By the mid-twenties, physicists started associating waves to all known particles, the so-called **matter waves** (the first proposal was made by French physicist Louis De Broglie in 1924 in his doctoral dissertation).

The **photoelectric effect** (observed by Hertz in 1887) showed that an atom hit by a beam of light may absorb it, causing some electrons to transition to a higher-energy orbital (i.e., farther from the nucleus). Later on, the absorbed energy may be released in the form of emitted light, causing the excited electrons to revert to a lower orbital. What the photoelectric effect unraveled was that light-matter transactions *always* occur through discrete packets of energy, known as **photons** (the concept was introduced by Einstein in his seminal 1905 paper, as a way to account for the photoelectric effect). Photons act as genuine particles that can get absorbed and emitted, one at a time.

Further experimental evidence from many quarters accumulated over time, strongly suggesting that the old duality particle–wave theory must be replaced by a new theory of the microscopic world in which *both matter and light manifest a particle-like and a wave-like behavior*. Time was ripe for the conceptual framework of quantum mechanics.

In Chapter 3 we met a toy version of the double-slit experiment; as it turns out, this was an actual experiment, indeed an entire series of related experiments, the first one being carried out with light by the English polymath Thomas Young around 1801. Before we move on, it is worth our while to revisit it briefly, as it contains most of the main ingredients that make up quantum's magic.

One shines light at a boundary with two slits that are very close to each other. The pattern of the light to the right of the boundary will have certain regions that are dark and certain others that are bright, as depicted in Figure 4.1.

The reason why there are regions on the screen with no light is that light waves are interfering with each other. Light is propagating as a single wave from its source;

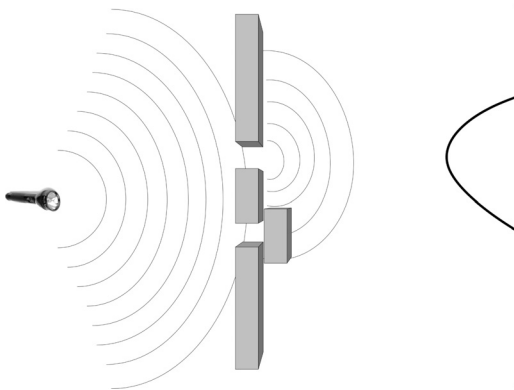


Figure 4.2. Young's double-slit experiment with one slit closed.

the two slits cause this wave to split into two independent ones, which can then interfere with each other when reaching the screen. Some regions are going to be darker, others are going to be brighter, depending on whether the two waves are in phase (positive interference) or out of phase (negative interference).

What would happen if we closed off one of the slits? In that case, there is no splitting and therefore no interference pattern whatsoever (Figure 4.2).

Two remarks on this seminal experiment are in order:

- As we have already pointed out in Chapter 3, the double-slit experiment can be done with just one photon at a time. Rather than spotting patterns of lighter or darker light on the screen, we are now looking for which region is more or less likely for the single photon to land. The same pattern can then be viewed as describing the probability for a certain region to get hit by the photon. The natural question then is, if there is a single photon why would there be any interference pattern? Yet, experiments have shown that such a pattern is there. Our photon is a true chameleon: sometimes it behaves as a particle and sometimes as a wave, depending on how it is observed.
- The double-slit experiment is not only about light: one can perform it equally well with electrons, protons, and even atomic nuclei, and they will all exhibit exactly the same interference behavior.³ Once again, this clearly indicates that the rigid distinction between waves and particles as a paradigm of description of the physical world is untenable at the quantum level.

In the rest of this section, we are going to introduce the basic mathematical description of a quantum physical system. We shall restrict ourselves to two simple examples, to illustrate the basic machinery:

- a particle confined to a set of discrete positions on a line
- a single-particle spin system

³ Such experiments have indeed been performed, only much later than Young's original version of the double-slit experiment. We invite you to read about this fascinating slice of experimental physics in Rodgers (2002).

Consider a subatomic particle on a line; moreover, let us suppose that it can only be detected at one of the equally spaced points $\{x_0, x_1, \dots, x_{n-1}\}$, where $x_1 = x_0 + \delta x$, $x_2 = x_1 + \delta x$, \dots , with δx some fixed increment.

$$\begin{array}{ccccccc}
 x_0 & & x_1 & & \dots & & x_i & & \dots & & x_{n-1} \\
 & \bullet & \text{---} & \bullet & \text{---} & \bullet & \text{---} & \bullet & \text{---} & \bullet &
 \end{array} \tag{4.1}$$

In real life, a particle can of course occupy any of the points of the line, not just a finite subset thereof. However, if we followed this route, the state space of our system would be infinite dimensional, requiring a considerably larger mathematical apparatus than the one covered in the last chapters. Whereas such an apparatus is vital for quantum mechanics, it is not needed for an exposition of quantum computing.⁴ For our current exposition, we can thus assume that the set $\{x_0, x_1, \dots, x_{n-1}\}$ is composed of a great many points (n large) and that δx is tiny, thereby providing a reasonably good approximation of a continuous system.

We are now going to associate to the current **state** of the particle an n -dimensional complex column vector $[c_0, c_1, \dots, c_{n-1}]^T$.

The particle being at the point x_i shall be denoted as $|x_i\rangle$, using the Dirac **ket** notation. (Do not worry about the funny symbol: it will be explained momentarily.) To each of these n basic states, we shall associate a column vector:

$$\begin{aligned}
 |x_0\rangle &\mapsto [1, 0, \dots, 0]^T \\
 |x_1\rangle &\mapsto [0, 1, \dots, 0]^T \\
 &\vdots \\
 |x_{n-1}\rangle &\mapsto [0, 0, \dots, 1]^T.
 \end{aligned} \tag{4.2}$$

Observe that these vectors form the canonical basis of \mathbb{C}^n . From the standpoint of classical mechanics, the basic states in Equation (4.2) are all we shall ever need. Not so in quantum mechanics: experimental evidence testifies to the fact that the particle can be in a strange fuzzy blending of these states (think again of the double-slit!). To catch up with Nature, we shall make a bold leap by positing that *all* vectors in \mathbb{C}^n represent a legitimate physical state of the particle.

What can all this possibly mean?

An arbitrary state, which we shall denote as $|\psi\rangle$, will be a linear combination of $|x_0\rangle, |x_1\rangle, \dots, |x_{n-1}\rangle$, by suitable complex weights, c_0, c_1, \dots, c_{n-1} , known as **complex amplitudes**.⁵

$$|\psi\rangle = c_0|x_0\rangle + c_1|x_1\rangle + \dots + c_{n-1}|x_{n-1}\rangle. \tag{4.3}$$

⁴ We mention in passing that in computer simulation one must always turn a continuous physical system (classical or quantum) into a discrete one: computers cannot deal with infinities.

⁵ This name comes from the fact that $|\psi\rangle$ is indeed a (complex) wave when we study its time evolution, as we shall see at the end of Section 4.3. Waves are characterized by their amplitude (think of the intensity of a sound wave) – hence the name above – as well as by their frequency (in case of sound waves, their pitch). As it turns out, the frequency of $|\psi\rangle$ plays a key role in the particle's momentum. You can think of Equation (4.3) as describing $|\psi\rangle$ as the overlap of n waves, the $|x_i\rangle$, each contributing with intensity c_i .

Thus, every state of our system can be represented by an element of \mathbb{C}^n as

$$|\psi\rangle \mapsto [c_0, c_1, \dots, c_{n-1}]^T. \quad (4.4)$$

We say that the state $|\psi\rangle$ is a **superposition** of the basic states. $|\psi\rangle$ represents the particle as being simultaneously in all $\{x_0, x_1, \dots, x_{n-1}\}$ locations, or a blending of all the $|x_i\rangle$. There are, however, different possible blendings (much like in the recipe for baking an apple pie you can vary the proportions of the ingredients and obtain different flavors). The complex numbers c_0, c_1, \dots, c_{n-1} tell us precisely which superposition our particle is currently in. The norm square of the complex number c_i divided by the norm squared of $|\psi\rangle$ will tell us the probability that, *after* observing the particle, we will detect it at the point x_i :

$$p(x_i) = \frac{|c_i|^2}{||\psi\rangle|^2} = \frac{|c_i|^2}{\sum_j |c_j|^2}. \quad (4.5)$$

Observe that $p(x_i)$ is always a positive real number and $0 \leq p(x_i) \leq 1$, as any genuine probability should be.

When $|\psi\rangle$ is observed, we will find it in one of the basic states. We might write it as⁶

$$|\psi\rangle \rightsquigarrow |x_i\rangle. \quad (4.6)$$

Example 4.1.1 Let us assume that the particle can only be at the four points $\{x_0, x_1, x_2, x_3\}$. Thus, we are concerned with the state space \mathbb{C}^4 . Let us also assume that now the state vector is

$$|\psi\rangle = \begin{bmatrix} -3-i \\ -2i \\ i \\ 2 \end{bmatrix}. \quad (4.7)$$

We shall calculate the probability that our particle can be found at position x_2 . The norm of $|\psi\rangle$ is given by

$$||\psi\rangle| = \sqrt{|-3-i|^2 + |-2i|^2 + |i|^2 + |2|^2} = 4.3589. \quad (4.8)$$

The probability is therefore

$$\frac{|i|^2}{(4.3589)^2} = 0.052624. \quad (4.9)$$

□

⁶ The wiggly line is used throughout this chapter to denote the state of a quantum system before and after measurement.

Exercise 4.1.1 Let us assume that the particle is confined to $\{x_0, x_1, \dots, x_5\}$ and the current state vector is

$$|\psi\rangle = [2 - i, 2i, 1 - i, 1, -2i, 2]^T. \quad (4.10)$$

What is the likelihood of finding the particle at position x_3 ? ■

Kets can be added: if

$$|\psi\rangle = c_0|x_0\rangle + c_1|x_1\rangle + \dots + c_{n-1}|x_{n-1}\rangle = [c_0, c_1, \dots, c_{n-1}]^T \quad (4.11)$$

and

$$|\psi'\rangle = c'_0|x_0\rangle + c'_1|x_1\rangle + \dots + c'_{n-1}|x_{n-1}\rangle = [c'_0, c'_1, \dots, c'_{n-1}]^T, \quad (4.12)$$

then

$$\begin{aligned} |\psi\rangle + |\psi'\rangle &= (c_0 + c'_0)|x_0\rangle + (c_1 + c'_1)|x_1\rangle + \dots + (c_{n-1} + c'_{n-1})|x_{n-1}\rangle \\ &= [c_0 + c'_0, c_1 + c'_1, \dots, c_{n-1} + c'_{n-1}]^T. \end{aligned} \quad (4.13)$$

Also, for a complex number $c \in \mathbb{C}$, we can scalar multiply a ket by c :

$$c|\psi\rangle = cc_0|x_0\rangle + cc_1|x_1\rangle + \dots + cc_{n-1}|x_{n-1}\rangle = [cc_0, cc_1, \dots, cc_{n-1}]^T. \quad (4.14)$$

What happens if we add a ket to itself?

$$\begin{aligned} |\psi\rangle + |\psi\rangle &= 2|\psi\rangle = [c_0 + c_0, c_1 + c_1, \dots, c_j + c_j, \dots, c_{n-1} + c_{n-1}]^T \\ &= [2c_0, 2c_1, \dots, 2c_j, \dots, 2c_{n-1}]^T. \end{aligned} \quad (4.15)$$

The sum of the moduli squared is

$$\begin{aligned} S' &= |2c_0|^2 + |2c_1|^2 + \dots + |2c_{n-1}|^2 = 2^2|c_0|^2 + 2^2|c_1|^2 + \dots + 2^2|c_{n-1}|^2 \\ &= 2^2(|c_0|^2 + |c_1|^2 + \dots + |c_{n-1}|^2). \end{aligned} \quad (4.16)$$

For the state $2|\psi\rangle$, the chance that the particle will be found in position j is

$$\begin{aligned} p(x_j) &= \frac{|2c_j|^2}{S'} = \frac{2^2|c_j|^2}{2^2(|c_0|^2 + |c_1|^2 + \dots + |c_{n-1}|^2)} \\ &= \frac{|c_j|^2}{|c_0|^2 + |c_1|^2 + \dots + |c_{n-1}|^2}. \end{aligned} \quad (4.17)$$

In other words, the ket $2|\psi\rangle$ describes *the same physical system* as $|\psi\rangle$. Notice that we could replace 2 with an arbitrary $c \in \mathbb{C}$ and get the same results. Geometrically, the vector $|\psi\rangle$ and all its complex scalar multiples $c|\psi\rangle$, i.e., *the entire subspace generated by $|\psi\rangle$* , describe the same physical state. The length of $|\psi\rangle$ does not matter as far as physics goes.

Exercise 4.1.2 Let $|\psi\rangle$ be $[c_0, c_1, \dots, c_{n-1}]^T$. Check that multiplying $|\psi\rangle$ by any complex number c will not alter the calculation of probabilities. (Hint: Factor out c in the ratio.) ■

Example 4.1.2 The vectors

$$|\psi_1\rangle = \begin{bmatrix} 1+i \\ i \end{bmatrix} \quad \text{and} \quad |\psi_2\rangle = \begin{bmatrix} 2+4i \\ 3i-1 \end{bmatrix} \quad (4.18)$$

differ by the factor $3+i$ (verify it!), and are thus representatives of the same quantum state. \square

Exercise 4.1.3 Do the vectors $[1+i, 2-i]^T$ and $[2+2i, 1-2i]^T$ represent the same state? \blacksquare

As we can multiply (or divide) a ket by any (complex) number and still have a representation of the same physical state, we may as well work with a **normalized** $|\psi\rangle$, i.e.,

$$\frac{|\psi\rangle}{||\psi||} \quad (4.19)$$

which has length 1.⁷

Example 4.1.3 The vector $[2-3i, 1+2i]^T$ has length given by

$$\sqrt{|2-3i|^2 + |1+2i|^2} = 4.2426. \quad (4.20)$$

We can normalize it by simply dividing by its length:

$$\frac{1}{4.2426} [2-3i, 1+2i]^T = [0.4714 - 0.7071i, 0.2357 + 0.4714i]^T. \quad (4.21)$$

\square

Exercise 4.1.4 Normalize the ket

$$|\psi\rangle = [3-i, 2+6i, 7-8i, 6.3+4.9i, 13i, 0, 21.1]^T. \quad (4.22)$$

\blacksquare

Exercise 4.1.5 (a) Verify that the two state vectors $[\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]^T$ and $[\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}]^T$ are each of length 1 in \mathbb{C}^2 . (b) Find the vector on the unit ball of \mathbb{C}^2 representing the superposition (addition) of these two states. \blacksquare

Given a normalized ket $|\psi\rangle$, the denominator of Equation (4.5) is 1, and hence, the equation reduces to

$$p(x_i) = |c_i|^2. \quad (4.23)$$

We are now done with our first motivating example. Let us move on to the second one. In order to talk about it, we need to introduce a property of subatomic particles called **spin**. As it turns out, spin will play a major role in our story, because it is the prototypical way to implement quantum bits of information, or qubits, which we shall encounter in Section 5.1.

⁷ In Section 3.3, we limited ourselves to normalized complex vectors. Now you see why!

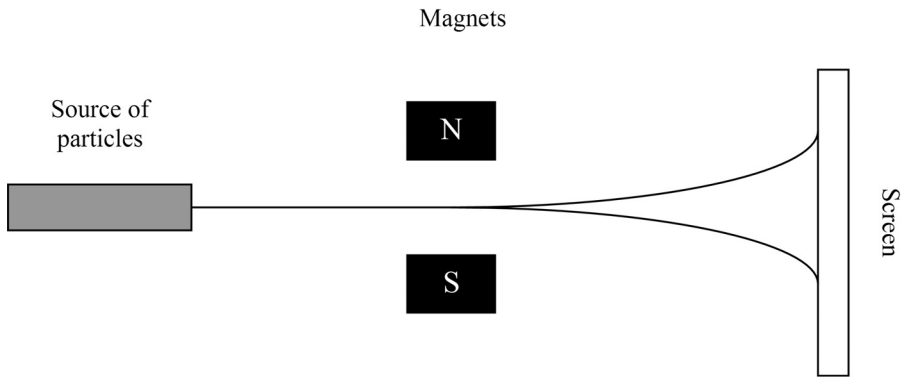


Figure 4.3. The Stern–Gerlach experiment.

What is spin? The **Stern–Gerlach experiment** (first performed in 1922) showed that an electron in the presence of a magnetic field will behave as if it were a charged spinning top: it will act as a small magnet and strive to align itself to the external field. The Stern–Gerlach experiment (as shown in Figure 4.3) consists of shooting a beam of electrons through a nonhomogeneous magnetic field oriented in a certain direction, say, vertically (z direction). As it happens, the field splits the beam into two streams, with opposite spin. Certain electrons will be found spinning one way, and certain others spinning the opposite way.

With respect to a classical spinning top, there are two striking differences:

- First, the electron does not appear to have an internal structure, by itself it is just a charged point. It acts as a spinning top but it is no top! Spin is therefore a new property of the quantum world, with no classical analog.
- Secondly, and quite surprisingly, all our electrons can be found either at the top of the screen or at the bottom, none in between. But, we had not prepared the “spinning” electrons in any way before letting them interact with the magnetic field. Classically, one would have expected them to have different magnetic components along the vertical axis, and therefore to be differently pulled by the field. There should be some in the middle of the screen. But there isn’t. Conclusion: when the spinning particle is measured in a given direction, it can only be found in two states: it spins either clockwise or anticlockwise (as shown in Figure 4.4).

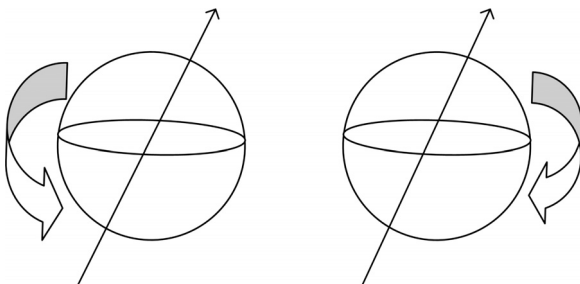


Figure 4.4. Particles with spin.

For each given direction in space, there are only two basic spin states. For the vertical axis, these states have a name: spin **up** $|\uparrow\rangle$ and spin **down** $|\downarrow\rangle$. The generic state will then be a superposition of up and down, or

$$|\psi\rangle = c_0|\uparrow\rangle + c_1|\downarrow\rangle. \quad (4.24)$$

Just like before, c_0 is the amplitude of finding the particle in the up state, and similarly for c_1 .

Example 4.1.4 Consider a particle whose spin is described by the ket

$$|\psi\rangle = (3 - 4i)|\uparrow\rangle + (7 + 2i)|\downarrow\rangle. \quad (4.25)$$

The length of the ket is

$$\sqrt{|3 - 4i|^2 + |7 + 2i|^2} = 8.8318. \quad (4.26)$$

Therefore, the probability of detecting the spin of the particle in the up direction is

$$p(\uparrow) = \frac{|3 - 4i|^2}{8.8318^2} = \frac{25}{78}. \quad (4.27)$$

The probability of detecting the spin of the particle in state down is

$$p(\downarrow) = \frac{|7 + 2i|^2}{8.8318^2} = \frac{53}{78}. \quad (4.28)$$

□

Exercise 4.1.6 Let the spinning electron's current state be $|\psi\rangle = 3i|\uparrow\rangle - 2|\downarrow\rangle$. Find the probability that it will be detected in the up state. ■

Exercise 4.1.7 Normalize the ket given in Equation (4.25). ■

In Chapter 2, the inner product was introduced as an abstract mathematical idea. This product turned a vector space into a space with a *geometry*: angles, orthogonality, and distance were added to the canvas. Let us now investigate its physical meaning. The inner product of the state space gives us a tool to compute complex numbers known as **transition amplitudes**, which in turn will enable us to determine how likely the state of the system *before* a specific measurement (start state), will change to another (end state), *after* measurement has been carried out. Let

$$|\psi\rangle = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} \quad \text{and} \quad |\psi'\rangle = \begin{bmatrix} c'_0 \\ c'_1 \\ \vdots \\ c'_{n-1} \end{bmatrix} \quad (4.29)$$

be two normalized states. We can extract the transition amplitude between state $|\psi\rangle$ and state $|\psi'\rangle$ by the following recipe: $|\psi\rangle$ will be our start state. The end state will be a row vector whose coordinates will be the complex conjugate of $|\psi'\rangle$ coordinates.

Such a state is called a **bra**, and will be denoted $\langle\psi'|$, or equivalently

$$\langle\psi'| = |\psi'\rangle^\dagger = [\overline{c'_0}, \overline{c'_1}, \dots, \overline{c'_{n-1}}]. \quad (4.30)$$

To find the transition amplitude we multiply them *as matrices* (notice that we put them side by side, forming a bra–ket, or bra(c)ket, i.e., their inner product):

$$\langle\psi'|\psi\rangle = [\overline{c'_0}, \overline{c'_1}, \dots, \overline{c'_{n-1}}] \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = \overline{c'_0} \times c_0 + \overline{c'_1} \times c_1 + \dots + \overline{c'_{n-1}} \times c_{n-1}. \quad (4.31)$$

We can represent the start state, the ending state, and the amplitude of going from the first to the second as the decorated arrow:

$$|\psi\rangle \overset{\langle\psi|\psi'\rangle}{\rightsquigarrow} |\psi'\rangle. \quad (4.32)$$

This recipe is, of course, none other than the inner product of Section 2.4. What we have done is simply split the product into the bra–ket form. Although this is mathematically equivalent to our previous definition, it is quite handy for doing calculations, and moreover opens up an entirely new vista: it shifts the focus from *states to state transitions*.⁸

Note: The transition amplitude between two states may be zero. In fact, that happens precisely when the two states are orthogonal to one another. This simple fact hints at the physical content of orthogonality: orthogonal states are as far apart as they can possibly be. We can think of them as *mutually exclusive alternatives*: for instance, an electron can be in an arbitrary superposition of spin up and down, but after we measure it in the z direction, it will always be *either* up *or* down, never both up *and* down. If our electron was already in the up state before the z direction measurement, it will never transition to the down state as a result of the measurement.

Assume that we are given a normalized start state $|\psi\rangle$ and an orthonormal basis $\{|b_0\rangle, |b_1\rangle, \dots, |b_{n-1}\rangle\}$, representing a maximal list of mutually exclusive end states associated with some specific measurement of the system. In other words, we know beforehand that the result of our measurement will necessarily be one or the other of the states in the basis, but never a superposition of any of them. We show in Section 4.3 that for every complete measurement of a quantum system there is an associated orthonormal basis of all its possible outcomes.

⁸ This line of thought has been pursued by some researchers, in the ambitious attempt to provide a satisfactory interpretation of quantum mechanics. For instance, Yakhir Aharonov and his colleagues have in recent years proposed a model called the two-vector formalism, in which the single vector description is replaced with the full bra–ket pair. The interested reader can consult Aharonov's recent book *Quantum Paradoxes* (Aharonov and Rohrlich, 2005).

We can express $|\psi\rangle$ in the basis $\{|b_0\rangle, |b_1\rangle, \dots, |b_{n-1}\rangle\}$ as

$$|\psi\rangle = b_0|b_0\rangle + b_1|b_1\rangle + \dots + b_{n-1}|b_{n-1}\rangle. \quad (4.33)$$

We invite you to check that $b_i = \langle b_i|\psi\rangle$ and that $|b_0|^2 + |b_1|^2 + \dots + |b_{n-1}|^2 = 1$.

It is thus natural to read Equation (4.33) in the following way: each $|b_i|^2$ is the *probability* of ending up in state $|b_i\rangle$ after a measurement has been made.

$$\begin{array}{c}
 |b_0\rangle \\
 \nearrow \langle b_0|\psi\rangle \\
 |b_1\rangle \\
 \nearrow \langle b_1|\psi\rangle \\
 \vdots \\
 |b_i\rangle \\
 \nearrow \langle b_i|\psi\rangle \\
 \vdots \\
 |b_{n-1}\rangle \\
 \nearrow \langle b_{n-1}|\psi\rangle
 \end{array}
 \quad (4.34)$$

Exercise 4.1.8 Check that the set $\{|x_0\rangle, |x_1\rangle, \dots, |x_{n-1}\rangle\}$ is an orthonormal basis for the state space of the particle on the line. Similarly, verify that $\{|\uparrow\rangle, |\downarrow\rangle\}$ is an orthonormal basis of the one-particle spin system. ■

From now on, we shall use the row-column and the bra-ket notation introduced earlier interchangeably, as we deem fit.⁹

Let us work through a couple of examples together.

Example 4.1.5 Let us compute the bra corresponding to the ket $|\psi\rangle = [3, 1 - 2i]^T$. It is quite easy; we take the complex conjugate of all the entries, and list them: $\langle\psi| = [3, 1 + 2i]$. □

Example 4.1.6 Let us now compute the amplitude of the transition from $|\psi\rangle = \frac{\sqrt{2}}{2}[1, i]^T$ to $|\phi\rangle = \frac{\sqrt{2}}{2}[i, -1]^T$. We first need to write down the bra corresponding to the end state:

$$\langle\phi| = \frac{\sqrt{2}}{2}[-i, -1]. \quad (4.35)$$

⁹ An historical note is in order: the bra-ket notation, which is now ubiquitous in quantum mechanics, was introduced by the great physicist Paul A.M. Dirac around 1930.

Now we can take their inner product:

$$\langle \phi, \psi \rangle = -i. \quad (4.36)$$

□

Exercise 4.1.9 Calculate the bra corresponding to the ket $|\psi\rangle = [3 + i, -2i]^T$. ■

Exercise 4.1.10 Calculate the amplitude of the transition from $\frac{\sqrt{2}}{2}[i, -1]^T$ to $\frac{\sqrt{2}}{2}[1, -i]^T$. ■

Observe that in the calculation of transition amplitudes via the inner product, the requirement that the representatives be normalized states can be easily removed by simply dividing the hermitian product by the product of the length of the two vectors (or equivalently, normalizing your states first, and then computing their inner product). Here is an example.

Example 4.1.7 Let us calculate the amplitude of the transition from $|\psi\rangle = [1, -i]^T$ to $|\phi\rangle = [i, 1]^T$. Both vectors have norm $\sqrt{2}$.

We can take their inner product first:

$$\langle \phi | \psi \rangle = [-i, 1][1, -i]^T = -2i. \quad (4.37)$$

and then divide it by the product of their norm:

$$\frac{-2i}{\sqrt{2} * \sqrt{2}} = -i. \quad (4.38)$$

Equivalently, we can first normalize them, and then take their product:

$$\left\langle \frac{1}{\sqrt{2}}\phi \middle| \frac{1}{\sqrt{2}}\psi \right\rangle = \left[\frac{-i}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right] \left[\frac{1}{\sqrt{2}}, \frac{-i}{\sqrt{2}} \right]^T = -i. \quad (4.39)$$

The result is, of course, the same. We can concisely indicate it as

$$\frac{\langle \phi | \psi \rangle}{||\phi\rangle| \ ||\psi\rangle|}. \quad (4.40)$$

□

Let us pause one moment, and see where we are.

- We have learned to associate a vector space to a quantum system. The dimension of this space reflects the amount of basic states of the system.
- States can be superposed, by adding their representing vectors.
- A state is left unchanged if its representing vector is multiplied by a complex scalar.
- The state space has a geometry, given by its inner product. This geometry has a physical meaning: it tells us the likelihood for a given state to transition into another one after being measured. States that are orthogonal to one another are mutually exclusive.

Before moving on to the next sections, we invite you to write a simple computer simulation.

Programming Drill 4.1.1 *Write a program that simulates the first quantum system described in this section. The user should be able to specify how many points the particle can occupy (warning: keep the max number low, or you will fairly quickly run out of memory). The user will also specify a ket state vector by assigning its amplitudes. The program, when asked the likelihood of finding the particle at a given point, will perform the calculations described in Example 4.1.1. If the user enters two kets, the system will calculate the probability of transitioning from the first ket to the second, after an observation has been made.*

4.2 OBSERVABLES

Physics is, by and large, about observations: physical quantities like mass, momentum, velocity, etc., make sense only insofar as they can be observed in a quantifiable way. We can think of a physical system as specified by a double list: on the one hand, its state space, i.e., the collection of all the states it can possibly be found in (see the previous section), and on the other hand, the set of its **observables**, i.e., the physical quantities that can be observed in each state of the state space.

Each observable may be thought of as a specific question we pose to the system: if the system is currently in some given state $|\psi\rangle$, which values can we possibly observe?

In our quantum dictionary, we need to introduce the mathematical analog of an observable:

Postulate 4.2.1 *To each physical observable there corresponds a hermitian operator.*

Let us see what this postulate actually entails. First of all, an observable is a linear operator, which means that it maps states to states. If we apply the observable Ω to the state vector $|\psi\rangle$, the resulting state is now $\Omega|\psi\rangle$.

Example 4.2.1 Let $|\psi\rangle = [-1, -1 - i]^T$ be the start state in the two-dimensional spin state space. Now, let

$$\Omega = \begin{bmatrix} -1 & -i \\ i & 1 \end{bmatrix}. \quad (4.41)$$

This matrix acts as an operator on \mathbb{C}^2 . Therefore, we can apply it to $|\psi\rangle$. The result is the vector $\Omega|\psi\rangle = [i, -1 - 2i]^T$. Observe that $|\psi\rangle$ and $\Omega|\psi\rangle$ are *not* scalar multiples of one another, and thus they do not represent the same state: Ω has modified the state of the system. \square

Secondly, as we already know from Chapter 2, the eigenvalues of a hermitian operator are all real. The physical meaning of this fact is established by the following:

Postulate 4.2.2 *The eigenvalues of a hermitian operator Ω associated with a physical observable are the only possible values observable can take as a result of measuring it on any given state. Furthermore, the eigenvectors of Ω form a basis for the state space.*

As we have said before, observables can be thought of as legitimate questions we can pose to quantum systems. *Each question admits a set of answers: the eigenvalues of the observable.* We learn in the next section how to compute the likelihood that one specific answer will come up out of the entire set.

Before delving into the subtler properties of observables, let us mention some real-life ones. In the case of the first quantum system of Section 4.1, namely, the particle on the line, the most obvious observable is **position**. As we have stated already, each observable represents a specific question we pose to the quantum system. Position asks: “Where can the particle be found?” Which hermitian operator corresponds to position? We are going to tell first how it acts on the basic states:

$$P(|\psi\rangle) = P(|x_i\rangle) = x_i |x_i\rangle. \quad (4.42)$$

In plain words, P acts as multiplication by position.

As the basic states form a basis, we can extend Equation 4.42 to arbitrary states:

$$P\left(\sum c_i |x_i\rangle\right) = \sum x_i c_i |x_i\rangle. \quad (4.43)$$

Here is the matrix representation of the operator in the standard basis:

$$P = \begin{bmatrix} x_0 & 0 & \cdots & 0 \\ 0 & x_1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{n-1} \end{bmatrix}. \quad (4.44)$$

P is simply the diagonal matrix whose entries are the x_i coordinates. Observe that P is trivially hermitian, its eigenvalues are the x_i values, and its normalized eigenvectors are precisely the basic state vectors that we met at the beginning of Section 4.1: $|x_0\rangle, |x_1\rangle, \dots, |x_{n-1}\rangle$.

Exercise 4.2.1 Verify the last statement. [Hint: Do it by brute force (start with a generic vector, multiply it by the position operator, and assume that the result vector is a scalar multiple of the original one. Conclude that it must be one of the basis vectors).] ■

There is a second natural question one may ask of our particle: What is your velocity? Actually, physicists ask a slightly different question: “What is your momentum?” where **momentum** is defined classically as *velocity times mass*. There is a quantum analog of this question, which is represented in our discrete model by the

following operator (recall that δx is the increment of distance on the line):

$$M(|\psi\rangle) = -i * \hbar * \frac{|\psi(x + \delta x)\rangle - |\psi(x)\rangle}{\delta x}. \quad (4.45)$$

In words, momentum is, up to the constant $-i * \hbar$, the rate of change of the state vector from one point to the next.¹⁰

The constant \hbar (pronounced **h bar**) that we have just met is a universal constant in quantum mechanics, known as the **reduced Planck constant**. Although it plays a fundamental role in modern physics (it is one of the universal constants of nature), for the purpose of the present discussion it can be safely ignored.

As it turns out, position and momentum are the most elementary questions we can ask of the particle: there are of course many more, such as energy, angular momentum, etc., but these two are in a sense the basic building blocks (most observables can be expressed in terms of position and momentum). We shall meet again position and momentum at the end of the next section.

Our second example of observables comes from the spin system. The typical question we might pose to such a system is: given a specific direction in space, in which way is the particle spinning? We can, for instance, ask: is the particle spinning up or down in the z direction? Left or right in the x direction? In or out in the y direction? The three spin operators corresponding to these questions are

$$S_z = \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad S_y = \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad S_x = \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (4.46)$$

Each of the three spin operators comes equipped with its orthonormal basis. We have already met up and down, the eigenbasis of S_z . S_x has eigenbasis $\{| \leftarrow \rangle, | \rightarrow \rangle\}$, or left and right, and S_y has $\{| \swarrow \rangle, | \searrow \rangle\}$, or in and out.

Exercise 4.2.2 Consider a particle in initial spin up. Apply S_x to it and determine the probability that the resulting state is still spin up. ■

.....
Reader Tip. The remainder of this section, although quite relevant for general quantum theory, is tangential to quantum computation, and can thus be safely skipped in a first reading (just take a look at the summary at the end of this section and proceed to Section 4.3). ♥

We are going to make some calculations with the operators described before in a little while; first, though, we need a few additional facts on observables and their associated hermitian matrices.

Up to this point, the collection of physical observables on a given quantum system is just a set. However, even an informal acquaintance with elementary physics

¹⁰ The calculus-enabled reader would have easily recognized a one-step discrete version of the derivative in the momentum. Indeed, if δx goes to zero, momentum is precisely the derivative with respect to position of ψ times the scalar $-i * \hbar$.

teaches us that observable quantities can be added, multiplied, or multiplied by a scalar number to form other meaningful physical quantities, i.e., other examples of observables abound: think of momentum as mass times velocity, work as force times displacement, total energy of a particle in motion as the sum of its kinetic and potential energies, etc. We are thus naturally concerned with the following issue: to what extent can we manipulate quantum observables to obtain yet other observables?

Let us start our investigation from the first step, namely, multiplying an observable by a number (i.e., a scalar). There is no problem with carrying out this operation: indeed, if we scalar multiply a hermitian matrix by a real scalar (i.e., we multiply all its entries), the result is still hermitian.

Exercise 4.2.3 Verify the last statement. ■

Exercise 4.2.4 What about complex scalars? Try to find a hermitian matrix and a complex number such that their product fails to be hermitian. ■

Let us make the next move. What about the addition of two hermitian matrices? Suppose we are looking at two physical observables, represented respectively by the hermitians Ω_1 and Ω_2 . Again, no problem: their sum $\Omega_1 + \Omega_2$ is the observable whose representative is the sum of the corresponding hermitian operators, $\Omega_1 + \Omega_2$, which happens to be hermitian.

Exercise 4.2.5 Check that the sum of two arbitrary hermitian matrices is hermitian. ■

From these two facts it ensues that the set of hermitian matrices of fixed dimension forms a real (but not a complex) vector space.

How about products? It is quite tempting to conclude that the product of two physical quantities, represented respectively by the hermitians Ω_1 and Ω_2 , is an observable whose representative is the product (i.e., matrix composition) of Ω_1 and Ω_2 . There are two substantial difficulties here. First, the order in which operators are applied to state vectors matters. Why? Well, simply because matrix multiplication, unlike multiplication of ordinary numbers or functions, is *not*, in general, a commutative operation.

Example 4.2.2 Let

$$\Omega_1 = \begin{bmatrix} 1 & -1-i \\ -1+i & 1 \end{bmatrix} \quad \text{and} \quad \Omega_2 = \begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix}. \quad (4.47)$$

Their product $\Omega_2 \star \Omega_1$ is equal to

$$\Omega_2 \star \Omega_1 = \begin{bmatrix} 1-i & -1 \\ -3+2i & 3+i \end{bmatrix}, \quad (4.48)$$

whereas $\Omega_1 \star \Omega_2$ is equal to

$$\Omega_1 \star \Omega_2 = \begin{bmatrix} 1+i & -3-2i \\ -1 & 3-i \end{bmatrix}. \quad (4.49)$$

□

Exercise 4.2.6 Let $\Omega_1 = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}$ and $\Omega_2 = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$. Verify that both are hermitian. Do they commute with respect to multiplication? ■

The second difficulty is just as serious: in general, the product of hermitian operators is not guaranteed to be hermitian. Let us now investigate in a more rigorous way what it takes for the product of two hermitian operators to be hermitian. Notice that we have

$$\langle \Omega_1 \star \Omega_2 \phi, \psi \rangle = \langle \Omega_2 \phi, \Omega_1 \psi \rangle = \langle \phi, \Omega_2 \star \Omega_1 \psi \rangle, \quad (4.50)$$

where the first equality comes from the fact that Ω_1 is hermitian and the second equality comes from the fact that Ω_2 is hermitian. For $\Omega_1 \star \Omega_2$ to be hermitian, we would need that

$$\langle \Omega_1 \star \Omega_2 \phi, \psi \rangle = \langle \phi, \Omega_1 \star \Omega_2 \psi \rangle. \quad (4.51)$$

This in turn implies

$$\Omega_1 \star \Omega_2 = \Omega_2 \star \Omega_1, \quad (4.52)$$

or equivalently, the operator

$$[\Omega_1, \Omega_2] = \Omega_1 \star \Omega_2 - \Omega_2 \star \Omega_1 \quad (4.53)$$

must be the zero operator (i.e., the operator that sends every vector to the zero vector).

The operator $[\Omega_1, \Omega_2]$ is so important that it deserves its own name; it is called the **commutator** of Ω_1 and Ω_2 . We have just learned that if the commutator is zero then the product (in whichever order) is hermitian. We are going to meet the commutator again in a little while. Meanwhile, let us familiarize ourselves with the commutator through a simple and very important example.

Example 4.2.3 Let us calculate the commutators of the three spin matrices (we shall deliberately ignore the constant factor $\frac{\hbar}{2}$):

$$[S_x, S_y] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 2i \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (4.54)$$

$$[S_y, S_z] = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = 2i \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad (4.55)$$

$$[S_z, S_x] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = 2i \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \quad (4.56)$$

A bit more concisely,

$$[S_x, S_y] = 2i S_z, \quad [S_y, S_z] = 2i S_x, \quad [S_z, S_x] = 2i S_y. \quad (4.57)$$

As we have just seen, none of the commutators are zero. The spin operators do not commute with each other. \square

Now it is your turn.

Exercise 4.2.7 Explicitly calculate the commutator of the operators of Example 4.2.2. \blacksquare

Note: A moment's thought shows that the product of a hermitian operator with itself always commutes and so does the exponent operation. Therefore, given a single hermitian Ω , we automatically get the entire algebra of polynomials over Ω , i.e., all operators of the form

$$\Omega' = \alpha_0 + \alpha_1 \Omega + \alpha_2 \Omega^2 + \cdots + \alpha_{n-1} \Omega^{n-1}. \quad (4.58)$$

All such operators commute with one another.

Exercise 4.2.8 Show that the commutator of two hermitian matrices is a hermitian matrix. \blacksquare

If the commutator of two hermitian operators is zero, or equivalently, the two operators commute, there is no difficulty in assigning their product (in whatever order) as the mathematical equivalent of the physical product of their associated observables. But what about the other cases, when the two operators do *not* commute? The Heisenberg's uncertainty principle, which we are going to meet at the end of this section, will provide an answer.

There is yet another aspect of the association between observables and hermitian operators that can provide substantial physical insight: we know from Chapter 2 that hermitian operators are precisely those operators that behave well with respect to the inner product, i.e.,

$$\langle \Omega \phi, \psi \rangle = \langle \phi, \Omega \psi \rangle \quad (4.59)$$

for each pair $|\psi\rangle, |\phi\rangle$.

From this fact, it immediately derives that $\langle \Omega \psi, \psi \rangle$ is a real number for each $|\psi\rangle$, which we shall denote as $\langle \Omega \rangle_\psi$ (the subscript points to the fact that this quantity depends on the state vector). We can attach a physical meaning to the number $\langle \Omega \rangle_\psi$.

Postulate 4.2.3 $\langle \Omega \rangle_\psi$ is the **expected value** of observing Ω repeatedly on the same state ψ .

This postulate states the following: suppose that

$$\lambda_0, \lambda_1, \dots, \lambda_{n-1} \quad (4.60)$$

is the list of eigenvalues of Ω . Let us prepare our quantum system so that it is in state $|\psi\rangle$ and let us observe the value of Ω . We are going to obtain one or another of the aforementioned eigenvalues. Now, let us start all over again many times, say, n times, and let us keep track of what was observed each time. At the end of our experiment, the eigenvalue λ_i has been seen p_i times, where $0 \leq p_i \leq n$ (in statistical jargon, its frequency is p_i/n). Now perform the calculation

$$\lambda_0 \times \frac{p_0}{n} + \lambda_1 \times \frac{p_1}{n} + \dots + \lambda_{n-1} \times \frac{p_{n-1}}{n}. \quad (4.61)$$

If n is sufficiently large, this number (known in statistics as the estimated expected value of Ω) will be very close to $\langle \Omega \psi, \psi \rangle$.

Example 4.2.4 Let us calculate the expected value of the position operator on an arbitrary normalized state vector: let

$$|\psi\rangle = c_0|x_0\rangle + c_1|x_1\rangle + \dots + c_{n-1}|x_{n-1}\rangle \quad (4.62)$$

be our state vector and

$$\langle P\psi, \psi \rangle = |c_0|^2 \times x_0 + |c_1|^2 \times x_1 + \dots + |c_{n-1}|^2 \times x_{n-1}, \quad (4.63)$$

where

$$|c_0|^2 + |c_1|^2 + \dots + |c_{n-1}|^2 = 1. \quad (4.64)$$

In particular, if $|\psi\rangle$ happens to be just $|x_i\rangle$, we simply get x_i (verify it!). In other words, the expected value of position on any of its eigenvectors $|x_i\rangle$ is the corresponding position x_i on the line. \square

Example 4.2.5 Let $|\psi\rangle = \left[\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}i \right]^T$ and $\Omega = \begin{bmatrix} 1 & -i \\ i & 2 \end{bmatrix}$.

Let us calculate $\Omega(|\psi\rangle)$:

$$\Omega(|\psi\rangle) = \begin{bmatrix} 1 & -i \\ i & 2 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2}i \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ \frac{3}{2}\sqrt{2}i \end{bmatrix}. \quad (4.65)$$

The bra associated with $\Omega|\psi\rangle$ is $\left[\sqrt{2}, -\frac{3}{2}\sqrt{2}i \right]$. The scalar product $\langle \Omega\psi | \psi \rangle$, i.e., the average value of Ω on $|\psi\rangle$, is thus equal to

$$\left[\sqrt{2}, -\frac{3}{2}\sqrt{2}i \right] \left[\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}i \right]^T = 2.5. \quad (4.66)$$

\square

Exercise 4.2.9 Repeat the steps of the previous example where

$$|\psi\rangle = \left| \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right\rangle \quad (4.67)$$

and

$$\Omega = \begin{bmatrix} 3 & 1 + 2i \\ 1 - 2i & -1 \end{bmatrix}. \quad (4.68)$$

■

We now know that the result of observing Ω repeatedly on a given state will be a certain frequency distribution on the set of its eigenvalues. In plain words, sooner or later we will encounter all its eigenvalues, some more frequently and some less. In the next section we compute the probability that a given eigenvalue of Ω will actually be observed on a given state. For now, we may be interested in knowing the spread of the distribution around its expected value, i.e., the **variance** of the distribution. A small variance will tell us that most of the eigenvalues are very close to the mean, whereas a large variance means just the opposite. We can define the variance in our framework in a few stages. First, we introduce the hermitian operator

$$\Delta_\psi(\Omega) = \Omega - \langle \Omega \rangle_\psi I \quad (4.69)$$

(I is the identity operator). The operator $\Delta_\psi(\Omega)$ acts on a generic vector $|\phi\rangle$ in the following fashion:

$$\Delta_\psi(\Omega)|\phi\rangle = \Omega|\phi\rangle - (\langle \Omega \rangle_\psi)|\phi\rangle. \quad (4.70)$$

So $\Delta_\psi(\Omega)$ just subtracts the mean from the result of Ω . What then is the mean of $\Delta_\psi(\Omega)$ itself on the normalized state $|\psi\rangle$? A simple calculation shows that it is precisely zero: $\Delta_\psi(\Omega)$ is the *demeaned* version of Ω .

Exercise 4.2.10 Verify the last statement. ■

We can now define the variance of Ω at $|\psi\rangle$ as the expectation value of $\Delta_\psi(\Omega)$ squared (i.e., the operator $\Delta_\psi(\Omega)$ composed with itself):

$$\text{Var}_\psi(\Omega) = \langle (\Delta_\psi(\Omega)) \star (\Delta_\psi(\Omega)) \rangle_\psi. \quad (4.71)$$

Admittedly, the definition looks at first sight rather obscure, although it is not so bad if we remember the usual definition of the variance of a random variable X as

$$\text{Var}(X) = E((X - \mu)^2) = E((X - \mu)(X - \mu)), \quad (4.72)$$

where E is the expected value function. The best course is to turn to a simple example to get a concrete feel for it.

Example 4.2.6 Let Ω be a 2-by-2 diagonal matrix with real entries:

$$\Omega = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

and let

$$\langle |\psi\rangle = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \quad (4.73)$$

Let us denote by μ (pronounced “mu”) the mean of Ω on $|\psi\rangle$.

$$\Delta_\psi(\Omega) = \Omega - \langle \Omega \rangle_\psi = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} - \begin{bmatrix} \mu & 0 \\ 0 & \mu \end{bmatrix} = \begin{bmatrix} \lambda_1 - \mu & 0 \\ 0 & \lambda_2 - \mu \end{bmatrix}. \quad (4.74)$$

Now we calculate $\Delta_\psi(\Omega) \star \Delta_\psi(\Omega)$:

$$\begin{aligned} \Delta_\psi(\Omega) \star \Delta_\psi(\Omega) &= \begin{bmatrix} \lambda_1 - \mu & 0 \\ 0 & \lambda_2 - \mu \end{bmatrix} \begin{bmatrix} \lambda_1 - \mu & 0 \\ 0 & \lambda_2 - \mu \end{bmatrix} \\ &= \begin{bmatrix} (\lambda_1 - \mu)^2 & 0 \\ 0 & (\lambda_2 - \mu)^2 \end{bmatrix}. \end{aligned} \quad (4.75)$$

Finally, we can compute the variance:

$$\begin{aligned} \langle (\Delta_\psi(\Omega))(\Delta_\psi(\Omega)) \rangle_\psi &= \begin{bmatrix} \overline{c_1} & \overline{c_2} \end{bmatrix} \begin{bmatrix} (\lambda_1 - \mu)^2 & 0 \\ 0 & (\lambda_2 - \mu)^2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ &= |c_1|^2 \times (\lambda_1 - \mu)^2 + |c_2|^2 \times (\lambda_2 - \mu)^2. \end{aligned} \quad (4.76)$$

We are now able to see that if both λ_1 and λ_2 are very close to μ , the term in the equation will be close to zero. Conversely, if either of the two eigenvalues is far from μ (it is immaterial whether above or below it, because we are taking squares), the variance will be a big real number. Conclusion: the variance does indeed inform us about the spread of the eigenvalues around their mean. \square

Our reader may still be a bit unsatisfied after this example: after all, what it shows is that the definition of variance given above works as it should in the case of diagonal matrices. Actually, it is a known fact that all hermitian matrices can be diagonalized by switching to a basis of eigenvectors, so the example is comprehensive enough to legitimize our definition.

Example 4.2.7 Let us calculate the variance of the operator described in Example 4.2.5:

$$\Delta_\psi(\Omega) = \Omega - \langle \Omega \rangle_\psi = \begin{bmatrix} 1 & -i \\ i & 2 \end{bmatrix} - \begin{bmatrix} 2.5 & 0 \\ 0 & 2.5 \end{bmatrix} = \begin{bmatrix} -1.5 & -i \\ i & -0.5 \end{bmatrix}. \quad (4.77)$$

We now compute $\Delta_\psi(\Omega) \star \Delta_\psi(\Omega)$:

$$\Delta_\psi(\Omega) \star \Delta_\psi(\Omega) = \begin{bmatrix} -1.5 & -i \\ i & -0.5 \end{bmatrix} \begin{bmatrix} -1.5 & -i \\ i & -0.5 \end{bmatrix} = \begin{bmatrix} 3.25 & 2i \\ -2i & 1.25 \end{bmatrix}. \quad (4.78)$$

Hence the variance is

$$\langle (\Delta_\psi(\Omega))(\Delta_\psi(\Omega)) \rangle_\psi = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}i \end{bmatrix} \begin{bmatrix} 3.25 & 2i \\ -2i & 1.25 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} = 0.25. \quad (4.79)$$

□

Exercise 4.2.11 Calculate the variance of the position operator. Show that the variance of position on any of its eigenvectors is zero. ■

Exercise 4.2.12 Calculate the variance of S_z on a generic spin state. Show that the variance of S_z reaches a maximum on the state $\frac{\sqrt{2}}{2}(|\downarrow\rangle + |\uparrow\rangle)$. ■

Note: The variance of the same hermitian varies from state to state: In particular, on an eigenvector of the operator the variance is zero, and the expected value is just the corresponding eigenvalue: we can say that an observable is sharp on its eigenvectors (no ambiguity on the outcome).

Exercise 4.2.13 Prove the preceding statement. (Hint: Work out some examples first.) ■

We have built all the machinery needed to introduce a fundamental theorem of quantum mechanics, known as Heisenberg's uncertainty principle. Let us begin with two observables, represented by the two hermitians Ω_1 and Ω_2 , and a given state, say, $|\psi\rangle$. We can compute the variance of Ω_1 and Ω_2 on $|\psi\rangle$, obtaining $\text{Var}_\psi(\Omega_1)$ and $\text{Var}_\psi(\Omega_2)$. Do these two quantities relate in any way, and if so, how?

Let us see what the question actually means. We have two observables, and our hope would be to simultaneously minimize their variances, thereby getting a sharp outcome for both. If there were no correlation in the variances, we could expect a very sharp measure of both observables on some convenient state (such as a common eigenvector, if any such existed). Alas, this is not the case, as shown by the following.

Theorem 4.2.1 (Heisenberg’s Uncertainty Principle). *The product of the variances of two arbitrary hermitian operators on a given state is always greater than or equal to one-fourth the square of the expected value of their commutator. In formulas:*

$$\text{Var}_\psi(\Omega_1) \times \text{Var}_\psi(\Omega_2) \geq \frac{1}{4} |\langle [\Omega_1, \Omega_2] \rangle_\psi|^2. \quad (4.80)$$

As promised, we have found our commutator once again. Heisenberg’s principle tells us that the commutator *measures how good a simultaneous measure of two observables can possibly be*. In particular, if the commutator happens to be zero (or equivalently, if the observables commute), there is no limit (at least in principle) to our accuracy. In quantum mechanics, however, there are plenty of operators that do not commute: in fact, we have seen that the directional spin operators provide one such example.

Exercise 4.2.14 Use the calculation of the commutator in Example 4.2.3 and Heisenberg’s principle to give an estimate of how accurate a simultaneous observation of spin in the z and x directions can be. ■

Another typical example, related to our first quantum system, is given by the pair position–momentum, which we have also met in the last section. So far $|\psi\rangle$ for the particle on the line has been described in terms of its position eigenbasis, i.e., the collection $\{|x_i\rangle\}$. $|\psi\rangle$ can be written in many other orthonormal bases, corresponding to different observables. One of those is the momentum eigenbasis. This basis comes up when we think of $|\psi\rangle$ as a wave (a bit like a wave hovering over the line). We can thus decompose it into its basic frequencies, just as we can resolve a sound into its basic pure tones. These pure tones are precisely the elements of the momentum eigenbasis.

The image of $|\psi\rangle$ in the position basis is as different as it can possibly be from the one associated with the momentum eigenbasis. The position eigenbasis is made of “peaks,” i.e., vectors that are zero everywhere except at a point (**Dirac’s deltas**, in math jargon). Therefore, $|\psi\rangle$ is decomposed into a weighted sum of peaks. The momentum eigenbasis, on the other hand, is made of sinusoids, whose position is totally undetermined.

The commutator of the position–momentum pair captures well this inherent dissimilarity: it is not zero, and therefore our hope to keep the comforting traditional picture of a particle as a tiny billiard ball moving around in space is dashed. If we can pin down the particle position at a given point in time (i.e., if the variance of its position operator is very small), we are at a loss as to its momentum (i.e., the variance of its momentum operator is very big), and vice versa.

Let us sum up:

- Observables are represented by hermitian operators. The result of an observation is always an eigenvalue of the hermitian.
- The expression $\langle \psi | \Omega | \psi \rangle$ represents the expected value of observing Ω on $|\psi\rangle$.
- Observables in general do not commute. This means that the order of observation matters. Moreover, if the commutator of two observables is not zero, there is an intrinsic limit to our capability of measuring their values simultaneously.

Programming Drill 4.2.1 *Continue your simulation of a quantum system by adding observables to the picture: the user will input a square matrix of the appropriate size, and a ket vector. The program will verify that the matrix is hermitian, and if so, it will calculate the mean value and the variance of the observable on the given state.*

4.3 MEASURING

The act of carrying out an observation on a given physical system is called **measuring**. Just as a single observable represents a specific question posed to the system, measuring is the process consisting of *asking a specific question and receiving a definite answer*.

In classical physics, we implicitly assumed that

- the act of measuring would leave the system in whatever state it already was, at least in principle; and
- the result of a measurement on a well-defined state is predictable, i.e., if we know the state with absolute certainty, we can anticipate the value of the observable on that state.

Both these assumptions proved wrong, as research in the subatomic scale has repeatedly shown: systems *do* get perturbed and modified as a result of measuring them. Furthermore, only the probability of observing specific values can be calculated: measurement is inherently a nondeterministic process.

Let us briefly recapitulate what we know: an observable can only assume one of its eigenvalues as the result of an observation. So far though, nothing tells us how frequently we are going to see a specific eigenvalue, say, λ . Moreover, our framework does not tell us yet what happens to the state vector if λ is actually observed. We need an additional postulate to handle concrete measures:

Postulate 4.3.1 *Let Ω be an observable and $|\psi\rangle$ be a state. If the result of measuring Ω is the eigenvalue λ , the state after measurement will always be an eigenvector corresponding to λ .*

Example 4.3.1 Let us go back to Example 4.2.1: It is easy to check that the eigenvalues of Ω are $\lambda_1 = -\sqrt{2}$ and $\lambda_2 = \sqrt{2}$ and the corresponding normalized eigenvectors are $|e_1\rangle = [-0.923i, -0.382]^T$ and $|e_2\rangle = [-0.382i, 0.923]^T$.

Now, let us suppose that after an observation of Ω on $|\psi\rangle = \frac{1}{2}[1, 1]^T$, the actual value observed is λ_1 . The system has “collapsed” from $|\psi\rangle$ to $|e_1\rangle$. □

Exercise 4.3.1 Find all the possible states the system described in Exercise 4.2.2 can transition into after a measurement has been carried out. ■

What is the probability that a normalized start state $|\psi\rangle$ will transition to a specific eigenvector, say, $|e\rangle$? We must go back to what we said in Section 4.1: the probability of the transition to the eigenvector is given by the square of the inner product of the two states: $|\langle e|\psi\rangle|^2$. This expression has a simple meaning: it is the *projection* of $|\psi\rangle$ along $|e\rangle$.

We are ready for a new insight into the real meaning of $\langle \Omega \rangle_\psi$ of the last section: first, let us recall that the normalized eigenvectors of Ω constitute an orthogonal basis of the state space. Therefore, we can express $|\psi\rangle$ as a linear combination in this basis:

$$|\psi\rangle = c_0|e_0\rangle + c_1|e_1\rangle + \cdots + c_{n-1}|e_{n-1}\rangle. \quad (4.81)$$

Now, let us compute the mean:

$$\langle \Omega \rangle_\psi = \langle \Omega \psi, \psi \rangle = |c_0|^2 \lambda_0 + |c_1|^2 \lambda_1 + \cdots + |c_{n-1}|^2 \lambda_{n-1}. \quad (4.82)$$

(Verify this identity!)

As we can now see, $\langle \Omega \rangle_\psi$ is precisely the mean value of the probability distribution

$$(\lambda_0, p_0), (\lambda_1, p_1), \dots, (\lambda_{n-1}, p_{n-1}), \quad (4.83)$$

where each p_i is the square of the amplitude of the collapse into the corresponding eigenvector.

Example 4.3.2 Let us go back to Example 4.3.1 and calculate the probabilities that our state vector will fall into one of the two eigenvectors:

$$p_1 = |\langle \psi | e_1 \rangle|^2 = 0.5 \quad \text{and} \quad p_2 = |\langle \psi | e_2 \rangle|^2 = 0.5. \quad (4.84)$$

Now, let us compute the mean value of the distribution:

$$p_1 \times \lambda_1 + p_2 \times \lambda_2 = 0.00, \quad (4.85)$$

which is precisely the value we obtained by directly calculating $\langle \psi | \Omega | \psi \rangle$. □

Exercise 4.3.2 Perform the same calculations as in the last example, using Exercise 4.3.1. Then draw the probability distribution of the eigenvalues as in the previous example. ■

Note: As a result of the foregoing discussion, an important fact emerges. Suppose we ask a specific question (i.e., we choose an observable) and perform a measurement once. We get an answer, say, λ , and the system transitions to the corresponding eigenvector. Now, let us ask the same question immediately thereafter. What is going to happen? The system will give exactly the *same* answer, and stay where it is. All right, you may say. But, what about changing the question? The following example will clarify matters.

Example 4.3.3 Until now we have dealt with measurements relative to only one observable. What if there were more than one observable involved? With each observable there is a different set of eigenvectors the system can possibly collapse to after a measurement has taken place. As it turns out, the answers we get will *depend on which order we pose our questions*, i.e., which observable we measure first.

There is an intriguing experiment that one can easily perform in order to see some of these ideas in action (and have some fun in the process). Suppose you shoot a beam of light. Light is also a wave, and like all waves it vibrates during its journey (think of sea waves). There are two possibilities: either it vibrates along all possible

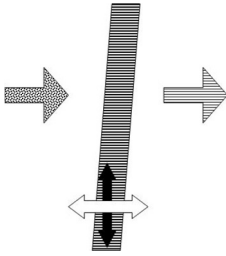


Figure 4.5. Light partially passing through one polarization sheet.

planes orthogonal to its line of propagation, or it does it only in a specific one. In the second case we say that light is **polarized**.¹¹ What kind of questions we can ask concerning polarization? We can set a specific plane, and ask: is light vibrating along this plane or is orthogonal?

For our experiment we need thin plastic semitransparent polarization sheets (they are fairly easy to obtain). Polarization sheets do two things: once you orient them in a specific direction, they measure the polarization of light in the orthogonal basis corresponding to that direction (let us call it the vertical–horizontal basis), and then filter out those photons that collapsed to one of the elements of the basis (Figure 4.5).

What if we had two sheets? If the two sheets were oriented in the same direction, there would be no difference whatsoever (why? because we are asking the same question; the photon will give once more the same exact answer). However, if we rotated the second sheet by 90° , then no light would pass through both sheets (Figure 4.6).

Placing the sheets orthogonal to each other ensures that the permitted half that passes through the left sheet is filtered out by the right sheet.

What happens if we add a *third* sheet? Placing a third sheet to the left or to the right of the other two sheets does not have any effect whatsoever. No light was permitted before and none will be allowed through the additional sheet. However, placing the third sheet in-between the other two *at an angle*, say, 45° , does have a remarkable effect (Figure 4.7).

Light will pass through all the three sheets! How can this be? Let us see what is going on here. The left sheet measures all the light relative to the up–down basis. The polarized light in the vertical polarization state that goes through is then considered to be a superposition with respect to the diagonal middle sheet measuring basis. The middle sheet recollapses the permitted half, filters some, and passes some through. But what is passed through is now in a diagonal polarization state. When this light passes through the right sheet, it is again in a superposition of the vertical–horizontal basis, and so it must collapse once more. Notice that only one-eighth of the original light passes through all three sheets. \square

¹¹ Polarization is a familiar phenomenon: fancy sun glasses are made on the basis of light polarization.

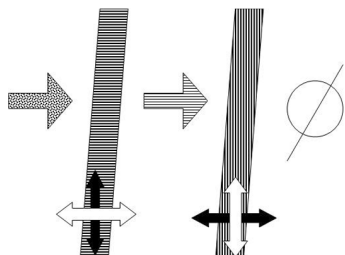


Figure 4.6. No light passing through two polarization sheets at orthogonal angles.

A brief summary is in order:

- The end state of the measurement of an observable is always one of its eigenvectors.
- The probability for an initial state to collapse into an eigenvector of the observable is given by the length squared of the projection.
- When we measure several observables, the order of measurements matters.

We have come a long way. We now have three main ingredients to cook up quantum dishes. We need one more, dynamics.

Programming Drill 4.3.1 *Next step in the simulation: when the user enters an observable and a state vector, the program will return the list of eigenvalues of the observable, the mean value of the observable on the state, and the probability that the state will transition to each one of the eigenstates. Optional: plot the corresponding probability distribution.*

4.4 DYNAMICS

Thus far, we have been concerned with static quantum systems, i.e., systems that do not evolve over time. To be sure, changes could still occur as a result of one or possibly many measurements, but the system itself was not time-dependent. In reality, of course, quantum systems do evolve over time, and we thus need to add a new hue to the canvas namely **quantum dynamics**. Just as hermitian operators represent physical observables, unitary operators introduce dynamics in the quantum arena.

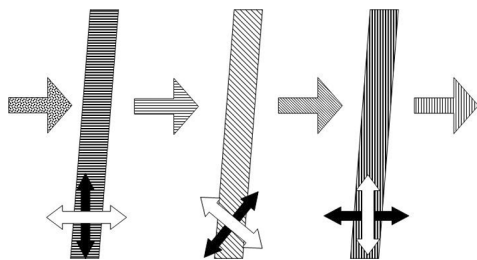


Figure 4.7. Light partially passing through three polarization sheets.

Postulate 4.4.1 *The evolution of a quantum system (that is not a measurement) is given by a unitary operator or transformation.*

That is, if U is a unitary matrix that represents a unitary operator and $|\psi(t)\rangle$ represents a state of the system at time t , then

$$|\psi(t+1)\rangle = U|\psi(t)\rangle \quad (4.86)$$

will represent the system at time $t+1$.

An important feature of unitary transformations is that they are closed under composition and inverse, i.e., the product of two arbitrary unitary matrices is unitary, and the inverse of a unitary transformation is also unitary. Finally, there is a multiplicative identity, namely, the identity operator itself (which is trivially unitary). In math jargon, one says that the set of unitary transformations constitutes a **group of transformations** with respect to composition.

Exercise 4.4.1 Verify that

$$U_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad U_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} \quad (4.87)$$

are unitary matrices. Multiply them and verify that their product is also unitary. ■

We are now going to see how dynamics is determined by unitary transformations: assume we have a rule, \mathcal{U} , that associates with each instant of time

$$t_0, t_1, t_2, \dots, t_{n-1} \quad (4.88)$$

a unitary matrix

$$\mathcal{U}[t_0], \mathcal{U}[t_1], \dots, \mathcal{U}[t_{n-1}]. \quad (4.89)$$

Let us start with an initial state vector $|\psi\rangle$. We can apply $\mathcal{U}[t_0]$ to $|\psi\rangle$, then apply $\mathcal{U}[t_1]$ to the result, and so forth. We will obtain a sequence of state vectors

$$\mathcal{U}[t_0]|\psi\rangle, \quad (4.90)$$

$$\mathcal{U}[t_1]\mathcal{U}[t_0]|\psi\rangle, \quad (4.91)$$

$$\vdots \quad (4.92)$$

$$\mathcal{U}[t_{n-1}]\mathcal{U}[t_{n-2}] \cdots \mathcal{U}[t_0]|\psi\rangle. \quad (4.93)$$

Such a sequence is called the **orbit**¹² of $|\psi\rangle$ under the action of $\mathfrak{U}[t_i]$ at the time clicks t_0, t_1, \dots, t_{n-1} .

$$\begin{array}{ccccccc}
 & \xrightarrow{\mathfrak{U}[t_0]} & & \xrightarrow{\mathfrak{U}[t_1]} & & \xrightarrow{\mathfrak{U}[t_2]} & \\
 |\psi\rangle & & \mathfrak{U}[t_0]|\psi\rangle & & \mathfrak{U}[t_1]\mathfrak{U}[t_0]|\psi\rangle & & \mathfrak{U}[t_2]\mathfrak{U}[t_1]\mathfrak{U}[t_0]|\psi\rangle \\
 & \xleftarrow{\mathfrak{U}[t_0]^\dagger} & & \xleftarrow{\mathfrak{U}[t_1]^\dagger} & & \xleftarrow{\mathfrak{U}[t_2]^\dagger} & \\
 & \longrightarrow & \cdots & \longrightarrow & \mathfrak{U}[t_{n-1}]\mathfrak{U}[t_{n-2}]\cdots\mathfrak{U}[t_0]|\psi\rangle. & &
 \end{array} \tag{4.94}$$

Observe that one can always go back, just like running a movie backward, simply by applying the inverses of $\mathfrak{U}[t_0], \mathfrak{U}[t_1], \dots, \mathfrak{U}[t_{n-1}]$ in reverse order: evolution of a quantum system is *symmetric* with respect to time.

We can now preview how a quantum computation will look. A quantum computer shall be placed into an initial state $|\psi\rangle$, and we shall then apply a sequence of unitary operators to the state. When we are done, we will measure the output and get a final state. The next chapters are largely devoted to working out these ideas in detail.

Here is an exercise for you on dynamics:

Exercise 4.4.2 Go back to Example 3.3.2 (quantum billiard ball), keep the same initial state vector $[1, 0, 0, 0]^T$, but change the unitary map to

$$\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{i}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & 0 & \frac{i}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}. \tag{4.95}$$

Determine the state of the system after three time steps. What is the chance of the quantum ball to be found at point 3? ■

The reader may wonder how the sequence $\mathfrak{U}[t_i]$ of unitary transformations is actually selected in real-life quantum mechanics. In other words, given a concrete quantum system, how is its dynamics determined? How does the system change? The answer lies in an equation known as the **Schrödinger equation**:¹³

$$\frac{|\psi(t + \delta t)\rangle - |\psi(t)\rangle}{\delta t} = -i \frac{2\pi}{\hbar} \mathcal{H}|\psi(t)\rangle. \tag{4.96}$$

¹² A small warning: one commonly thinks of an orbit as closed (a typical example is the orbit of the moon around the earth). In dynamics, this is not always the case: an orbit can be open or closed.

¹³ The version shown here is actually the discretized version of the original equation, which is a differential equation obtained from the above by letting δt become infinitesimal. It is this discretized version (or variants thereof) that is usually employed in computer simulation of quantum systems.

A complete discussion of this fundamental equation goes beyond the scope of this introductory chapter. However, without going into technical details, we can at least convey its spirit. Classical mechanics taught physicists that the global energy of an isolated system is preserved throughout its evolution.¹⁴ Energy is an observable, and therefore for a concrete quantum system it is possible to write down a hermitian matrix representing it (this expression will of course vary from system to system). This observable is called the **hamiltonian** of the system, indicated by \mathcal{H} in Equation (4.96).

The Schrödinger equation states that the rate of variation of the state vector $|\psi(t)\rangle$ with respect to time at the instant t is equal (up to the scalar factor $\frac{2\pi}{h}$) to $|\psi(t)\rangle$ multiplied by the operator $-i * \mathcal{H}$. By solving the equation with some initial conditions one is able to determine the evolution of the system over time.

Time for a small recap:

- Quantum dynamics is given by unitary transformations.
- Unitary transformations are invertible; thus, all closed system dynamics are reversible in time (as long as no measurement is involved).
- The concrete dynamics is given by the Schrödinger equation, which determines the evolution of a quantum system whenever its hamiltonian is specified.

Programming Drill 4.4.1 *Add dynamics to your computer simulation of the particle on a grid: the user should input a number of time steps n , and a corresponding sequence of unitary matrices U_n of the appropriate size. The program will then compute the state vector after the entire sequence U_n has been applied.*

4.5 ASSEMBLING QUANTUM SYSTEMS

The opening section of this chapter described a simple quantum system: a particle moving in a confined one-dimensional grid (the set of points $\{x_0, x_1, \dots, x_{n-1}\}$). Now, let us suppose that we are dealing with two particles confined to the grid. We shall make the following assumption: the points on the grid that can be occupied by the first particle will be $\{x_0, x_1, \dots, x_{n-1}\}$. The second particle can be at the points $\{y_0, y_1, \dots, y_{m-1}\}$.

x_0 x_1 \dots x_{n-1} y_0 y_1 \dots y_{m-1}

\bullet — \bullet — \dots — \bullet \bullet — \bullet — \dots — \bullet

(4.97)

Can we lift the description we already have to this new setup? Yes. The details will keep us busy in this section.

Our answer will not be confined to the aforementioned system. Instead, it will provide us with a quantum version of a building block game, i.e., a way of assembling more complex quantum systems starting from simpler ones. This procedure

¹⁴ For instance, a stone dropped from a height falls down in such a way that its kinetic energy plus its potential energy plus energy dissipated from attrition is constant.

lies at the very core of modern quantum physics: it enables physicists to model multiparticle quantum systems.¹⁵

We need one last expansion of our quantum dictionary: *assembling quantum systems means tensoring the state space of their constituents*.

Postulate 4.5.1 Assume we have two independent quantum systems Q and Q' , represented respectively by the vector spaces \mathbb{V} and \mathbb{V}' . The quantum system obtained by merging Q and Q' will have the tensor product $\mathbb{V} \otimes \mathbb{V}'$ as a state space.

Notice that the postulate above enables us to assemble as many systems as we like. The tensor product of vector spaces is associative, so we can progressively build larger and larger systems:

$$\mathbb{V}_0 \otimes \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_k. \quad (4.98)$$

Let us go back to our example. To begin with, there are $n \times m$ possible basic states:

$|x_0\rangle \otimes |y_0\rangle$, meaning the first particle is at x_0 and the second particle at y_0 .
 $|x_0\rangle \otimes |y_1\rangle$, meaning the first particle is at x_0 and second particle at y_1 .
 \vdots
 $|x_0\rangle \otimes |y_{m-1}\rangle$, meaning the first particle is at x_0 and the second particle at y_{m-1} .
 $|x_1\rangle \otimes |y_0\rangle$, meaning the first particle is at x_1 and the second particle at y_0 .
 \vdots
 $|x_i\rangle \otimes |y_j\rangle$, meaning the first particle is at x_i and the second particle at y_j .
 \vdots
 $|x_{n-1}\rangle \otimes |y_{m-1}\rangle$, meaning the first particle is at x_{n-1} and the second particle at y_{m-1} .

Now, let us write the generic state vector as a superposition of the basic states:

$$|\psi\rangle = c_{0,0}|x_0\rangle \otimes |y_0\rangle + \cdots + c_{i,j}|x_i\rangle \otimes |y_j\rangle + \cdots + c_{n-1,m-1}|x_{n-1}\rangle \otimes |y_{m-1}\rangle, \quad (4.99)$$

which is a vector in the $(n \times m)$ -dimensional complex space $\mathbb{C}^{n \times m}$.

The quantum amplitude $|c_{i,j}|$ squared will give us the probability of finding the two particles at positions x_i and y_j , respectively, as shown by the following example.

Example 4.5.1 Assume $n = 2$ and $m = 2$ in the above. We are thus dealing with the state space \mathbb{C}^4 whose standard basis is

$$\{|x_0\rangle \otimes |y_0\rangle, |x_0\rangle \otimes |y_1\rangle, |x_1\rangle \otimes |y_0\rangle, |x_1\rangle \otimes |y_1\rangle\}. \quad (4.100)$$

¹⁵ By thinking of fields such as the electromagnetic field as systems composed of infinitely many particles, this procedure makes field theory amenable to the quantum approach.

Now, let us consider the state vector for the two-particle system given by

$$|\psi\rangle = i|x_0\rangle \otimes |y_0\rangle + (1-i)|x_0\rangle \otimes |y_1\rangle + 2|x_1\rangle \otimes |x_0\rangle + (-1-i)|x_1\rangle \otimes |x_1\rangle. \quad (4.101)$$

What is the probability of finding the first particle at location x_1 and the second one at y_1 ? We look at the last amplitude in the list given before, and use the same recipe as in the one-particle system:

$$p(x_1, y_1) = \frac{|-1-i|^2}{|i|^2 + |1-i|^2 + |2|^2 + |-1-i|^2} = 0.2222. \quad (4.102)$$

□

Exercise 4.5.1 Redo the steps of the last example when $n = m = 4$ and $c_{0,0} = c_{0,1} = \dots = c_{3,3} = 1 + i$. ■

The same machinery can be applied to any other quantum system. For instance, it is instructive to generalize our spin example of Section 4.1 to a system where many particles are involved. You can try yourself.

Exercise 4.5.2 Write down the generic state vector for the system of two particles with spin. Generalize it to a system with n particles (this is important: it will be the physical realization for quantum registers!). ■

Now that we are a bit familiar with quantum assemblage, we are ready for the final puzzling surprise of quantum mechanics: **entanglement**. Entanglement will force us to abandon one last comforting faith, namely, that assembled complex systems can be understood completely in terms of their constituents.

The basic states of the assembled system are just the tensor product of basic states of its constituents. It would be nice if *each* generic state vector could be rewritten as the tensor product of two states, one coming from the first quantum subsystem and the other one from the second. It turns out that this is not true, as is easily shown by this example.

Example 4.5.2 Let us work on the simplest nontrivial two-particle system: each particle is allowed only two points. Consider the state

$$|\psi\rangle = |x_0\rangle \otimes |y_0\rangle + |x_1\rangle \otimes |y_1\rangle. \quad (4.103)$$

In order to clarify what is left out, we might write this as

$$|\psi\rangle = 1|x_0\rangle \otimes |y_0\rangle + 0|x_0\rangle \otimes |y_1\rangle + 0|x_1\rangle \otimes |y_0\rangle + 1|x_1\rangle \otimes |y_1\rangle. \quad (4.104)$$

Let us see if we can write $|\psi\rangle$ as the tensor product of two states coming from the two subsystems. Any vector representing the first particle on the line can be written as

$$c_0|x_0\rangle + c_1|x_1\rangle. \quad (4.105)$$

Similarly, any vector representing the second particle on the line can be written as

$$c'_0|y_0\rangle + c'_1|y_1\rangle. \quad (4.106)$$

Therefore, if $|\psi\rangle$ came from the tensor product of the two subsystems, we would have

$$(c_0|x_0\rangle + c_1|x_1\rangle) \otimes (c'_0|y_0\rangle + c'_1|y_1\rangle) = c_0c'_0|x_0\rangle \otimes |y_0\rangle + c_0c'_1|x_0\rangle \otimes |y_1\rangle \\ + c_1c'_0|x_1\rangle \otimes |y_0\rangle + c_1c'_1|x_1\rangle \otimes |y_1\rangle. \quad (4.107)$$

For our $|\psi\rangle$ in Equation (4.104) this would imply that $c_0c'_0 = c_1c'_1 = 1$ and $c_0c'_1 = c_1c'_0 = 0$. However, these equations have no solution. We conclude that $|\psi\rangle$ cannot be rewritten as a tensor product.

Let us go back to $|\psi\rangle$ and see what it physically means. What would happen if we measured the first particle? A quick calculation will show that the first particle has a 50–50 chance of being found at the position x_0 or at x_1 . So, what if it is, in fact, found in position x_0 ? Because the term $|x_0\rangle \otimes |y_1\rangle$ has a 0 coefficient, we know that there is no chance that the second particle will be found in position y_1 . We must then conclude that the second particle can only be found in position y_0 . Similarly, if the first particle is found in position x_1 , then the second particle must be in position y_1 . Notice that the situation is perfectly symmetrical with respect to the two particles, i.e., it would be the same if we measured the second one first. The individual states of the two particles are intimately related to one another, or **entangled**. The amazing side of this story is that the x_i 's can be light years away from the y_j 's. Regardless of their actual distance in space, a measurement's outcome for one particle will always determine the measurement's outcome for the other one.

The state $|\psi\rangle$ is in sharp contrast to other states like

$$|\psi'\rangle = |x_0\rangle \otimes |y_0\rangle + |x_0\rangle \otimes |y_1\rangle + |x_1\rangle \otimes |y_0\rangle + |x_1\rangle \otimes |y_1\rangle. \quad (4.108)$$

Here, finding the first particle at a particular position does not provide any clue as to where the second particle will be found (check it!). \square

States that can be broken into the tensor product of states from the constituent subsystems (like $|\psi'\rangle$) are called **separable states**, whereas states that are unbreakable (like $|\psi\rangle$) are referred to as **entangled states**.

Exercise 4.5.3 Assume the same scenario as in Example 4.5.2 and let

$$|\phi\rangle = |x_0\rangle \otimes |y_1\rangle + |x_1\rangle \otimes |y_1\rangle. \quad (4.109)$$

Is this state separable? \blacksquare

A clear physical case of entanglement is in order. We must revert to spin. Just as there are laws of conservation of momentum, angular momentum, energy-mass, and other physical properties, so too there is a law of conservation of total spin of a quantum system. This means that in an isolated system the total amount of spin must stay the same. Let us fix a specific direction, say, the vertical one (z axis), and the corresponding spin basis, up and down. Consider the case of a quantum system,

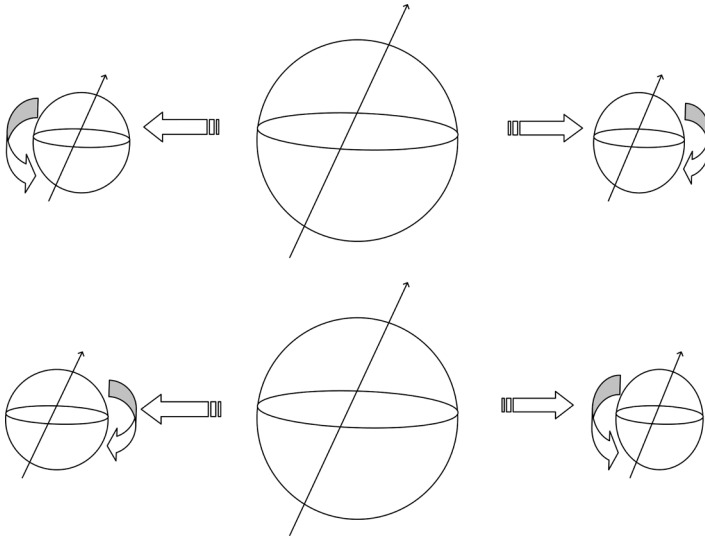


Figure 4.8. Two possible scenarios of a composite system where the total spin is zero.

such as a composite particle, whose total spin is zero. This particle might split up at some point in time into two other particles that do have spin (Figure 4.8).

The spin states of the two particles will now be entangled. The law of conservation of spin stipulates that because we began with a system of total spin zero, the sum of the spins of the two particles must cancel each other out. This amounts to the fact that if we measure the spin of the left particle along the z axis and we find it in state $|\uparrow_L\rangle$ (where the subscript is to describe which particle we are dealing with), then it must be that the spin of the particle on the right will be $|\downarrow_R\rangle$. Similarly, if the state of the left particle is $|\downarrow_L\rangle$, then the spin of the right particle must be $|\uparrow_R\rangle$.

We can describe this within our notation. In terms of vector spaces, the basis that describes the left particle is $\mathcal{B}_L = \{|\uparrow_L\rangle, |\downarrow_L\rangle\}$ and the basis that describes the right particle is $\mathcal{B}_R = \{|\uparrow_R\rangle, |\downarrow_R\rangle\}$. The basis elements of the entire system are

$$\{|\uparrow_L \otimes \uparrow_R\rangle, |\uparrow_L \otimes \downarrow_R\rangle, |\downarrow_L \otimes \uparrow_R\rangle, |\downarrow_L \otimes \downarrow_R\rangle\}. \quad (4.110)$$

In such a vector space, our entangled particles can be described by

$$\frac{|\uparrow_L \otimes \downarrow_R\rangle + |\downarrow_L \otimes \uparrow_R\rangle}{\sqrt{2}}, \quad (4.111)$$

similar to Equation (4.104). As we said before, the combinations $|\uparrow_L \otimes \uparrow_R\rangle$ and $|\downarrow_L \otimes \downarrow_R\rangle$ cannot occur because of the law of conservation of spin. When one measures the left particle and it collapses to the state $|\uparrow_L\rangle$ then *instantaneously* the right particle will collapse to the state $|\downarrow_R\rangle$, *even if the right particle is millions of light-years away*.

How will entanglement arise in the tale that we are telling? We find in Chapter 6 that it plays a central role in algorithm design. It is also used extensively in Chapter 9 while discussing cryptography (Section 9.4) and teleportation (Section 9.5). Entanglement makes a final appearance in Chapter 11, in connection with decoherence.

What have we learned?

- We can use the tensor product to build complex quantum systems out of simpler ones.
- The new system cannot be analyzed simply in terms of states belonging to its subsystems. An entire set of new states has been created, which cannot be resolved into their constituents.

Programming Drill 4.5.1 *Expand the simulation of the last sections by letting the user choose the number of particles.*

.....
References: There are many elementary introductions to quantum mechanics that are very readable. Here is a list of some of them: Chen (2003), Gillespie (1974), Martin (1982), Polkinghorne (2002), and White (1966).

Special mention must be made of the classic introduction by P.A.M. Dirac (1982). Seventy years after its first publication, it remains a classic that is worth reading.

For a more advanced and modern presentation see, e.g., Volume III of Feynman (1963), Hannabuss (1997), Sakurai (1994), or Sudbery (1986).

For a short history of the early development of quantum mechanics, see Gamow (1985).