

## Chapter 9

# Quantum mechanics

By the end of the 19<sup>th</sup> century, physicists thought they had figured out most of what there is to know about the laws of nature. Newton's laws of mechanics described the motion of objects in space, and Maxwell's equations described electricity and magnetism. Wave phenomena—including the propagation of sound, light, and waves on the surface of liquids—were also well understood. Only a few small inconsistencies between theory and experiments with atoms and radiation remained unsolved.

“[...] it seems probable that most of the grand underlying principles have now been firmly established and that further advances are to be sought chiefly in the rigorous application of these principles to all the phenomena which come under our notice.”

—Albert A. Michelson in 1894

Physicists like Michelson were worried about the future of physics research. It was as if they were wondering, “What are we going to do now that we've figured everything out?” Little did they know about the quantum storm that was about to hit physics, and with it, the complete rewrite of our understanding of nature at the smallest scale.

Understanding the structure of atoms—the smallest constituents of matter known at the time—was no trivial task. Describing the absorption of electromagnetic radiation by metals also turned out to be quite complicated. In both cases, the physical theories of the time predicted that the energy of physical systems could take on any value; yet experimental observations showed *discrete* energy levels. Imagine you throw a (very-very tiny) ball, and the laws of physics force you to choose an initial velocity for the ball from a list of “allowed” values: 0 m/s, 1 m/s, 2 m/s, 3 m/s, and so forth. That would

be weird, no? Weird indeed, and this is the situation physicists were facing in the beginning of the 20<sup>th</sup> century: their theories described the energy levels of atoms as real numbers  $E \in \mathbb{R}$ , but experiments showed that only a discrete set of energy levels exist. For example, the energy levels that the electrons of the hydrogen atom can take on are:

$$E \in \{ 21.8 \times 10^{-19} \text{ J}, 5.44 \times 10^{-19} \text{ J}, 2.42 \times 10^{-19} \text{ J}, \\ 13.6 \times 10^{-20} \text{ J}, 8.71 \times 10^{-20} \text{ J}, 6.05 \times 10^{-20} \text{ J}, \dots \}.$$

Other experimental observations suggested that electromagnetic radiation is not a continuous wave, but comes in discrete “wave packets,” which we call *photons* today. The theory of *quantum mechanics* was born out of a need to explain these observations. The term *quantum*, from the Latin *quantus* for quantity, was coined to describe the discrete nature of the phenomena that physicists were trying to explain.

During the first half of the 20<sup>th</sup> century, in experiment after experiment, quantum principles were used to correctly predict many previously-unexplained observations. During the second half of the 20<sup>th</sup> century, biologists, chemists, engineers, and physicists applied quantum principles to all areas of science. This process of “upgrading” classical models to quantum models led to a better understanding of the laws of nature, and the discovery of useful things like transistors and lasers.

The fundamental principles of quantum mechanics can be explained in the space on the back of an envelope. Understanding quantum mechanics is a matter of combining a little knowledge of linear algebra (vectors, inner products, projections) with some probability theory (Chapter 8). In this chapter, we’ll take a little excursion to the land of physics to learn about the ideas of great scientists like Bohr, Planck, Dirac, Heisenberg, and Pauli. Your linear algebra skills will allow you to learn about some fascinating 20<sup>th</sup>-century discoveries. This chapter is totally optional reading, reserved for readers who *insist* on learning about the quantum world. If you’re not interested in quantum mechanics, it’s okay to skip this chapter, but I recommend you check out Section 9.3 on *Dirac notation* for vectors and matrices. Learning Dirac notation serves as an excellent review of the core concepts of linear algebra.

## 9.1 Introduction

The principles of quantum mechanics have far-reaching implications for many areas of science: physics, chemistry, biology, engineering,

philosophy, and many other fields of study. Each field of study has its own view on quantum mechanics, and has developed a specialized language for describing quantum concepts. We'll formally introduce the postulates of quantum mechanics in Section 9.5, but before we get there, let's look at some of the disciplines where quantum principles are used.

**Physics** Physicists use the laws of quantum mechanics as a toolbox to understand and predict the outcomes of atomic-scale physics experiments. By “upgrading” classical physics models to reflect the ideas of quantum mechanics, physicists (and chemists) obtain more accurate models that lead to better predictions.

For example, in a *classical* physics model, the motion of a particle is described by its position  $x(t)$  and velocity  $v(t)$  as functions of time:

$$\text{classical state} = (x(t), v(t)), \text{ for all times } t.$$

At any given time  $t$ , the particle is at position  $x(t)$  and moving with velocity  $v(t)$ . Using Newton's laws of motion and calculus, we can predict the position and the velocity of a particle at all times.

In a quantum description of the motion of a particle in one dimension, the state of a particle is represented by a *wave function*  $|\psi(x, t)\rangle$ , which is a complex-valued function of position  $x$  and time  $t$ :

$$\text{quantum state} = |\psi(x, t)\rangle, \text{ for all times } t.$$

At any given time  $t$ , the state of the particle corresponds to a complex-valued function of a real variable  $|\psi(x)\rangle \in \{\mathbb{R} \rightarrow \mathbb{C}\}$ . The wave function  $|\psi(x)\rangle$  is also called the *probability-amplitude* function. The probability of finding the particle at position  $x_a$  is proportional to the value of the squared norm of the wave function:

$$\Pr(\{\text{particle position} = x_a\}) \propto ||\psi(x_a)\rangle|^2.$$

Instead of having a definite position  $x(t)$  as in the classical model, the position of the particle in a quantum model is described by a probability distribution calculated from its wave function  $|\psi(x)\rangle$ . Instead of having a definite momentum  $p(t)$ , the momentum of a quantum particle is another function calculated based on its wave function  $|\psi(x)\rangle$ .

Classical models provide accurate predictions for physics problems involving macroscopic objects, but fail to predict the physics of atomic-scale phenomena. Much of 20<sup>th</sup>-century physics research efforts were dedicated to the study of quantum concepts like ground states, measurements, spin angular momentum, polarization, uncertainty, entanglement, and non-locality.

**Computer science** Computer scientists understand quantum mechanics using principles of information. Quantum principles impose a fundamental change to the “data types” used to represent information. Classical information is represented as *bits*, elements of the finite field of size two  $\mathbb{Z}_2$ :

$$\text{bit: } x = 0 \text{ or } x = 1.$$

In the quantum world, the fundamental unit of information is the *qubit*, which is a two-dimensional unit vector in a complex inner product space:

$$\text{qubit: } |x\rangle = \alpha|0\rangle + \beta|1\rangle.$$

This change to the underlying information model requires reconsidering fundamental information processing tasks like computation, data compression, encryption, and communication.

**Philosophy** Philosophers have also updated their conceptions of the world to incorporate the laws of quantum mechanics. Observations of physics experiments forced them to reconsider the fundamental question, “What are things made of?” Another interesting question philosophers have considered is whether the quantum state  $|\psi\rangle$  of a physical system really exists, or if  $|\psi\rangle$  is a representation of our knowledge about the system.

A third central philosophy concept that quantum mechanics calls into question is *determinism*—the clockwork-model of the universe, where each effect has a cause we can trace, like the connections between gears in a mechanical clock. The laws of physics tell us that the next state of the universe is determined by the current state of the universe, and the state changes according to the equations of physics. However, representing the universe as a quantum state has implications for our understanding of how the universe “ticks.” Clockwork (deterministic) models of the universe are not wrong—they just require a quantum upgrade.

Many scientists are also interested in the philosophical aspects of quantum mechanics. Physicists call these types of questions *foundations* or *interpretations*. Since different philosophical interpretations of quantum phenomena cannot be tested experimentally, these questions are considered outside the scope of physics research. Nevertheless, these questions are so deep and fascinating that physicists continue to pursue them, and contribute interesting philosophical work.

[ Philosophical issues in quantum theory ]

<http://plato.stanford.edu/entries/qt-issues/>

## Physical models of the world

We'll situate our discussion of quantum mechanics within two conceptual worlds:

- The **real world** is where physical experiments are performed.
- The **mathematical world** is a purely theoretical construct that aims to model certain aspects of the real world.

The better the mathematical model, the more closely its predictions correspond to the behaviour of real-world systems. Sill, no math or physics model can ever predict real-world outcomes with 100% accuracy. When we say that a certain mathematical model is better than another, we mean it can predict the outcomes of controlled experiments with greater accuracy. Physicists are very open to new theories—anyone can be a physicist! You can start by constructing any crazy mathematical model for describing nature, and if your model correctly predicts the outcomes of experiments, other physicists will start using it.

We can make a further distinction among mathematical models, classifying them into two categories depending on the type of math they use:

- **Classical models** describe the world in terms of real variables like positions and velocities.
- **Quantum models** describe systems in terms of vectors in complex vector spaces.

Table 9.1 compares the objects used in the two types of mathematical models of the real world. In physics, classical models describe the motion of particles using trajectories  $\vec{r}(t)$ , whereas quantum models use wave functions  $|\psi(\vec{r}, t)\rangle$ . In computer science, classical information is stored in bits  $i \in \{0, 1\}$ , whereas quantum information is stored in qubits  $|x\rangle \in \mathbb{C}^2$ .

**Example** Let's analyze the difference between classical and quantum models of the real world using an example. Consider a photon (a particle of light) going through an optical circuit that consists of several lenses, mirrors, and other optical instruments. A photon detector is placed at position  $x_f$  at the end of the circuit. The objective of the experiment is to predict if the photon will arrive at the detector and cause it to “click.” The two possible outcomes of the experiment are `click` (photon arrives at detector) or `noclick` (photon doesn't arrive at detector).<sup>1</sup>

---

<sup>1</sup>We're assuming the detector has 100% efficiency (detects every photon that arrives at it) and generates zero noise (no false-positive clicks).

**Real world:**

- The motion of a ball thrown in the air
- The motion of an electron through space
- The paths of light particles moving through optical circuits
- The electric current flowing though a superconducting loop

**Classical models:**

- $x(t) \in \{\mathbb{R} \rightarrow \mathbb{R}\}$
- $\vec{r}(t) \in \{\mathbb{R} \rightarrow \mathbb{R}^3\}$
- $i \in \mathbb{Z}_2 \equiv \{0, 1\}$  (a bit)
- $j \in \mathbb{Z}_d$

**Quantum models:**

- $|\psi(x, t)\rangle \in \{\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}\}$
- $|\psi(\vec{r}, t)\rangle \in \{\mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{C}\}$
- $|x\rangle \in \mathbb{C}^2$  (a qubit)
- $|y\rangle \in \mathbb{C}^d$

**Table 9.1:** Examples of the math used in classical and quantum models.

A classical model of the motion of the photon calculates the photon's position at all times  $x(t)$  and leads to the prediction  $i = 1$  (click) if  $x_f = x(t)$ , for some  $t$ . On the other hand, if the detector does not lie on the photon's trajectory, then the classical model will predict  $i = 0$  (noclick).

A quantum model would describe the photon's trajectory through the circuit as a linear combination of two different possible paths:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad \text{where } |\alpha|^2 + |\beta|^2 = 1.$$

Here  $|1\rangle$  describes paths that arrive at the detector, and  $|0\rangle$  describes paths that don't. The coefficients  $\alpha$  and  $\beta$  describe the relative "weights" of the different paths. Using the quantum model, we can obtain a probabilistic prediction of whether the detector will click or not:

$$\Pr(\text{noclick}) = |\alpha|^2 \quad \text{and} \quad \Pr(\text{click}) = |\beta|^2.$$

For this example, both the classical and the quantum models describe the same real-world phenomenon. We can test the validity of both models by comparing the models' predictions with what happens in reality.

Note that the two models make very different assumptions about reality. The classical model assumes the photon follows a single path through the circuit, whereas the quantum model assumes the photon can take multiple paths through the circuit. Despite the difference in the mathematical substrate of the models and their fundamentally different views of reality, we can compare the two models' predictions on the same footing. Note it doesn't make sense to say one

model is more real than the other. The only thing that is real is the photon in the optical circuit, and the photon doesn't care whether you use classical or quantum models to describe its path.

## Quantum model peculiarities

We'll now comment on the relative “intuitiveness” of classical and quantum models and introduce the concept of *quantum measurement*, which is of central importance in quantum mechanics.

Classical models have the advantage of being more intuitively understandable than quantum models. The variables in classical models often correspond to measurable aspects of real-world systems. We can identify the position variable in a classical model with the position of a particle in the real world. Velocity and momentum are harder to understand intuitively, but we have some general intuition about motion and collisions from everyday life. In general, we can understand classical models more readily because it's easier for us to think about a mechanistic, clockwork-like universe, in which objects push on each other with clearly defined cause and effect, like a clock that goes click, click, click.

In contrast, we do not enjoy such intuitive interpretation of quantum models, since our senses cannot directly perceive movement and interaction at the quantum level. Because quantum models' states do not directly correspond to observable aspects in the real world, quantum models are often described as mysterious and counterintuitive. Quantum models are harder to understand in part because they use complex vector quantities to represent systems, and complex numbers are more difficult to visualize. For example, visualizing the complex-valued state of a photon  $|\psi\rangle$  is difficult, since you must think about both the real part and the imaginary part of  $|\psi\rangle$ . Even though we can't see what  $|\psi\rangle$  looks like, we can describe it using an equation, and do math calculations with it. In particular, we can compare the predictions obtained from calculations based on the quantum state  $|\psi\rangle$  to measurements performed in the real world.

The process of *quantum measurement* is how we map the predictions of the quantum model to observable quantities. A quantum measurement acts on a particle's wave function  $|\psi\rangle$  to produce a classical outcome. **Performing quantum measurements is like asking questions of particles, and the measurement outcomes are the answers to these questions.**

What is your position?  $\Leftrightarrow$  position( $|\psi\rangle$ ) =  $x \in \mathbb{R}$

What is your momentum?  $\Leftrightarrow$  momentum( $|\psi\rangle$ ) =  $p \in \mathbb{R}$

What is your spin momentum?  $\Leftrightarrow$  spin <sub>$\uparrow\downarrow$</sub> ( $|\psi\rangle$ ) =  $s \in \{\uparrow, \downarrow\}$

Since measurement outcomes correspond to real-world quantities that can be measured, we can judge the merits of quantum models the same way we judge the merits of classical models—in terms of the quality of their predictions.

## Chapter overview

In the next section, we'll describe a tabletop experiment involving lasers and polarization lenses, with an outcome that's difficult to explain using classical physics. The remainder of the chapter will introduce the tools needed to explain the outcome of this experiment in terms of quantum physics. We'll start by introducing a special notation for vectors that is used to describe quantum phenomena (Section 9.3).

In Section 9.5, we'll formally define the “rules” of quantum mechanics, also known as the *postulates* of quantum mechanics. We'll learn the “rules of the game” using the simplest possible quantum systems (qubits), and define how quantum systems are prepared, how we manipulate them using *quantum operations*, and how we extract information from them using *quantum measurements*. This part of the chapter is based on the notes from the introductory lectures of a graduate-level quantum information course, so don't think you'll be getting some watered-down, hand-wavy version of quantum mechanics. You'll learn the real stuff, because I know you can handle it.

In Section 9.6 we'll apply the quantum formalism to the polarizing lenses experiment, showing that a quantum model leads to the correct qualitative and quantitative prediction for the observed outcome. We'll close the chapter with short explanations of different applications of quantum mechanics with pointers for further exploration about each topic.

Throughout the chapter, we'll focus on *matrix* quantum mechanics and use computer science language to describe quantum phenomena. A computer science approach allows us to discuss the fundamental aspects of quantum theory without introducing all the physics required to understand atoms. Finally, I just might throw in a sample calculation using the wave function of the hydrogen atom, to give you an idea of what that's like.

## 9.2 Polarizing lenses experiment

Let's run through a simple tabletop experiment that illustrates the limitations of classical, deterministic reasoning. The outcome of the



experiment will highlight the need for careful consideration of the measurements used in scientific experiments.

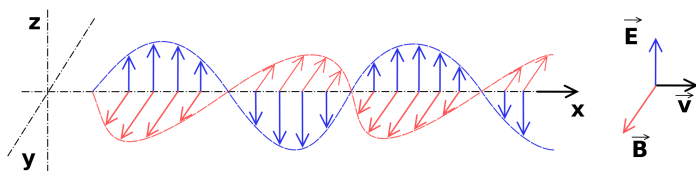
We'll describe the experiment using words and diagrams, but you can easily reproduce the experiment in your own "lab," since it requires only simple equipment. I encourage you to try it yourself. You'll need three polarizing lenses, a laser pointer, a piece of paper, and three binder clips for holding the lenses upright. You can buy polarizing lenses on the cheap from a second-hand camera shop—any polarizing lens will do.

## Background

In photography, polarizing lenses are used to filter out undesirable light reflections, like reflections that occur from water surfaces or glass windows. To better understand the experiment, we need to introduce some basic notions about the physics of light, specifically the concept of light polarization.

Light consists of photons. Photons are travelling pulses of electromagnetic energy. Electromagnetic energy can travel through space in the form of a wave. Polarization refers to the orientation of the electric field  $\vec{E}$  of a propagating electromagnetic wave.

Light is normally unpolarized, meaning it corresponds to a mixture of photons that have electric and magnetic components of random orientation. A light beam is *polarized* if all its photons have the same orientation of their electric field.

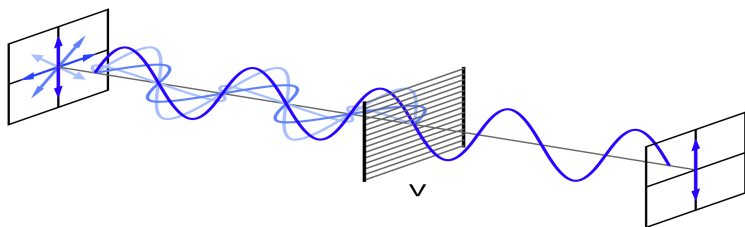


**Figure 9.1:** A photon is a pulse of electromagnetic energy. The energy of a photon travels in the form of a wave that has an electric component  $\vec{E}$  and a magnetic component  $\vec{B}$ . The figure shows a photon travelling in the positive  $x$ -direction with its electric component along the  $z$ -axis.

Light reflected from flat surfaces, like the surface of a lake or a glass window, becomes polarized, which means the electric components of all the reflected photons become aligned.

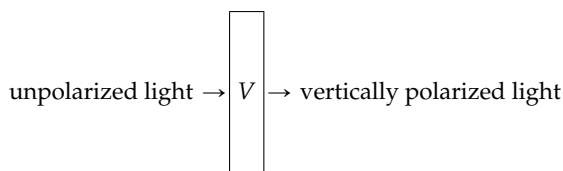
Photographers use this fact to selectively filter out light with a particular polarization. A *polarizing filter* or *polarizing lens* has a special coating which conducts electricity in one direction, but not in the other. You can think of a polarizing filter as a surface covered

by tiny conductive bands that interact with the electric component of incoming light particles. Light rays that hit the filter will either pass through or be reflected depending on their polarization. Light particles with a polarization perpendicular to the conductive bands pass through the filter, while light particles with polarization parallel to the conductive bands are reflected. This is because the filter's surface has different conductive properties in the parallel and perpendicular directions.



**Figure 9.2:** Incoming photons interact with the horizontal conductive bands of a polarizing filter. The horizontal bands of the filter reflect the horizontal component of the photons's electric field. Vertically-polarized photons pass through the filter because the conductive bands are perpendicular to their electric field. Thus, a vertically polarizing filter denoted  $V$  allows only vertically polarized light to pass through.

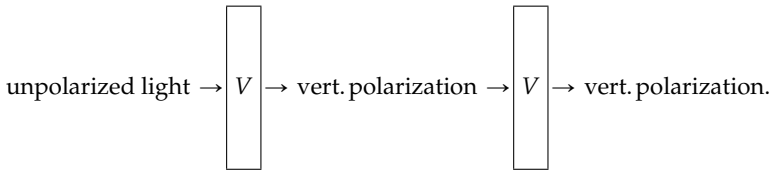
Consider the illustration in Figure 9.3. The effect of a vertically polarizing lens on a beam of light is to only allow vertically polarized light to pass through.



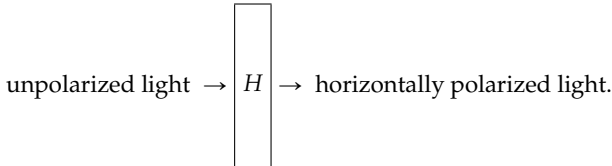
**Figure 9.3:** A vertically polarizing lens ( $V$ ) allows only vertically polarized light particles to pass through.

In Figure 9.4 we see another aspect of polarizing lenses. If the light is already vertically polarized, adding a second vertically polarizing lens will not affect the beam. All light that passes through the first lens will also pass through the second.

Taking a vertically polarizing lens and rotating it by 90 degrees turns it into a horizontally polarizing lens. See Figure 9.5. Note that horizontally polarizing lenses and vertically polarizing lenses are complementary: vertically polarized light will not pass through

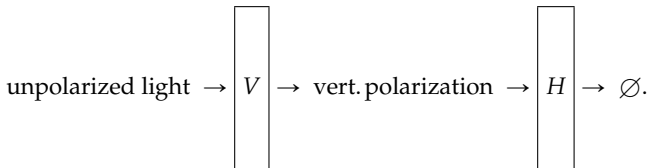


**Figure 9.4:** A second vertically polarizing lens has no further effect since light is already vertically polarized.



**Figure 9.5:** A horizontally polarizing lens (H) allows only horizontally polarized light particles to pass through.

a horizontally polarizing lens. This situation is illustrated in Figure 9.6.



**Figure 9.6:** Placing a horizontally polarizing lens after the vertically polarizing lens has the effect of filtering all light. Zero photons make it through both filters, which we indicate with the empty set symbol  $\emptyset$ .

The previous examples can familiarize you with the properties of polarizing lenses, in case you don't have actual lenses to play with. If you do have polarizing lenses, you can shine a laser pointer through them, observing when light passes through the filter and when light is filtered out. Use binder clips to position the lenses on a flat surface, and reproduce the setup in Figure 9.4. Don't worry about finding the exact orientation for "vertical." Any orientation of the lenses will do, as long as the first and the second polarizing lens have the same orientation. Next, you can rotate the second lens by  $90^\circ$  to obtain the setup shown in Figure 9.6, where the second lens has a perpendicular orientation and thus rejects all light.

**Example** Polarized sunglasses leverage the properties of light polarization to make outdoor activities more enjoyable. When a beam of light bounces off the surface of a lake, it becomes horizontally polarized. This polarization effect is due to the interaction of light's electric field at the surface of the water. A person wearing vertically polarizing lenses (polarized sunglasses) cannot see the sun's reflection off the water surface because the *V*-polarizing lenses filter out the horizontally polarized light reflected from the surface of the lake. This effect is useful for people who are often outdoors in bright sunlight, as it reduces the blinding effect of the sun's reflection.

## Classical physics paradigm

Before we describe the outcome of the polarizing lenses experiment, let's take a moment to describe the assumptions about the world that 19<sup>th</sup>-century physicists held. Understanding this classical world view will explain why the outcomes of the polarizing lenses experiment are so surprising.

The classical laws of physics are deterministic, meaning they do not allow randomness. According to the classical school, experimental outcomes depend on *definite* variables, like the properties of particles. Physicists assume they can predict the outcome of any experiment given they know the properties of the particles involved. If some outcome cannot be predicted, it must be because the value of some property of the particles is unknown. In other words, everything happens for a reason. Another key assumption classical physicists make is that a photon's properties are immutable, meaning we cannot change them. Classical physicists assume their **experiments correspond to passive observations** that cannot change the system's properties.

A 19<sup>th</sup>-century physicist would expect the outcomes of polarizing lenses experiments to be fully determined by the polarization property of photons. Each photon carries a tag *H* or *V* that describes its polarization. In the setup shown in Figure 9.3, each photon that passes through the lens must have `tag="V"`, because we know by definition that a *V*-polarizing lens only allows vertically polarized photons to pass through. Readers familiar with SQL syntax will recognize the action of the vertically polarizing lens as the following query:

```
SELECT photon FROM photons WHERE tag="V";
```

In other words, out of all the incoming photons, only the vertically polarized photons pass through the lens. Similarly, for the *H*-polarizing

lens shown in Figure 9.5, the filtering process can be understood as the query:

```
SELECT photon FROM photons WHERE tag="H";
```

In both cases, classical physicists would assume that whether or not a photon passes through a lens is predetermined, and is dependent only on the photon's tag.

For the purpose of our discussion, we'll restrict our attention to photons with either horizontal ( $\text{tag}=\text{"H"}$ ) or vertical ( $\text{tag}=\text{"V"}$ ) polarization. There are other possible polarization directions, but we'll focus on the tags H and V because they're mutually exclusive—if a photon is horizontally polarized, we know a vertically polarizing lens will reject it. We can assert that photons that pass through an *H*-polarizing lens are not vertically polarized; if they were, these photons would reflect off the lens instead of passing through.

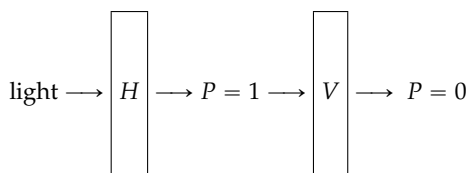
## Polarizing lenses experiment

The physics experiment we'll describe consists of sending photons through an optical circuit made of polarizing lenses and observing how many photons pass through the circuit. We describe the number of photons that reach any point in the circuit in qualitative terms, by referring to the *optical power* reaching that point, denoted  $P$ . We choose the light intensity of the beam after it passes through the first polarizing lens as our reference, and call it  $P = 1$  (full brightness). You can think of optical power as brightness. If you were to insert a piece of paper somewhere in the optical circuit, the optical power would measure the brightness of the spot of light on the paper at that particular location in the circuit. When power  $P = 1$ , the spot of light is fully bright. If  $P = 0.5$  the spot is half as bright as when  $P = 1$ . The case  $P = 0$  corresponds to zero brightness and occurs when no photons hit the piece of paper.

The initial setup for the experiment consists of an *H*-polarizing lens followed by a *V*-polarizing lens, as shown in Figure 9.7.

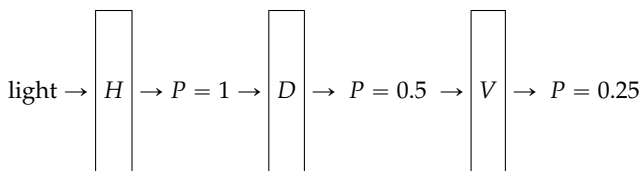
We know the photons that pass through the first lens are horizontally polarized. It's no surprise that when this light hits the second lens, none of the photons make it through, since a *V*-polarizing lens rejects *H*-polarized photons.

**Adding a third lens** We now introduce a third lens between the first two lenses, and we orient the middle lens differently from the other two—in the diagonal direction, for example. The result is shown in Figure 9.8. Suddenly, light appears at the end of the circuit! How



**Figure 9.7:** The initial setup for the polarizing lenses experiment consists of an  $H$ -polarizing lens followed by a  $V$ -polarizing lens. Only photons with tag="H" can pass through the first lens, so no photons with tag="V" pass through the first lens. No photons can pass through both lenses since the  $V$ -polarizing lens accepts only photons with tag="V".

does this make sense? You tell me if this is crazy or not. Intuitively, adding more filtering only reduces the amount of light passing through the circuit; yet the amount of light that passes through the circuit increases when we add the middle filter. How can adding more filtering increase light intensity? What is going on?



**Figure 9.8:** Adding an additional polarizing filter in the middle of the circuit causes light to appear at the end of the optical circuit.

We pick the middle lens to be a diagonally polarizing lens  $D$ . A diagonally polarizing lens is obtained by rotating any polarizing lens by  $45^\circ$ . The exact choice for the middle lens is not crucial for the experiment to work; so long as its polarization is different from the  $H$ - and  $V$ -polarizing lenses that surround it.

For a demonstration of the three-polarizing lenses experiment, see this Youtube video: <https://youtu.be/PJHCADY-Bio?t=6m14s>.

## Classical analysis

The experimental observations illustrated in Figure 9.8 are difficult to explain using the classical way of thinking, in which particle properties are immutable tags, and experiments are passive observations.

We have evidence that a particle's state can change during the process of measurement. Let's examine this more closely. We'll trace the path of the photons through the optical circuit in Figure 9.8, keeping track of what we know about the photons at each stage. First, all

photons that pass through the first  $H$ -polarizing lens are known to be horizontally polarized (tag="H"). We're sure no  $V$ -polarized photons pass through the first lens, because an  $H$ -polarizing lens is guaranteed to reject all  $V$ -polarized photons. Yet, after passing through the second lens (the  $D$  lens), these same photons appear to be vertically polarized since they pass through the third lens. Is something wrong with the tagging system? It seems the photons' tag states are affected by the measurements performed on them. This fact is difficult to explain for classical physicists since they assume measurements correspond to passive observations. In the classical paradigm, measuring a photon's  $D$ -polarization using the middle lens should not affect its H and V tags.

In Section 9.5 we'll revisit the polarizing lenses experiment after we have learned the postulates of quantum mechanics. We'll see we can explain the outcome of the experiment by describing the photon's polarization in terms of vector-like states; and that our understanding of vectors can even predict the final light intensity of  $P = 0.25$  we observed during the experiment. Before we discuss the postulates of quantum mechanics (Section 9.5), we'll need to introduce some new notation for describing quantum states.

## 9.3 Dirac notation for vectors

This section is a quick primer on Dirac notation, which is a precise and concise language for talking about vectors and matrices. The Dirac notation for vectors  $|v\rangle$  is an alternative to the usual notations for vectors like  $\vec{v}$  or  $\mathbf{v}$ .

This new notation will look weird initially, but once you get the hang of it, I guarantee you'll like it. Learning Dirac notation is an excellent way to review essential linear algebra concepts like bases, vectors, inner products, and matrices. Understanding Dirac notation is essential if you're interested in learning quantum physics, but it's also worth learning "just because"—it's *that* cool.

We'll now discuss several vector topics you're familiar with, and compare standard vector notation  $\vec{v}$  with the equivalent Dirac notation  $|v\rangle$ .

### The standard basis

Consider a  $d$ -dimensional complex vector space  $\mathbb{C}^d$ . We refer to complex vector spaces as Hilbert spaces, in honour of David Hilbert, who contributed prolific developments to math and physics.

To understand any vector space, it is essential to construct a basis for the space. A natural choice for a basis is the standard basis, which we'll denote  $\{|0\rangle, |1\rangle, |2\rangle, \dots, |d-1\rangle\}$ . The basis vectors are defined as:

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad |1\rangle \equiv \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad |d-1\rangle \equiv \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Note the indices are shifted by one so the first basis vector has index 0, not index 1. This zero-based indexing is chosen to make certain links between quantum theory and computer science more apparent.

One benefit of Dirac notation is that it doesn't require writing subscripts. To refer to a vector associated with properties  $a$ ,  $b$ , and  $c$ , we can write  $|a, b, c\rangle$ , instead of the more convoluted expression  $\vec{v}_{a,b,c}$ .

We'll now focus solely on the two-dimensional complex vector space  $\mathbb{C}^2$ ; however, the results and definitions presented below also apply to vectors of any dimension.

## Vectors

In Dirac notation, a vector in  $\mathbb{C}^2$  is denoted as a *ket*:

$$|v\rangle = \alpha|0\rangle + \beta|1\rangle \quad \Leftrightarrow \quad \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

where  $\alpha \in \mathbb{C}$  and  $\beta \in \mathbb{C}$  are the *components* of  $|v\rangle$  and  $\{|0\rangle, |1\rangle\}$  is the standard basis for  $\mathbb{C}^2$ :

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Why do we call the angle-bracket thing a “ket,” you ask? Let me tell you about the *bra* part, and then it will start to make sense.

The Hermitian transpose of the ket-vector  $|v\rangle = \alpha|0\rangle + \beta|1\rangle$  is the *bra*-vector  $\langle v|$ :

$$\langle v| = \bar{\alpha}\langle 0| + \bar{\beta}\langle 1| \quad \Leftrightarrow \quad [\bar{\alpha}, \bar{\beta}] = \bar{\alpha}[1, 0] + \bar{\beta}[0, 1].$$

Recall that the Hermitian transpose, also called the complex transpose, is the combination of the regular transpose ( $\vec{v} \rightarrow \vec{v}^T$ ) and the complex conjugation of each component of the vector ( $v_i \rightarrow \bar{v}_i$ ), and is denoted as the dagger operator “ $\dagger$ ” (see page 364 for a refresher on the dagger operator). Now observe how much simpler the *bra*



notation for the Hermitian transpose of a vector is compared to the other notations we've seen so far:

$$\langle v| \quad \Leftrightarrow \quad \vec{v}^\dagger \equiv \overline{(\vec{v}^\top)} \equiv (\vec{v})^\top.$$

When you put a bra  $\langle v|$  next to a ket  $|w\rangle$ , they become a bra-ket  $\langle v|w\rangle$ , which looks very similar to the brackets used to denote the inner product between two vectors. Observe the relative notational simplicity of calculating the inner product between the vectors  $|v\rangle = v_0|0\rangle + v_1|1\rangle$  and  $|w\rangle = w_0|0\rangle + w_1|1\rangle$  in Dirac notation:

$$\langle v|w\rangle = \overline{v_0}w_0 + \overline{v_1}w_1 \quad \Leftrightarrow \quad \vec{v}^\dagger \vec{w} \equiv \vec{v} \cdot \vec{w}.$$

Complex conjugation was already applied to the coordinates of  $|v\rangle$  when transforming it into a bra-vector  $\langle v|$ , thus we can simply “put together” the bra and the ket to compute the inner product. Because the bra notation  $\langle v|$  contains the Hermitian transpose, it removes the need for the dagger symbol. For instance, the inner product of  $|v\rangle = \alpha|0\rangle + \beta|1\rangle$  with itself is

$$\begin{aligned} \langle v|v\rangle &= (\overline{\alpha}\langle 0| + \overline{\beta}\langle 1|, \alpha|0\rangle + \beta|1\rangle) \\ &= \overline{\alpha}\alpha\langle 0|0\rangle + \overline{\alpha}\beta \underbrace{\langle 0|1\rangle}_0 + \overline{\beta}\alpha \underbrace{\langle 1|0\rangle}_0 + \overline{\beta}\beta\langle 1|1\rangle \\ &= |\alpha|^2 + |\beta|^2. \end{aligned}$$

This small simplification of the notation for inner products is quite useful, since inner products are the workhorse for calculating vector coordinates, finding projections, and performing change-of-basis transformations. The concise notation enables us to dig deeper into these aspects of linear algebra without getting overwhelmed by notational complexity. Moreover, Dirac's bra-ket notation is sufficiently simple to use in equations without needing to define new variables  $v_i$  for vector coordinates. We'll look at this more closely in the next section.

### Vector coordinates

The *coordinates*  $v_i$  of a vector  $\vec{v}$  with respect to an orthonormal basis  $\{\hat{e}_i\}$  are computed using the inner product  $v_i \equiv \hat{e}_i^\dagger \vec{v}$ . In Dirac notation, the coordinates of  $|v\rangle$  with respect to the standard basis  $\{|0\rangle, |1\rangle\}$  can be written as  $\langle i|v\rangle$ . We can write any vector  $|v\rangle$  as a linear combination of kets, with bra-kets as coefficients:

$$|v\rangle = \underbrace{\langle 0|v\rangle}_{v_0} |0\rangle + \underbrace{\langle 1|v\rangle}_{v_1} |1\rangle.$$

The expression  $\langle i|v\rangle$  explicitly defines the  $i^{\text{th}}$  coordinate of  $|v\rangle$ ; therefore, we don't need to define the variable  $v_i$ .

Another basis for the vector space  $\mathbb{C}^2$  is the *Hadamard basis*, which corresponds to the standard basis rotated by  $45^\circ$  in the counter-clockwise direction:

$$\begin{aligned}|+\rangle &\equiv \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle, \\ |-\rangle &\equiv \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle.\end{aligned}$$

The Hadamard basis, henceforth denoted  $B_h \equiv \{|+\rangle, |-\rangle\}$ , is an orthonormal basis:

$$\langle +|+\rangle = 1, \quad \langle +|-\rangle = 0, \quad \langle -|+\rangle = 0, \quad \langle -|-\rangle = 1.$$

Since the Hadamard basis  $B_h$  is an orthonormal basis, we can express any vector  $|v\rangle \in \mathbb{C}^2$  as a linear combination of kets:

$$|v\rangle = \underbrace{\langle +|v\rangle}_{v_+} |+\rangle + \underbrace{\langle -|v\rangle}_{v_-} |-\rangle.$$

Note that the coefficients of the linear combination are computed using the inner product with the corresponding basis vector  $\langle +|v\rangle$  and  $\langle -|v\rangle$ . The bra-ket notation allows us to refer to the coordinates of  $|v\rangle$  with respect to  $\{|+\rangle, |-\rangle\}$ , without the need to define variables  $v_+$  and  $v_-$ .

It's not uncommon to see vector coordinates with respect to different bases used in calculations. With the usual vector notation, we must specify which basis we're using as a subscript. For example, the same vector  $\vec{v}$  can be expressed as a coordinate vector  $\vec{v} = (v_0, v_1)_{B_s}^T$  with respect to the standard basis  $B_s$ , or as a coordinate vector  $\vec{v} = (v_+, v_-)_{B_h}^T$  with respect to the Hadamard basis. In the bra-ket notation, the coordinates with respect to  $B_s$  are  $\langle 0|v\rangle$  and  $\langle 1|v\rangle$ , and the coordinate with respect to  $B_h$  are  $\langle +|v\rangle$  and  $\langle -|v\rangle$ , making the choice of basis evident.

## Change of basis

Consider the task of finding the *change-of-basis* matrix  $_{B_h}[\mathbb{1}]_{B_s}$  that converts vectors from the standard basis  $B_s = \{(1,0), (0,1)\}$  to the Hadamard basis  $B_h = \{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})\}$ .

Using the standard approach for finding change-of-basis matrices discussed in Section 4.3 (page 240), we know the columns of  $_{B_h}[\mathbb{1}]_{B_s}$  contain the coordinates of  $(1,0)$  and  $(0,1)$  as expressed with respect

the basis  $B_h$ :

$${}_{B_h}[\mathbb{1}]_{B_s} = {}_{B_h} \begin{bmatrix} \langle +|0\rangle & \langle +|1\rangle \\ \langle -|0\rangle & \langle -|1\rangle \end{bmatrix} = {}_{B_h} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

We can now compute the coordinates of any vector  $\vec{v} \equiv (v_0, v_1)_{B_s}^T$  with respect to the Hadamard basis, by multiplying  $(v_0, v_1)_{B_s}^T$  by the change-of-basis matrix:

$$\begin{aligned} \begin{bmatrix} v_+ \\ v_- \end{bmatrix}_{B_h} &= {}_{B_h}[\mathbb{1}]_{B_s} \begin{bmatrix} v_0 \\ v_1 \end{bmatrix}_{B_s} \\ &= {}_{B_h} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \end{bmatrix}_{B_s} \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}}(v_0 + v_1) \\ \frac{1}{\sqrt{2}}(v_0 - v_1) \end{bmatrix}_{B_h}. \end{aligned}$$

This is the usual approach for computing the coordinates of a vector in one basis in terms of the coordinates of another basis using matrix notation.

Consider now the same change-of-basis operation, but with calculations carried out in Dirac notation. Given the vector  $\vec{v} = (v_0, v_1)_{B_s} = v_0|0\rangle + v_1|1\rangle$  expressed as coordinates with respect to the standard basis  $B_s$ , we want to find  $\vec{v} = (v_+, v_-)_{B_h} = v_+|+\rangle + v_-|-\rangle$ . Starting from the definition of  $v_+$  and  $v_-$ , we obtain

$$\begin{aligned} \vec{v} &\equiv (v_+, v_-)_{B_h}^T \\ &\equiv \langle +|v\rangle|+\rangle + \langle -|v\rangle|-\rangle \\ &= \langle +|(v_0|0\rangle + v_1|1\rangle)|+\rangle + \langle -|(v_0|0\rangle + v_1|1\rangle)|-\rangle \\ &= (v_0\langle +|0\rangle + v_1\langle +|1\rangle)|+\rangle + (v_0\langle -|0\rangle + v_1\langle -|1\rangle)|-\rangle \\ &= \underbrace{\frac{1}{\sqrt{2}}(v_0 + v_1)}_{v_+}|+\rangle + \underbrace{\frac{1}{\sqrt{2}}(v_0 - v_1)}_{v_-}|-\rangle. \end{aligned}$$

Working from the definitions of vectors and their coordinates, and using only basic algebra rules, we can perform the change-of-basis operation without explicitly constructing the change-of-basis matrix.

## Outer products

Recall the *outer product* operation for vectors that we introduced in Section 5.2 (see page 291). The expression  $|0\rangle\langle 0|$  is equivalent to the

projection onto the subspace spanned by the vector  $|0\rangle$ :

$$|0\rangle\langle 0| \quad \Leftrightarrow \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

We can verify this by considering the product of  $|0\rangle\langle 0|$  and an arbitrary vector  $|v\rangle = \alpha|0\rangle + \beta|1\rangle$ :

$$|0\rangle\langle 0| (\alpha|0\rangle + \beta|1\rangle) = \alpha|0\rangle \underbrace{\langle 0|0\rangle}_1 + \beta|0\rangle \underbrace{\langle 0|1\rangle}_0 = \alpha|0\rangle.$$

The ability to easily express outer products is another win for Dirac notation. For example, the projection onto the  $|+\rangle$ -subspace is  $|+\rangle\langle +|$ .

## Matrices

Now get ready for some crazy stuff. It turns out outer-product expressions are useful not only for projections, but can in fact represent any matrix. Consider the linear operator  $A : \mathbb{C}^2 \rightarrow \mathbb{C}^2$  and its matrix representation with respect to the standard basis:

$${}_{B_s}[A]_{B_s} \equiv {}_{B_s} \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}_{B_s}.$$

Instead of positioning the entries in an array, we can represent  $A$  as a linear combination of outer products:

$$A \equiv a_{00}|0\rangle\langle 0| + a_{01}|0\rangle\langle 1| + a_{10}|1\rangle\langle 0| + a_{11}|1\rangle\langle 1|.$$

Consider the matrix vector product  $A|x\rangle = |y\rangle$ . The matrix entry  $a_{10}$  describes the multiplication factor that connects the  $|0\rangle$ -component in the input  $|x\rangle$  to the  $|1\rangle$ -component of the output vector  $|y\rangle$ . The expression  $a_{10}|1\rangle\langle 0|$  is a concise description of the same story. The  $\langle 0|$  in this expression will “select” only the  $|0\rangle$  component of the input, and the  $|1\rangle$  indicates that this term contributes to the  $|1\rangle$  component of the output.

The entries of the matrix representation  ${}_{B_s}[A]_{B_s}$  depend on the choice of bases for the input and output spaces. The value of the entry  $a_{ij}$  in the matrix representation is computed by “probing” the matrix with the  $j^{\text{th}}$  basis vector of the input basis, and observing the value of the  $i^{\text{th}}$  component in the resulting output. We can express the entire “probing procedure” easily in Dirac notation:  $a_{ij} \equiv \langle i|A|j\rangle$ . Thus, we can write the matrix entries as follows:

$${}_{B_s} \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}_{B_s} = {}_{B_s} \begin{bmatrix} \langle 0|A|0\rangle & \langle 0|A|1\rangle \\ \langle 1|A|0\rangle & \langle 1|A|1\rangle \end{bmatrix}_{B_s}.$$

In fact, we don't need matrix notation or the entries  $a_{ij}$  at all. Instead, we can express  $A$  as a linear combination of outer products, with appropriately chosen coefficients:

$$A \equiv \langle 0|A|0\rangle|0\rangle\langle 0| + \langle 0|A|1\rangle|0\rangle\langle 1| + \langle 1|A|0\rangle|1\rangle\langle 0| + \langle 1|A|1\rangle|1\rangle\langle 1|.$$

Let's verify the formula for the  $a_{10} = \langle 1|A|0\rangle$ -entry of  $A$ , to see how this linear combination of outer products thing works. We start from the definition  $A \equiv a_{00}|0\rangle\langle 0| + a_{01}|0\rangle\langle 1| + a_{10}|1\rangle\langle 0| + a_{11}|1\rangle\langle 1|$ , and multiply  $A$  by  $|0\rangle$  from the right and  $\langle 1|$  from the left:

$$\begin{aligned} \langle 1|A|0\rangle &= \langle 1|\left(a_{00}|0\rangle\langle 0| + a_{01}|0\rangle\langle 1| + a_{10}|1\rangle\langle 0| + a_{11}|1\rangle\langle 1|\right)|0\rangle \\ &= \langle 1|\left(a_{00}|0\rangle \underbrace{\langle 0|0\rangle}_1 + a_{01}|0\rangle \underbrace{\langle 1|0\rangle}_0 + a_{10}|1\rangle \underbrace{\langle 0|0\rangle}_1 + a_{11}|1\rangle \underbrace{\langle 1|0\rangle}_0\right) \\ &= \langle 1|\left(a_{00}|0\rangle + a_{10}|1\rangle\right) \\ &= a_{00} \underbrace{\langle 1|0\rangle}_0 + a_{10} \underbrace{\langle 1|1\rangle}_1 = a_{10}. \end{aligned}$$

Indeed,  $\langle 1|A|0\rangle$  is the same as  $a_{10}$ . In fact, we'll rarely use the notation  $a_{10}$ , since  $\langle 1|A|0\rangle$  is just as easy to write, and much more intuitive: the  $a_{10}$ -entry of  $A$  is what you obtain when you "sandwich" the matrix  $A$  between the vectors  $\langle 1|$  on the left and  $|0\rangle$  on the right.

In Dirac notation, the basis appears explicitly in expressions for entries of a matrix. We can define the entries of  $A$  in any other basis easily and precisely. The representation of  $A$  with respect to the Hadamard basis  $B_h = \{|+\rangle, |-\rangle\}$  is

$${}_{B_h}[A]_{B_h} \equiv {}_{B_h} \begin{bmatrix} \langle +|A|+\rangle & \langle +|A|-\rangle \\ \langle -|A|+\rangle & \langle -|A|-\rangle \end{bmatrix}_{B_h},$$

or equivalently,

$$A = \langle +|A|+\rangle|+\rangle\langle +| + \langle +|A|-\rangle|+\rangle\langle -| + \langle -|A|+\rangle|-\rangle\langle +| + \langle -|A|-\rangle|-\rangle\langle -|.$$

The coefficient  $\langle +|A|+\rangle$  is the  $a_{++}$ -entry of the matrix representation of  $A$  with respect to the Hadamard basis  $B_h$ .

**Summary** Dirac notation is a convenient way to represent linear algebra concepts: vectors  $|v\rangle$  and their Hermitian conjugates  $\langle v|$ , vector coordinates  $\langle i|v\rangle$ , inner products  $\langle v|w\rangle$ , outer products  $|v\rangle\langle w|$ , and matrix entries  $\langle i|A|j\rangle$ . Because of this expressiveness, Dirac notation is widely used when discussing quantum mechanical topics in modern chemistry, physics, and computer science. In particular, Dirac notation is a core component in the study of quantum physics. In fact, we could say that if you understand Dirac notation, you already understand half of quantum mechanics.

## Exercises

**E9.1** Consider the vectors  $|u\rangle = \alpha|0\rangle + \beta|3\rangle$  and  $|v\rangle = a|1\rangle + b|2\rangle$ , which are elements of  $\mathbb{C}^4$ . Express the ket  $|u\rangle$  and the bra  $\langle v|$  as four-dimensional coordinate vectors. Indicate whether your answers correspond to row vectors or column vectors.

**9.1**  $\vec{u} = (\alpha, 0, 0, \beta)^T$  (a column vector);  $\vec{v}^T = (0, \bar{a}, \bar{b}, 0)$  (a row vector).

**E9.2** Express the vector  $\vec{w} = (1, i, -1)^T \in \mathbb{C}^3$  as a ket  $|w\rangle$  and as a bra  $\langle w|$ . Compute its length  $\|\vec{w}\| \equiv \sqrt{\langle w|w\rangle}$ .

**9.2**  $|w\rangle = 1|0\rangle + i|1\rangle - |2\rangle$ ;  $\langle w| = 1\langle 0| - i\langle 1| - \langle 2|$ ;  $\|\vec{w}\| = \sqrt{3}$ .

**E9.3** Find the determinant of the following matrix:

$$A = 1|0\rangle\langle 0| + 2|0\rangle\langle 1| + 3|1\rangle\langle 0| + 4|1\rangle\langle 1|.$$

**9.3**  $\det(A) = -2$ .

**E9.4** Consider the vectors  $|u\rangle = |0\rangle$ ,  $|v\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{i}{\sqrt{2}}|1\rangle$ , and  $|w\rangle = \frac{1}{\sqrt{5}}|0\rangle - \frac{2}{\sqrt{5}}|1\rangle$ . Compute the following expressions:

$$\mathbf{a}) |u\rangle + |v\rangle \quad \mathbf{b}) \| |v\rangle \| + \| |w\rangle \| \quad \mathbf{c}) \langle u|v\rangle \quad \mathbf{d}) \langle v|w\rangle \quad \mathbf{e}) \langle w|v\rangle \quad \mathbf{f}) |u\rangle\langle u| \\ \mathbf{g}) \Pi_{|u\rangle}(|v\rangle) \quad \mathbf{h}) |v\rangle\langle v| \quad \mathbf{i}) \Pi_{|v\rangle}(|u\rangle) \quad \mathbf{j}) \Pi_{|v\rangle}(|w\rangle)$$

Hint: The projection matrix onto a unit vector  $\hat{a}$  is equal to  $\hat{a}\hat{a}^\dagger$ .

**9.4** **a)**  $(1 + \frac{1}{\sqrt{2}})|0\rangle + \frac{i}{\sqrt{2}}|1\rangle$ ; **b)** 2; **c)**  $\frac{\sqrt{2}}{2}$ ; **d)**  $\frac{\sqrt{10}}{10} + \frac{\sqrt{10}i}{5}$ ; **e)**  $\frac{\sqrt{10}}{10} - \frac{\sqrt{10}i}{5}$ ; **f)**  $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ ; **g)**  $\frac{\sqrt{2}}{2}|0\rangle$ ; **h)**  $\begin{bmatrix} \frac{1}{2} & -\frac{i}{2} \\ \frac{i}{2} & \frac{1}{2} \end{bmatrix}$ ; **i)**  $\frac{1}{2}|0\rangle + \frac{i}{2}|1\rangle$ ; **j)**  $(\frac{\sqrt{5}}{10} + \frac{\sqrt{5}i}{5})|0\rangle - (\frac{\sqrt{5}}{5} - \frac{\sqrt{5}i}{10})|1\rangle$ .

**E9.5** Given the matrix  $A = 1|0\rangle\langle 0| + 3|0\rangle\langle 1| + 4|1\rangle\langle 0| + 5|1\rangle\langle 1|$  and the vectors  $|v\rangle = 1|0\rangle + 3|1\rangle$  and  $|w\rangle = -3|0\rangle - 2i|1\rangle$ , compute:

$$\mathbf{a}) A|v\rangle \quad \mathbf{b}) \langle v|A \quad \mathbf{c}) A|w\rangle \quad \mathbf{d}) \langle v|A|v\rangle \quad \mathbf{e}) \langle w|A \quad \mathbf{f}) \langle w|A|w\rangle$$

**9.5** **a)**  $A|v\rangle = 10|0\rangle + 19|1\rangle$ ; **b)**  $\langle v|A = 13\langle 0| + 18\langle 1|$ ; **c)**  $A|w\rangle = (-3 - 6i)|0\rangle - (12 + 10i)|1\rangle$ ; **d)**  $\langle v|A|v\rangle = 67$ ; **e)**  $\langle w|A = (-3 + 8i)\langle 0| + (-9 + 10i)\langle 1|$ ; **f)**  $\langle w|A|w\rangle = 29 - 6i$ .

**E9.6** Express the linear transformation  $T(x, y) = (2x + y, -3y)$  as a matrix that acts on an input vector of the form  $x|0\rangle + y|1\rangle$ .

**9.6**  $M_T = 2|0\rangle\langle 0| + 1|0\rangle\langle 1| - 3|1\rangle\langle 1|$ .

## 9.4 Quantum information processing

Digital technology is sought after because of the computational, storage, and communication advantages of manipulating digital information instead of continuous-time signals. Similarly, quantum technology enables certain new advances for computational and communication tasks. This section will equip you with a mental model for thinking about quantum information processing tasks in analogy to a digital information processing pipeline you're already familiar with: the digitization, compression, and playback of sound recording that we discussed in Section 7.11 (page 465).

The use of quantum technology for information processing tasks is no more mysterious than the use of digital technology for information processing tasks. Playing a digital recording on your mp3 player involves a number of processing, conversion, and signal amplification steps. Similarly, using a quantum computer involves several conversion, processing, and measurement steps. In both cases you input some data into a machine, and wait for the machine to process the data and output the answer.

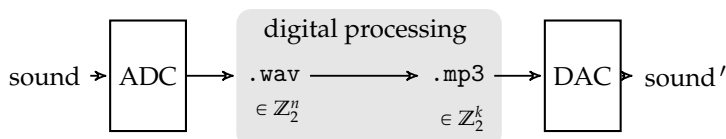
We can think of both digital and quantum technology as black box processes, with internal workings that we can't access directly. In both cases, the intermediate representation of data is in a format that is unintelligible: we can't understand what is encoded in quantum states any more than we can understand what is encoded in digital data. For instance, an mp3 file contains ones and zeros; but, unless we're unusually gifted, it's impossible to tell which artist plays the song 010100101010101000111... just by looking at the raw, digital data. To understand information processing in the digital and quantum worlds, we must study the “adaptors”—the processes used to convert the internal data representation into signals we can intelligibly perceive.

To further highlight the parallel structure between digital information processing and quantum information processing, we'll now review the mp3 compression task as an example of a digital information processing pipeline.

### Digital signal processing

A *sound card* is a computer component that converts between analog signals that we can hear and digital signals that computers understand. The sound card digitizes sound using an analog-to-digital converter (ADC). For music playback, the sound card uses a digital-to-analog converter (DAC), which transforms digital sounds into analog sound vibrations to be played through speakers. The ADC re-

ceives signal via the sound card's line-in and microphone jacks; the DAC outputs sound via the sound card's line-out and headphone jacks.



**Figure 9.9:** A digital information processing pipeline for sound recording and playback. Sound vibrations are captured by a microphone and converted to digital form using an analog-to-digital converter (ADC). Next the digital *wav* file is converted to the more compact *mp3* format using digital processing. In the last step, sound is converted back into analog sound vibrations by a digital-to-analog converter (DAC).

Figure 9.9 illustrates a full digital information processing pipeline for sound. We use an *analog-to-digital converter* (ADC) to transform the analog sound into digital form; the sound's digital representation is then processed, and finally a *digital-to-analog converter* (DAC) transforms the digital signal into analog form. If the music encoding, processing, and playback steps are successful, the final output will sound like the original sound recorded.

The grey-shaded region in Figure 9.9 corresponds to digital data. The example of *mp3* compression is just one of many uses of digital processing that become possible once we convert analog information into digital form. Information stored in digital form allows for countless other affordances: the long-term storage of the *mp3* file, the file's transmission over the internet, or the computation of a digital fingerprint of the song. Mathematically speaking, we can describe any digital information processing task as a function,  $f : \mathbb{Z}_2^n \rightarrow \mathbb{Z}_2^k$ , that takes  $n$ -dimensional binary vectors as inputs and produces  $k$ -dimensional binary vectors as outputs.

## Quantum information processing

In order to understand the context where quantum information processing has applications, we'll describe a hypothetical scenario. Suppose you're trying to calculate the output of some mathematical function  $f(x)$  for the input  $x$ . Your first approach is to express the math function  $f$  as a function in code  $\mathfrak{f}$ , but you quickly notice that running the code of the function  $\mathfrak{f}$  takes too long for the types of inputs you're interested in. Computing  $\mathfrak{f}(x)$  would take a century on your laptop, and even if you ran your code on a powerful cluster of computers, it might still take years to compute the output. Normally

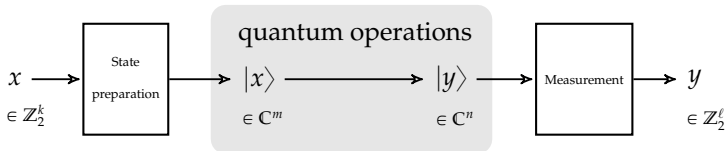


people quit at this point.

You're not a quitter though, so instead of waiting for the classical computer to finish computing  $f(x)$ , you decide to use a quantum computer. Figure 9.10 shows the sequence of steps needed to operate a quantum computer. Starting from the input  $x$  expressed as a classical bitstring  $x \in \mathbb{Z}_2^k$ , you use state preparation to encode the input into a quantum vector  $|x\rangle \in \mathbb{C}^m$  inside the quantum computer. You can then perform quantum operations on the input state to transform it to an output state  $|y\rangle \in \mathbb{C}^n$ . Finally, you perform a measurement on the output state  $|y\rangle$  to obtain a classical bitstring  $y \in \mathbb{Z}_2^\ell$ , which is hopefully the answer you're looking for.

The main difference between classical and quantum computation is the underlying data types used in the computation. Instead of working with *bits* and using digital functions  $f: \mathbb{Z}_2^k \rightarrow \mathbb{Z}_2^\ell$ , quantum computers work with *qubits* and use *quantum operations*  $T: \mathbb{C}^m \rightarrow \mathbb{C}^n$  to perform computations. Quantum computation is more general than classical computation since quantum computers can make use of quantum effects like constructive interference.

The other big difference between classical and quantum computation is that the quantum computer outputs the state  $|y\rangle$  only once. You can think of **quantum measurement as asking a question** about the state  $|y\rangle$ . You're free to perform any measurement, but **you can ask only one question**, since quantum measurements disrupt the state of a system.



**Figure 9.10:** A quantum information processing pipeline. A classical bitstring  $x$  of length  $k$  is used as instructions for preparing an  $m$ -dimensional quantum state  $|x\rangle$ . Next, quantum operations are performed on the state  $|x\rangle$  to convert it to the output state  $|y\rangle$ . Finally, the state  $|y\rangle$  is *measured* to obtain the classical bitstring  $y$  as output.

Quantum processing pipelines are analogous to digital information processing pipelines. The process of *state preparation* in quantum processing is analogous to the analog-to-digital conversion step in digital processing. In both cases we convert the input to the format required for the processing step. Similarly, quantum *measurements* correspond to the digital-to-analog conversion step. In both cases we must convert the output to a format we can understand.

In the next section, we'll discuss the components of the quantum information processing pipeline in more detail. We'll introduce the

four *postulates* of quantum mechanics, which specify how quantum systems are represented and what we can do with them. The postulates of quantum mechanics roughly correspond to the conversion steps illustrated in Figure 9.10. One postulate defines how quantum states are prepared, another postulate describe the types of operations we can perform on quantum states, and a third postulate formally defines the process of quantum measurement. The next section is the “quantum mechanics explained in the space on the back of an envelope” part alluded to in the introduction of this chapter. We’ve set the scene, introduced Dirac notation, and now we can finally discuss the details of the quantum formalism.

## 9.5 Postulates of quantum mechanics

The *postulates* of quantum mechanics dictate the rules for working within the “quantum world.” The four postulates define:

- What quantum states are
- Which quantum operations can be performed on quantum states
- How to extract information from quantum systems by measuring them
- How to represent composite quantum systems

These postulates specify the structure that all quantum theories must have. Together, the four postulates are known as the *quantum formalism*, and describe the math structure common to all fields that use quantum mechanics: physics, chemistry, engineering, and quantum information. Note the postulates are not provable or derivable from a more basic theory: scientists simply take the postulates as facts and make sure their theories embody these principles.

### Quantum states

Quantum states are modelled as special types of vectors. The *state* of a  $d$ -dimensional quantum system is a unit vector  $|\psi\rangle$ , in a  $d$ -dimensional complex inner product vector space  $\mathbb{C}^d$ , which we call the Hilbert space. This is the first postulate that belongs on the back of the envelope.

**Postulate 1.** To every isolated quantum system is associated a complex inner product space (Hilbert space) called the *state space*. A state is described by a unit vector in state space.

The following comments apply to the description of quantum systems:

- The requirement that state vectors must have length one will become important when we discuss the probabilistic nature of quantum measurements.
- The *global phase* of a quantum state vector doesn't matter, which means  $|\psi\rangle \equiv e^{i\theta}|\psi\rangle, \forall \theta \in \mathbb{R}$ . The vectors  $|\psi\rangle$ ,  $-|\psi\rangle$ , and  $i|\psi\rangle$  all represent the same quantum state.
- Each physical system corresponds to its own Hilbert space, usually denoted with the same label as the system,  $|\psi\rangle_1 \in V_1$  and  $|\psi\rangle_2 \in V_2$ .

In general, the quantum states of physical systems are represented as vectors in  $d$ -dimensional or sometimes even infinite-dimensional Hilbert spaces. To keep things simple, we'll focus on two-dimensional quantum systems.

**The qubit** In analogy with two-state classical bits  $\in \{0, 1\}$  we call two-dimensional quantum systems *qubits*  $\in \mathbb{C}^2$ , which is short for *quantum bit*. Many physical systems, like the polarization of a photon or the spin of an electron, can be represented as qubits. A qubit is a unit vector in a two-dimensional Hilbert space  $\mathbb{C}^2$ :

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

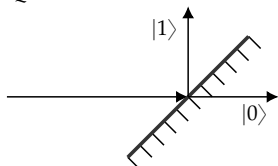
where

$$\begin{aligned} \alpha &\in \mathbb{R}, & (\text{global phase doesn't matter}) \\ \beta &\in \mathbb{C}, \\ |\alpha|^2 + |\beta|^2 &= 1. \end{aligned}$$

Recall that  $|0\rangle$  and  $|1\rangle$  are the vectors of the standard basis for  $\mathbb{C}^2$ :

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

The restriction that  $\alpha$  must be a real number follows from the fact that the global phase of a quantum state can be ignored. The condition that a quantum state must have length one is equivalent to the constraint  $|\alpha|^2 + |\beta|^2 = 1$ .



**Figure 9.11:** A photon encounters a half-silvered mirror, which is also known as a beam splitter. The photon can take one of the two possible paths, so we describe it as the superposition  $|\gamma\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ , where  $|0\rangle$  describes the photon passing through the mirror, and  $|1\rangle$  describes the photon being reflected.

Though the notion of a qubit is an abstract concept, many physical systems can embody it. In Figure 9.11, the information of a qubit is encoded in the path a photon takes after it encounters a half-silvered mirror. Qubits are a device-independent representation of quantum information, similar to how classical bits are device-independent representations for classical information. A bit is a bit, regardless of whether it is transmitted over the network, stored in RAM, or stored on a hard drive. Similarly, a qubit is a qubit, regardless of whether it's encoded in the polarization of a photon, an electron's spin, or in the direction of magnetic flux of a superconducting loop.

**Quantum state preparation** The operation of encoding some classical information into a quantum system is called *state preparation*. Imagine an apparatus that prepares quantum systems in one of several possible quantum states, depending on the position of the “control switch”  $x$  of the machine.

$$x \longrightarrow \boxed{\text{state preparation}} \longrightarrow |x\rangle$$

**Figure 9.12:** The classical input  $x$  is used to prepare a quantum system. The quantum state  $|x\rangle$  is produced when the classical input is  $x$ .

An example of quantum state preparation is a machine that can produce photons in two different polarizations  $|H\rangle$  and  $|V\rangle$ . If the input  $x = 0$  is specified, the machine will produce the state  $|H\rangle$ , and if  $x = 1$  is specified as the input, the machine will produce  $|V\rangle$ .

The quantum state preparation step is analogous to the analog-to-digital conversion step; it's how we get classical information into a quantum computer. Once we have converted the input  $x$  into a quantum state  $|x\rangle$ , the natural next question to ask is what we can *do* with  $|x\rangle$ . What quantum operations are we allowed to perform on quantum states?

## Quantum operations

The second definition we can write on the back of the envelope is about identifying *quantum operations* with unitary transformations acting on quantum states. Every quantum operation can be represented as a unitary operation  $U$  applied to the input state  $|\psi\rangle$  to produce the output state  $|\psi'\rangle$ .

$$|\psi\rangle \longrightarrow \boxed{U} \longrightarrow |\psi'\rangle$$

The requirement that all quantum operations be unitary is codified in the second postulate of quantum mechanics.

**Postulate 2.** Time evolution of an isolated quantum system is unitary. If the state at time  $t$  is  $|\psi\rangle$  and at time  $t'$  is  $|\psi'\rangle$ , then there exists a unitary operator  $U$  such that  $|\psi'\rangle = U|\psi\rangle$ .

This is a major piece of the quantum puzzle, so let's analyze what it means: mathematically, physically, and computationally.

First let's get the math out of the way. Recall that a unitary matrix  $U$  obeys  $U^\dagger U = U U^\dagger = \mathbb{1}$ . Postulate 2 ensures that quantum states will maintain their unit-length property after quantum operations are performed on them. Assume the quantum system starts from a state  $|\psi\rangle$  that has length one,  $\|\psi\|^2 = \langle\psi|\psi\rangle = 1$ . After the unitary  $U$  is applied, the state after evolution will be  $|\psi'\rangle = U|\psi\rangle$  and its squared length will be  $\|\psi'\|^2 = \|U|\psi\rangle\|^2 = \langle\psi|U^\dagger U|\psi\rangle = \langle\psi|\mathbb{1}|\psi\rangle = \langle\psi|\psi\rangle = 1$ . In other words, **quantum operations are length-preserving**.

The second postulate refers explicitly to a time variable and the state of the system at different times, which is the physics way of thinking about quantum operations. For example, applying a quantum operation to a real physical system could correspond to applying a certain magnetic field to that system. If applying the magnetic field for one second applies the rotation operation  $R_\theta$ , then applying the magnetic field for two seconds performs the rotation  $R_{2\theta}$ ; therefore the time variable is of central importance to the experiment.

The computer science perspective on the second postulate is more abstract. Instead of counting seconds, computer scientists model quantum operations as discrete “gates” that can be applied to quantum states. Computer scientists describe their quantum algorithms in terms of unitary operators  $U$ , and they assume there exists a specific physical operation that performs the quantum operation  $U$  on a real physical qubit system.

In this chapter, we'll take the computer science approach and think of quantum information processing tasks in terms of applying

different gates to quantum states. We'll now define several quantum gates that perform useful unitary operations on qubits.

**Example 1: phase gate** The  $Z$  operator is defined by its action on the vectors of the standard basis.

$$Z|0\rangle = |0\rangle, \quad Z|1\rangle = -|1\rangle.$$

The  $Z$  operator leaves the  $|0\rangle$  unchanged but flips the phase of  $|1\rangle$ .

Given knowledge of the actions of the  $Z$  operator on the vectors of the standard basis, we can construct its matrix representations:

$$Z = {}_{B_s} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}_{B_s} \equiv |0\rangle\langle 0| - |1\rangle\langle 1|.$$

**Example 2: NOT gate** The  $X$  operator is defined by the following actions on the vectors of the standard basis:

$$X|0\rangle = |1\rangle, \quad X|1\rangle = |0\rangle.$$

The  $X$  operator acts as a “quantum NOT gate,” changing  $|0\rangle$ s into  $|1\rangle$ s, and  $|1\rangle$ s into  $|0\rangle$ s. The matrix representation of the  $X$  operator is:

$$X = {}_{B_s} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}_{B_s} \equiv |0\rangle\langle 1| + |1\rangle\langle 0|.$$

**Example 3: Hadamard gate** The Hadamard operator takes the vectors of the standard basis to the vectors of the Hadamard basis  $|+\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$  and  $|-\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ :

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \equiv |+\rangle, \quad H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \equiv |-\rangle.$$

You can also think of the  $H$  operator as a  $45^\circ$  counter-clockwise rotation. The matrix representation of the Hadamard gate is

$$H = {}_{B_s} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}_{B_s}.$$

By linearity, we can deduce the effects of the operators  $Z$ ,  $X$ , and  $H$  on an arbitrary qubit  $\alpha|0\rangle + \beta|1\rangle$ :

$$\begin{aligned} Z(\alpha|0\rangle + \beta|1\rangle) &= \alpha|0\rangle - \beta|1\rangle, \\ X(\alpha|0\rangle + \beta|1\rangle) &= \beta|0\rangle + \alpha|1\rangle, \\ H(\alpha|0\rangle + \beta|1\rangle) &= \frac{\alpha+\beta}{\sqrt{2}}|0\rangle + \frac{\alpha-\beta}{\sqrt{2}}|1\rangle. \end{aligned}$$

**Example 4** The effect of the operator  $XZ$  corresponds to the combination of the effects of the  $Z$  and  $X$  operators. We can understand the action of  $XZ$  either by applying it to an arbitrary qubit  $\alpha|0\rangle + \beta|1\rangle$ :

$$XZ(\alpha|0\rangle + \beta|1\rangle) = -\beta|0\rangle + \alpha|1\rangle,$$

or by multiplying together the operator's matrix representations:

$$XZ = {}_{B_s} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} {}_{B_s} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = {}_{B_s} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} {}_{B_s}.$$

In general, it is possible to perform many other quantum operations and combinations of operations when working with quantum states. The examples presented above represent the most commonly used operations. Note that unitary time evolution is invertible: for every quantum gate  $G$  there exists an inverse gate  $G^{-1}$  such that  $G^{-1}G = \mathbb{1}$ .

## Links

[ Wikipedia article on quantum gates ]

[https://en.wikipedia.org/wiki/Quantum\\_gate](https://en.wikipedia.org/wiki/Quantum_gate)

## Exercises

**E9.7** The Hadamard gate is defined as  $H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ . Compute the effect of the operator  $HH$  on the vectors of the standard basis  $\{|0\rangle, |1\rangle\}$ .

**9.7**  $HH|0\rangle = |0\rangle$  and  $HH|1\rangle = |1\rangle$ .

**9.7** Let's first see what happens to  $|0\rangle$  when we apply the operator  $HH$ . The result of the first  $H$  applied to  $|0\rangle$  is  $H|0\rangle = |+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ . Applying the second  $H$  operator, we get  $HH|0\rangle = H|+\rangle = \frac{1}{\sqrt{2}}(H|0\rangle + H|1\rangle)$ . Applying the  $H$  operation gives  $\frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) + \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle))$ , which simplifies to  $\frac{1}{\sqrt{2}}(\frac{2}{\sqrt{2}}|0\rangle) = |0\rangle$ . So  $HH|0\rangle = |0\rangle$ . A similar calculation shows that  $HH|1\rangle = |1\rangle$ .

**E9.8** Compute  $XX$ ,  $XZ$ , and  $ZX$ .

**9.8**  $XX = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ,  $XZ = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ ,  $ZX = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ .

**9.8** Use matrix multiplication. Note  $XZ$  is not the same as  $ZX$ .

\*\*\*

So far, we've filled half the space on the back of the envelope. It's time to talk about the third fundamental idea in quantum mechanics: quantum measurements.

## Quantum measurements

A quantum measurement performed on a quantum system corresponds to a collection of projection operators  $\{\Pi_i\}$  that act on the Hilbert space. A measurement with  $n$  possible outcomes is represented by  $n$  projection operators  $\{\Pi_1, \Pi_2, \dots, \Pi_n\}$ . The projection operators form a *decomposition of the identity*, meaning their sum is equal to the identity matrix:

$$\sum_{i=1}^n \Pi_i = \mathbb{1}.$$

Intuitively, the  $n$  different projection operators correspond to  $n$  different alternatives for the evolution of the quantum system. Performing the measurement is like asking “which way is it going to be?” and letting the system decide which path it wants to take. *Born’s rule* is used to assign probabilities to different measurement outcomes.

**Postulate 3.** A quantum measurement is modelled by a collection of projection operators  $\{\Pi_i\}$  that act on the state space of the system being measured and satisfy  $\sum_i \Pi_i = \mathbb{1}$ . The index  $i$  labels the different measurement outcomes.

The probability of outcome  $i$  when performing measurement  $\{\Pi_i\}$  on a quantum system in the state  $|\psi\rangle$  is given by the squared length of the state after applying the  $i^{\text{th}}$  projection operator:

$$\Pr(\{\text{outcome } i \text{ given state } |\psi\rangle\}) \equiv \|\Pi_i|\psi\rangle\|^2 \quad (\text{Born’s rule}).$$

When outcome  $i$  occurs, the post-measurement state of the system is

$$|\psi'_i\rangle \equiv \frac{\Pi_i|\psi\rangle}{\|\Pi_i|\psi\rangle\|}.$$

Let’s unpack this definition to see what is going on.

**Born’s rule** For the measurement defined by the projection operators  $\{\Pi_1, \Pi_2, \dots, \Pi_n\}$ , Born’s rule states that the probability of outcome  $i$  is  $\langle\psi|\Pi_i|\psi\rangle$ .

This expression for the squared norm of the overlap between  $|\psi\rangle$  and  $\Pi_i$  can be written in several equivalent ways:

$$\|\Pi_i|\psi\rangle\|^2 = (\Pi_i|\psi\rangle, \Pi_i|\psi\rangle) = \langle\psi|\Pi_i\Pi_i|\psi\rangle = \langle\psi|\Pi_i|\psi\rangle,$$

where the last equality follows from the idempotence property of projectors  $\Pi_i = \Pi_i^2$ . The last expression, where the projection opera-



tor is “sandwiched” by two copies of the quantum state, is the physicist’s usual way of expressing Born’s rule by defining  $\Pr(\{\text{outcome } i\}) \equiv \langle \psi | \Pi_i | \psi \rangle$ . For the class of projective measurements we’re discussing here, the two definitions are equivalent.

The set of projection operators  $\{\Pi_1, \Pi_2, \dots, \Pi_n\}$  forms a decomposition of the identity  $\mathbb{1} = \sum_i \Pi_i$ . This guarantees that the probabilistic distribution of the different outcomes is normalized:

$$\begin{aligned} 1 &= \|\mathbb{1}|\psi\rangle\|^2 = \|(\Pi_1 + \dots + \Pi_n)|\psi\rangle\|^2 \\ &\stackrel{\text{py}}{=} \|\Pi_1|\psi\rangle\|^2 + \dots + \|\Pi_n|\psi\rangle\|^2 \\ &= \Pr(\{\text{outcome } 1\}) + \dots + \Pr(\{\text{outcome } n\}). \end{aligned}$$

That’s good to check; otherwise Kolmogorov would be angry with us. Note the equality labelled  $\stackrel{\text{py}}{=}$  follows from Pythagoras’ theorem; we’re using the fact that the operators  $\{\Pi_i\}$  are mutually orthogonal.

**Post-measurement state** When outcome  $i$  occurs, Postulate 3 tells us that the state of the quantum system becomes  $|\psi'_i\rangle \equiv \frac{\Pi_i|\psi\rangle}{\|\Pi_i|\psi\rangle\|}$ , which is the result of applying the projection  $\Pi_i$  to obtain  $\Pi_i|\psi\rangle$ , and then normalizing the state so that  $\| |\psi'_i\rangle \| = 1$ .

Measurements are not passive observations! Quantum measurement is an invasive procedure that typically changes the state of the system being measured. In general, the quantum state after the measurement will not be the same as the state before the measurement:

$$|\psi'_i\rangle \neq |\psi\rangle,$$

and we say the state is *disturbed* by the measurement, (though it’s possible that  $|\psi'_i\rangle = |\psi\rangle$  when the input state lives entirely within the image space of one of the projection operators  $|\psi'_i\rangle = \Pi_i|\psi\rangle = |\psi\rangle$ ).

A quantum measurement is an interaction that creates classical information and destroys quantum information. By measuring, we obtain the classical information  $i$  that tells us which outcome occurred, but we disturb the initial state  $|\psi\rangle$ , forcing the quantum state into the “aligned with  $\Pi_i$ ”-state  $|\psi'_i\rangle$ . We can still carry out further experiments with the post-measurement state  $|\psi'_i\rangle$ , but it’s not the same as the initial state  $|\psi\rangle$ . Specifically, we’ve lost all the information about  $|\psi\rangle$  that used to exist in the subspace  $(\mathbb{1} - \Pi_i)$ .

Another way to describe what happens during a quantum measurement is to say the state  $|\psi\rangle$  *collapses* into the state  $|\psi'_i\rangle$ . This is the terminology used by physicists to describe the effects of the projection that occurs during a measurement. Before the measurement, the quantum state  $|\psi\rangle$  could be any vector, but after we observe the

outcome  $i$ , we know the state has “collapsed” and is now confined to the image subspace of the projection operator  $\Pi_i$ . We won’t use the terminology of state collapse in this chapter, because the term “projection” more accurately describes what’s happening.

Historically, the founders of quantum mechanics chose some rather poor terminology to describe quantum measurements. In addition to “collapse,” they also used the term “observer” to describe whomever observes the outcome  $i$  of a quantum measurement experiment. Trying to understand quantum mechanics using this terminology often leads to magical thinking and anthropomorphizing of the underlying physical events. Proponents of pop-psychology ideas have latched onto the term “observer” to suggest that the collapse of quantum states that occurs during quantum measurements is caused by the presence of human consciousness. Watch out for that fluff.

A more grounded way to think about quantum measurement is in terms of *interaction* between the quantum state of a particle and a classical measurement apparatus that can be in one of  $n$  possible states. Due to the relative size of the two interacting systems (tiny quantum system and large measurement apparatus), the state  $|\psi\rangle$  is forced to “align” with one of the  $n$  possible states of the measurement apparatus.

**Example 4** In Figure 9.13 a state vector  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  is measured with photo detectors modelled as projectors given by

$$\begin{aligned}\Pi_0 &= |0\rangle\langle 0| & \sum_j \Pi_j &= \mathbb{1} \\ \Pi_1 &= |1\rangle\langle 1|\end{aligned}$$

$$\begin{aligned}\Pr(\{0\}|\psi) &\equiv \Pr(\{\text{outcome } 0 \text{ given state } |\psi\rangle\}) \\ &\equiv \langle \psi | \Pi_0 | \psi \rangle \\ &= \langle \psi | |0\rangle\langle 0| | \psi \rangle \\ &= (\bar{\alpha}\langle 0| + \bar{\beta}\langle 1|) |0\rangle\langle 0| (\alpha|0\rangle + \beta|1\rangle) \\ &= \bar{\alpha}\alpha \\ &= |\alpha|^2.\end{aligned}$$

The probability of outcome 1 is  $\Pr(\{1\}|\psi) \equiv \langle \psi | \Pi_1 | \psi \rangle = |\beta|^2$ . After the measurement, the quantum system exists in one of two possible states:  $|\psi'_0\rangle = |0\rangle$  or  $|\psi'_1\rangle = |1\rangle$ .

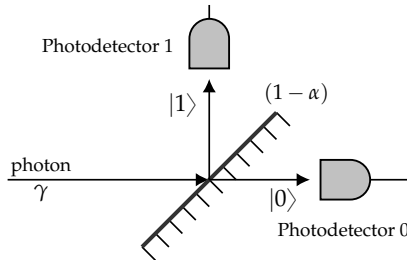
**Example 5** Consider the measurement  $\{\Pi_+, \Pi_-\}$  that consists of the projectors onto the Hadamard basis:

$$\begin{aligned}\Pi_+ &= |+\rangle\langle +| & \Pi_+ + \Pi_- &= \mathbb{1} \\ \Pi_- &= |-\rangle\langle -|\end{aligned}$$

Given the quantum state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , the probability of outcome “+” is given by

$$\begin{aligned}
 \Pr(\{+\}|\psi) &\equiv \|\Pi_+|\psi\rangle\|^2 \\
 &= \||+\rangle\langle+|\psi\rangle\|^2 \\
 &= \||+\rangle\langle+|(\alpha|0\rangle + \beta|1\rangle)\|^2 \\
 &= \||+\rangle(\alpha\langle+|0\rangle + \beta\langle+|1\rangle)\|^2 \\
 &= \||+\rangle\left(\alpha\frac{1}{\sqrt{2}} + \beta\frac{1}{\sqrt{2}}\right)\|^2 \\
 &= \frac{|\alpha + \beta|^2}{2} \||+\rangle\|^2 \\
 &= \frac{|\alpha + \beta|^2}{2}.
 \end{aligned}$$

The probability of outcome “−” is  $\Pr(\{-\}|\psi) = \frac{|\alpha - \beta|^2}{2}$ . After the measurement, the quantum system is in one of two possible states:  $|\psi'_+\rangle = |+\rangle$  or  $|\psi'_-\rangle = |-\rangle$ .



**Figure 9.13:** The state of a photon after encountering a  $(1 - \alpha)$ -silvered mirror is  $|\gamma\rangle = \alpha|0\rangle + \beta|1\rangle$ . The probability that the horizontal photodetector “clicks” is  $|\alpha|^2$ , and is obtained by projecting  $|\gamma\rangle$  on the subspace  $|0\rangle\langle 0|$ . The probability that the top photodetector clicks is equal to  $|\beta|^2$ , and is obtained by projecting  $|\gamma\rangle$  on the subspace  $|1\rangle\langle 1|$ .

The measurement process is a fundamental aspect of quantum models. You’ll need to acclimate to the idea that measurements change systems’ states. It’s not magic; it’s a phenomenon that occurs due to the relative size of the systems (tiny quantum particles and huge measurement apparatus), and the fact that measurements force quantum and classical systems to interact.

## Composite quantum systems

So far we discussed state preparation, quantum operations, and quantum measurements of individual qubits. There’s just enough room

on the back of our envelope to discuss quantum models for systems made of multiple qubits.

Classically, if we have two bits  $b_1 \in \{0, 1\}$  and  $b_2 \in \{0, 1\}$ , we can concatenate them to obtain a bit string  $b_1 b_2 \in \{0, 1\}^2$ , which can have one of four possible values: 00, 01, 10, and 11. The combined state of two qubits  $|\varphi_1\rangle \in \mathbb{C}^2$  and  $|\varphi_2\rangle \in \mathbb{C}^2$  is the *tensor product state*  $|\varphi_1\rangle \otimes |\varphi_2\rangle$  in the four-dimensional *tensor product space*  $\mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^4$ . A basis for the tensor product space can be obtained by taking all possible combinations of the basis vectors for the individual qubits:  $\{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}$ .

**Postulate 4.** The state space of a composite quantum system is equal to the tensor product of the state spaces of the individual systems. If systems  $1, 2, \dots, n$  exist in states  $|\varphi_1\rangle, |\varphi_2\rangle, \dots, |\varphi_n\rangle$ , then the state of the composite system is  $|\varphi_1\rangle \otimes |\varphi_2\rangle \otimes \dots \otimes |\varphi_n\rangle$ .

Postulate 4 tells us how we can combine the state spaces of different quantum systems to describe a composite system. Many interesting quantum applications involve operations on multiple qubits and are described by vectors in a tensor product space, so let's look more closely at this “ $\otimes$ ”-thing.

**Tensor product space** If you'd never heard of *tensor products* before, don't worry—the only scary part is the tensor product symbol “ $\otimes$ ,” which we'll explain in this section. A *tensor product space* consists of all possible combinations of the basis vectors for the two subspaces. For example, consider two qubits  $|\varphi_1\rangle \in V_1 \equiv \mathbb{C}^2$  and  $|\varphi_2\rangle \in V_2 \equiv \mathbb{C}^2$ . We'll denote the standard basis for  $V_1$  as  $B_1 \equiv \{|0\rangle_1, |1\rangle_1\}$  and the standard basis for  $V_2$  as  $B_2 \equiv \{|0\rangle_2, |1\rangle_2\}$ . The tensor product space  $B_{12} \equiv V_1 \otimes V_2$  is four-dimensional and has the following basis:

$$B_{12} \equiv \{|0\rangle_1 \otimes |0\rangle_2, |0\rangle_1 \otimes |1\rangle_2, |1\rangle_1 \otimes |0\rangle_2, |1\rangle_1 \otimes |1\rangle_2\}.$$

This level of subscripts and the explicit use of the symbol  $\otimes$  hurts the eyes (and the hand if you must use this notation to solve problems). It's therefore customary to drop the subscripts, omit the tensor product symbol, and draw a single ket that contains a “string” of indices:

$$|a\rangle_1 \otimes |b\rangle_2 = |a\rangle \otimes |b\rangle = |a\rangle |b\rangle = |ab\rangle.$$

The basis for the tensor product space  $B_{12} \equiv V_1 \otimes V_2$  looks much nicer in the simplified notation:

$$B_{12} \equiv \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}.$$

**Tensor product of two vectors** Suppose we're given two qubits with states described by the following vectors:

$$|\varphi_a\rangle_1 = \alpha_1|0\rangle_1 + \beta_1|1\rangle_1, \quad |\varphi_b\rangle_2 = \alpha_2|0\rangle_2 + \beta_2|1\rangle_2.$$

Note the subscripts indicate which system we're describing:  $|0\rangle_1$  is the state  $|0\rangle$  for the first qubit, while  $|0\rangle_2$  is the state  $|0\rangle$  of the second qubit.

The state of the combined system is the tensor product state  $|\varphi_{ab}\rangle_{12} = |\varphi_a\rangle_1 \otimes |\varphi_b\rangle_2$ , which is computed by combining all possible combinations of the components of  $|\varphi_a\rangle_1$  and the components of  $|\varphi_b\rangle_2$ :

$$(\alpha_1, \beta_1)_{B_1} \otimes (\alpha_2, \beta_2)_{B_2} = (\alpha_1\alpha_2, \alpha_1\beta_2, \beta_1\alpha_2, \beta_1\beta_2)_{B_{12}}.$$

The notion of “all possible combinations” is easier to see by considering the tensor product operation in terms of the basis vectors:

$$\begin{aligned} |\varphi_{ab}\rangle_{12} &= |\varphi_a\rangle_1 \otimes |\varphi_b\rangle_2 \\ &= (\alpha_1|0\rangle_1 + \beta_1|1\rangle_1) \otimes (\alpha_2|0\rangle_2 + \beta_2|1\rangle_2) \\ &= \alpha_1\alpha_2|0\rangle_1|0\rangle_2 + \alpha_1\beta_2|0\rangle_1|1\rangle_2 + \beta_1\alpha_2|1\rangle_1|0\rangle_2 + \beta_1\beta_2|1\rangle_1|1\rangle_2 \\ &= \alpha_1\alpha_2|00\rangle + \alpha_1\beta_2|01\rangle + \beta_1\alpha_2|10\rangle + \beta_1\beta_2|11\rangle \\ &= (\alpha_1\alpha_2, \alpha_1\beta_2, \beta_1\alpha_2, \beta_1\beta_2)_{B_{12}}, \end{aligned}$$

where  $B_{12} \equiv \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$  is the standard basis for the tensor product space.

**State spaces and dimension counting** A quantum state that consists of  $n$  qubits can represent any unit vector in  $\mathbb{C}^{2^n}$ . That's an insanely big state space—a huge  $2^n$ -dimensional playground. In comparison, a classical bitstring of length  $n$  can take on one of  $2^n$  values. Let's compare the state space of a two-bit classical register  $R$  with the state space of a two-qubit quantum register  $|R\rangle$ . The classical register can take on one of four possible values:

$$R \in \{00, 01, 10, 11\},$$

whereas the quantum register  $|R\rangle$  can be any unit vector in  $\mathbb{C}^4$ . Similar to a classical registers, the standard basis for the state space of a two-qubit quantum register also consists of four basis vectors:

$$|00\rangle \equiv (1, 0, 0, 0), |01\rangle \equiv (0, 1, 0, 0), |10\rangle \equiv (0, 0, 1, 0), |11\rangle \equiv (0, 0, 0, 1),$$

but the quantum register can also represent superpositions of the basis states like  $|R\rangle = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0)$  or  $|R\rangle = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0)$ .

Using a very large vector space to represent states does not necessarily make a model more powerful, but the large dimension of

the tensor product space suggests many new possibilities. Much of the recent excitement in the area of quantum computing is based on the promise of using the qubits of a quantum computer to perform computations in very large quantum state spaces. We shouldn't get carried away with enthusiasm, because with great state space comes great noise. It's easy to imagine  $n$  qubits in a row in a mathematical model, but building a physical system that can store  $n$  qubits and protect them from noise is a much more difficult task. Another bottleneck in quantum computing is the difficulty of extracting information from quantum systems. The quantum state space of  $n$  qubits is  $\mathbb{C}^{2^n}$ , but projective measurements of the form  $\{\Pi_1, \Pi_2, \dots, \Pi_m\}$  can only obtain *one* answer to a question with  $m$  possible classical outcomes ( $m \leq 2^n$ ). We'll learn more about theoretical and practical considerations for quantum computing in Section 9.8.

## Exercises

**E9.9** Show that the quantum state  $|0\rangle|1\rangle - |1\rangle|0\rangle$  is equal to the quantum state  $|+\rangle|-\rangle - |-\rangle|+\rangle$ .

Hint: Express  $|0\rangle$  and  $|1\rangle$  in the basis  $\{|+\rangle, |-\rangle\}$ .

**9.9** Using the definition of  $|+\rangle$  and  $|-\rangle$ , we obtain expressions for the vectors of the standard basis written in terms of the Hadamard basis:  $|0\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle$ , and  $|1\rangle = \frac{1}{\sqrt{2}}|+\rangle - \frac{1}{\sqrt{2}}|-\rangle$ . The remainder of the calculations require arithmetic and simplifications:

$$\begin{aligned} |01\rangle - |10\rangle &= |0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle \\ &= \left(\frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle\right)\left(\frac{1}{\sqrt{2}}|+\rangle - \frac{1}{\sqrt{2}}|-\rangle\right) - \left(\frac{1}{\sqrt{2}}|+\rangle - \frac{1}{\sqrt{2}}|-\rangle\right)\left(\frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle\right) \\ &= \frac{1}{2}(\cancel{|+\rangle|+\rangle} - |+\rangle|-\rangle + |-\rangle|+\rangle - \cancel{|-\rangle|-\rangle}) - \frac{1}{2}(\cancel{|+\rangle|+\rangle} - \cancel{|+\rangle|-\rangle} + |+\rangle|-\rangle + \cancel{|-\rangle|+\rangle} - \cancel{|-\rangle|-\rangle}) \\ &= \frac{1}{2}(-2|+\rangle|-\rangle + 2|-\rangle|+\rangle) = -(|+\rangle|-\rangle - |-\rangle|+\rangle) = |+\rangle|-\rangle - |-\rangle|+\rangle. \end{aligned}$$

The last equality holds since the state's global phase can be ignored.

## Quantum entanglement

At the risk of veering further off-topic for a linear algebra book, we'll now briefly describe *entangled* quantum states. In particular, we'll discuss the properties of the *entangled state*  $|\Psi_-\rangle \equiv \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ . Entanglement is some really crazy stuff, and some of the most fascinating results in quantum information science make use of pre-shared entangled states.

In Section 7.9 we discussed how a secret key shared between two parties, Alice and Bob, can function as a *communication resource* used

to achieve private communication (via the one-time pad cryptosystem). You can think of entanglement as another type of communication resource: a stronger-than-classical correlation between two parts of a quantum system. One half of the system is controlled by Alice, the other half is controlled by Bob. When the collective state of a two-qubit quantum system is in the entangled state  $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ , measuring the individual qubits will produce anti-correlated results *in any basis*.

**Example 7** The Einstein-Podolsky-Rosen (EPR) state is a two-qubit quantum state with interesting *nonlocal* properties. We assume Alice controls one half of the quantum state, and Bob controls the other half of the following state:

$$|\Psi_{-}\rangle^{AB} \equiv \frac{1}{\sqrt{2}}|0\rangle^A|1\rangle^B - \frac{1}{\sqrt{2}}|1\rangle^A|0\rangle^B.$$

Note the use of superscripts to denote which party controls each part of the system.

Let's analyze the different measurements Alice and Bob can perform on this state. If Alice measures her system in the basis  $\{|0\rangle, |1\rangle\}$ , the projection operators that correspond to the two outcomes are

$$\Pi_0^{AB} = |0\rangle\langle 0|^A \otimes \mathbb{I}^B \quad \text{and} \quad \Pi_1^{AB} = |1\rangle\langle 1|^A \otimes \mathbb{I}^B.$$

Since only Alice's half of the state is measured, the measurement acts like the identity operator on Bob's half of the state. There's a 50-50 chance of outcomes 0 and 1:

$$\Pr(0|\Psi_{-}) = \langle \Psi_{-} | \Pi_0^{AB} | \Psi_{-} \rangle = \frac{1}{2}, \quad \Pr(1|\Psi_{-}) = \langle \Psi_{-} | \Pi_1^{AB} | \Psi_{-} \rangle = \frac{1}{2}.$$

Depending on the outcome, the post-measurement state of the system will be either  $|0\rangle^A|1\rangle^B$  or  $|1\rangle^A|0\rangle^B$ . If Bob then measures his half of the system, he'll obtain the outcome opposite Alice's. In other words, the measurement outcomes that Alice and Bob obtain are perfectly anti-correlated.

What if Alice and Bob choose to measure their respective halves of the EPR state  $|\Psi_{-}\rangle^{AB}$  in the basis  $\{|+\rangle, |-\rangle\}$ ? Using some basic calculations (see Exercise E9.9), we can express the EPR state  $|\Psi_{-}\rangle^{AB}$  in terms of the basis  $\{|+\rangle, |-\rangle\}$  as follows:

$$\frac{1}{\sqrt{2}}|0\rangle^A|1\rangle^B - \frac{1}{\sqrt{2}}|1\rangle^A|0\rangle^B = |\Psi_{-}\rangle^{AB} = \frac{1}{\sqrt{2}}|+\rangle^A|-\rangle^B - \frac{1}{\sqrt{2}}|-\rangle^A|+\rangle^B.$$

Observe that the state  $|\Psi_{-}\rangle^{AB}$  has the same structure in the Hadamard basis as in the standard basis. Thus, Alice and Bob's measurement

outcomes will also be perfectly anti-correlated when measuring in the Hadamard basis.

A state  $|\psi\rangle^{AB}$  is called *entangled* if it cannot be written as a tensor product of quantum states  $|\phi\rangle^A \otimes |\varphi\rangle^B$ , where  $|\phi\rangle^A$  describes the state held by Alice and  $|\varphi\rangle^B$  the state held by Bob. The EPR state  $|\Psi_{-}\rangle^{AB}$  is *entangled*, which means it cannot be written as a tensor product of the quantum states of individual qubits:

$$|0\rangle\otimes|1\rangle - |1\rangle\otimes|0\rangle \neq (a|0\rangle + b|1\rangle)\otimes(c|0\rangle + d|1\rangle),$$

for any  $a, b, c, d \in \mathbb{C}$ . Since we cannot describe the EPR state  $|\Psi_{-}\rangle^{AB}$  as the tensor product of two local states  $|\phi\rangle^A$  and  $|\varphi\rangle^B$ , we say it requires a *nonlocal* description, which is another way of saying  $|\Psi_{-}\rangle^{AB}$  is entangled.

There is something strange about the EPR state. If Alice measures her half of the state and finds  $|0\rangle$ , then we know immediately that Bob's state will be  $|1\rangle$ . The collapse in the superposition on Alice's side immediately causes a collapse of the superposition on Bob's side. Note that Bob's collapse will occur *immediately*, no matter how distant Bob's system is from Alice's. This is what the authors Einstein, Podolsky, and Rosen called "spooky action at a distance." How could Bob's system "know" to always produce the opposite outcome even at the other end of the universe?

Now imagine you have a whole bunch of physical systems prepared in the EPR state. Alice controls one half of each of the EPR pairs, while Bob controls the other half. This is a communication resource called *shared entanglement*. Many of the quantum information protocols make use of shared entanglement between Alice and Bob, to achieve novel communication tasks.

A useful perspective for thinking about quantum entanglement is to consider the information encoded in the quantum state. This is called the *information theory* perspective on quantum physics. From the point of view of information, we can say a system is entangled whenever we know more about the system as a whole than about its parts. We have complete certainty about the state of the composite system  $|\Psi_{-}\rangle^{AB}$ , and complete uncertainty about the states of the individual subsystems controlled by Alice and Bob. When Alice measures her half of the entangled state, the two outcomes of the measurement are equally likely. Her measurement outcome essentially corresponds to a completely random bit, which encodes zero information. I would like to tell you more about quantum information theory, but let's not go further off course.

Instead of discussing theory, let's describe a practical scenario in which it is possible to know more about a whole system than about its constituent parts.



**Physics example** We'll now describe a physical process that leads to the creation of an entangled quantum state. Consider a quantum particle  $p$  that *decays* into two quantum subparticles,  $p_a$  and  $p_b$ . The decay process obeys various physics conservation laws; in particular, the total spin angular momentum before and after the decay must be conserved. Supposing the particle  $p$  has zero spin angular momentum before the decay, then the law of conservation of angular momentum dictates the subparticles  $p_a$  and  $p_b$  must have opposite spin. The spin angular momentum of each particle is in an unknown direction (up, down, left, right, in, or out), but whichever spin direction we measure for  $p_a$ , the spin angular momentum of  $p_b$  immediately takes on the opposite direction. This way, regardless of the direction of their individual spins, their combined spin angular momentum will be zero, as required by the law of conservation of angular momentum.

The general scenario discussed above describes what would happen if a Helium atom were to explode, the two electrons in its ground state flying off to distant sides of the universe. The two electrons have opposite spins, but we don't know the directions of the individual spins. The only thing we know is that their total spin equals zero, since the ground state of the Helium atom has spin zero. This is how the "anti-correlation in *any* basis" aspect of quantum entanglement arises.

\* \* \*

## Summary

We can summarize the concepts of quantum mechanics we learned in this chapter, and relate them to the concepts of linear algebra:

quantum state  $\Leftrightarrow$  vector  $|v\rangle \in \mathbb{C}^d$   
 evolution  $\Leftrightarrow$  unitary operations  
 measurement  $\Leftrightarrow$  projections  
 composite system  $\Leftrightarrow$  tensor product

The quantum formalism embodied in the four postulates of quantum mechanics has been applied in describing many physical phenomena. Using complex vectors to represent quantum states leads to useful models and predictions for experimental outcomes. In the next section, we'll use the quantum formalism to analyze the outcomes of the polarizing lenses experiment.

In addition to the applications of quantum principles, studying the structure of quantum states and operations is an interesting field

on its own. An example of a quantum idea that is fundamentally new is the existence of *entangled* quantum states. Later in this chapter, we'll discuss an interesting application of quantum entanglement as part of the *quantum teleportation* protocol, illustrated in Figure 9.23 (page 572).

## Exercises

**E9.10** Find the matrix representation of the projection matrices  $\Pi_+ \equiv |+\rangle\langle+|$  and  $\Pi_- \equiv |-\rangle\langle-|$ . Show that  $\Pi_+ + \Pi_- = \mathbb{1}$ .

**9.10**  $M_{\Pi_+} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$  and  $M_{\Pi_-} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$ .

**E9.11** Compute the probability of outcome “−” for the measurement  $\{\Pi_+, \Pi_-\} = \{|+\rangle\langle+|, |-\rangle\langle-|\}$  performed on the quantum state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ .

**9.11**  $\Pr(\{-\}|\psi) = \frac{|\alpha - \beta|^2}{2}$ .

**E9.12** Given the state  $|\theta\rangle = (\frac{1}{\sqrt{2}}, \frac{e^{i\theta}}{\sqrt{2}})^T$ , find a quantum state  $|\theta^\perp\rangle$  that is orthogonal to  $|\theta\rangle$ . Find the projection operators  $\Pi_\theta$  and  $\Pi_{\theta^\perp}$  that correspond to the measurements in the basis  $\{|\theta\rangle, |\theta^\perp\rangle\}$ . Verify that  $\Pi_\theta + \Pi_{\theta^\perp} = \mathbb{1}$ . Compute the probability of outcome  $\theta$  when performing the measurement  $\{\Pi_\theta, \Pi_{\theta^\perp}\}$  on the state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ .

Hint: The state  $|\theta^\perp\rangle$  satisfies  $\langle\theta^\perp|\theta\rangle = 0$ .

**9.12**  $|\theta^\perp\rangle = (\frac{1}{\sqrt{2}}, \frac{e^{-i(\pi-\theta)}}{\sqrt{2}})^T$ .  $M_{\Pi_\theta} = \begin{bmatrix} \frac{1}{2} & \frac{e^{-i\theta}}{2} \\ \frac{e^{i\theta}}{2} & \frac{1}{2} \end{bmatrix}$  and  $M_{\Pi_{\theta^\perp}} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2}e^{i(\pi-\theta)} \\ \frac{1}{2}e^{-i(\pi-\theta)} & \frac{1}{2} \end{bmatrix}$ .  
 $\Pr(\{-\}|\psi) = \sqrt{\left|\frac{\alpha}{2} + \frac{\beta}{2}e^{-i\theta}\right|^2 + \left|\frac{\alpha}{2}e^{i\theta} + \frac{\beta}{2}\right|^2}$ .

**E9.13** Given the two qubits  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  and  $|\phi\rangle = \gamma|0\rangle + \delta|1\rangle$ , compute the tensor product state  $|\psi\rangle \otimes |\phi\rangle$ .

**9.13**  $|\psi\rangle \otimes |\phi\rangle = \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle$ .

## Links

[ Compact set of notes on QM written by a physicist ]  
[http://graybits.biz/notes/quantum\\_mechanics/preface](http://graybits.biz/notes/quantum_mechanics/preface)

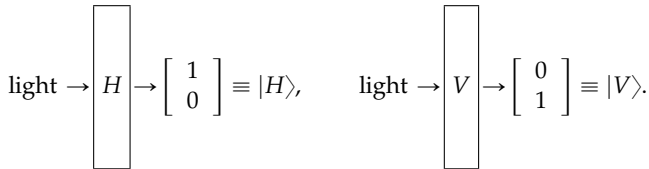
[ Lecture series on QM written by a computer scientist ]  
<http://scottaaronson.com/democritus/lec9.html>

[ Quantum mechanics summary written by a philosopher ]  
<http://plato.stanford.edu/entries/qm/>

## 9.6 Polarizing lenses experiment revisited

Let's revisit the polarizing lenses experiment, this time modelling photons' polarization states as two-dimensional complex vectors. We define the state of a horizontally polarized photon as  $|H\rangle \equiv |0\rangle \equiv (1, 0)^T$ , and the state of a vertically polarized photon as  $|V\rangle \equiv |1\rangle \equiv (0, 1)^T$ . This choice corresponds to the observation that horizontal and vertical polarizations are complementary. We shall interpret the polarizing lenses in the optical circuit both as state preparation and measurement steps. The measurement outcome we observe is whether photons pass through the complete optical circuit.

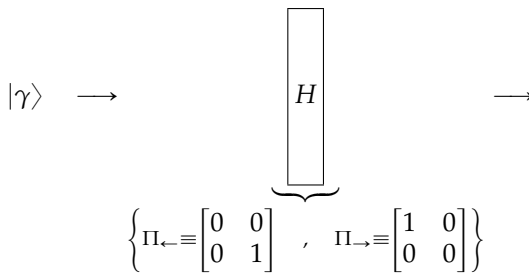
Figure 9.14 illustrates the state preparation step we can use in experiments with photons. We can prepare photons in the states  $|H\rangle$  or  $|V\rangle$  by starting from unpolarized photons, and passing them through either an  $H$ -polarizing lens or a  $V$ -polarizing lens.



**Figure 9.14:** State preparation procedure for photons with horizontal or vertical polarization. Depending on our choice of  $H$ - or  $V$ -polarizing lens, we can prepare different quantum states.

Placing a polarizing lens in the path of a photon performs the measurement  $\{\Pi_{\leftarrow}, \Pi_{\rightarrow}\}$  of the photon's polarization state. Using the lens to measure a photon's polarization state yields two possible outcomes: "photon passes through lens" or "photon is reflected by lens." The polarizing lens is the measurement apparatus, and the measurement corresponds to whether each photon that hits the lens is horizontally or vertically polarized.

Figure 9.15 shows the projection operators that correspond to a measurement using a horizontally polarizing lens. The outcome "passes through" corresponds to the projection operator  $\Pi_{\rightarrow} = |H\rangle\langle H| \equiv \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ . The outcome "is reflected" corresponds to the projection matrix  $\Pi_{\leftarrow} = |V\rangle\langle V| \equiv \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ .

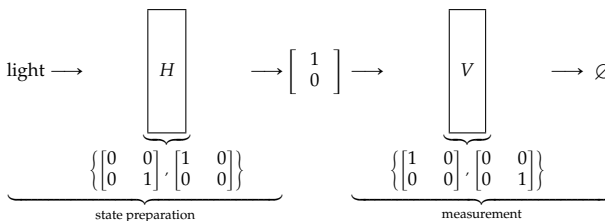


**Figure 9.15:** A horizontally polarizing lens corresponds to the quantum measurement  $\{\Pi_{\leftarrow}, \Pi_{\rightarrow}\}$ . An incoming photon in the state  $|\gamma\rangle$  is asked to choose one of the two alternative paths. With probability  $\|\Pi_{\rightarrow}|\gamma\rangle\|^2$ , the photon passes through the  $H$ -polarizing lens and becomes horizontally polarized  $|\gamma'\rangle = (1, 0)^T$ . With probability  $\|\Pi_{\leftarrow}|\gamma\rangle\|^2$ , the photon is reflected.

Recall that the probability of outcome  $i$  in a quantum measurement  $\{\Pi_i\}$  is given by the expression  $\|\Pi_i|\gamma\rangle\|^2$ , where  $|\gamma\rangle$  is the state of the incoming photon. Knowing the projection operators that correspond to the “passes through” and “is reflected” outcomes allows us to predict the probability that photons with a given state will pass through the lens. Figure 9.16 illustrates a two-step experiment in which photons prepared in the horizontally polarized state  $|H\rangle = (1, 0)^T$  arrive at a  $V$ -polarizing lens. The probability that photons pass through the  $V$ -polarizing lens is

$$\begin{aligned} \Pr(\{\text{pass through } V \text{ lens given state } |H\rangle\}) &= \|V \chi V |H\rangle\|^2 \\ &= \left\| \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|^2 \\ &= \left\| \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\|^2 = 0. \end{aligned}$$

Indeed, we observe the same result in the lab—all  $|H\rangle$  photons are rejected by the  $V$ -polarizing lens.



**Figure 9.16:** Photons prepared in the state  $|H\rangle = (1, 0)^T$  are rejected by the  $V$ -polarizing lens because the horizontally polarized state has zero overlap with the projector  $\Pi_{\rightarrow} = |V\rangle\langle V|$ .

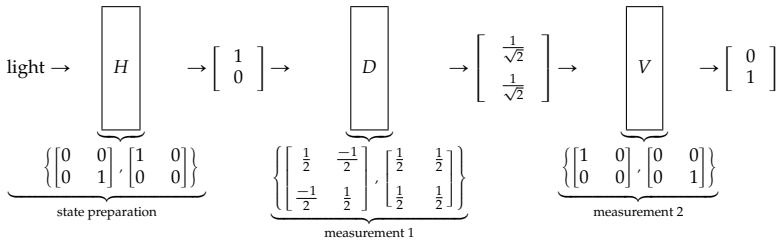
Let's now use the quantum formalism to analyze the results of the three-lenses experiment we discussed earlier in the chapter. Figure 9.17 shows the optical circuit that consists of a state preparation step and two measurement steps. The diagonally polarizing lens placed in the middle of the circuit only allows photons with  $45^\circ$ -diagonal polarization to pass through:  $|D\rangle \equiv (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$ . The projection operator associated with the “passes through” outcome of the diagonal polarization measurement is

$$\Pi_{\rightarrow} = |D\rangle\langle D| = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

The probability of  $|H\rangle$  photons passing through the diagonal lens is

$$\begin{aligned} \Pr(\{\text{pass through } D \text{ lens given state } |H\rangle\}) &= \|\langle D|\langle D||H\rangle\|^2 \\ &= \left\| \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|^2 \\ &= \left\| \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \right\|^2 \\ &= \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}. \end{aligned}$$

The post-measurement state of photons that make it through the diagonally polarizing lens is  $|D\rangle \equiv (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$ .



**Figure 9.17:** Photons prepared in the state  $|H\rangle$  are subjected to two sequential measurements: a diagonal polarizing measurement  $D$  followed by a vertical polarizing measurement  $V$ . The projection operators for the “is reflected” and “passes through” outcomes are indicated in each step.

Photons that pass through the middle lens are in the state  $|D\rangle$ . The probability of these photons passing through the  $V$ -polarizing lens

is

$$\begin{aligned}
 \Pr(\{\text{pass through } V \text{ lens given state } |D\rangle\}) &= \|V \chi V |D\rangle\|^2 \\
 &= \left\| \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right\|^2 \\
 &= \left\| \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right\|^2 \\
 &= 0^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}.
 \end{aligned}$$

The overall probability of a photon passing through both measurement lenses is  $\frac{1}{4}$ .

This probability is consistent with the light intensity observations we see in Figure 9.8 (page 522). Note the new interpretation that quantum physics provides for the same observations. We previously referred to the optical power  $P = 0.5$  after the first measurement and optical power  $P = 0.25$  after the second measurement. The quantum formalism predicts the probability  $p = 0.5$  for photons to pass through the first lens and the probability  $p = 0.25$  for the photons to reach the end of the circuit.

## Discussion

The polarizing lenses experiment illustrates some key aspects of the quantum formalism: the modelling of quantum states as vectors, the process of state preparation, and the effects of measurements on systems. Mainly, the experiment shows how a probabilistic approach for describing the light intensity in the circuit successfully predicts the experiment's outcome.

Let's compare the quantum interpretation of this experiment with a classical interpretation. It's possible to explain all the observations of the polarizing lenses experiment using the classical theory of electromagnetic waves. Modelling the light beam as a classical wave allows us to understand the projections to the "passes through" orientation of the polarizing lenses. Because the light beam consists of many photons, it behaves like a continuous quantity that can be split: part of the wave passes through the lens while another part is reflected. Classical wave theory correctly predicts the qualitative aspects of the experiment shown in Figure 9.17, but the light intensity is described in terms of optical power  $P = \frac{1}{4}$ , which is the average squared amplitude of an electromagnetic wave. In contrast, quantum theory models photons as discrete packets of energy, and explains the outcome of the polarizing lenses experiment as the prob-

ability  $p = \frac{1}{4}$  that a photon will pass through the circuit. For this tabletop experiment, both the classical model of light (as a wave that can be infinitely subdivided) and the quantum model of light (as discrete particles) predict the same outcome. Therefore, the polarizing lenses experiment performed with a laser pointer does not serve as proof for the necessity of a quantum mechanical description of reality.

As we discussed in the introduction to this chapter, to really reveal quantum effects, we need to look at the very small-scale or very low-power regimes. It's possible to reproduce the polarizing lenses experiment using a *single photon source*. A single photon source behaves like a super-weak laser pointer that emits only one photon at a time. When discussing the single photon regime, the classical theory of electromagnetic waves hits a wall, since a single photon cannot subdivide into parts—it is a quantum of light; a small, indivisible bundle of energy. Classical wave theory can correctly predict average optical power, but it can't provide a clear picture of what exactly happens when individual photons hit a polarizing lens.

The polarizing lenses experiment is inspired by the famous *Stern–Gerlach experiment*, which is performed with the magnetic spin of silver atoms, and which involves a similar demonstration, comparable observed outcomes, and analogous reasoning. I encourage you to learn more about the original Stern–Gerlach experiment.

[ The Stern–Gerlach experiment ]

[https://en.wikipedia.org/wiki/Stern-Gerlach\\_experiment](https://en.wikipedia.org/wiki/Stern-Gerlach_experiment)

<https://youtube.com/watch?v=rg4Fnag4V-E>

## 9.7 Quantum physics is not that weird

Without a doubt, you've heard that quantum mechanics is weird, mysterious, and generally “magical.” Well, unless vector operations count as magic, it's not *that* magical. In this section, we'll single out three so-called “weird” aspects of quantum mechanics: superposition, interference, and the fact that quantum measurements affect the states of systems being measured.

### Quantum superposition

Classical binary variables (bits) can have one of two possible values: 0 or 1. Examples of physical systems that behave like bits are electric switches that can be either open or closed, digital transistors that either conduct or don't conduct electricity, and capacitors that are either charged or discharged.

A quantum bit (qubit), can be both 0 and 1 *at the same time*. Wow! Said this way, it surely sounds impressive and mystical, no? But if we use the term *linear combination* instead of “at the same time,” the quantum reality doesn’t seem so foreign. A quantum state is a linear combination of the basis states. This isn’t so crazy. The *superposition principle* is a general notion in physics that is not specific to quantum phenomena, but applies to all systems described by differential equations. Indeed, superpositions exist in many classical physics problems, too.

**Example** Consider a mass attached to a spring that undergoes simple harmonic motion. The differential equation that governs the motion of the mass is  $x''(t) + \omega^2 x(t) = 0$ . This equation has two solutions:  $x_0(t) = \sin(\omega t)$  and  $x_1(t) = \cos(\omega t)$ , corresponding to two different starting points of the oscillation. Since both  $x_0(t)$  and  $x_1(t)$  satisfy the equation  $x''(t) + \omega^2 x(t) = 0$ , any linear combination of  $x_0(t)$  and  $x_1(t)$  is also a solution. Thus, the most general solution to the differential equation is of the form:

$$x(t) = \alpha x_0(t) + \beta x_1(t) = \alpha \sin(\omega t) + \beta \cos(\omega t).$$

Usually we combine the sin and cos terms and describe the equation of motion for the mass-spring system in the equivalent form  $x(t) = A \cos(\omega t + \phi)$ , where  $A$  and  $\phi$  are computed from  $\alpha$  and  $\beta$ . The mass-spring system might be described as undergoing both sin motion and cos motion “at the same time,” but do you see how ridiculous and incomplete this sounds?

\* \* \*

The notion of *quantum superposition* is simply a consequence of the general superposition principle for differential equations. If the quantum states  $|0\rangle$  and  $|1\rangle$  both represent valid solutions to a quantum differential equation, then the state of the system can be described as a linear combination of these two solutions:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle.$$

The observation that “ $|\psi\rangle$  is both  $|0\rangle$  and  $|1\rangle$  at the same time” is not wrong; it’s just not very useful. It’s much more precise to describe the quantum state  $|\psi\rangle$  as a linear combination.

## Interference

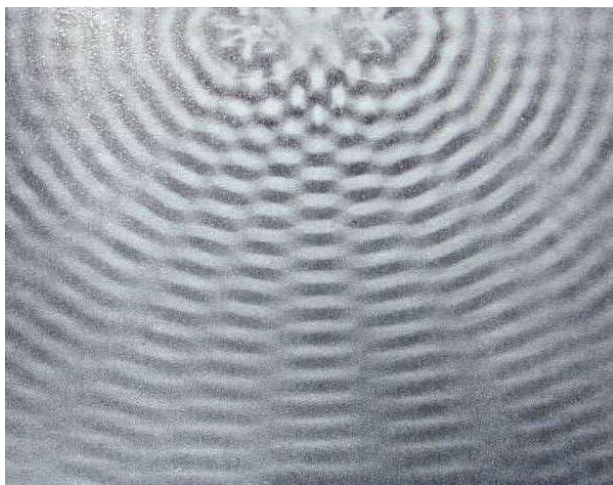
Unlike particles that bounce off each other, waves can co-exist in the same place. The resulting wave pattern is the sum of the constituent



waves. Quantum particles behave similarly to waves in certain experiments, and this can lead to *interference* between quantum systems.

The prototypical example of interference is Young's double-split experiment, in which particles passing through two thin slits interact with each other, causing an interference pattern of alternating bright and dark spots on a screen. Classical physics models assume particles behave like tiny point-like balls that bounce off each other whenever they come in contact. A classical model predicts that particles will appear on the screen in two bright peaks, directly facing the two slits.

In contrast, the quantum model of a particle describes it as a travelling energy pulse that exhibits wave-like properties.<sup>2</sup> In a quantum model, the particles passing through the slits behave like waves and can combine constructively or destructively, depending on the relative distances travelled by the particles. Similar interference patterns occur whenever waves combine, as in the example of waves on the surface of a liquid, or sound waves.



**Figure 9.18:** The waves emitted by two synchronized sources form an interference pattern. Observe the stripes of destructive interference where the waves meet “out of sync” (peak to trough) and cancel each other out.

Performing Young's double-split experiment reveals a pattern of bright and dark stripes (called fringes) on the screen in support of the quantum model. The locations of the dark fringes correspond exactly to the places where particles passing through the two slits

---

<sup>2</sup>This is where the name *wave function* comes from.

arrive “out of sync,” and combine destructively:

$$|\psi\rangle - |\psi\rangle = 0.$$

The locations where destructive interference occurs correspond to the dark fringes on the screen, where no particles arrive.

The idea that one wave can cancel another wave is not new. What *is* new is the observation that particles behave like waves that can interfere with each other. That’s definitely new. Indeed, interference was one of the first puzzling effects of quantum systems that was observed. Observations from interference experiments forced physicists to attribute wave-like properties to particles.

[ Video demonstration of Young’s double-split experiment ]

<https://youtube.com/watch?v=qCmtegdq00A>

## Measurement of a system affects the system’s state

Another seemingly weird aspect of quantum mechanics is the notion that quantum measurements can affect the states of the systems being measured. This phenomenon is not attributable to some sort of quantum magic, but is rather due to the energy scale and the size of systems where quantum physics comes into play. Let’s see why.

When we think about physical systems on the scale of individual atoms, we can no longer consider ourselves (and our physical measurement apparatuses) as passive observers of these systems. Instead, we need to account for the *interactions* between quantum systems and the measurement apparatuses used to observe them. The fact that measurements affect the state of the very systems they measure is not some magical process, but rather a consequence of the natural properties of the particles themselves. The particles we observe are affected by our measurement methods.

## Wave functions

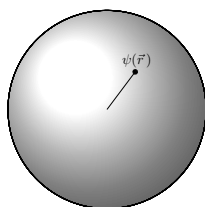
The quantum mechanics techniques we discussed in this chapter are useful for modelling physical systems that have discrete sets of states. In *matrix quantum mechanics*, quantum states are described by vectors in finite-dimensional, complex inner product spaces. Other physics problems require the use of *wave function quantum mechanics*, in which quantum states are represented as complex-valued functions of space coordinates  $\vec{r} = (x, y, z)$ . Instead of the dot product between vectors, the inner product for wave functions is  $\langle f(\vec{r}), g(\vec{r}) \rangle = \iiint_{\mathbb{R}^3} \overline{f(\vec{r})} g(\vec{r}) d^3\vec{r}$ . This may seem like a totally new ball game, but actually calculations using wave functions are not too different from

the inner product calculations we used to compute Fourier transformations in Section 7.11.

It's beyond the scope of the current presentation to discuss *wave functions* in detail, but I want to show you an example of a calculation with wave functions, so you won't say I didn't show you some proper physics stuff. The ground state of the hydrogen atom is described by the wave function  $\psi(\vec{r}) = \frac{1}{\sqrt{\pi a^3}} \exp(-r/a)$ , where  $r = \|\vec{r}\|$ . The probability of finding the electron at position  $\vec{r}$  from the proton is described by the inner product  $\psi(\vec{r})\psi(\vec{r}) \equiv |\psi(\vec{r})|^2$ :

$$\Pr(\{\text{finding electron at } \vec{r}\}) = |\psi(\vec{r})|^2.$$

Since  $\psi(\vec{r})$  depends only on the distance  $r$ , we know the wave function has a spherically symmetric shape, as illustrated in Figure 9.19.



**Figure 9.19:** The  $s$  orbital of an electron is spherically symmetric.

We'll now check whether  $|\psi(\vec{r})|^2$  is a properly normalized probability density function. Integrating the probability density function  $|\psi(\vec{r})|^2$  over all of  $\mathbb{R}^3$  should give a total probability of one. Instead of Cartesian coordinates  $(x, y, z)$ , we'll use spherical coordinates  $(r, \phi, \theta)$  to solve this problem. In spherical coordinates, the volume of a thin slice from the surface of a sphere of width  $d\theta$ , height  $d\phi$ , and thickness  $dr$  is given by  $r^2 \sin \phi \, d\phi \, d\theta \, dr$ . If you haven't seen spherical coordinates before, don't worry about this expression too much. The conversion factor  $r^2 \sin \phi$  is a trick for converting the "small piece of volume"  $d^3\vec{r} = dx \, dy \, dz$  to an equivalent small piece of volume in spherical coordinates  $d^3\vec{r} = r^2 \sin \phi \, d\phi \, d\theta \, dr$ . See P3.26 for the derivation.

We split the triple integral into two parts: an integral that depends only on the radius  $r$ , and double integral over the angles  $\phi$  and  $\theta$ . The total probability of finding the electron somewhere in space is

$$p_{\text{total}} = \iiint_{\mathbb{R}^3} |\psi(\vec{r})|^2 \, d^3\vec{r}$$

$$\begin{aligned}
&= \int_0^\infty \int_0^{2\pi} \int_0^\pi |\psi(r)|^2 r^2 \sin \phi \, d\phi \, d\theta \, dr \\
&= \left( \int_0^\infty \frac{1}{\pi a^3} \exp(-2r/a) r^2 \, dr \right) \left( \int_0^{2\pi} \int_0^\pi \sin \phi \, d\phi \, d\theta \right) \\
&= \left( \int_0^\infty \frac{1}{\pi a^3} \exp(-2r/a) r^2 \, dr \right) (4\pi) \\
&= \underbrace{\int_0^\infty \frac{4}{a^3} \exp\left(-\frac{2r}{a}\right) r^2 \, dr}_{p(r)}.
\end{aligned}$$

The expression  $p(r) = \frac{4}{a^3} \exp(-2r/a) r^2$  describes the probability of finding the electron at a distance  $r$  from the centre of the nucleus. We can complete the calculation of total probability by taking the integral of  $p(r)$  from  $r = 0$  to  $r = \infty$ :

$$\begin{aligned}
p_{\text{total}} &= \int_0^\infty p(r) \, dr \\
&= \frac{4}{a^3} \int_0^\infty \exp\left(-\frac{2r}{a}\right) r^2 \, dr \\
&= 1.
\end{aligned}$$

The purpose of working through this wave function calculation is to give you an idea of the complex calculations physicists regularly perform using the wave function formalism. In comparison, the matrix formalism for quantum mechanics is much simpler, involving only basic linear algebra calculations.

## 9.8 Quantum mechanics applications

What can we accomplish using quantum physics that we can't do using classical physics? What can we compute with qubits that we can't compute with bits? You've learned about the quantum formalism; but how useful is it? In this section, we'll explore some areas of physics and computer science that wouldn't exist without the laws of quantum mechanics.

### Particle physics

The basic quantum mechanics formalism we learned in this chapter is not appropriate for describing the behaviour of high energy particles. The best current model for describing high energy physics is called *quantum field theory*. Quantum field theory is a generalization

of basic quantum mechanics that models fundamental particles as disturbances in various *particle fields*. Recall that photons, the quanta of light, are described as disturbances of the electromagnetic field. The equations of quantum field theory describe how elementary particles and antiparticles are created and destroyed, and the various interactions between different particle fields. Quantum field theory is a bit like chemistry, where different combinations of atoms are transformed into other combinations of atoms—only instead of atoms, we have elementary particles like quarks and leptons. In the same way *Mendeleev's periodic table* is a catalogue of all available atoms, the *Standard Model* of particle physics is a catalogue of all available elementary particles. These elementary particles combine to form other, more complex particles, like protons and neutrons. The transformations of these particles are only observed in high energy physics experiments performed in particle accelerators.

High energy physics becomes important at the extreme end of an energy continuum, where quantum field theory is most applicable. In contrast, the rules of chemistry dominate the low-energy end of the energy continuum. Chemical reactions describe how molecules transform into other molecules; and molecules essentially represent the various ways electrons are shared between groups of atoms. At higher energies, atoms are “stripped” of their electrons; the atoms have so much energy, they’re no longer bound to the nucleus. At this point, the laws of chemistry are no longer relevant, since electrons and molecules move freely. Enter nuclear physics, which studies the combinations of protons and neutrons that form the nuclei of different atoms. A nuclear reaction is like a chemical reaction, but instead of chemical molecules, the reactant and products are various types of nuclei. An example of nuclear reaction is the fusion of two heavy hydrogen nuclei to form a helium nucleus. At higher energy still, even protons and neutrons can break apart, and the analysis shifts to interactions between elementary particles like leptons, bosons, neutrinos, quarks, and photons. This is the domain of high energy physics.

The basic postulates of quantum mechanics still apply in quantum field theory, but the models become more complicated since we assume even the interactions between particles are quantized. You can think of the basic quantum mechanics described in this chapter as learning the alphabet, and quantum field theory as studying Shakespeare, including the invention of new words. Studying quantum field theory requires new math tools like path integrals, new intuitions like symmetry observations, and new computational tricks like *renormalization*.

The essential way of thinking about photons, electrons, and the interactions between them can be obtained by reading Richard Feyn-

man's short book titled *QED*, which stands for *quantum electrodynamics*. In this tiny book, Feynman uses the analogy of a "tiny clock" attached to each particle to explain the phase  $e^{i\theta}$  of a wave function. From this simple analogy, the author builds to explain complex concepts (path integrals, for instance) at the graduate level of quantum field theory. I highly recommended this book; it's a wonderful chance to learn from one of the great scientists in the field and one of the best physics teachers of all times.

[ The Standard Model of particle physics ]

[https://en.wikipedia.org/wiki/Standard\\_Model](https://en.wikipedia.org/wiki/Standard_Model)

[ Nuclear fusion is how energy is generated inside stars ]

[https://en.wikipedia.org/wiki/Nuclear\\_fusion](https://en.wikipedia.org/wiki/Nuclear_fusion)

[BOOK] Richard P. Feynman. *QED: The strange theory of light and matter*. Princeton University Press, 2006, ISBN 0691125759.

## Solid state physics

Physicists have sought to understand the inner structure of materials since the first days of physics. As they've learned, they've developed numerous applications, from semiconductors to lasers, photovoltaic batteries (solar panels), light emitting diodes (LEDs). These applications all depend on materials with specially engineered conductivity properties. Indeed, working with the conductivity of materials gives us insight into their other properties. We can classify materials into the following general conductivity-type groups: insulators, metals, and semi-conductors. These categories correspond to materials with different *energy band structures*.

Insulators are the most boring type of material, because their energy band structure doesn't permit any interesting chemical interactions. Take glass, for instance—just a clump of silica ( $\text{SiO}_2$ ). The term *glass* is used in physics to describe any material made of randomly oriented molecules that lack a specific crystal structure.

Conductors are more interesting. A hand-wavy explanation of conductivity would be to say the electrons in conductors like aluminum and copper are "free to move around." Solid state physics allows for a more precise understanding of the phenomenon. Using quantum mechanical models, we can determine the energy levels that electrons can occupy, and predict how many electrons will be available to conduct electricity.

Semiconductors are the most interesting type of material since they can switch between conductive and non-conductive states. The *transistor*, the invention that makes all electronics possible, consists

of a sandwich of three different types of semiconductors. The voltage applied to the middle section of a transistor is called the *gate voltage*, and it controls the amount of current that can flow through the transistor. If the gate voltage is set to ON (think 1 in binary), then semiconducting material is biased; free electrons are available in its conduction band, and current can flow through. If the gate voltage is set to OFF (think 0 in binary), then the conduction band is depleted and the transistor won't conduct electricity. The improvements in semiconductor technologies—specifically the ability to pack billions of transistors into a tiny microprocessor chip—have been fuelling the ongoing computer revolution pretty much since transistors were first commercialized. In summary, no solid state physics = no mobile phones.

Quantum mechanics is used so much in solid state physics that we could suitably nickname the field “applied quantum physics.”

[ Simple explanation of energy band structure and conductivity ]

[https://wikipedia.org/wiki/Electrical\\_resistivity\\_and\\_conductivity](https://wikipedia.org/wiki/Electrical_resistivity_and_conductivity)

## Superconductors

Certain materials exhibit surprising physical properties at very low temperatures. By low temperatures, I mean *really low*, like  $-272^{\circ}\text{C}$ ; a temperature close to *absolute zero*, the temperature at which all molecules stop. You'd exhibit surprising properties too if you were placed in an environment this cold! Take for example the properties of different metals, which are generally considered to have low electric resistance. There are regular conductors like aluminum that have low resistance, and high-end conductors like silver that have even lower resistance, and then there are *superconductors* which have zero resistance. Superconductors are an example of a purely quantum phenomenon that cannot be explained by classical physics.

Some of the most iconic landmarks of modern scientific progress, like magnetic resonance imaging (MRI) machines and magnetically levitating bullet trains, are made possible by superconductor technology. Superconductors offer zero resistance to electric current, which means they can support much stronger currents than regular conductors like aluminum and silver. All these applications require extensive refrigeration equipment to keep the materials at a temperature where they behave like superconductors.

[ Superconductivity ]

<https://en.wikipedia.org/wiki/Superconductivity>

## Quantum optics

Classical optics deals with beams of light that contain quintillions of photons. A *quintillion* is  $10^{18}$ , which is more than a lot. When working with this many photons, it's possible to model light beams as continuous electromagnetic waves, and use classical electromagnetic theory and optics to understand experiments. Quantum optics comes into play when we perform optics experiments that involve far fewer photons, including experiments with single photons. When a single photon travels through an optical circuit, it cannot “split” like a continuous wave. For example, when a beam of light hits a half-silvered mirror, we say the beam is partially reflected. However, we can't say the same for a single photon, since the photon cannot be split. Instead, the state of the photon after it has encountered a half-silvered mirror is best described as a superposition of the “passed through” and “reflected” states, as shown in Figure 9.11 (page 536).

An example of a quantum optics effect is the *spontaneous down-conversion effect*, in which a single photon is absorbed by a material and then reemitted as two photons with an entangled polarization state:

$$|\Psi_{-}\rangle = \frac{1}{\sqrt{2}}|H\rangle|V\rangle - \frac{1}{\sqrt{2}}|V\rangle|H\rangle.$$

By the crystal's properties, we know one of the two emitted photons has horizontal polarization and the other has vertical polarization, but we don't know which is which. Such entangled photons can be used as starting points for other experiments that involve entanglement. Another interesting aspect of quantum optics are the so-called *squeezed states* that can be detected more accurately than regular (unsqueezed) photons.

Quantum optics is a field of active research. Scientists in academia and industry study exotic photon generation, advanced photon detection schemes, and generally explore how photons can most efficiently carry information.

[ Basic principles in physics of light ]

<https://materialford.wordpress.com/introduction-to-research-light/>

## Quantum cryptography

Performing a quantum measurement on the state  $|\psi\rangle$  tends to disturb the state. From the perspective of experimental physics, this disturbance is an obstacle since it gives us limited, one-time access to the quantum state  $|\psi\rangle$ , making the study of quantum states more difficult. From the perspective of cryptography however, the state-disturbing aspect of quantum measurement is an interesting and po-



tentially valuable property. If Alice transmits a secret message to Bob encoded in the state of a quantum system, it would be impossible for an eavesdropper Eve to “listen in” on the state unnoticed because Eve’s measurement would disturb the state. The *BB84 protocol*, named after its inventors Charles Bennett and Gilles Brassard, is based on this principle.

The standard basis  $B_s = \{|0\rangle, |1\rangle\}$  and the Hadamard basis  $B_h = \{|+\rangle, |-\rangle\}$  are *mutually unbiased bases*, which means that a basis vector from one basis lies exactly halfway between the vectors from the other basis. If we measure the state  $|+\rangle$  (or  $|-\rangle$ ) in the basis  $B_s$ , the outcomes 0 and 1 are equally unlikely. Similarly, measuring  $|0\rangle$  (or  $|1\rangle$ ) in the basis  $B_h$  produces outcomes “+” and “-” with equal probability. The use of mutually unbiased bases is central to the security of the BB84 protocol, which we’ll describe step by step:

1. Alice starts with  $2n$  random candidate bits which she sends to Bob. Roughly half of the candidate bits will live on to become the final shared secret key, while the other half will be discarded. She chooses one of the bases  $B_s$  or  $B_h$  at random when encoding each bit of information she wants to send. Bob chooses to perform his measurement randomly, either in the standard basis or in the Hadamard basis. The information is transmitted correctly whenever Bob happens to pick the same basis as Alice, which happens about half the time. Whenever Bob’s basis for measurement is different from the basis Alice uses for encoding, Bob’s output is completely random.
2. Alice and Bob publicly announce the basis they used for each transmission and discard the bits where different bases were used. This leaves Alice and Bob with roughly  $n$  candidate bits of secret key.
3. Alice and Bob then publicly reveal  $\alpha n$  of the candidate bits, which we’ll call the check bits. Assuming the quantum communication channel between Alice and Bob does not introduce any noise, Alice and Bob’s copies of the check bits should be identical, since they used the same basis. If they observe many check bits that disagree, Alice and Bob will immediately abort the protocol.
4. If the  $\alpha n$  check bits agree, then Alice and Bob can be sure the remaining  $(1 - \alpha)n$  bits they share are known only to them.

Consider what happens if the eavesdropper Eve tries to intercept the messages between Alice and Bob. Eve can measure the quantum state  $|\psi\rangle$  sent by Alice, then forward to Bob the post measurement state  $|\psi'\rangle$ . Eve is forced to choose a basis for the measurement

she performs, and her measurement disturbs the state  $|\psi\rangle$  whenever she picks a basis different from the one used by Alice. Since it's not possible to measure quantum systems without disturbing them, the eavesdropper Eve reveals her presence by introducing errors in the transmitted data. Some of the *check bits* Alice and Bob compare in Step 3 will disagree, and thus Alice and Bob will know that someone is eavesdropping on them. Though quantum mechanics does not prevent eavesdropping, it does give Alice and Bob the ability to detect when an eavesdropper is present.

The BB84 protocol established the beginning of a new field at the intersection of computer science and physics that studies *quantum key distribution* protocols. The field has developed rapidly from theory to research, and today there are even commercial quantum cryptography systems. It's interesting to compare quantum cryptography with the public key cryptography systems discussed in Section 7.9. The security of the RSA public-key encryption is based on the computational difficulty of factoring large numbers. The security of quantum cryptography is guaranteed by the laws of quantum mechanics.

[ Bennett–Brassard quantum cryptography protocol from 1984 ]  
<https://en.wikipedia.org/wiki/BB84>

[ Using quantum phenomena to distribute secret keys ]  
[https://en.wikipedia.org/wiki/Quantum\\_key\\_distribution](https://en.wikipedia.org/wiki/Quantum_key_distribution)

## Quantum computing

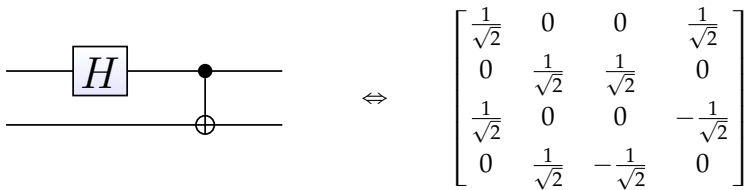
The idea of quantum computing has existed since the early days of quantum physics. Richard Feynman originally proposed the idea of a *quantum simulator* in 1982, which is a quantum apparatus that can simulate the quantum behaviour of another physical system. Imagine a device that can simulate the behaviour of physical systems that would otherwise be too difficult and expensive to build. The quantum simulator would be much better at simulating quantum phenomena than any simulation of quantum physics on a classical computer.

Another possible application of a quantum simulator could be to encode classical mathematical optimization problems as constraints in a quantum system, then let the quantum evolution of the system “search” for good solutions. Using a quantum simulator in this way, it might be possible to find solutions to optimization problems much faster than any classical optimization algorithm could.

Once computer scientists started thinking about quantum computing, they weren't satisfied with studying optimization problems

alone, and they set out to qualify and quantify all the computational tasks that are possible with qubits. A quantum computer stores and manipulates information that is encoded as quantum states. It's possible to perform certain computational tasks on a quantum computer much faster than on any classical computer. We'll discuss *Grover's search algorithm* and *Shor's factoring algorithm* below, but first let's introduce the basic notions of quantum computing.

**Quantum circuits** Computer scientists like to think of quantum computing tasks as series of “quantum gates,” in analogy with the logic gates used to construct classical computers. Figure 9.20 shows an example of a quantum circuit that takes two qubits as inputs and produces two qubits as outputs.



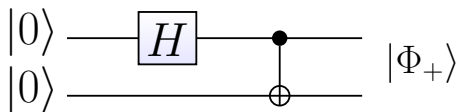
**Figure 9.20:** A quantum circuit that applies the Hadamard gate to the first qubit, then applies the controlled-NOT gate from the first qubit to the second qubit.

This circuit in Figure 9.20 is the combination of two quantum gates. The first operation is to apply the Hadamard gate  $H$  on the first qubit, leaving the second qubit untouched. This operation is equivalent to multiplying the input state by the matrix  $H \otimes \mathbb{1}$ . The second operation is called the *controlled-NOT* (or controlled- $X$ ) gate, which applies the  $X$  operator (also known as the NOT gate) to the second qubit whenever the first qubit is  $|1\rangle$ , and does nothing otherwise:

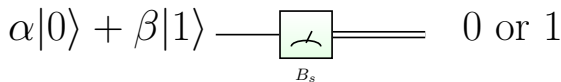
$$\text{CNOT}(|0\rangle \otimes |\varphi\rangle) = |0\rangle \otimes |\varphi\rangle, \quad \text{CNOT}(|1\rangle \otimes |\varphi\rangle) = |1\rangle \otimes X|\varphi\rangle.$$

The circuit illustrated in Figure 9.20 can be used to create entangled quantum states. If we input the quantum state  $|00\rangle \equiv |0\rangle \otimes |0\rangle$  into the circuit, we obtain the maximally entangled state  $|\Phi_+\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$  as output, as depicted in Figure 9.21.

Quantum circuits can also represent quantum measurements. Figure 9.22 shows how a quantum measurement in the standard basis is represented.



**Figure 9.21:** Inputting  $|0\rangle\otimes|0\rangle$  into the circuit produces an EPR state  $|\Phi_+\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$  on the two output wires of the circuit.



**Figure 9.22:** Measurement in the standard basis  $B_s = \{|0\rangle, |1\rangle\}$ . The projectors of this measurement are  $\Pi_0 = |0\rangle\langle 0|$  and  $\Pi_1 = |1\rangle\langle 1|$ .

We use double lines to represent the flow of classical information in the circuit.

**Quantum registers** Consider a quantum computer with a single register  $|R\rangle$  that consists of three qubits. The quantum state of this quantum register is a vector in  $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ :

$$|R\rangle = (\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle) \otimes (\alpha_3|0\rangle + \beta_3|1\rangle),$$

where the tensor product  $\otimes$  is used to combine the quantum states of the individual qubits. We'll call this the "physical representation" of the register and use 0-based indexing for the qubits. Borrowing language from classical computing, we'll call the rightmost qubit the *least significant* qubit, and the leftmost qubit the *most significant* qubit.

The tensor product of three vectors with dimension two is a vector with dimension eight. The quantum register  $|R\rangle$  is thus a vector in an eight-dimensional vector space. The quantum state of a three-qubit register can be written as:

$$|R\rangle = a_0|0\rangle + a_1|1\rangle + a_2|2\rangle + a_3|3\rangle + a_4|4\rangle + a_5|5\rangle + a_6|6\rangle + a_7|7\rangle,$$

where  $a_i$  are complex components. We'll call this eight-dimensional vector space the "logical representation" of the quantum register. Part of the excitement about quantum computing is the huge size of the "logical space" where quantum computations take place. The logical space of a 10-qubit quantum register has dimension  $2^{10} = 1024$ . That's 1024 complex components we're talking about. That's a big state space for just a 10-qubit quantum register. Compare this

with a 10-bit classical register, which can store one of  $2^{10} = 1024$  discrete values.

We won't discuss quantum computing further here, but I still want to show you some examples of single-qubit quantum operations and their effect on the tensor product space, so you'll have an idea of the craziness that is possible.

**Quantum gates** Let's say you've managed to construct a quantum register; what can you do with it? Recall the single-qubit quantum operations  $Z$ ,  $X$ , and  $H$  we described earlier. We can apply any of these operations on individual qubits in the quantum register. For example, applying the  $X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  gate to the first (most significant) qubit of the quantum register corresponds to the following quantum operation:

$$\begin{array}{c} \text{---} \\ \boxed{X} \\ \text{---} \end{array} \Leftrightarrow X \otimes \mathbb{1} \otimes \mathbb{1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

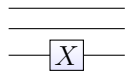
The operator  $X \otimes \mathbb{1} \otimes \mathbb{1}$  “toggles” the first qubit in the register while leaving all other qubits unchanged.

Yes, I know the tensor product operation is a bit crazy, but that's the representation of composite quantum systems and operations so please get used to it. What if we apply the  $X$  operator to the middle qubit?

$$\begin{array}{c} \text{---} \\ \boxed{X} \\ \text{---} \end{array} \Leftrightarrow \mathbb{1} \otimes X \otimes \mathbb{1} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Compare the structure of the operators  $X \otimes \mathbb{1} \otimes \mathbb{1}$  and  $\mathbb{1} \otimes X \otimes \mathbb{1}$ . See how the action of  $X$ s affects different parts of the tensor product space  $\mathbb{C}^8$ ?

To complete the picture, let's also see the effects of applying the  $X$  gate to the third (least significant) qubit in the register:



The diagram shows three horizontal lines representing qubits. The bottom line has a box labeled 'X' on it, indicating an X gate is applied to the third qubit.

$$\Leftrightarrow \mathbb{1} \otimes \mathbb{1} \otimes X = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Crazy stuff, right? Don't worry, in time you'll get used to the space-within-a-space structure concept.

### Okay, so what?

Quantum computers give us access to a very large state space. The fundamental promise of quantum computing is that a small set of simple quantum operations (quantum gates) can be used to perform interesting computational tasks. Sure it's difficult to interact with and manipulate quantum systems, but the space is so damn big that it's worth checking out what kind of computing you can do in there. It turns out there are already several useful things you can do using a quantum computer. The two flagship applications for quantum computing are Grover's search algorithm and Shor's factoring algorithm.

**Grover's search algorithm** Suppose you're given an unsorted list of  $n$  items and you want to find a particular item in that list. This is called an *unstructured search problem*. This is a hard problem to solve for a classical computer since the algorithm must parse through the entire list, which takes roughly  $n$  steps. In contrast, the unstructured problem can be solved in roughly  $\sqrt{n}$  steps on a quantum computer using *Grover's search algorithm*.

The quantum speed for the unstructured search problem sure is nice, but it's really nothing to get excited about. The real money-maker for the field of quantum computing has been Shor's factoring algorithm for factoring products of prime numbers.

**Shor's factoring algorithm** The security of the RSA cryptosystem we discussed in Section 7.9 is based on the assumption that factoring products of large prime numbers is computationally intractable. Given the product  $de$  of two unknown prime numbers  $d$  and  $e$ , it is

computationally difficult to find the factors  $e$  and  $d$ . No classical algorithm is known that can factor large numbers; even the letter agencies will have a hard time finding the factors of  $de$  when  $d$  and  $e$  are chosen to be sufficiently large prime numbers. Thus, if an algorithm that could quickly factor large numbers existed, attackers would be able to break many of the current security systems. *Shor's factoring algorithm* fits the bill, theoretically speaking.

Shor's algorithm reduces the factoring problem to the problem of *period finding*, which can be solved efficiently using the quantum Fourier transform. Shor's algorithm can factor large numbers efficiently (in polynomial time). This means RSA encryption would be easily hackable by running Shor's algorithm on a sufficiently large, and sufficiently reliable quantum computer. The letter agencies are excited about this development since they'd love to be able to hack all present-day cryptography. Can you imagine not being able to log in securely to any website because Eve is listening in, hacking your crypto using her quantum computer?

Currently, Shor's algorithm is only a *theoretical* concern. Despite considerable effort, no quantum computers exist today that can manipulate quantum registers with thousands of qubits.

## Discussion

Quantum computing certainly presents interesting possibilities, but it's a little early to imagine a quantum computing revolution in tomorrow's newspaper. As with startup ventures, it's the implementation that counts—not the idea. The current status of quantum computing as a technology is mixed. On one hand, certain quantum algorithms performed in logical space are very powerful; on the other hand, the difficulty of building a quantum computer is not to be underestimated.

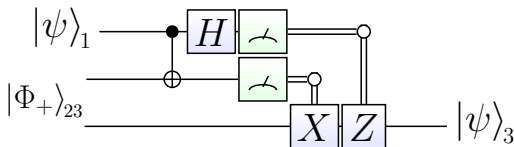
It's also important to keep in mind that quantum computers are not better at solving arbitrary computational problems than the computers we already use. The problems that may benefit from a quantum speedup have a particular structure, which can be tackled with a choreographed pattern of constructive and destructive interference in quantum registers. Yet not all computationally hard problems have this structure. Quantum computing technology is at a crossroad: it could become a revolutionary development, or it could turn out that building a large-scale quantum computer is not worth the engineering challenge. So although it's cool we can execute certain tasks faster on a quantum computer, don't throw out your classical computer just yet.

Even if the quest to build a quantum computer doesn't pan out,

we're certain to learn many interesting things about fundamental physics along the way. Besides, learning about the fundamental nature of quantum information is more scientifically valuable than trying to hack people's email. In the next section, we'll give an example of a new communication task that was discovered through the study of quantum information science.

**Quantum teleportation** Figure 9.23 illustrates a surprising aspect of quantum information: we can “teleport” a quantum state  $|\psi\rangle$  from one lab to another. The quantum state  $|\psi\rangle$  starts in the first qubit of the register, which is held by Alice, and ends in the third qubit, which is in Bob's lab, but there is no quantum communication channel between the two labs. This is why the term “quantum teleportation” was coined to describe this communication task, since the state  $|\psi\rangle$  seems to materialize in Bob's lab like the teleportation machines used in Star Trek.

The communication resources required for the quantum teleportation protocol are one maximally entangled state shared between Alice's and Bob's labs, and two bits of classical communication from Alice to Bob. We can express the quantum teleportation protocol as a quantum circuit.



**Figure 9.23:** The first two qubits are in Alice's lab. The state of the first qubit  $|\psi\rangle_1$  is transferred into the third qubit  $|\psi\rangle_3$ , which Bob controls. We say  $\psi$  is “teleported” from Alice's lab to Bob's lab because the quantum state ends up in Bob's lab, but there is no quantum communication channel connecting the labs. The state teleportation happens thanks to the pre-shared entanglement and the two bits of classical information.

The quantum teleportation protocol requires that Alice and Bob pre-share a maximally-entangled state  $|\Phi_+\rangle$ . By meeting in a central location, Alice and Bob can produce an entangled state using the circuit shown in Figure 9.21 (page 568). Alice and Bob then bring their respective halves of the entangled state to their labs. Note Bob's lab could be very far from Alice's lab; in another building, another city, or even on the other side of the world.

The initial state for the quantum teleportation protocol is

$$|\psi\rangle_1 \otimes |\Phi_+\rangle_{23} = \left( \alpha|0\rangle_1 + \beta|1\rangle_1 \right) \otimes \left( \frac{1}{\sqrt{2}}|00\rangle_{23} + \frac{1}{\sqrt{2}}|11\rangle_{23} \right).$$



Alice has two qubits in her lab, the state  $|\psi\rangle_1 = \alpha|0\rangle_1 + \beta|1\rangle_1$  and half of the entangled state, and Bob has the third qubit, which is the other half of the entangled state. At the end of the teleportation protocol, information about the state  $\psi$  appears in Bob's lab:  $|\psi\rangle_3 = \alpha|0\rangle_3 + \beta|1\rangle_3$ .

Without quantum communication, it seems impossible for Alice to communicate the components  $\alpha$  and  $\beta$  to Bob. The pre-shared entanglement between Alice and Bob somehow enables this feat. The quantum information about the state  $\psi$  becomes available in Bob's lab as soon as Alice performs the measurement of her two qubits. But before Bob obtains the state information, he must apply a *recovery operation*, which is one of  $\mathbb{I}$ ,  $X$ ,  $Z$ , or  $ZX$ . The state information will remain unknown until Bob learns which of the four recovery operations he must perform on the state. The information about which operation Bob should perform can be transmitted by classical means: Alice can shout the result to Bob if he's next door, tell him the results by phone, or send him a text message. After applying the needed recovery operation, Bob obtains the state  $|\psi\rangle_3 = \alpha|0\rangle_3 + \beta|1\rangle_3$ , which is the state that originated in the first qubit controlled by Alice.

The need for pre-shared entanglement  $|\Phi_+\rangle$  between Alice and Bob is analogous to how Alice and Bob needed to pre-share a secret key  $\tilde{k}$  in order to use the one-time pad encryption protocol. Indeed, pre-shared entangled states are a prime resource in quantum information science. The *superdense coding protocol* is another surprising application of quantum entanglement. With this protocol, Alice can communicate two bits of classical information to Bob by sending him a single qubit and consuming one pre-shared entangled state.

## Links

[ Quantum simulators and practical implementations ]  
[https://en.wikipedia.org/wiki/Quantum\\_simulator](https://en.wikipedia.org/wiki/Quantum_simulator)

[ Some data about the difficulty of RSA factoring ]  
[https://en.wikipedia.org/wiki/RSA\\_numbers](https://en.wikipedia.org/wiki/RSA_numbers)

[ An introduction to quantum computing ]  
<http://arxiv.org/abs/0708.0261v1/>

[ Video tutorials on quantum computing by Michael Nielsen ]  
<http://michaelnielsen.org/blog/quantum-computing-for-the-determined/>

[ Grover's algorithm for unstructured search ]  
[https://en.wikipedia.org/wiki/Grover's\\_algorithm](https://en.wikipedia.org/wiki/Grover's_algorithm)

[ Shor's algorithm for factoring products of prime integers ]  
[https://en.wikipedia.org/wiki/Shor's\\_algorithm](https://en.wikipedia.org/wiki/Shor's_algorithm)

[ Emerging insights on limitations of quantum computing ]  
<https://archive.siam.org/pdf/news/100.pdf>

## Quantum error-correcting codes

Quantum states are finicky things. Every interaction of a qubit with its environment corrupts the quantum information the qubit stores. In the previous section we talked about quantum computing in the abstract, assuming the existence of an ideal noiseless quantum computer. Since the real world is a noisy place, constructing a practical quantum computer is a much greater challenge.

Recall that errors caused by noise are also a problem for classical computers. If classical computers can be made robust to errors using error-correcting codes, can we use error-correcting codes on quantum computers too? Indeed it's possible to use *quantum error-correcting codes* to defend against the effects of quantum noise. Keep in mind, quantum error-correcting codes are more complicated to build than their classical counterparts, so it's not an obvious thing to do, but it can be done.

We won't go into too much detail, but it's worth pointing out the following interesting fact about quantum error correction. Building quantum error-correcting codes that can defend against a finite set of errors is sufficient to defend against all possible types of errors. The use of quantum error-correcting schemes is analogous to the classical error-correcting schemes we saw in Section 7.10. We encode  $k$  qubits of data that we want to protect from noise into a larger  $n$ -qubit state. The encoded state can support some number of errors before losing the data. The error-correcting procedure involves a syndrome measurement on a portion of the state, and "correction" operators applied to the remaining portion. I encourage you to follow the links below to learn more about this topic.

Building reliable quantum gates is a formidably complicated task due to the difficulty of protecting qubits from noise while simultaneously enabling quantum operations and strong interactions between qubits. It is the author's opinion that Feynman's original idea of building quantum simulators for physical systems will be the first useful applications in quantum computing.

[ More on quantum error-correcting codes ]  
[https://en.wikipedia.org/wiki/Quantum\\_error\\_correction](https://en.wikipedia.org/wiki/Quantum_error_correction)

## Quantum information theory

Classical information theory studies problems like the compression of information and the transmission of information through noisy

communication channels. Quantum information theory studies the analogous problems of compression of quantum information and communication over noisy quantum channels.

The appearance of the word “theory” in “quantum information theory” should indicate that this is a mostly theoretical area of research that studies problems in the abstract. The main results of information theory are abstract theorems that may not have direct bearing on practical communication scenarios. For now, applications of quantum information theory remain in the far-off future, but that’s how it is with theory subjects in general. The classical information theorems proven in the 1970s probably looked a bit useless, too; but these theorems serve as the basis of all modern wireless communications. Perhaps the purely theoretical quantum information theorems of today will solve the practical communication problems of the future.

Current efforts in quantum information theory aim to establish capacity results for quantum channels. Some of the existing results are directly analogous to classical capacity results. Other problems in quantum information theory, like the use of entanglement-assisted codes, have no classical counterparts and require a completely new way of thinking about communication problems. The book *From Classical to Quantum Shannon Theory* by Mark M. Wilde is an excellent guide to the field.

Recently, quantum theory has been applied to novel communication systems, and there is a growing interest from the communications industry to develop applications that push optical communication channels to the bounds of their theoretical efficiency. Essentially, quantum networks are being invented in parallel with quantum computers, so that when we finally build quantum computers, we’ll be able to connect them together, presumably so they can share funny cat videos. What else would we use them for?

[BOOK] Mark M. Wilde. *From Classical to Quantum Shannon Theory*, Cambridge University Press, Second edition, ISBN 1107176166, <http://arxiv.org/abs/1106.1445>.

## Conclusion

With this chapter, I wanted to bring you closer to the fascinating subject of quantum mechanics. I hope the material helped you understand the basic principles of quantum mechanics and clarified some of the sensational mythology surrounding the “mysteries” of the quantum world. While there’s still much to discover, there’s nothing too counterintuitive about quantum mechanics; it’s just linear algebra, right?

One hundred years ago, quantum mechanics was seen as a foreign subject not to be trusted. In time, physicists developed good models, found better ways to explain experiments, wrote good books, and even started teaching the subject to undergraduate students. This gives me hope for humanity that we can handle even the most complex and uncertain topics when we put our minds to it.

Today we face many complex problems: consolidated corporate control of innovation, cartels, corruption, eroding democratic government systems, the militarization of everything, and conflicting ideologies. We have Sunni and Shia brothers shooting at each other and red gang versus blue gang brothers shooting at each other, and all of this for no good reason. We have all kinds of other bullshit divisions between us.

Let's hope that one hundred years from now, we'll have learned to limit the violent and corrupt aspects of human nature, so that we can realize the potential of every child born anywhere in the world. Right now it seems like it won't be an easy change, but this is how it seemed when people were trying to figure out quantum mechanics, too. All it takes is a critical mass of people who realize and truly internalize the fact we're all on the same team, and all the divisions we see between us are pure bullshit. If we make sure that two generations of kids grow up without economic strife or bullets flying by their heads, then—I guarantee you—they will be able to figure out the rest. Together, an educated citizenry armed with the knowledge of math, history, science, and technology is more powerful than a dozen Systems combined. All the System has is an outdated hierarchical power structure and capital to pay people and make them do as they're told. We've got six billion people, the internet, and printing presses on our side. Who do you think will win?

## 9.9 Quantum mechanics problems

Let's recap what just happened here. Did we really cover all the topics of an introductory quantum mechanics course? Yes, we did! Thanks to your solid knowledge of linear algebra, learning the postulates of quantum mechanics took only a few dozen pages. Sure we went quickly and skipped the more physics-y topics, but we covered all the core ideas of quantum theory.

But surely it's impossible to learn quantum mechanics in such a short time? Well, you tell me. You're here. The problems are here. Prove to me you've really learned quantum mechanics by tackling the practice problems presented in this section like a boss. It's the end of the book, so don't be saving your energy. Solve these problems and then you're done.

**P9.1** You work in a quantum computing startup and your boss asks you to implement the quantum gate  $Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ . Can you do it?

Hint: Recall the requirements for quantum gates.

**9.1** No, since  $Q$  is not unitary.

**9.1** Quantum gates correspond to unitary operators. Since  $Q^\dagger Q \neq \mathbb{1}$ ,  $Q$  is not unitary and it cannot be implemented by any physical device. The boss is *not* always right!

**P9.2** The  $Y$  gate is defined as  $Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$ . Compute the effect of the operator  $YY$  on the vectors of the standard basis  $\{|0\rangle, |1\rangle\}$ .

**9.2**  $YY|0\rangle = |0\rangle$  and  $YY|1\rangle = |1\rangle$ .

**9.2** Let's first see what happens to  $|0\rangle$  when we apply the operator  $YY$ . The result of the first  $Y$  applied to  $|0\rangle$  is  $Y|0\rangle = i|1\rangle$ . Then, applying the second  $Y$  operator, we get  $YY|0\rangle = Y(i|1\rangle) = iY|1\rangle = i(-i)|0\rangle = |0\rangle$ . So  $YY|0\rangle = |0\rangle$ . A similar calculation shows that  $YY|1\rangle = |1\rangle$ .

**P9.3** Compute  $XHHY(\alpha|0\rangle + \beta|1\rangle)$ .

Hint: Use the Hadamard gate's properties to simplify the calculation.

**9.3**  $i\alpha|0\rangle - i\beta|1\rangle$ .

**9.3** Since  $HH = \mathbb{1}$ , we obtain  $XHHY = XY = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$ , then multiply this operator by  $\alpha|0\rangle + \beta|1\rangle$  to obtain the answer.

**P9.4** Specifying an arbitrary vector  $\alpha|0\rangle + \beta|1\rangle \in \mathbb{C}^2$  requires four parameters: the real and imaginary parts of  $\alpha$  and  $\beta$ . Thus one might think that qubits have four *degrees of freedom*. However, the unit-length requirement and the fact that we can ignore the global phase of a qubit correspond to additional constraints that reduce the number of degrees of freedom. How many parameters are required to specify a general quantum state  $|\psi\rangle \in \mathbb{C}^2$ ?

**9.4** Two parameters.

**9.4** Starting from the four degrees of freedom for general two-dimensional complex vectors, we must subtract one degree of freedom for each of the constraints: one because we're ignoring global phase, and one because we require  $|\alpha|^2 + |\beta|^2 = 1$ :

$$4 \text{ d.f.} - \alpha \text{ real} - \{\|\psi\rangle\| = 1\} = 2 \text{ d.f.}$$

A qubit  $|\psi\rangle$  has only two degrees of freedom. In other words, two parameters are sufficient to describe any qubit.

**P9.5** We can write any qubit using only two real parameters:

$$|\psi\rangle = \alpha|0\rangle + \sqrt{1 - \alpha^2}e^{i\varphi}|1\rangle,$$

where  $\alpha \in \mathbb{R}$  and  $\varphi \in \mathbb{R}$ . What are the ranges of values for  $\alpha$  and  $\varphi$  such that all qubits can be represented?

**9.5**  $\alpha \in [0, 1]$  and  $\varphi \in [0, 2\pi]$ .

**9.5** Since the squared-magnitudes of the components of  $|\psi\rangle$  must be a probability distribution,  $\alpha$  is restricted in the range  $[0, 1]$ . The phase of the second component can be chosen arbitrarily;  $\varphi \in [0, 2\pi]$ .

**P9.6** Consider the parametrization for qubits using two angles  $\theta$  and  $\varphi$ :

$$|\psi\rangle = \cos(\theta/2)|0\rangle + \sin(\theta/2)e^{i\varphi}|1\rangle.$$

What values should  $\theta$  and  $\varphi$  have for all qubits to be represented?

**9.6**  $\theta \in [0, \pi]$  and  $\varphi \in [0, 2\pi]$ .

**9.6** These are the angles of the *Bloch sphere*, which is a useful way to visualize qubit quantum states.

**P9.7** Compute the products of the quantum gates  $HXH$  and  $HZH$ .

**9.7**  $HXH = Z$  and  $HZH = X$ .

**P9.8** Consider the state  $|v\rangle = (a, b)^T$  and its orthogonal complement  $|v^\perp\rangle = (\bar{b}, -\bar{a})^T$ . The projection operators  $\Pi_v$  and  $\Pi_{v^\perp}$  correspond to the measurements in the basis  $\{|v\rangle, |v^\perp\rangle\}$ . Compute the probability of outcome  $v$  in the measurement  $\{\Pi_v, \Pi_{v^\perp}\}$  on the state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ .

**9.8**  $\Pr(\{v\}|\psi) = |\alpha\bar{a} + \beta\bar{b}|^2$ .

**P9.9** When we measure a quantum system  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  in the basis  $\{|0\rangle, |1\rangle\}$ , the Born rule tells us the probability of outcome 0 is equal to  $\langle\psi|\Pi_0|\psi\rangle$ . Consider this calculation that involves the trace operation:

$$\begin{aligned} \Pr(\{0\}|\psi) &= \langle\psi|\Pi_0|\psi\rangle = \langle\psi|\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}|\psi\rangle = \text{Tr}\left\{\langle\psi|\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}|\psi\rangle\right\} \\ &\stackrel{(c)}{=} \text{Tr}\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}|\psi\rangle\langle\psi|\right\} = \text{Tr}\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\underbrace{\begin{bmatrix} |\alpha|^2 & \bar{\beta}\alpha \\ \bar{\alpha}\beta & |\beta|^2 \end{bmatrix}}_{\text{density matrix}}\right\} \\ &= \text{Tr}\left\{\begin{bmatrix} |\alpha|^2 & \bar{\beta}\alpha \\ 0 & 0 \end{bmatrix}\right\} = |\alpha|^2. \end{aligned}$$

The equality labelled (c) follows from the cyclic property of the trace operation  $\text{Tr}\{ABC\} = \text{Tr}\{BCA\}$ . The above calculation suggests an alternative approach for computing the probabilities of different outcomes of quantum

measurements,  $\Pr(\{x\}|\psi) = \text{Tr}\{\Pi_x \rho\}$ , where  $\rho$  is the *density matrix* representation of the quantum state. The density matrix of the quantum state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  corresponds to the outer product  $\rho = |\psi\rangle\langle\psi|$ .

Calculate the probability of outcome  $\{1\}$  using  $\Pr(\{1\}|\psi) = \text{Tr}\{\Pi_1 \rho\}$ . Calculate the probabilities of the two outcomes of a measurement in the Hadamard basis  $\Pr(\{+\}|\psi) = \text{Tr}\{\Pi_+ \rho\}$  and  $\Pr(\{-\}|\psi) = \text{Tr}\{\Pi_- \rho\}$ .

**9.9**  $\Pr(\{0\}|\psi) = |\beta|^2$ ;  $\Pr(\{+\}|\psi) = \frac{1}{2}|\alpha + \beta|^2$ ;  $\Pr(\{-\}|\psi) = \frac{1}{2}|\alpha - \beta|^2$ .

**9.9** All three calculations require computing the product of the appropriate projection matrix with the density matrix  $\rho = \begin{bmatrix} |\alpha|^2 & \bar{\alpha}\beta \\ \bar{\alpha}\beta & |\beta|^2 \end{bmatrix}$ , then taking the trace. The density matrix representation is used in many domains of physics, as well as in quantum information theory.

**P9.10** This problem explores the operation of the quantum teleportation circuit shown in Figure 9.23 (see page 572). The initial state of the three-qubits register is  $|\psi\rangle_1 \otimes |\Phi_+\rangle_{23}$ , where  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  is a quantum state Alice wants to send to Bob, and where  $|\Phi_+\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$  is a maximally entangled state shared between Alice and Bob.

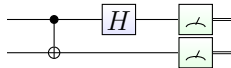
a) Show the following equation holds by expanding the tensor product:

$$|\psi\rangle_1 \otimes |\Phi_+\rangle_{23} = \frac{1}{\sqrt{2}} [\alpha|000\rangle_{123} + \beta|100\rangle_{123} + \alpha|011\rangle_{123} + \beta|111\rangle_{123}].$$

b) The expression from part a) can be written as a linear combination of the four Bell states:  $|\Phi_+\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ ,  $|\Phi_-\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$ ,  $|\Psi_+\rangle \equiv \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ , and  $|\Psi_-\rangle \equiv \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ . Verify the equation

$$\begin{aligned} & \frac{1}{\sqrt{2}} [\alpha|000\rangle_{123} + \beta|100\rangle_{123} + \alpha|011\rangle_{123} + \beta|111\rangle_{123}] \\ &= \frac{1}{2} [|\Phi_+\rangle_{12} |\psi\rangle_3 + |\Phi_-\rangle_{12} Z |\psi\rangle_3 + |\Psi_+\rangle_{12} X |\psi\rangle_3 + |\Psi_-\rangle_{12} XZ |\psi\rangle_3]. \end{aligned}$$

c) A *Bell measurement* consists of the combination of a controlled-NOT gate and a Hadamard gate on the first qubit, followed by measurements of both qubits in the standard basis, as illustrated in the circuit below.



Using the definition of the controlled-NOT gate (page 567) and the Hadamard gate (page 538), show that the Bell measurement performed on the state  $|\Phi_+\rangle$  produces the classical measurement outcome 00. Similarly, show that measuring  $|\Phi_-\rangle$  produces 10, measuring  $|\Psi_+\rangle$  produces 01, and measuring  $|\Psi_-\rangle$  produces 11.

d) After Alice performs the Bell measurement on the two qubits under her control, the state of Bob's qubit will be one of the following:  $|\psi\rangle_3$  if the measurement outcome is 00,  $Z|\psi\rangle_3$  if outcome is 10,  $X|\psi\rangle_3$  if outcome is 01, or  $XZ|\psi\rangle_3$  if outcome is 11. Indicate the *recovery operation* Bob must apply in order to recover the state  $|\psi\rangle_3$  in each case.

**9.10 a)** Expand the expressions for  $|\psi\rangle_1$  and  $|\Phi_+\rangle_{23}$  and compute the tensor product. **b)** Using the definitions of the Bell states  $|\Phi_+\rangle$ ,  $|\Phi_-\rangle$ ,  $|\Psi_+\rangle$ ,  $|\Psi_-\rangle$ , and the  $X$  and  $Z$  operators, we expand the expression on the right side of the equation and show it equals the expression on the left side, which in turn equals  $|\psi\rangle_1 \otimes |\Phi_+\rangle_{23}$ . **c)** Applying the CNOT gate to the state  $|\Phi_+\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$  results in  $\frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle$ . Next, applying the Hadamard gate on the first qubit leaves us with the state  $|00\rangle$ , which leads to the measurement outcome 00 when both qubits are measured in the standard basis. The analysis for the other bell states is similar. **d)** The recovery operation that Bob must perform is determined by the measurement outcome. Since the  $X$  and  $Z$  gates are self-inverse, the recovery operations that Bob must perform are described by the following mapping:  $00 \rightarrow \mathbb{I}$ ,  $10 \rightarrow Z$ ,  $01 \rightarrow X$ ,  $11 \rightarrow ZX$ . See this video for more details about the quantum teleportation protocol: [youtu.be/3wZ35c3oYUE](https://youtu.be/3wZ35c3oYUE).

**P9.11** The wave function of the electron of the hydrogen atom is  $\psi(\vec{r}) = \frac{1}{\sqrt{\pi a^3}} \exp(-r/a)$ . The electron's distance from the centre is described by the random variable  $R$  with probability distribution  $p_R(r) = \frac{4}{a^3} \exp(-2r/a)r^2$ . Calculate the expected distance of the electron  $\mathbb{E}_R[R] = \int_{r=0}^{\infty} r p_R(r) dr$ .

Hint: You can solve this problem using integration by parts once.

**9.11**  $\mathbb{E}_R[R] = \frac{3}{2}a$ .

**9.11** Using integration by parts with  $u = r^3$  and  $dv = \exp(-2r/a)r^2 dr$  simplifies the integral to the form  $\frac{3a}{2} (\frac{4}{a^3} \int_0^{\infty} \exp(-\frac{2r}{a})r^2 dr)$ . Instead of continuing with two more steps of integration by parts, we can recognize the expression inside the brackets to be equal to one, since  $p_R(r) = \frac{4}{a^3} \exp(-2r/a)r^2$  is a probability distribution. This calculation supports the classical chemical viewpoint, which describes electrons as living in an “electron cloud” or “orbital” of radius roughly  $\frac{3}{2}a$ . Note this is somewhat misleading since the actual wave function  $\psi(\vec{r})$  drops off as  $\exp(-r)$ .

**P9.12** Show that the functions  $\psi_1(x) = 2x - 1$  and  $\psi_2(x) = 6x^2 - 6x + 1$  are orthogonal with respect to the inner product  $\langle f, g \rangle = \int_0^1 f(x)g(x) dx$ .

**9.12**  $\langle \psi_1, \psi_2 \rangle = 0$ .

**9.12** The inner product  $\langle \psi_1, \psi_2 \rangle = \int_0^1 \overline{\psi_1(x)} \psi_2(x) dx$  corresponds to the integral  $\int_0^1 (2x - 1)(6x^2 - 6x + 1) dx = \int_0^1 (12x^3 - 18x^2 + 8x - 1) dx$ . Using the formula  $\int x^n dx = \frac{x^{n+1}}{n+1}$ , we find  $\langle \psi_1, \psi_2 \rangle = \left[ \frac{12}{4}x^4 - \frac{18}{3}x^3 + \frac{8}{2}x^2 - x \right]_0^1 = 0$ . The functions  $\psi_1(x)$  and  $\psi_2(x)$  are called the *Shifted Legendre polynomials*.

**P9.13** Consider a model of a particle in a one-dimensional box of width one. The state of the particle is described by the wave function  $\psi(x)$ , where  $x \in [0, 1]$ . Find the probability of observing  $x$  in the first quarter of the box ( $x$  between 0 and  $\frac{1}{4}$ ) for the following wave functions: **a)**  $\psi_a(x) = \sqrt{3}(2x - 1)$ , **b)**  $\psi_b(x) = \sqrt{5}(6x^2 - 6x + 1)$ , **c)** a constant wave function  $\psi_c$ .



Hint: The probability of finding the particle somewhere in the interval  $[a, b]$  is computed using the integral  $\Pr(\{a \leq x \leq b\}|\psi) = \int_a^b |\psi(x)|^2 dx$ .

**9.13**  $\Pr(\{0 \leq x \leq \frac{1}{4}\}|\psi_a) = \frac{7}{16}$ ;  $\Pr(\{0 \leq x \leq \frac{1}{4}\}|\psi_b) = \frac{79}{256}$ ;  $\Pr(\{0 \leq x \leq \frac{1}{4}\}|\psi_c) = \frac{1}{4}$ .

**9.13** You can perform the required integrals by hand or use the SymPy command `integrate((sqrt(3)*(2*x-1))**2, (x, 0, 1/4))` for part **a**). For part **b**), use the command `integrate((sqrt(5)*(6*x**2-6*x+1))**2, (x, 0, 1/4))`. For part **c**), the constant wave function is  $\psi_c(x) = 1$ , and the interval  $[0, \frac{1}{4}]$  contains exactly  $\frac{1}{4}$  of its probability mass.

**This was an excerpt from the  
No Bullshit Guide to Linear Algebra  
by Ivan Savov**

**The book is available on [lulu.com](https://lulu.com), the amazons,  
and select bookstores.**

**see table of contents on next pages:**

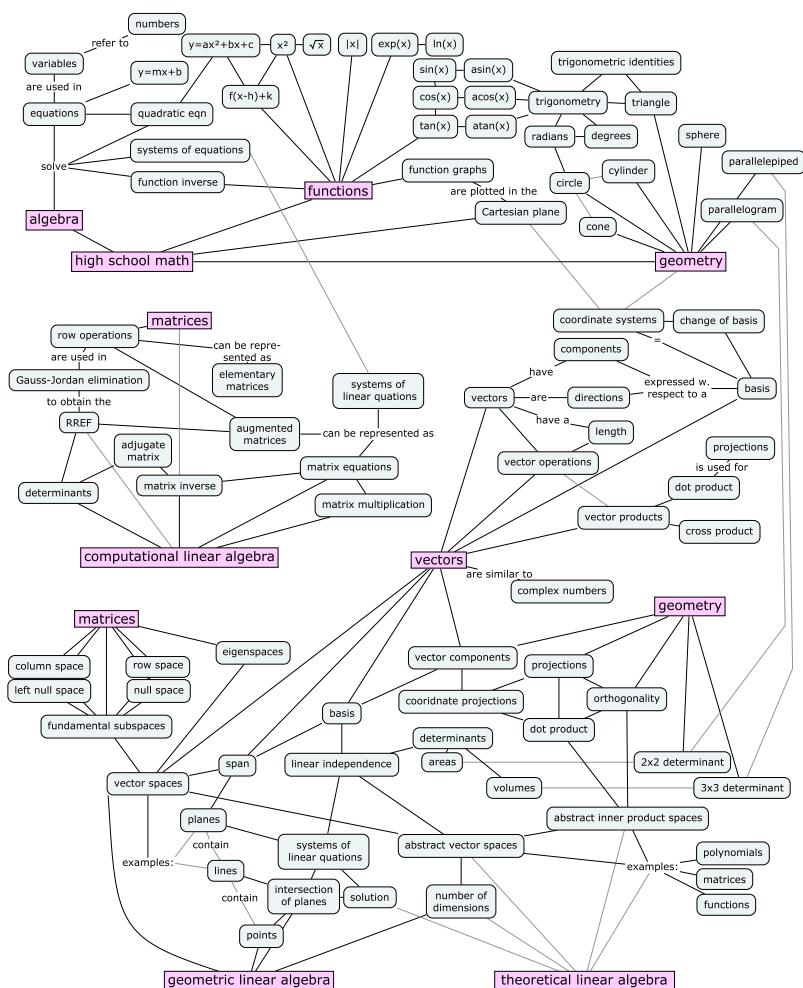
# Contents

<b>Preface</b>	<b>ix</b>
<b>Introduction</b>	<b>1</b>
<b>1 Math fundamentals</b>	<b>9</b>
1.1 Solving equations . . . . .	10
1.2 Numbers . . . . .	12
1.3 Variables . . . . .	18
1.4 Functions and their inverses . . . . .	20
1.5 Basic rules of algebra . . . . .	23
1.6 Solving quadratic equations . . . . .	30
1.7 The Cartesian plane . . . . .	35
1.8 Functions . . . . .	38
1.9 Functions reference . . . . .	49
1.10 Geometry . . . . .	64
1.11 Trigonometry . . . . .	69
1.12 Trigonometric identities . . . . .	76
1.13 Vectors . . . . .	79
1.14 Complex numbers . . . . .	94
1.15 Solving systems of linear equations . . . . .	101
1.16 Set notation . . . . .	107
1.17 Math problems . . . . .	123
<b>2 Intro to linear algebra</b>	<b>133</b>
2.1 Definitions . . . . .	133
2.2 Vector operations . . . . .	140
2.3 Matrix operations . . . . .	148
2.4 Linearity . . . . .	154
2.5 Overview of linear algebra . . . . .	158
2.6 Introductory problems . . . . .	162

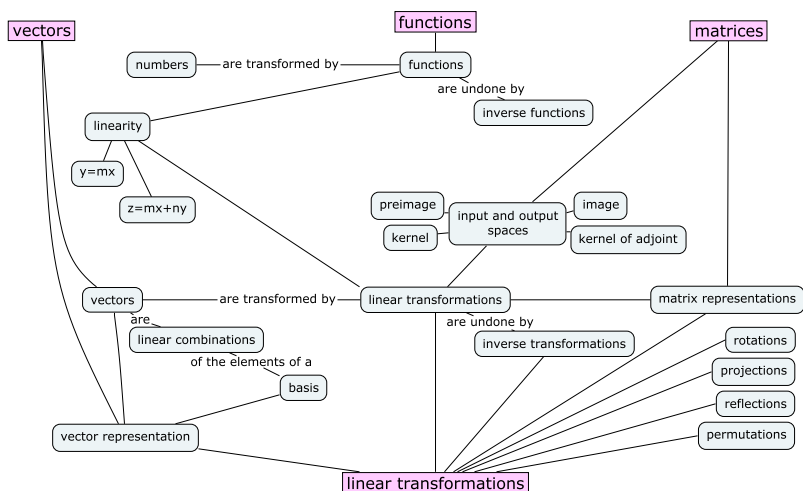
<b>3</b>	<b>Computational linear algebra</b>	<b>167</b>
3.1	Reduced row echelon form . . . . .	168
3.2	Matrix equations . . . . .	181
3.3	Matrix multiplication . . . . .	185
3.4	Determinants . . . . .	190
3.5	Matrix inverse . . . . .	201
3.6	Computational problems . . . . .	209
<b>4</b>	<b>Geometric aspects of linear algebra</b>	<b>219</b>
4.1	Lines and planes . . . . .	219
4.2	Projections . . . . .	230
4.3	Coordinate projections . . . . .	236
4.4	Vector spaces . . . . .	243
4.5	Vector space techniques . . . . .	255
4.6	Geometric problems . . . . .	265
<b>5</b>	<b>Linear transformations</b>	<b>275</b>
5.1	Linear transformations . . . . .	275
5.2	Finding matrix representations . . . . .	287
5.3	Change of basis for matrices . . . . .	300
5.4	Invertible matrix theorem . . . . .	306
5.5	Linear transformations problems . . . . .	313
<b>6</b>	<b>Theoretical linear algebra</b>	<b>319</b>
6.1	Eigenvalues and eigenvectors . . . . .	320
6.2	Special types of matrices . . . . .	335
6.3	Abstract vector spaces . . . . .	342
6.4	Abstract inner product spaces . . . . .	346
6.5	Gram–Schmidt orthogonalization . . . . .	352
6.6	Matrix decompositions . . . . .	356
6.7	Linear algebra with complex numbers . . . . .	363
6.8	Theory problems . . . . .	380
<b>7</b>	<b>Applications</b>	<b>395</b>
7.1	Balancing chemical equations . . . . .	396
7.2	Input–output models in economics . . . . .	398
7.3	Electric circuits . . . . .	400
7.4	Graphs . . . . .	405
7.5	Fibonacci sequence . . . . .	408
7.6	Linear programming . . . . .	411
7.7	Least squares approximate solutions . . . . .	412
7.8	Computer graphics . . . . .	422
7.9	Cryptography . . . . .	435
7.10	Error-correcting codes . . . . .	446
7.11	Fourier analysis . . . . .	455
7.12	Applications problems . . . . .	470

<b>8</b>	<b>Probability theory</b>	<b>479</b>
8.1	Probability distributions . . . . .	479
8.2	Markov chains . . . . .	490
8.3	Google's PageRank algorithm . . . . .	497
8.4	Probability problems . . . . .	504
<b>9</b>	<b>Quantum mechanics</b>	<b>509</b>
9.1	Introduction . . . . .	510
9.2	Polarizing lenses experiment . . . . .	516
9.3	Dirac notation for vectors . . . . .	523
9.4	Quantum information processing . . . . .	531
9.5	Postulates of quantum mechanics . . . . .	534
9.6	Polarizing lenses experiment revisited . . . . .	551
9.7	Quantum physics is not that weird . . . . .	555
9.8	Quantum mechanics applications . . . . .	560
9.9	Quantum mechanics problems . . . . .	576
	<b>End matter</b>	<b>583</b>
	Conclusion . . . . .	583
	Social stuff . . . . .	585
	Acknowledgements . . . . .	585
	General linear algebra links . . . . .	586
<b>A</b>	<b>Answers and solutions</b>	<b>587</b>
<b>B</b>	<b>Notation</b>	<b>619</b>
	Math notation . . . . .	619
	Set notation . . . . .	620
	Vectors notation . . . . .	620
	Complex numbers notation . . . . .	621
	Vector space notation . . . . .	621
	Abstract vector spaces notation . . . . .	622
	Notation for matrices and matrix operations . . . . .	622
	Notation for linear transformations . . . . .	623
	Matrix decompositions . . . . .	623
	<b>Index</b>	<b>625</b>

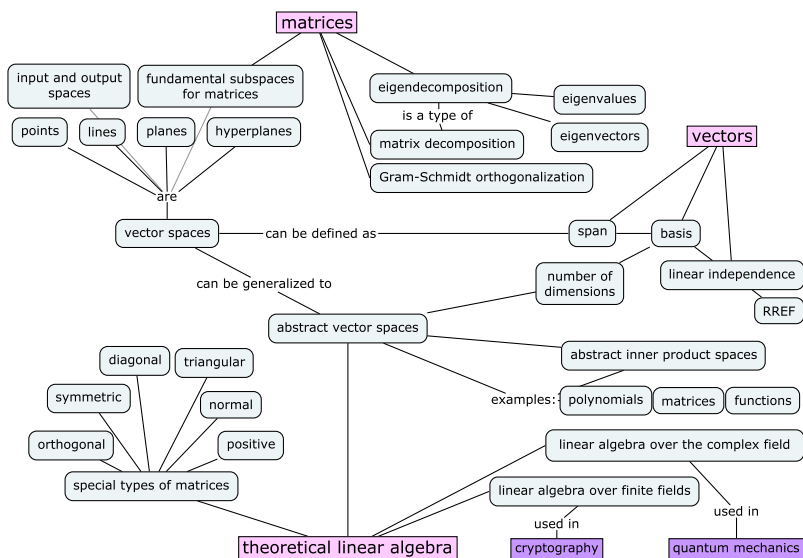
# Concept maps



**Figure 1:** This concept map shows all the topics and concepts covered in this book and illustrates the connections between them. This book is all about linking the concepts together. Since it's a lot of stuff, we'll start slowly by reviewing prerequisite topics of high school math (Chapter 1), and gradually build your knowledge from there.



**Figure 2:** Chapter 5 is about linear transformations and their properties.



**Figure 3:** Chapter 6 covers theoretical aspects of linear algebra.

