Lecture 1

Systems and Experiments

Lenny and Art wander into Hilbert's Place.

Art: *What is this, the Twilight Zone? Or some kind of fun house? I can't get my bearings.*

Lenny: *Take a breath. You'll get used to it.*

Art: *Which way is up?*

1.1 Quantum Mechanics Is Different

What is so special about quantum mechanics? Why is it so hard to understand? It would be easy to blame the "hard mathematics," and there may be some truth in that idea. But that can't be the whole story. Lots of nonphysicists are able to master classical mechanics and field theory, which also require hard mathematics.

Quantum mechanics deals with the behavior of objects so small that we humans are ill equipped to visualize them at all. Individual atoms are near the upper end of this scale in terms of size. Electrons are frequently used as objects of study. *Our sensory organs are simply not built to perceive the motion of an electron. The best we can do is to try to understand electrons and their motion as mathematical abstractions.*

"So what?" says the skeptic. "Classical mechanics is filled to the brim with mathematical abstractions—point masses, rigid bodies, inertial reference frames, positions, momenta, fields, waves—the list goes on and on. There's nothing new about mathematical abstractions." This is actually a fair point, and indeed the classical and quantum worlds have some important things in common. Quantum mechanics, however, is different in two ways:

1. *Different Abstractions*. Quantum abstractions are fundamentally different from classical ones. For example, we'll see that the idea of a state in quantum mechanics is conceptually very different from its classical counterpart. States are represented by different mathematical objects and have a different logical structure.

2. *States and Measurements*. In the classical world, the relationship between the state of a system and the result of a measurement on that

system is very straightforward. In fact, it's trivial. The labels that describe a state (the position and momentum of a particle, for example) are the *same* labels that characterize measurements of that state. To put it another way, one can perform an experiment to determine the state of a system. In the quantum world, this is not true. States and measurements are two different things, and the relationship between them is subtle and nonintuitive.

These ideas are crucial, and we'll come back to them again and again.

1.2 Spins and Qubits

The concept of spin is derived from particle physics. Particles have properties in addition to their location in space. For example, they may or may not have electric charge, or mass. An electron is not the same as a quark or a neutrino. But even a specific type of particle, such as an electron, is not completely specified by its location. Attached to the electron is an extra degree of freedom called its *spin*. Naively, the spin can be pictured as a little arrow that points in some direction, but that naive picture is too classical to accurately represent the real situation. The spin of an electron is about as quantum mechanical as a system can be, and any attempt to visualize it classically will badly miss the point.

We can and will abstract the idea of a spin, and forget that it is attached to an electron. The quantum spin is a system that can be studied in its own right. In fact, the quantum spin, isolated from the electron that carries it through space, is both the simplest and the most quantum of systems.

The isolated quantum spin is an example of the general class of simple systems we call qubits—quantum bits—that play the same role in the quantum world as logical bits play in defining the state of your computer. Many systems—maybe even all systems—can be built up by combining qubits. Thus in learning about them, we are learning about a great deal more.

1.3 An Experiment

Let's make these ideas concrete, using the simplest example we can find. In the first lecture of *Volume I*, we began by discussing a very simple deterministic system: a coin that can show either heads (H) or tails (T). We can call this a two-state system, or a bit, with the two states being H and T. More formally we invent a "degree of freedom" called σ that can take on two values, namely +1 and −1. The state H is replaced by

$$\sigma = +1$$

and the state *T* by

$$\sigma = -1.$$

Classically, that's all there is to the space of states. The system is either in state $\sigma = +1$ or $\sigma = -1$ and there is nothing in between. In quantum mechanics, we'll think of this system as a qubit.

Volume I also discussed simple evolution laws that tell us how to update the state from instant to instant. The simplest law is just that nothing happens. In that case, if we go from one discrete instant (n) to the next ($n + 1$), the law of evolution is

$$\sigma(n+1) \;=\; \sigma(n)\,. \qquad (1.1)$$

Let's expose a hidden assumption that we were careless about in *Volume I*. An experiment involves more than just a system to study. It also involves an apparatus A to make measurements and record the results of the measurements. In the case of the two-state system, the apparatus interacts with the system (the spin) and records the value of σ. Think of the apparatus as a black box[1] with a window that displays the result of a measurement. There is also a "this end up" arrow on the apparatus. The up-arrow is important because it shows how the apparatus is oriented in space, and its direction will affect the outcomes of our measurements. We begin by pointing it along the z axis (Fig. 1.1). Initially, we have no knowledge of whether $\sigma = +1$ or $\sigma = -1$. Our purpose is to do an experiment to find out the value of σ.

Before the apparatus interacts with the spin, the window is blank (labeled with a question mark in our diagrams). After it measures σ, the window shows a +1 or a −1. By looking at the apparatus, we determine the value of σ. That whole process constitutes a very simple experiment designed to measure σ.

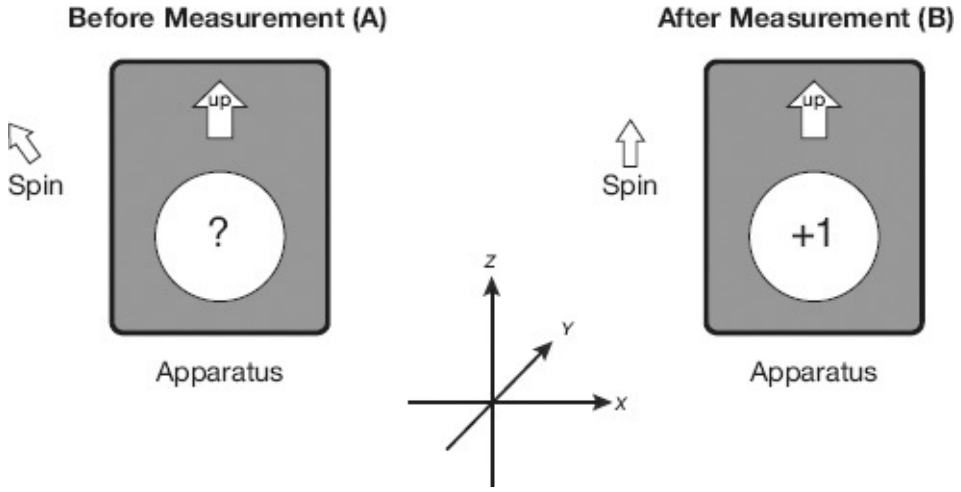

Figure 1.1: (A) Spin and cat-free apparatus before any measurement is made. (B) Spin and apparatus after one measurement has been made, resulting in $\sigma_z = +1$. The spin is now prepared in the $\sigma_z = +1$ state. If the spin is not disturbed and the apparatus keeps the same orientation, all subsequent measurements will give the same result. Coordinate axes show our convention for labeling the directions of space.

Now that we've measured σ, let's reset the apparatus to neutral and, without disturbing the spin, measure σ again. Assuming the simple law of Eq. 1.1, we should get the same answer as we did the first time. The result $\sigma = +1$ will be followed by $\sigma = +1$. Likewise for $\sigma = -1$. The same will be true for any number of repetitions. This is good because it allows us to confirm the result of an experiment. We can also say this in the following way: The first interaction with the apparatus A *prepares* the system in one of the two states. Subsequent experiments *confirm* that state. So far, there is no difference between classical and quantum physics.

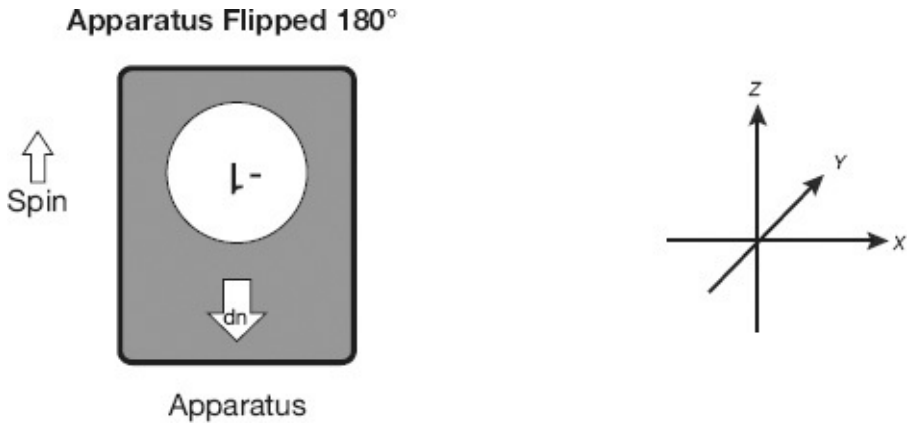

Figure 1.2: The apparatus is flipped without disturbing the previously measured spin. A new measurement results in $\sigma_z = -1$.

Now let's do something new. After preparing the spin by measuring it with A, we turn the apparatus upside down and then measure σ again (Fig.

1.2). What we find is that if we originally prepared $\sigma = +1$, the upside down apparatus records $\sigma = -1$. Similarly, if we originally prepared $\sigma = -1$, the upside down apparatus records $\sigma = +1$. In other words, turning the apparatus over interchanges $\sigma = +1$ and $\sigma = -1$. From these results, we might conclude that σ is a degree of freedom that is associated with a sense of direction in space. For example, if σ were an oriented vector of some sort, then it would be natural to expect that turning the apparatus over would reverse the reading. A simple explanation is that the apparatus measures the component of the vector along an axis embedded in the apparatus. Is this explanation correct for all configurations?

If we are convinced that the spin is a vector, we would naturally describe it by three components: σ_z, σ_x, and σ_y. When the apparatus is upright along the z axis, it is positioned to measure σ_z.

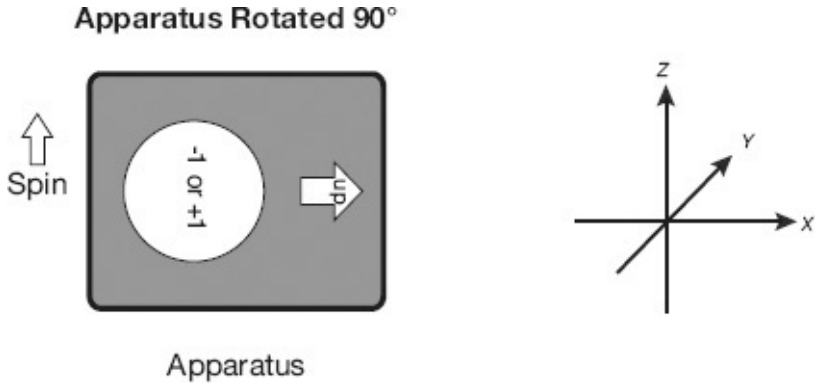

Figure 1.3: The apparatus rotated by 90°. A new measurement results in $\sigma_x = -1$ with 50 percent probability.

So far, there is still no difference between classical physics and quantum physics. The difference only becomes apparent when we rotate the apparatus through an arbitrary angle, say $\frac{\pi}{2}$ radians (90 degrees). The apparatus begins in the upright position (with the up-arrow along the z axis). A spin is prepared with $\sigma = +1$. Next, rotate A so that the up-arrow points along the x axis (Fig. 1.3), and then make a measurement of what is presumably the x component of the spin, σ_x.

If in fact σ really represents the component of a vector along the up-arrow, one would expect to get zero. Why? Initially, we confirmed that σ was directed along the z axis, suggesting that its component along x must be zero. But we get a surprise when we measure σ_x: Instead of giving $\sigma_x = 0$, the apparatus gives either $\sigma_x = +1$ or $\sigma_x = -1$. A is very stubborn—no matter which way it is oriented, it refuses to give any answer other than $\sigma = \pm 1$. If the spin really is a vector, it is a very peculiar one indeed.

Nevertheless, we do find something interesting. Suppose we repeat the operation many times, each time following the same procedure, that is:

- Beginning with $\boldsymbol{A}$ along the z axis, prepare $\boldsymbol{\sigma = +1}$.
- Rotate the apparatus so that it is oriented along the x axis.
- Measure $\boldsymbol{\sigma}$.

The repeated experiment spits out a random series of plus-ones and minus-ones. Determinism has broken down, but in a particular way. If we do many repetitions, we will find that the numbers of $\boldsymbol{\sigma = +1}$ events and $\boldsymbol{\sigma = -1}$ events are statistically equal. In other words, the average value of $\boldsymbol{\sigma}$ is zero. Instead of the classical result—namely, that the component of $\boldsymbol{\sigma}$ along the x axis is zero—we find that the *average of these repeated measurements* is zero.

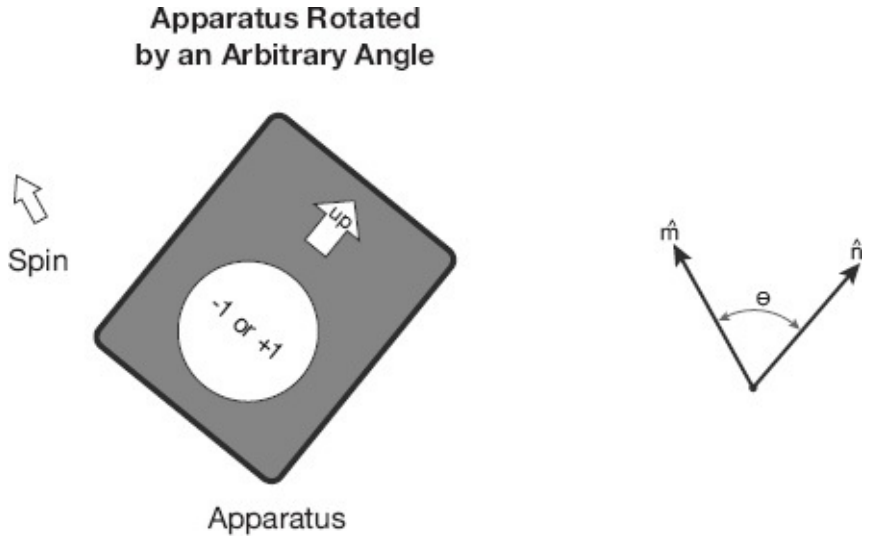

Figure 1.4: The apparatus rotated by an arbitrary angle within the $\boldsymbol{x - z}$ plane. Average measurement result is $\boldsymbol{\hat{n} \cdot \hat{m}}$.

Now let's do the whole thing over again, but instead of rotating $\boldsymbol{A}$ to lie on the x axis, rotate it to an arbitrary direction along the unit vector[2] $\boldsymbol{\hat{n}}$. Classically, if $\boldsymbol{\sigma}$ were a vector, we would expect the result of the experiment to be the component of $\boldsymbol{\sigma}$ along the $\boldsymbol{\hat{n}}$ axis. If $\boldsymbol{\hat{n}}$ lies at an angle $\boldsymbol{\theta}$ with respect to z, the classical answer would be $\boldsymbol{\sigma}$ = cos $\boldsymbol{\theta}$. But as you might guess, each time we do the experiment we get $\boldsymbol{\sigma = +1}$ or $\boldsymbol{\sigma = -1}$. However, the result is statistically biased so that the average value is $\boldsymbol{\cos\theta}$.

The situation is of course more general. We did not have to start with $\boldsymbol{A}$ oriented along z. Pick any direction $\boldsymbol{\hat{m}}$ and start with the up-arrow pointing along $\boldsymbol{\hat{m}}$. Prepare a spin so that the apparatus reads +1. Then, without disturbing the spin, rotate the apparatus to the direction $\boldsymbol{\hat{n}}$, as shown in Fig. 1.4. A new experiment on the same spin will give random results ±1, but with

an average value equal to the cosine of the angle between $\hat{n}$ and $\hat{m}$. In other words, the average will be $\hat{n} \cdot \hat{m}$.

The quantum mechanical notation for the statistical average of a quantity Q is Dirac's bracket notation $\langle Q \rangle$. We may summarize the results of our experimental investigation as follows: If we begin with $\boldsymbol{A}$ oriented along $\hat{m}$ and confirm that $\sigma = +1$, then subsequent measurement with $\boldsymbol{A}$ oriented along $\hat{n}$ gives the statistical result

$$\langle \sigma \rangle = \hat{n} \cdot \hat{m}.$$

What we are learning is that quantum mechanical systems are not deterministic—the results of experiments can be statistically random—but if we repeat an experiment many times, average quantities can follow the expectations of classical physics, at least up to a point.

1.4 Experiments Are Never Gentle

Every experiment involves an outside system—an apparatus—that must interact with the system in order to record a result. In that sense, every experiment is invasive. This is true in both classical and quantum physics, but only quantum physics makes a big deal out of it. Why is that so? Classically, an ideal measuring apparatus has a vanishingly small effect on the system it is measuring. Classical experiments can be arbitrarily gentle and still accurately and reproducibly record the results of the experiment. For example, the direction of an arrow can be determined by reflecting light off the arrow and focusing it to form an image. While it is true that the light must have a small enough wavelength to form an image, there is nothing in classical physics that prevents the image from being made with arbitrarily weak light. In other words, the light can have an arbitrarily small energy content.

In quantum mechanics, the situation is fundamentally different. *Any* interaction that is strong enough to measure some aspect of a system is necessarily strong enough to disrupt some other aspect of the same system. Thus, you can learn nothing about a quantum system without changing something else.

This should be evident in the examples involving $\boldsymbol{A}$ and $\boldsymbol{\sigma}$. Suppose we begin with $\sigma = +1$ along the z axis. If we measure $\boldsymbol{\sigma}$ again with $\boldsymbol{A}$ oriented along z, we will confirm the previous value. We can do this over and over without changing the result. But consider this possibility: Between successive

measurements along the z axis, we turn $\boldsymbol{A}$ through 90 degrees, make an intermediate measurement, and turn it back to its original direction. Will a subsequent measurement along the z axis confirm the original measurement? The answer is no. The intermediate measurement along the x axis will leave the spin in a completely random configuration as far as the next measurement is concerned. There is no way to make the intermediate determination of the spin without completely disrupting the final measurement. One might say that measuring one component of the spin destroys the information about another component. In fact, one simply cannot simultaneously know the components of the spin along two different axes, not in a reproducible way in any case. There is something fundamentally different about the state of a quantum system and the state of a classical system.

1.5 Propositions

The space of states of a classical system is a mathematical set. If the system is a coin, the space of states is a set of two elements, H and T. Using set notation, we would write $\{H, T\}$. If the system is a six-sided die, the space of states has six elements labeled $\{1, 2, 3, 4, 5, 6\}$. The logic of set theory is called *Boolean* logic. Boolean logic is just a formalized version of the familiar classical logic of propositions.

A fundamental idea in Boolean logic is the notion of a truth-value. The truth-value of a proposition is either *true* or *false*. Nothing in between is allowed. The related set theory concept is a subset. Roughly speaking, a proposition is true for all the elements in its corresponding subset and false for all the elements not in this subset. For example, if the set represents the possible states of a die, one can consider the proposition

A: The die shows an odd-numbered face.

The corresponding subset contains the three elements $\{1, 3, 5\}$.

Another proposition states

B: The die shows a number less than 4.

The corresponding subset contains the states $\{1, 2, 3\}$.

Every proposition has its opposite (also called its negation).

For example,

not *A: The die does not show an odd-numbered face.*

The subset for this negated proposition is $\{2, 4, 6\}$.

There are rules for combining propositions into more complex propositions, the most important being **or, and,** and **not.** We just saw an example of **not,** which gets applied to a single subset or proposition. **And** is straightforward, and applies to a pair of propositions.[3] It says they are both true. Applied to two subsets, **and** gives the elements common to both, that is, the *intersection* of the two subsets. In the die example, the intersection of subsets *A* and *B* is the subset of elements that are both odd and less than 4. Fig. 1.5 uses a Venn diagram to show how this works.

The **or** rule is similar to **and**, but has one additional subtlety. In everyday speech, the word *or* is generally used in the exclusive sense—the exclusive version is true if one or the other of two propositions is true, but not both. However, Boolean logic uses the *inclusive* version of **or**, which is true if either or both of the propositions are true. Thus, according to the inclusive **or**, the proposition

Albert Einstein discovered relativity ***or*** *Isaac Newton was English*

is true. So is

Albert Einstein discovered relativity ***or*** *Isaac Newton was Russian*.

The inclusive **or** is only wrong if both propositions are false. For example,

Albert Einstein discovered America[4] ***or*** *Isaac Newton was Russian*.

The inclusive **or** has a set theoretic interpretation as the union of two sets: it denotes the subset containing anything in either or both of the component subsets. In the die example, (*A* **or** *B*) denotes the subset {1, 2, 3, 5}.

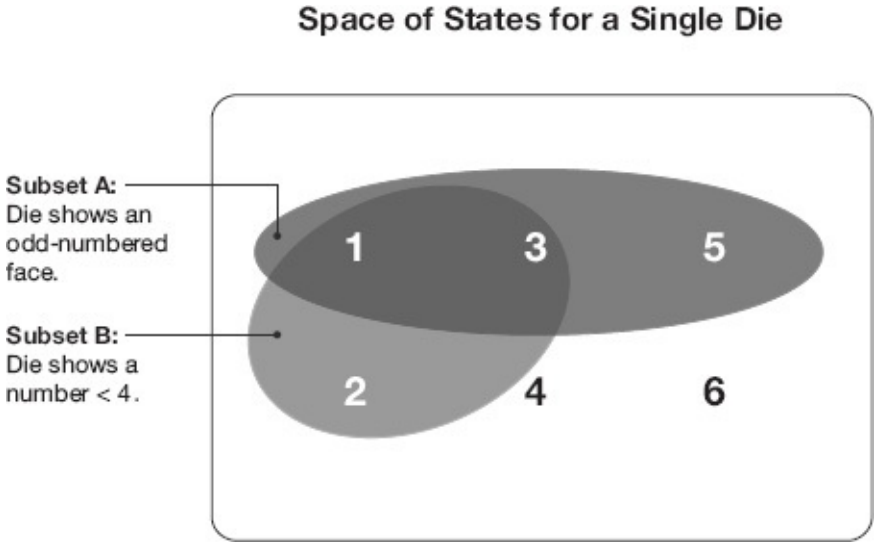

Figure 1.5: An Example of the Classical model of State Space. Subset *A* represents the proposition "the die shows an odd-numbered face." Subset *B*:

"The die shows a number < 4." Dark shading shows the intersection of *A* and *B*, which represents the proposition (*A* **and** *B*). White numbers are elements of the union of *A* with *B*, representing the proposition (*A* **or** *B*).

1.6 Testing Classical Propositions

Let's return to the simple quantum system consisting of a single spin, and the various propositions whose truth we could test using the apparatus ***A***. Consider the following two propositions:

A: The *z* component of the spin is +1.

B: The *x* component of the spin is +1.

Each of these is meaningful and can be tested by orienting ***A*** along the appropriate axis. The negation of each is also meaningful. For example, the negation of the first proposition is

not A: The *z* component of the spin is −1.

But now consider the composite propositions

(*A* **or** *B*): The *z* component of the spin is +1 **or** the *x* component of the spin is +1.

(*A* **and** *B*): The *z* component of the spin is +1 **and** the *x* component of the spin is +1.

Consider how we would test the proposition (*A* **or** *B*). If spins behaved classically (and of course they don't), we would proceed as follows:[5]

- Gently measure $\boldsymbol{\sigma_z}$ and record the value. If it is +1, we are finished: the proposition (*A* **or** *B*) is true. If $\boldsymbol{\sigma_z}$ is −1, continue to the next step.

- Gently measure $\boldsymbol{\sigma_x}$. If it is +1, then the proposition (*A* **or** *B*) is true. If not, this means that neither σ_z nor σ_x was equal to +1, and (*A* **or** *B*) is false.

There is an alternative procedure, which is to interchange the order of the two measurements. To emphasize this reversal of ordering, we'll call the new procedure (*B* **or** *A*):

- Gently measure $\boldsymbol{\sigma_x}$ and record the value. If it is +1 we are finished: The proposition (*B* **or** *A*) is true. If $\boldsymbol{\sigma_x}$ is −1 continue to the next step.

- Gently measure $\boldsymbol{\sigma_z}$. If it is +1, then (*B* **or** *A*) is true. If not, it means that neither σ_x nor $\boldsymbol{\sigma_z}$ was equal to +1, and (*B* **or** *A*) is false.

In classical physics, the two orders of operation give the same answer. The reason for this is that measurements can be arbitrarily gentle—so gentle that they do not affect the results of subsequent measurements. Therefore, the proposition (*A* **or** *B*) has the same meaning as the proposition (*B* **or** *A*).

1.7 Testing Quantum Propositions

Now we come to the quantum world that I described earlier. Let us imagine a situation in which someone (or something) unknown to us has secretly prepared a spin in the $\sigma_z = +1$ state. Our job is to use the apparatus A to determine whether the proposition (*A* **or** *B*) is true or false. We will try using the procedures outlined above.

We begin by measuring σ_z. Since the unknown agent has set things up, we will discover that $\sigma_z = +1$. It is unnecessary to go on: (*A* **or** *B*) is true. Nevertheless, we could test σ_x just to see what happens. The answer is unpredictable. We randomly find that $\sigma_x = +1$ or $\sigma_x = -1$. But neither of these outcomes affects the truth of proposition (*A* **or** *B*).

But now let's reverse the order of measurement. As before, we'll call the reversed procedure (*B* **or** *A*), and this time we'll measure σ_x first. Because the unknown agent set the spin to +1 along the *z* axis, the measurement of σ_x is random. If it turns out that $\sigma_x = +1$, we are finished: (*B* **or** *A*) is true. But suppose we find the opposite result, $\sigma_x = -1$. The spin is oriented along the −*x* direction. Let's pause here briefly, to make sure we understand what just happened. As a result of our first measurement, the spin is no longer in its original state $\sigma_z = +1$. It is in a new state, which is either $\sigma_x = +1$ or $\sigma_x = -1$. Please take a moment to let this idea sink in. We cannot overstate its importance.

Now we're ready to test the second half of proposition (*B* **or** *A*). Rotate the apparatus A to the *z* axis and measure σ_z. According to quantum mechanics, the result will be randomly ±1. This means that there is a 25 percent probability that the experiment produces $\sigma_x = -1$ and $\sigma_z = -1$. In other words, with a probability of $\frac{1}{4}$, we find that (*B* **or** *A*) is false; this occurs despite the fact that the hidden agent had originally made sure that $\sigma_z = +1$.

Evidently, in this example, the inclusive **or** is not symmetric. The truth of (*A* **or** *B*) may depend on the order in which we confirm the two propositions. This is not a small thing; it means not only that the laws of quantum physics are different from their classical counterparts, but that the very foundations of logic are different in quantum physics as well.

What about (*A* **and** *B*)? Suppose our first measurement yields $\sigma_z = +1$ and the second, $\sigma_x = +1$. This is of course a possible outcome. We would be inclined to say that (*A* **and** *B*) is true. But in science, especially in physics, the truth of a proposition implies that the proposition can be verified by subsequent observation. In classical physics, the gentleness of observations implies that subsequent experiments are unaffected and will confirm an earlier experiment. A coin that turns up Heads will not be flipped to Tails by the act of observing it—at least not classically. Quantum mechanically, the second measurement $(\sigma_x = +1)$ ruins the possibility of verifying the first. Once σ_x has been prepared along the *x* axis, another mesurement of σ_z will give a random answer. Thus (*A* **and** *B*) is not confirmable: the second piece of the experiment interferes with the possibility of confirming the first piece.

If you know a bit about quantum mechanics, you probably recognize that we are talking about the uncertainty principle. The uncertainty principle doesn't apply only to position and momentum (or velocity); it applies to many pairs of measurable quantities. In the case of the spin, it applies to propositions involving two different components of σ. In the case of position and momentum, the two propositions we might consider are:

A certain particle has position *x*.

That same particle has momentum *p*.

From these, we can form the two composite propositions

The particle has position *x* **and** the particle has momentum *p*.

The particle has position *x* **or** the particle has momentum *p*.

Awkward as they are, both of these propositions have meaning in the English language, and in classical physics as well. However, in quantum physics, the first of these propositions is completely meaningless (not even wrong), and the second one means something quite different from what you might think. It all comes down to a deep logical difference between the classical and quantum concepts of the state of a system. Explaining the quantum concept of state will require some abstract mathematics, so let's pause for a brief interlude on complex numbers and vector spaces. The need for complex quantities will become clear later on, when we study the mathematical representation of spin states.

1.8 Mathematical Interlude: Complex Numbers

Everyone who has gotten this far in the Theoretical Minimum series knows

about complex numbers. Nevertheless, I will spend a few lines reminding you of the essentials. Fig. 1.6 shows some of their basic elements.

A complex number z is the sum of a real number and an imaginary number. We can write it as

$$z = x + iy,$$

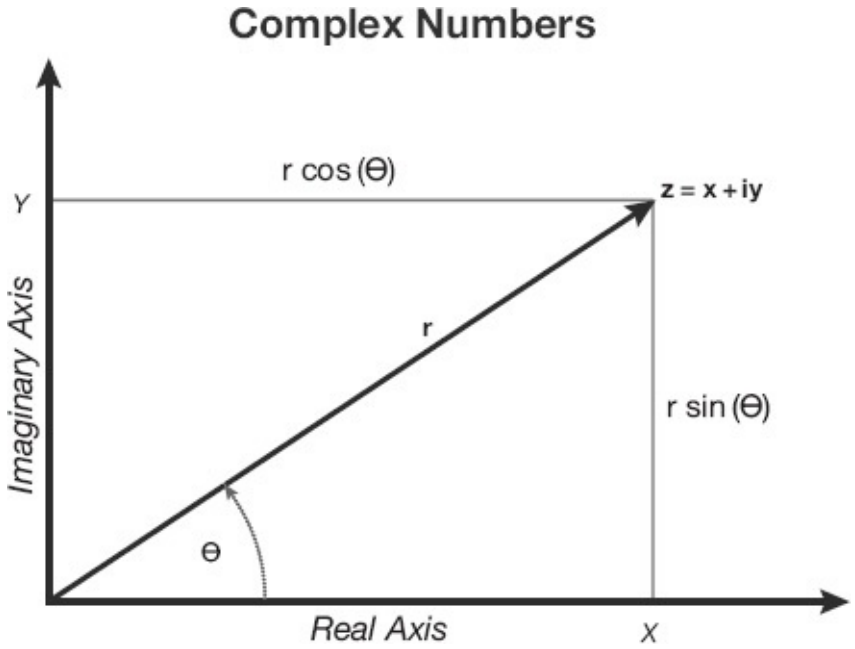

Figure 1.6: Two Common Ways to Represent Complex Numbers. In the Cartesian representation, x and y are the horizontal (real) and vertical (imaginary) components. In the polar representation, r is the radius, and θ is the angle made with the x axis. In each case, it takes two real numbers to represent a single complex number.

where x and y are real and $i^2 = -1$. Complex numbers can be added, multiplied, and divided by the standard rules of arithmetic. They can be visualized as points on the complex plane with coordinates x, y. They can also be represented in polar coordinates:

$$z = re^{i\theta} = r\left(\cos\theta + i\sin\theta\right).$$

Adding complex numbers is easy in component form: just add the components. Similarly, multiplying them is easy in their polar form: Simply multiply the radii and add the angles:

$$\left(r_1 e^{i\theta_1}\right)\ \left(r_2 e^{i\theta_2}\right) = \left(r_1 r_2\right) e^{i(\theta_1+\theta_2)}.$$

Every complex number z has a complex conjugate z^* that is obtained by simply reversing the sign of the imaginary part.

If

$$z = x + iy = re^{i\theta},$$

then

$$z^* = x - iy = re^{-i\theta}.$$

Multiplying a complex number and its conjugate always gives a positive real result:

$$z^* z = r^2.$$

It is of course true that every complex conjugate is itself a complex number, but it's often helpful to think of z and z^* as belonging to separate "dual" number systems. *Dual* here means that for every z there is a unique z^* and vice versa.

There is a special class of complex numbers that I'll call "phase-factors." A phase-factor is simply a complex number whose r-component is 1. If z is a phase-factor, then the following hold:

$$z^* z = 1$$
$$z = e^{i\theta}$$
$$z = \cos\theta + i\sin\theta.$$

1.9 Mathematical Interlude: Vector Spaces

1.9.1 Axioms

For a classical system, the space of states is a set (the set of possible states),

and the logic of classical physics is Boolean. That seems obvious and it is difficult to imagine any other possibility. Nevertheless, the real world operates along entirely different lines, at least whenever quantum mechanics is important. The space of states of a quantum system is *not* a mathematical set;[6] it is a *vector space*. Relations between the elements of a vector space are different from those between the elements of a set, and the logic of propositions is different as well.

Before I tell you about vector spaces, I need to clarify the term *vector*. As you know, we use this term to indicate an object in ordinary space that has a magnitude and a direction. Such vectors have three components, corresponding to the three dimensions of space. I want you to completely forget about that concept of a vector. From now on, whenever I want to talk about a thing with magnitude and direction in ordinary space, I will explicitly call it a *3-vector*. A mathematical vector space is an abstract construction that may or may not have anything to do with ordinary space. It may have any number of dimensions from 1 to ∞ and it may have components that are integers, real numbers, or even more general things.

The vector spaces we use to define quantum mechanical states are called *Hilbert spaces*. We won't give the mathematical definition here, but you may as well add this term to your vocabulary. When you come across the term *Hilbert space* in quantum mechanics, it refers to the space of states. A Hilbert space may have either a finite or an infinite number of dimensions.

In quantum mechanics, a vector space is composed of elements $|A\rangle$ called *ket-vectors* or just *kets*. Here are the axioms we will use to define the vector space of states of a quantum system (z and w are complex numbers):

1. The sum of any two ket-vectors is also a ket-vector:

$$|A\rangle + |B\rangle = |C\rangle\,.$$

2. Vector addition is commutative:

$$|A\rangle + |B\rangle = |B\rangle + |A\rangle\,.$$

3. Vector addition is associative:

$$\{|A\rangle + |B\rangle\} + |C\rangle = |A\rangle + \{|B\rangle + |C\rangle\}\,.$$

4. There is a unique vector 0 such that when you add it to any ket, it gives the same ket back:

$$|A\rangle + 0 = |A\rangle \,.$$

5. Given any ket $|A\rangle$, there is a unique ket $-|A\rangle$ such that

$$|A\rangle + (-|A\rangle) = 0.$$

6. Given any ket $|A\rangle$ and any complex number z, you can multiply them to get a new ket. Also, multiplication by a scalar is linear:

$$|zA\rangle = z\,|A\rangle = |B\rangle \,.$$

7. The distributive property holds:

$$z\{|A\rangle + |B\rangle\} = z\,|A\rangle + z\,|B\rangle$$
$$\{z + w\}\,|A\rangle = z\,|A\rangle + w\,|A\rangle \,.$$

Axioms 6 and 7 taken together are often called *linearity*.

Ordinary 3-vectors would satisfy these axioms except for one thing: Axiom 6 allows a vector to be multiplied by any complex number. Ordinary 3-vectors can be multiplied by real numbers (positive, negative, or zero) but multiplication by complex numbers is not defined. One can think of 3-vectors as forming a real vector space, and kets as forming a complex vector space. Our definition of ket-vectors is fairly abstract. As we will see, there are various concrete ways to represent ket-vectors as well.

1.9.2 Functions and Column Vectors

Let's look at some concrete examples of complex vector spaces. First of all, consider the set of continuous complex-valued functions of a variable x. Call the functions $A(x)$. You can add any two such functions and multiply them by

complex numbers. You can check that they satisfy all seven axioms. This example should make it obvious that we are talking about something much more general than three-dimensional arrows.

Two-dimensional column vectors provide another concrete example. We construct them by stacking up a pair of complex numbers, α_1 and α_2, in the form

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

and identifying this "stack" with the ket-vector $|A\rangle$. The complex numbers α are the components of $|A\rangle$. You can add two column vectors by adding their components:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \end{pmatrix}.$$

Moreover, you can multiply the column vector by a complex number z just by multiplying the components,

$$z \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} z\alpha_1 \\ z\alpha_2 \end{pmatrix}.$$

Column vector spaces of any number of dimensions can be constructed. For example, here is a five-dimensional column vector:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{pmatrix}.$$

Normally, we do not mix vectors of different dimensionality.

1.9.3 Bras and Kets

As we have seen, the complex numbers have a dual version: in the form of complex conjugate numbers. In the same way, a complex vector space has a dual version that is essentially the complex conjugate vector space. For every ket-vector $|A\rangle$, there is a "bra" vector in the dual space, denoted by $\langle A|$. Why the strange terms *bra* and *ket?* Shortly, we will define inner products of bras and kets, using expressions like $\langle B|A\rangle$ to form *bra-kets* or brackets. Inner products are extremely important in the mathematical machinery of quantum mechanics, and for characterizing vector spaces in general.

Bra vectors satisfy the same axioms as the ket-vectors, but there are two things to keep in mind about the correspondence between kets and bras:

1. Suppose $\langle A|$ is the bra corresponding to the ket $|A\rangle$, and $\langle B|$ is the bra corresponding to the ket $|B\rangle$. Then the bra corresponding to

$$|A\rangle + |B\rangle$$

is

$$\langle A| + \langle B|.$$

2. If z is a complex number, then it is *not* true that the bra corresponding to $z|A\rangle$ is $\langle A|z$. You have to remember to complex-conjugate. Thus, the bra corresponding to

$$z|A\rangle$$

is

$$\langle A|\, z^* .$$

In the concrete example where kets are represented by column vectors, the dual bras are represented by row vectors, with the entries being drawn from the complex conjugate numbers. Thus, if the ket $|A\rangle$ is represented by the column

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{pmatrix},$$

then the corresponding bra $\langle A|$ is represented by the row

$$\begin{pmatrix} \alpha_1^* & \alpha_2^* & \alpha_3^* & \alpha_4^* & \alpha_5^* \end{pmatrix}.$$

1.9.4 Inner Products

You are no doubt familiar with the dot product defined for ordinary 3-vectors. The analogous operation for bras and kets is the *inner product*. The inner product is always the product of a bra and a ket and it is written this way:

$$\langle B|A\rangle .$$

The result of this operation is a complex number. The axioms for inner products are not too hard to guess:

1. They are linear:

$$\langle C|\,\{\ |A\rangle + |B\rangle\ \} = \langle C|A\rangle + \langle C|B\rangle .$$

2. Interchanging bras and kets corresponds to complex conjugation:

$$\langle B|A\rangle = \langle A|B\rangle^* .$$

Exercise 1.1:

a) Using the axioms for inner products, prove

$$\{\langle \mathrm{A}| + \langle \mathrm{B}|\}\,|\mathrm{C}\rangle = \langle \mathrm{A}|\mathrm{C}\rangle + \langle \mathrm{B}|\mathrm{C}\rangle\,.$$

b) Prove $\langle \mathbf{A}|\mathbf{A}\rangle$ is a real number.

In the concrete representation of bras and kets by row and column vectors, the inner product is defined in terms of components:

$$\begin{aligned}\langle B|A\rangle &= \begin{pmatrix}\beta_1^* & \beta_2^* & \beta_3^* & \beta_4^* & \beta_5^*\end{pmatrix} \begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\\ \alpha_5\end{pmatrix} \\ &= \beta_1^*\alpha_1 + \beta_2^*\alpha_2 + \beta_3^*\alpha_3 + \beta_4^*\alpha_4 + \beta_5^*\alpha_5\end{aligned}$$

The rule for inner products is essentially the same as for dot products: add the products of corresponding components of the vectors whose inner product is being calculated.

Exercise 1.2: Show that the inner product defined by Eq. 1.2 satisfies all the axioms of inner products.

Using the inner product, we can define some concepts that are familiar from ordinary 3-vectors:

• *Normalized Vector*: A vector is said to be normalized if its inner product with itself is 1. Normalized vectors satisfy,

$$\langle A|A\rangle = 1.$$

For ordinary 3-vectors, the term *normalized vector* is usually replaced by *unit vector,* that is, a vector of unit length.

• *Orthogonal Vectors*: Two vectors are said to be orthogonal if their inner product is zero. $|A\rangle$ and $|B\rangle$ are orthogonal if

$$\langle B|A\rangle = 0.$$

This is the analog of saying that two 3-vectors are orthogonal if their dot product is zero.

1.9.5 Orthonormal Bases

When working with ordinary 3-vectors, it is extremely useful to introduce a set of three mutually orthogonal unit vectors and use them as a basis to construct any vector. A simple example would be the unit 3-vectors that point along the *x*, *y*, and *z* axes. They are usually called $\hat{\boldsymbol{i}}$, $\hat{\boldsymbol{j}}$, and $\hat{\boldsymbol{k}}$. Each is of unit length and orthogonal to the others. If you tried to find a fourth vector orthogonal to these three, there wouldn't be any—not in three dimensions anyway. However, if there were more dimensions of space, there would be more basis vectors. The dimension of a space can be defined as the maximum number of mutually orthogonal vectors in that space.

Obviously, there is nothing special about the particular axes *x*, *y*, and *z*. As long as the basis vectors are of unit length and are mutually orthogonal, they comprise an *orthonormal basis*.

The same principle is true for complex vector spaces. One can begin with any normalized vector and then look for a second one, orthogonal to the first. If you find one, then the space is at least two-dimensional. Then look for a third, fourth, and so on. Eventually, you may run out of new directions and there will not be any more orthogonal candidates. The maximum number of mutually orthogonal vectors is the dimension of the space. For column vectors, the dimension is simply the number of entries in the column.

Let's consider an *N*-dimensional space and a particular orthonormal basis

of ket-vectors labeled $|i\rangle$.[7] The label i runs from 1 to N. Consider a vector $|A\rangle$, written as a sum of basis vectors:

$$|A\rangle = \sum_i \alpha_i \, |i\rangle \, . \qquad (1.3)$$

The α_i are complex numbers called the *components* of the vector, and to calculate them we take the inner product of both sides with a basis bra $\langle j|$:

$$\langle j|A\rangle = \sum_i \alpha_i \, \langle j|i\rangle \, . \qquad (1.4)$$

Next, we use the fact that the basis vectors are orthonormal. This implies that $\langle j|i\rangle = 0$ if i is not equal to j, and $\langle j|i\rangle = 1$ if $i = j$. In other words, $\langle j|i\rangle = \delta_{ij}$. This makes the sum in Eq. 1.4 collapse to one term:

$$\langle j|A\rangle = \alpha_j . \qquad (1.5)$$

Thus, we see that the components of a vector are just its inner products with the basis vectors. We can rewrite Eq. 1.3 in the elegant form

$$|A\rangle = \sum_i |i\rangle \, \langle i|A\rangle \, .$$

---

[1] "Black box" means we have no knowledge of what's inside the apparatus or how it works. But rest assured, it does not contain a cat.

[2] The standard notation for a unit vector (one of unit length) is to place a "hat" above the symbol representing the vector.

[3] **And** may be defined for multiple propositions, but we'll only consider two. The same goes for **or.**

[4] OK, perhaps Einstein did discover America. But he was not the first.

[5] Recall that the classical meaning of $\boldsymbol{\sigma}$ is different from the quantum mechanical meaning. Classically, $\boldsymbol{\sigma}$ is a straightforward 3-vector; $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_z$ represent its spatial components.

[6]To be a little more precise, we will not focus on the set-theoretic properties of state spaces, even though they may of course be regarded as sets.

[7]Mathematically, basis vectors are not required to be orthonormal. However, in quantum mechanics they generally are. In this book, whenever we say *basis*, we mean an orthonormal basis.