Quantum Mechanics

A Spins-First, Experimental Approach

QuadriviumPress

[object Object]

© 2026 QuadriviumPress.

This book is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License (CC-BY-SA-4.0). You are free to share and adapt it, with attribution, under the same licence:

https://creativecommons.org/licenses/by-sa/4.0/

This book adapts openly licensed content from several sources. The per-chapter attribution ledger is `SOURCES.md`:

https://github.com/QuadriviumPress/quantumMechanics/blob/main/SOURCES.md

The interactive simulations reproduced here as screenshots are published by PhET Interactive Simulations (University of Colorado Boulder) and by the OpenLyceum project under their own licences; each caption links to the running simulation.

Contents

I Quantum states and dynamics **1**

1 Stern–Gerlach Experiments **2**
1.1 Stern–Gerlach experiments 2

2 States and Probability Amplitudes **19**
2.1 States and probability amplitudes 19

3 Operators and Measurement **38**
3.1 Operators and measurement 38

4 Time Evolution **55**
4.1 Time evolution 55

II Composite and continuous systems **73**

5 Composite Systems and Entanglement **74**
5.1 Composite systems and entanglement 74

6 Position and Momentum Representations **93**
6.1 Position and momentum representations 93

7 One-Dimensional Wave Mechanics **111**
7.1 One-dimensional wave mechanics 111

8 Three-Dimensional Wave Mechanics **126**
8.1 Three-dimensional wave mechanics 126

III Structure and applications **139**

9 Angular Momentum **140**
9.1 Angular momentum 140

10 Approximation Methods **153**
10.1 Approximation methods 153

11 Identical Particles **166**
11.1 Identical particles 166

12 Modern Quantum Applications **178**
12.1 Modern quantum applications 178

Part I

Quantum states and dynamics

Chapter 1

Stern–Gerlach Experiments

1.1 Stern–Gerlach experiments

Learning objectives

After this chapter, you should be able to:

- describe preparation, analysis, and detection in a Stern–Gerlach experiment;
- predict qualitative outcomes of sequential analyzers;
- calculate route probabilities and expected counts for analyzers at arbitrary relative angles;
- distinguish an incoherent mixture from a coherent quantum state;
- infer three empirical rules that a quantum model must reproduce;
- explain why spin is not adequately modeled as a tiny classical arrow;
- distinguish ideal probabilities from finite-sample frequencies; and
- design analyzer sequences that test competing physical explanations.

The final two objectives are developed in optional laboratory and inquiry extensions after the chapter's core conceptual sequence.

1.1 An experiment before a formalism

Quantum mechanics is easiest to trust when it grows out of a real experiment, rather than being handed to you as a list of abstract postulates. So this book begins with one of the simplest experiments in physics: the one performed by Otto Stern and Walther Gerlach in 1922. Their question was simple. Does an atom's magnetism point in only a few special directions, or can it point anywhere at all? The answer they found was strange enough to force physicists to rethink what it even means for a particle to have a direction.

Picture the apparatus in three stages. First, an oven heats a sample of silver until atoms boil off as a vapor. Second, that vapor escapes through a narrow opening, and a pair of slits downstream accepts only atoms traveling within a narrow range of directions; what emerges is a **collimated beam**. Collimation does not, by itself, select one exact speed: an oven beam retains a thermal spread of speeds unless a separate velocity selector is added. In the idealized diagrams below we suppress both spreads so that the spin-dependent splitting is easy to see. Third, the beam passes between the shaped poles of a magnet and then strikes a glass screen, where each atom leaves a visible mark.

One feature of the magnet matters more than any other: its poles are shaped so that the field is much stronger near one pole than the other. Without this asymmetry the experiment

would show nothing interesting, as the next few paragraphs will make clear.

To see why the asymmetry matters, picture each atom as a tiny bar magnet, described by a **magnetic moment** vector $\boldsymbol{\mu}$—the same quantity that tells a compass needle which way to twist. Just like a compass needle, the atom's energy depends on how $\boldsymbol{\mu}$ is oriented relative to the external field $\mathbf{B}$: lined up with the field is low energy, opposed to the field is high energy, and every angle in between interpolates smoothly. In symbols,

$$U = -\boldsymbol{\mu} \cdot \mathbf{B}. \tag{1.1}$$

A *uniform* field, however, cannot push the atom anywhere—it can only twist it. The force pulling one "pole" of the tiny magnet forward is exactly canceled by the force pushing the other pole backward, so the atom feels a torque but no net force. A push appears only when the field itself is stronger on one side of the atom than the other, that is, only when the field has a *gradient*. For a gradient along the z direction, the net force on the atom works out to

$$F_z = -\frac{\partial U}{\partial z} \approx \mu_z \frac{\partial B_z}{\partial z}. \tag{1.2}$$

Equation (1.2) is the whole operating principle of the device. It says that each atom is pushed by an amount proportional to μ_z, the component of its magnetic moment along the gradient direction. A larger μ_z means a larger deflection; a reversed μ_z means deflection the other way. In short, the magnet turns an otherwise invisible internal property of the atom—its magnetic orientation—into something we can see directly: a spot's position on the screen.

Now we can ask what classical physics predicts for the result. Nothing in classical physics picks out a preferred direction for $\boldsymbol{\mu}$, so the atoms emerging from the oven should have their magnetic moments pointing every which way. That means $\mu_z = \mu \cos\theta$ should take every value between $-\mu$ and $+\mu$, depending on each atom's orientation angle θ. Combined with equation (1.2), this predicts a continuous smear of deflections on the screen. For fixed-magnitude moments pointing isotropically, $\cos\theta$ is uniformly distributed, so the idealized classical pattern is an approximately uniform band between the two extreme deflections. Thermal speeds, finite slits, and detector resolution can broaden or reshape that band, but they do not turn it into two discrete spots.

That is not what Stern and Gerlach saw. Instead of the predicted continuous band, their deposited silver pattern separated into two traces. In the ideal two-outcome model used throughout this chapter, we represent those traces as two spots labeled $z+$ and $z-$.

What the apparatus measures

A Stern–Gerlach analyzer does not simply reveal the location of a pre-existing little arrow. It couples an internal degree of freedom to the atom's path, and the resulting spatial separation is what makes two possible outcomes visible at the detector.

The two-spot pattern tells us something remarkable: the component of the atom's magnetic moment that couples to the apparatus takes two discrete values rather than a continuum. Historically, Stern and Gerlach observed this space quantization before electron spin had been identified. In the modern description, a ground-state silver atom has closed inner shells and one unpaired outer electron with zero orbital angular momentum, so its relevant two-valued angular momentum is spin. No matter which axis we build the magnet along, an analyzer of this effective spin-$\frac{1}{2}$ system produces only two outcomes—never a value in between. The measured z component can only be

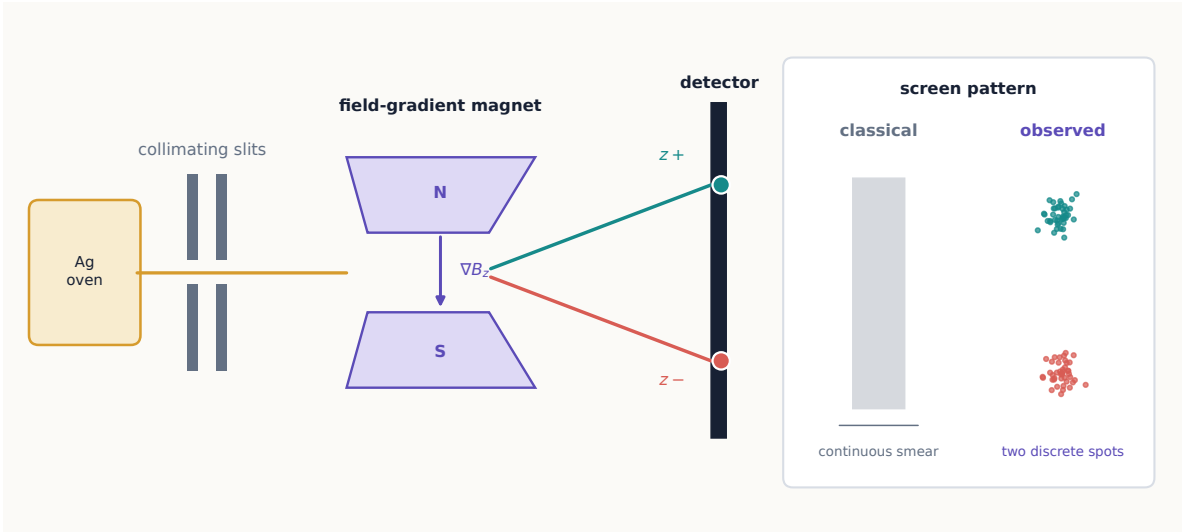

Figure 1.1: The Stern–Gerlach apparatus converts the spin component along the field gradient into a visible deflection. The decisive observation is not merely a deflected beam, but two discrete spots in place of the classical smear.

$$S_z = +\frac{\hbar}{2} \quad \text{or} \quad S_z = -\frac{\hbar}{2}. \tag{1.3}$$

This restriction—only two outcomes, no matter how the analyzer is oriented—is the first genuinely quantum fact in this book. Everything that follows is an attempt to build a mathematical theory around it.

From now on we abbreviate an analyzer oriented along z as SG_z. Something useful happens if we place a physical block in front of one of its two exits. Say we block the $z-$ exit and let only the $z+$ beam through. Every atom that makes it past the block is now guaranteed to have been recorded as $z+$—so the same device that just *measured* spin has become a device that *prepares* it. This dual role, one physical apparatus used first as a measurement and then as a preparation, will come up again and again in this chapter, so it is worth fixing firmly in mind now.

Ideal analyzers from this point onward

The original silver-beam experiment established the two-valued result, but the perfect filters, lossless beam routing, and chained analyzers used below are idealized devices. They isolate the logic of spin measurements. A laboratory implementation must also control the atoms' position, momentum, velocity spread, and accumulated phase.

Reading the experiment from left to right

To talk about these experiments precisely, it helps to separate every one of them into three stages:

1. A **preparation** specifies how the incoming ensemble was made. A source by itself may be unpolarized; a source followed by a filter prepares a selected spin state.

2. A **transformation** changes the state without yet recording a spin result. Examples include a uniform magnetic field or a coherent path recombiner.

3. A **measurement** ends in macroscopically distinguishable records, such as counts at the two exits of an analyzer.

This vocabulary heads off a common confusion. The symbol $z+$ gets used for three different things: an outcome recorded at a detector, the physical exit path leading to that detector, and the *state* that gets prepared when we select that path. Context usually makes clear which meaning is intended, but keep in mind that these are three related, not identical, ideas.

In the laboratory, what we actually observe are raw counts: some atoms land in the $z+$ spot, some in the $z-$ spot. If N_+ and N_- are the numbers of atoms detected at the two exits, the measured relative frequencies are

$$f_+ = \frac{N_+}{N_+ + N_-}, \qquad f_- = \frac{N_-}{N_+ + N_-}. \tag{1.4}$$

Quantum theory predicts the probabilities that these frequencies approach after many repetitions of the same experiment. It does not, in general, predict which exit any one particular atom will take.

1.2 Repeatability

Here is the simplest possible follow-up experiment: put a *second* SG_z analyzer right after the first one, oriented along the same z axis, and keep only the atoms that came out of the first analyzer's $z+$ exit. What happens at the second analyzer?

The answer is unambiguous: every atom that enters the second analyzer leaves through its $z+$ exit. None comes out $z-$.

Rule 1: repeatability
An immediate repetition of the same ideal measurement returns the same result with certainty.

This tells us that the first analyzer did more than sort atoms into two piles. Selecting the $z+$ output prepared something new: a beam in which every atom is now *definite* for a repeated S_z measurement, meaning that measuring it again along z can no longer produce a surprise.

Repeatability also gives us an operational definition of the label $|+z\rangle$, which will be introduced formally in Chapter 2: it means the preparation that produces $z+$ with certainty when immediately analyzed along z. We do not need to imagine what the atom is "really doing" between devices in order to define the state reproducibly.

Example 1.1: counting atoms

A source sends 12,000 atoms into SG_z. The source has no preferred orientation, so half emerge from each exit on average. The $z-$ exit is blocked, and the $z+$ beam enters a second SG_z.

About 6,000 atoms reach the second analyzer, and all 6,000 leave its $z+$ exit. Statistical fluctuations affect the first split, but an ideal repeat measurement does not create a second 50–50 split.

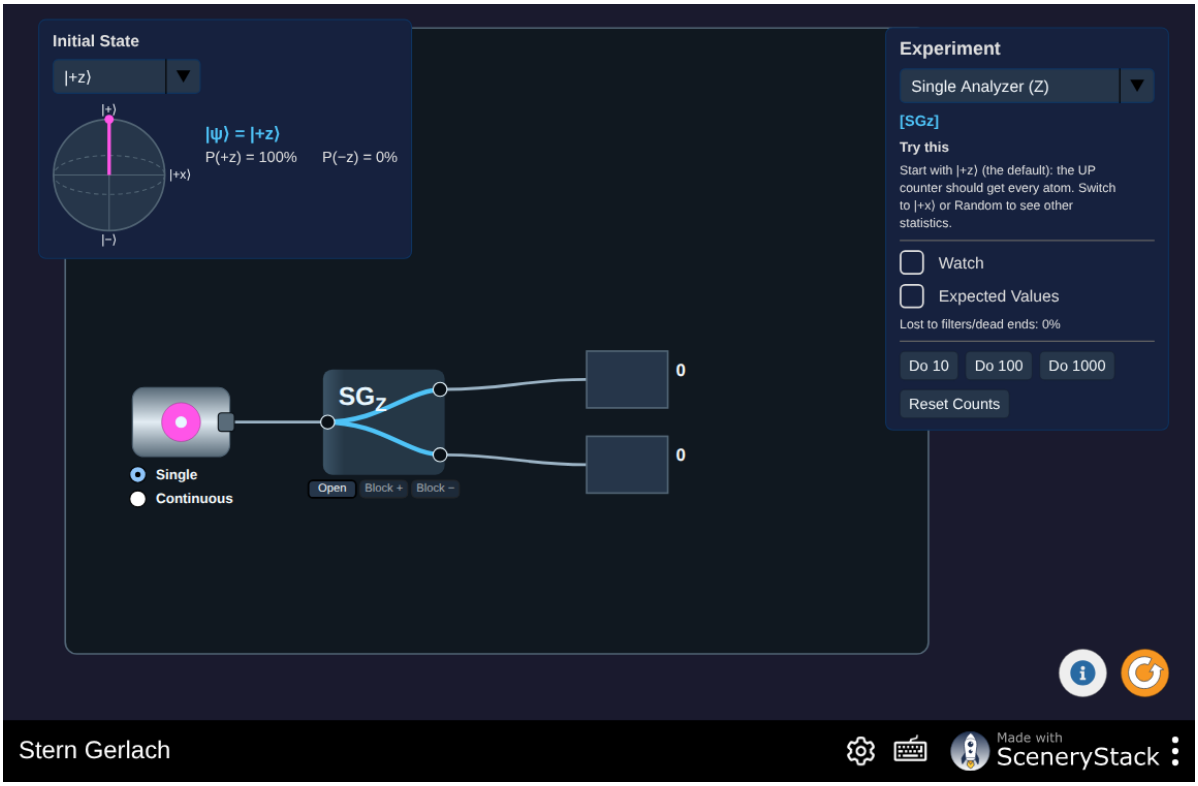

Figure 1.2: Assemble your own chain of ovens, analyzers, and counters, fire spin-$\frac{1}{2}$ or spin-1 atoms through it, and watch Monte Carlo counts accumulate next to the analytic prediction. Rebuild the repeated-analyzer setup from Example 1.1 before reading on, then try chaining analyzers at other angles. Interactive simulation: Stern Gerlach

1.3 Incompatible questions

Repeating a measurement along the *same* axis gave a simple, if important, result: nothing changes. The next natural question is what happens if we measure along a *different* axis instead. Prepare a $z+$ beam as before, but this time send it into SG_x, an analyzer whose field gradient points along x rather than z. The output divides equally:

$$P(x+ \mid z+) = P(x- \mid z+) = \frac{1}{2}. \tag{1.5}$$

So a $z+$ preparation is not definite for S_x: knowing that an atom is $z+$ tells us nothing about which way it will go in SG_x. Here is the more surprising part. Keep only the $x+$ output of that measurement, and send those atoms into a *third* analyzer, this one measuring S_z again. If the atom really still had a definite z value all along, we should recover $z+$ with certainty, exactly as in Section 1.2. Instead, both z outcomes now occur with equal probability, as if the intervening x measurement had erased what we thought we already knew:

$$z+ \longrightarrow x+ \longrightarrow \begin{cases} z+ & \text{with probability } 1/2, \\ z- & \text{with probability } 1/2. \end{cases} \tag{1.6}$$

This is the key point: measuring S_x changed what we could predict about S_z, even though nothing directly measured S_z in between. It is tempting to imagine that the atom secretly carried permanent values for *both* S_z and S_x, and that each analyzer simply reads the relevant value without changing anything else. The experiment rules out that specific, noninvasive **fixed-label model**: if the x analyzer merely revealed an unchanged x label, the final S_z measurement should still return $z+$ with certainty—but it does not. The sequence does not, by itself, rule out every possible hidden-variable account; a more elaborate model could allow the apparatus to alter hidden variables contextually. Bell's test in Chapter 5 will impose additional locality assumptions and reach a stronger conclusion.

Rule 2: incompatibility
Preparing a definite value of one spin component generally produces uncertain outcomes for a different component. An intervening measurement can erase the predictive certainty established by an earlier preparation.

Concept check 1.1

A $z+$ beam enters SG_x. The $x-$ exit is blocked, and the transmitted $x+$ beam enters another SG_x. What fraction of the original $z+$ beam reaches the final $x+$ detector?

Answer

One half. The first x measurement transmits half the atoms into $x+$. Repeating the S_x measurement then returns $x+$ with certainty.

What "disturbance" does and does not explain

One might instead say that the middle analyzer simply adds uncontrolled mechanical noise. Repeatability shows why that description is inadequate: if the device only randomized everything, its own S_x result would not be reliable on repetition. Yet an $x+$ selection is perfectly repeatable. The analyzer performs a reproducible transformation—it replaces a preparation definite

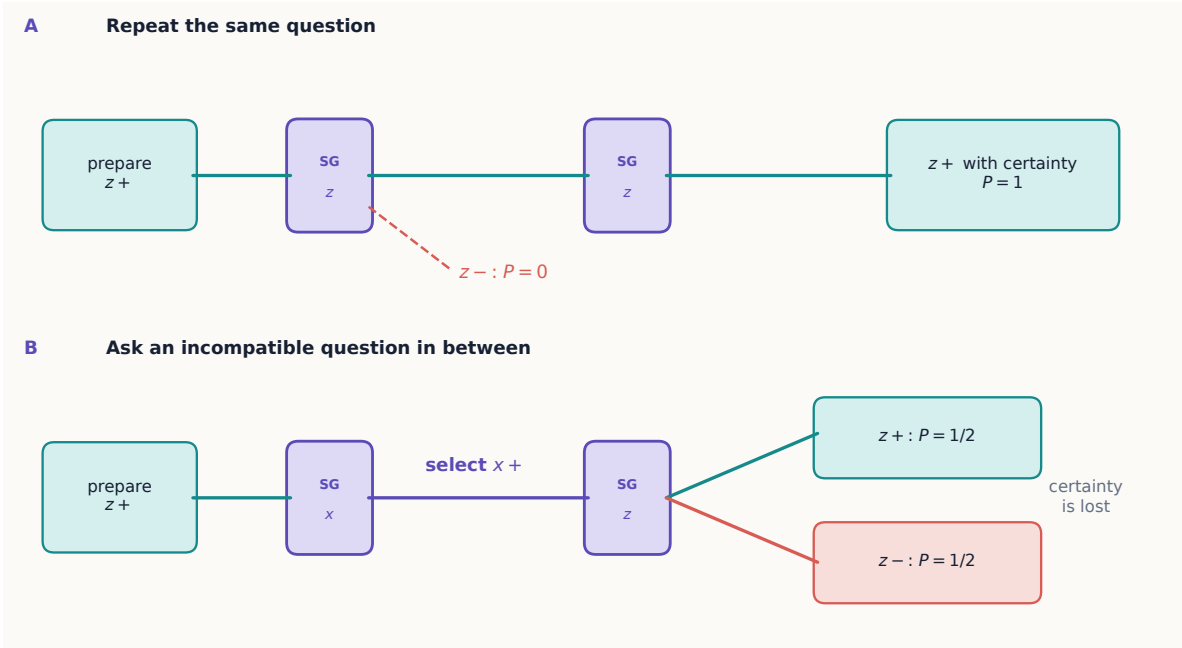

Figure 1.3: Aligned analyzers repeat a definite result, whereas an intervening analyzer along an incompatible axis replaces that certainty with a new preparation. The diagrams are read from left to right, just like the laboratory sequence.

for S_z with one definite for S_x—even though the later S_z result becomes uncertain. This does not prove that no underlying disturbance occurs; it shows that "mere random jostling" lacks the structure needed to explain the observations.

Order matters here in a way that has no classical analog. Compare

$$z+ \longrightarrow z+ \longrightarrow x \tag{1.7}$$

with

$$z+ \longrightarrow x+ \longrightarrow z. \tag{1.8}$$

In the first sequence, repeating the z measurement changes nothing, so the final x analyzer still sees a 50–50 split, exactly as it would from a fresh $z+$ beam. In the second sequence, the intervening $x+$ selection changes the outcome seen by the final z analyzer, turning certainty into a 50–50 split. Chapter 3 will express this order dependence precisely, as a statement about products of noncommuting matrices.

Analyzers at a general angle

So far we have only compared axes at 0° and 90°. What happens at angles in between? Nothing stops us from building an analyzer tilted at any angle relative to the preparation axis, and it seems reasonable to guess that the probability of recovering the original outcome should fall off smoothly as the analyzer tilts away from the preparation axis. If spin behaved like an ordinary arrow, we might guess that this probability varies as $\cos\alpha$, matching the projection of one vector onto another. Experiment gives a different, and at first surprising, answer. For a beam prepared spin-up along $\mathbf{a}$ and analyzed along $\mathbf{b}$, separated by angle α,

$$P(+\mathbf{b} \mid +\mathbf{a}) = \cos^2 \frac{\alpha}{2}, \qquad P(-\mathbf{b} \mid +\mathbf{a}) = \sin^2 \frac{\alpha}{2}. \tag{1.9}$$

Notice the *half*-angle, $\alpha/2$, rather than α itself. This is not a typo or a minor detail—it is an early hint that spin does not transform like an ordinary arrow in space. As Chapter 4 will show, a representative spinor acquires an overall minus sign under a 360° rotation and returns to the identical ket only after 720°. The minus sign does not change the physical state of an isolated spin, because overall phase is unobservable; it becomes detectable only relative to another coherent amplitude. The half-angle in equation (1.9) is an early fingerprint of this spinor transformation law.

It is worth checking that this formula reproduces everything found so far. At $\alpha = 0°$ (aligned analyzers), $\cos^2(0) = 1$: certainty of recovering the same outcome, matching Section 1.2. At $\alpha = 180°$ (oppositely aligned analyzers), $\cos^2(90°) = 0$: certainty of the *opposite* outcome. At $\alpha = 90°$ (perpendicular analyzers, like x versus z), $\cos^2(45°) = 1/2$: an equal split, matching Section 1.3. Equation (1.9) is therefore not a separate new rule; it is the general law that contains every case already found as a special case. For now we treat it as an empirical summary of the data. Chapter 2 will derive it from the mathematics of state vectors.

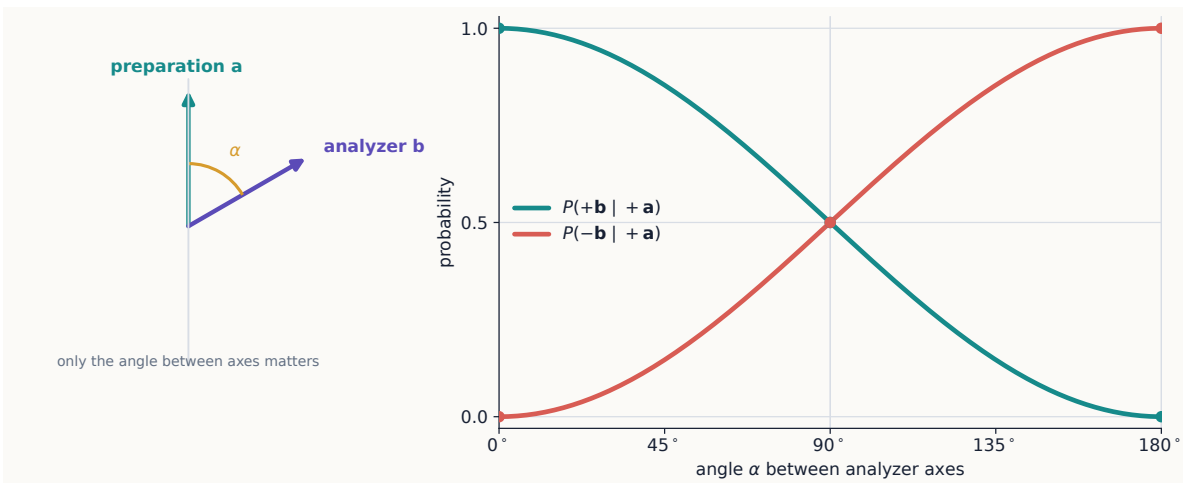

Figure 1.4: The half-angle rule smoothly joins repeatability at 0°, equal splitting at 90°, and certain reversal at 180°. At every angle the two exclusive probabilities add to one.

Example 1.2: three analyzer directions

A $z+$ beam passes through an analyzer whose axis $\mathbf{n}$ is 60° from $+z$, and only the $+\mathbf{n}$ output is retained. That beam is then measured along z. This means applying equation (1.9) twice in a row, using the *selected* output of the first analyzer as the input to the second.

At the middle analyzer, equation (1.9) applies with $\alpha = 60°$, so the half-angle appearing in the formula is $\alpha/2 = 30°$:

$$P(+\mathbf{n} \mid z+) = \cos^2 30° = \left(\frac{\sqrt{3}}{2}\right)^2 = \frac{3}{4}. \tag{1.10}$$

The selected middle output is a new $+\mathbf{n}$ preparation. Since the angle back to $+z$ is again 60°,

$$P(z+ \mid +\mathbf{n}) = \frac{3}{4}, \qquad P(z- \mid +\mathbf{n}) = \frac{1}{4}. \tag{1.11}$$

Thus 9/16 of the original atoms reach the final $z+$ detector and 3/16 reach the final $z-$ detector; the remaining quarter was blocked at the middle analyzer. Conditional probabilities must be multiplied along a selected path.

Concept check 1.2

A $z+$ beam is analyzed along an axis $\mathbf{n}$ only 10° from $+z$. One student predicts exactly $z+$ because the axes are nearly aligned; another predicts a 50–50 split because the axes are different. What does the experiment predict?

Answer

Neither. The probability of $+\mathbf{n}$ is $\cos^2(5°) \approx 0.9924$, while the probability of $-\mathbf{n}$ is about 0.0076. Quantum outcomes are discrete, but their probabilities vary continuously with analyzer orientation.

1.4 Recombining alternatives

Every experiment so far has ended with a measurement: an atom lands in one detector or another, and that is the end of the story. What happens if, instead of recording which path an atom took, we let two paths recombine before anything is measured?

What an ideal recombiner assumes

A textbook Stern–Gerlach recombiner is not merely a second ordinary magnet. It must reverse the path separation while erasing any correlation with position, momentum, or the environment and while preserving a controlled relative phase. Such devices are technically demanding for massive particles. Here the idealization lets us isolate the quantum rule for coherent alternatives.

Imagine a carefully built device that splits a $z+$ beam into two separate paths, one associated with $x+$ and one with $x-$, and then brings the two paths back together—coherently, without ever revealing which path a given atom traveled. Remarkably, the recombined beam turns out to be back in the original $z+$ state, exactly as if the device had never touched it. But if we block *either* path, so that only atoms from the other path get through, the later S_z statistics change completely.

This kills a natural but wrong guess: that the two paths behave like two ordinary, independent sub-beams, so that whatever happens to the recombined beam should just be the sum of what happens along each path separately. Quantum theory instead assigns a complex number—a **probability amplitude**—to each indistinguishable alternative. When alternatives are indistinguishable, we add the amplitudes *first*, and only afterward take the squared magnitude to get a probability.

Rule 3: interference of alternatives

When alternatives are indistinguishable, add their probability amplitudes. When a measurement makes the alternatives distinguishable, add their probabilities.

This distinction will later become the difference between a coherent superposition and an incoherent mixture.

The word **coherently** carries real experimental content. The apparatus must preserve a stable relation between the phases accumulated on the two paths. If vibrations, stray fields, or a path detector randomize that relation, the interference washes out even when nobody reads the detector. What matters is whether a durable physical record could distinguish the alternatives, not whether a person happens to look.

One can summarize the alternatives with two limiting arrangements:

- **Both paths open and recombined:** add the two complex amplitudes, including their relative phase, and then square the magnitude.
- **Paths detected separately or phase-randomized:** square the magnitudes for the alternatives separately and then add the probabilities.

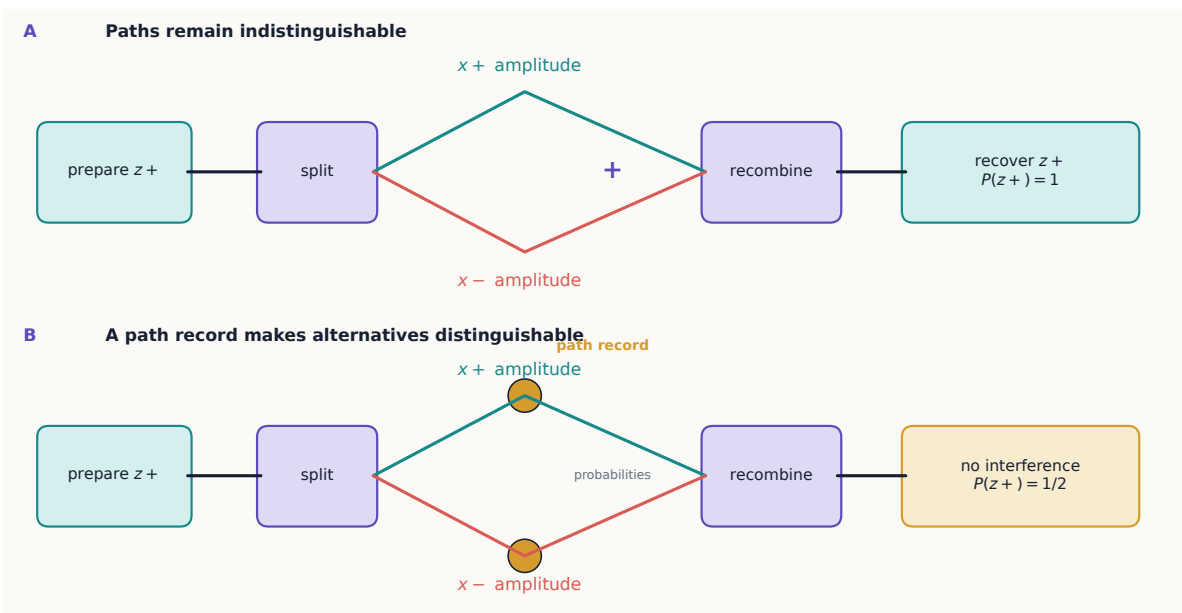

Figure 1.5: Recombination restores the original state only while the two alternatives remain coherent. A durable path record changes the calculation from adding amplitudes to adding probabilities, even if nobody reads the record.

This is why "half the atoms went through each arm" is not a complete description of the coherent experiment. The path amplitudes can cancel at one output and reinforce at another, even though each open arm alone sends atoms to both outputs.

Concept check 1.3

Why is a 50–50 collection of separately prepared $x+$ and $x-$ atoms not automatically equivalent to a coherent recombination of the two paths?

Answer

The collection records mutually exclusive preparations and is an incoherent mixture. A coherent recombination preserves the relative phase between the two amplitudes, so later probabilities can contain interference terms.

1.5 A minimal model

We now have three empirical rules—repeatability, incompatibility, and interference—plus the half-angle probability law. Before building a full mathematical theory, it helps to state in plain language exactly what that theory needs to accomplish:

- A preparation is represented by a **state**.
- A measurement has a discrete set of possible results.
- A state assigns an amplitude, not directly a probability, to each result.
- Measurement both produces a result and prepares a corresponding output state.
- Different analyzer orientations represent different, sometimes incompatible, physical questions.

For spin $\frac{1}{2}$, two independent states suffice to describe every possible pure spin state. We will represent them by vectors in a two-dimensional complex vector space. The word "two-dimensional" refers to state space, not ordinary physical space.

If "complex vector space" does not yet mean much to you, that is completely expected at this point—nothing so far has required it. Chapter 2 builds this idea from the ground up, starting with a review of ordinary arrows in the plane and of complex-number arithmetic itself, before ever writing down a spin state. The two-outcome rule, the incompatibility of S_z and S_x, and the interference rule found above are the experimental *reasons* that vector space is the right tool; the vocabulary comes second.

From laboratory rules to mathematical demands

Let's translate the experimental facts directly into mathematical requirements. Whatever structure we build, it needs:

- two mutually exclusive outcomes for each analyzer direction;
- a way to represent a preparation that is certain for one outcome;
- continuous families of other preparations, because the analyzer direction varies continuously;
- a rule that produces normalized probabilities;
- an update rule after an outcome is selected; and
- phase information capable of describing coherent recombination.

An ordinary list of two probabilities, say $(1/2, 1/2)$, cannot do this job: it records the *weight* of each alternative but throws away the *phase* that controls interference. A pair of **complex** numbers can do both at once, since a complex number naturally carries both a magnitude and a phase. This is exactly why Chapter 2 introduces state vectors with complex-number coefficients—called kets—before ever writing down a wavefunction.

1.6 What spin is—and is not

The name "spin" invites a picture: a tiny sphere spinning on its axis, like a planet. That picture is a useful crutch for the imagination, but it is not literally correct, and taking it too literally leads to wrong predictions. Spin does carry angular momentum, and often a magnetic moment, which is exactly why it interacts with the field in a Stern–Gerlach apparatus in the first place. But a rigid, rotating sphere would have a magnetic-moment component that varies continuously as its axis tilts—precisely the smooth spread that classical physics predicted and that the apparatus did *not* show. The name "spin" is a historical label; it is the experiments, not the name, that tell us what spin actually is.

There is a second way the arrow picture fails, beyond just the two discrete outcomes. A classical arrow pointing at a known angle α to the analyzer would have a projected component proportional to $\cos\alpha$, and that component would vary continuously as α varies. Real spin does not do this: the outcome is always $\pm\hbar/2$, and the angle only controls how *often* each outcome occurs, through equation (1.9).

One more warning about pictures. Chapter 2 introduces the Bloch sphere, which looks like a sphere covered in little arrows. Resist the urge to picture each arrow as a literal miniature compass needle sitting inside the particle. Each arrow is shorthand for an entire preparation recipe and the statistics it produces—not a claim that the particle contains a tiny spinning pointer.

1.7 A systematic way to predict analyzer chains

Until the vector formalism is available, a short operational recipe is enough for ideal sequential measurements:

1. Identify the state entering the first analyzer. If the source is unpolarized, the first analyzer produces equal counts.
2. At each analyzer, measure the angle from the **most recently selected preparation axis**, not necessarily from the original axis.
3. Use equation (1.9) for each branch.
4. Multiply conditional probabilities along one branch; add probabilities only when mutually exclusive detected branches are pooled.
5. If paths are recombined coherently, stop using this probability-tree recipe. Amplitudes and their phases are then required.

Example 1.3: keep all branches straight

An unpolarized source sends 16,000 atoms into SG_z. The $z+$ exit is selected, then the $x-$ exit of an SG_x analyzer is selected, and finally both exits of SG_z are counted.

The expected bookkeeping is

$$16{,}000 \xrightarrow{z+} 8{,}000 \xrightarrow{x-} 4{,}000 \xrightarrow{z} \begin{cases} 2{,}000 & z+, \\ 2{,}000 & z-\,. \end{cases} \tag{1.12}$$

The final counts are not 4,000 and zero. The state entering the last analyzer is $x-$, not the earlier $z+$ preparation. Writing the state beneath each segment of an apparatus sketch is a simple way to avoid this mistake.

1.8 Optional laboratory extension: from probabilities to data

The conceptual route to the state-vector formalism is complete at this point. This optional section connects ideal probabilities to the fluctuations and systematic effects encountered in laboratory data.

Equations such as (1.9) predict probabilities, whereas a real experiment produces integer counts. The link between the two is statistical.

Quick review: what a standard deviation tells you

If you flip a fair coin 100 times, you do not expect exactly 50 heads every time—you expect *something close to* 50, with the typical batch landing within roughly 5 of that number. The **standard deviation** is just a name for that typical amount of scatter: a way of saying, before the experiment is even run, "here is the size of wiggle you should expect around the predicted average, purely from chance." A result within one standard deviation of the prediction is unremarkable; a result many standard deviations away is a signal that something besides ordinary chance may be going on—an unfair coin, a miscounted trial, or, in the laboratory, a misaligned analyzer or a flawed theory. The formulas below simply make this everyday idea precise for the two-outcome case at hand.

If the preparation and analyzer are held fixed and N independent atoms are recorded, the number N_+ at the positive exit fluctuates around Np_+, where

$$p_+ = \cos^2 \frac{\alpha}{2}. \tag{1.13}$$

For ideal independent trials, the count follows a binomial distribution. Its standard deviation is

$$\sigma_{N_+} = \sqrt{Np_+(1-p_+)}, \tag{1.14}$$

and the measured frequency $f_+ = N_+/N$ has standard deviation

$$\sigma_{f_+} = \sqrt{\frac{p_+(1-p_+)}{N}}. \tag{1.15}$$

The $1/\sqrt{N}$ dependence matters. Recording four times as many atoms reduces the typical frequency fluctuation by only a factor of two. Quantum theory does not predict that every batch divides in exactly the theoretical ratio; it predicts the distribution of results from which those batches are drawn.

Example 1.4: is a split consistent with the angle rule?

A $z+$ source is analyzed along an axis 60° from $+z$. Among $N = 1{,}000$ detected atoms, $N_+ = 735$ leave the positive exit. The ideal prediction is $p_+ = 3/4$, so the expected count and its standard deviation are

$$Np_+ = 750, \qquad \sigma_{N_+} = \sqrt{1000\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)} \approx 13.7. \tag{1.16}$$

The observed count is 15 below the mean, about 1.1 standard deviations. Such a difference is ordinary statistical variation, not persuasive evidence against the angle rule. By contrast, a persistent discrepancy that grows in statistical significance as more data are collected suggests imperfect preparation, a misaligned analyzer, unequal detector efficiencies, or a failure of the model.

Counts require an experimental model
The binomial formulas assume independent trials, stable preparation, and equal detection conditions for the two exits. Background events, missed atoms, beam drift, and detector dead time can all change the count statistics. "The data look noisy" is not an explanation until the sources of noise are identified and tested.

Concept check 1.4

A run with 100 atoms gives 54 counts at $x+$ and 46 at $x-$. Has the expected 50–50 split failed?

Answer

No. For $N = 100$ and $p = 1/2$, the expected standard deviation of the positive count is $\sqrt{100(1/2)(1/2)} = 5$. A count of 54 is less than one standard deviation above the mean of 50 and is entirely typical.

Estimating an unknown analyzer angle

The same experiment can be used in reverse. For a known $z+$ input and an analyzer known to lie between 0 and π from $+z$, the observed positive frequency estimates the angle:

$$\alpha_{\text{est}} = 2\cos^{-1}\sqrt{f_+}. \tag{1.17}$$

If 81% of a large sample exits through $+\mathbf{n}$, then $\alpha_{\text{est}} = 2\cos^{-1}(0.9) \approx 51.7°$. Counts can therefore characterize an apparatus as well as test a prepared state. In later chapters, this reversal of viewpoint becomes quantum-state tomography: known analyzers are used to infer an unknown preparation.

1.9 Optional inquiry extension: separating competing explanations

This optional section turns the chapter's observations into a model-testing strategy: choose arrangements for which plausible explanations make different predictions.

A useful experiment does more than exhibit an effect; it distinguishes between models that would otherwise make the same prediction. A single z analyzer, on its own, cannot distinguish a genuine quantum state from a classical population of atoms carrying fixed z labels—both would give the same repeatable $z+$ result. Sequential analyzers are powerful precisely because they force competing stories to disagree.

Consider three hypotheses for a beam selected from the $z+$ exit:

1. **Fixed-label hypothesis:** every atom carries permanent values for all analyzer directions, and an analyzer only reads the relevant value.

2. **Classical-arrow hypothesis:** every atom carries an ordinary magnetic arrow, and the analyzer reports its component along the chosen axis.

3. **Quantum preparation hypothesis:** selecting an outcome prepares a state definite for that analyzer, while other components generally remain indefinite.

The sequence $z+ \to z$ is compatible with the first and third hypotheses, so repeatability alone does not decide between them. The sequence $z+ \to x+ \to z$, however, restores a 50–50 z distribution. That result conflicts with permanent, merely revealed z labels. The

fact that every analyzer still has only two exits conflicts with a classical arrow whose projected component can vary continuously. Coherent recombination then adds a third test: interference depends on a controllable relative phase that neither classical population model retains.

This pattern is a general method for reasoning in quantum mechanics:

- list the preparations and recorded outcomes without adding an unseen story;
- derive a quantitative prediction from each candidate model;
- choose a sequence for which those predictions differ; and
- repeat enough trials to separate systematic disagreement from count noise.

Example 1.5: measurement or mere beam splitting?

An ideal device separates a $z+$ input into $x+$ and $x-$ paths. Compare three arrangements.

- With both paths coherently recombined, a final z analyzer returns $z+$ with certainty.
- With the $x-$ path blocked, the final analyzer gives $z+$ and $z-$ equally.
- With a path marker present but both beams physically redirected toward the output, the final analyzer again gives the incoherent 50–50 statistics if the marker records perfectly distinguishable alternatives.

The last comparison shows why spatial overlap alone is not sufficient for interference. The relevant condition is whether the alternatives remain coherent—whether any physical degree of freedom carries usable which-path information. Chapter 2 will represent this distinction by retaining or discarding the cross terms between amplitudes.

Concept check 1.5

Why can the sequence $z+ \to x+ \to z$ test more than the sequence $z+ \to z$?

Answer

Both a fixed-label model and quantum theory can imitate immediate repeatability. They differ after an incompatible intermediate measurement: quantum theory predicts that selecting $x+$ prepares a new state and makes the final z result 50–50, whereas a device that merely reads a permanent z label should not erase it.

Summary

- A Stern–Gerlach analyzer correlates a spin component with one of two paths.
- Filtering an output prepares a repeatable state for that component.
- Orthogonal spin components are incompatible: preparing one makes another uncertain.
- Indistinguishable alternatives interfere through complex amplitudes.
- A spin-$\frac{1}{2}$ state requires a two-dimensional complex state space.
- Analyzer chains are sequences of conditional preparations and measurements; the most recent selected outcome determines the state entering the next device.
- Finite data fluctuate around theoretical probabilities with a characteristic scale proportional to $1/\sqrt{N}$.
- Sequential and coherent-recombination experiments are valuable because they force classical and quantum explanations to make different predictions.

Exercises

Exercises 1–6 rehearse the core analyzer logic, 7–11 consolidate angle rules and coherent alternatives, and 12–16 extend the chapter into data analysis and experimental design. Complete

the first group before choosing from the later two.

1. An unpolarized beam of 20,000 atoms passes through SG_z; the $z+$ output then passes through SG_x. Estimate the counts in the two final outputs.
2. Predict the possible final outcomes and their probabilities for $z+ \to x- \to z$.
3. Insert a second SG_x after the $x-$ selection in Exercise 2. Explain why its result differs qualitatively from the final S_z result.
4. A student says, "The first S_z analyzer tells us which atoms were already pointing up." Identify two observations in this chapter that this classical explanation fails to capture.
5. Explain operationally what must be changed to turn a spin analyzer into a state-preparation apparatus.
6. Design a sequence of ideal analyzers that begins with $z+$ and has a final $z-$ detection probability of 1/2. Give at least two sequences and explain them.
7. A $z+$ beam is sent to an analyzer tilted 120° from $+z$. Find both output probabilities using equation (1.9).
8. An unpolarized beam of 40,000 atoms passes through three analyzers. The selected outputs are first $z+$ and then $+\mathbf{n}$, where $\mathbf{n}$ is 60° from $+z$. Both outputs of the final z analyzer are recorded. Predict all final counts and account for atoms blocked earlier.
9. Draw probability trees for $z+ \to x+ \to z$ and $z+ \to x- \to z$. If the two selected x beams are later pooled incoherently, what are the final z probabilities?
10. A two-path apparatus gives deterministic $z+$ when both paths are open and recombined. With either path blocked, the final z analyzer gives both results. Explain why adding the two single-path probability distributions cannot reproduce the both-open result.
11. Classify each element as preparation, transformation, or measurement: an oven source, a $z+$ filter, a uniform magnetic field, an unobserved path marker, a coherent recombiner, and a two-channel detector. Some elements may play more than one role; explain your choices operationally.
12. A $z+$ beam is analyzed 90° from its preparation axis. In a run of 2,500 atoms, estimate the mean and standard deviation of the + count. Is a count of 1,290 surprising on this scale?
13. A large sample prepared in $z+$ gives $f_+ = 0.64$ at an analyzer whose axis is known to lie between 0° and 180° from $+z$. Estimate the angle.
14. Derive equation (1.15) from equation (1.14). Explain why the absolute uncertainty in a count grows with N while the uncertainty in a relative frequency shrinks.
15. Design a three-configuration experiment that distinguishes coherent recombination, an incoherent mixture of the two paths, and accidental blocking of one path. State the qualitative final z statistics expected in each configuration.
16. A source drifts so that its intensity falls by 10% during a run. Explain why raw output counts could then give a misleading comparison between two analyzer settings. Propose a frequency or normalization that reduces the problem.

Selected exercise guidance

Use these answers only after making a complete attempt.

Exercise 2

The $x-$ selection occurs with probability 1/2. Conditional on that selection, the final $z+$ and $z-$ probabilities are each 1/2, so each complete detected route contains 1/4 of the original beam.

Exercise 7

$P(+\mathbf{n}|z+) = \cos^2(60°) = 1/4$ and $P(-\mathbf{n}|z+) = 3/4$.

Exercise 12

The mean is 1,250 and the standard deviation is $\sqrt{2500(1/2)(1/2)} = 25$. A count of 1,290 is 1.6 standard deviations above the mean, so it is not especially surprising.

Exercise 13

$$\alpha = 2\cos^{-1}\sqrt{0.64} = 2\cos^{-1}(0.8) \approx 73.7°. \tag{1.18}$$

Chapter 2

States and Probability Amplitudes

2.1 States and probability amplitudes

Learning objectives

After this chapter, you should be able to:

- represent a spin-$\frac{1}{2}$ state using kets and column matrices;
- expand a state in any orthonormal basis;
- normalize a state and apply the Born rule;
- calculate selected routes through sequential analyzers;
- distinguish relative phase from overall phase;
- translate among the x, y, and z spin bases;
- distinguish a coherent pure state from a classical mixture;
- in the optional extension, construct spinors for an arbitrary analyzer direction; and
- add amplitudes for coherent alternatives and probabilities for exclusive recorded routes.

The subsections labeled "optional extension" provide compact matrix machinery for readers ready to go further; the core conceptual path does not depend on them.

2.1 States are vectors

Chapter 1 ended with a wish list: whatever mathematical object represents a spin state, it needs two mutually exclusive outcomes for every analyzer direction, a way to represent a preparation that is certain for one of them, continuous families of other preparations, and phase information that can produce interference. An ordinary table of probabilities cannot do all of this at once, because it has no room for phase. Complex vectors can, and this chapter makes that idea precise.

Prerequisite review: complex-number arithmetic

Skip this box if complex numbers are already familiar. If your last exposure to them was in an algebra class and feels distant, read it before continuing; the notation introduced here is used in the first state calculations below.

A complex number is a pair of real numbers, x and y, packaged as $z = x + iy$, where $i^2 = -1$. Complex numbers add component by component, while multiplication follows the ordinary rules of algebra with i^2 replaced by -1:

$$(2+3i)+(1-5i)=3-2i, \qquad (2+3i)(1-5i)=17-7i. \tag{2.1}$$

The **complex conjugate** of $z = x + iy$ is $z^* = x - iy$. Multiplying a number by its conjugate produces its nonnegative squared **modulus**:

$$z^*z = (x-iy)(x+iy) = x^2 + y^2 = |z|^2. \tag{2.2}$$

Thus $z = 3 + 4i$ has $z^* = 3 - 4i$ and $|z| = 5$. A probability built from a complex amplitude z is $|z|^2 = z^*z$, never z^2. For example, $|i/2|^2 = 1/4$, whereas $(i/2)^2 = -1/4$ cannot be a probability.

Geometrically, $z = x + iy$ is an arrow in a plane. Its length is $r = |z|$, and its angle ϕ from the real axis is its **phase**. Polar form records both:

$$z = x + iy = r\cos\phi + ir\sin\phi = re^{i\phi}, \tag{2.3}$$

where Euler's formula states $e^{i\phi} = \cos\phi + i\sin\phi$. Conjugation reverses the phase, $z^* = re^{-i\phi}$, without changing the magnitude.

The notation $|\cdot\rangle$, read "ket," simply names a vector in that complex vector space. There is no more mystery in the symbol than in writing $\vec{v}$ for an arrow in ordinary space—it is just a label, chosen to remind us that we are working with a state rather than an ordinary number.

Quick review: vectors, components, and dot products

If vectors and dot products are already second nature to you, skip ahead— everything here reappears below in the spin context. If not, here is the ordinary-space picture this chapter builds on.

An arrow $\vec{v}$ in the plane can be written as a sum of two perpendicular unit vectors, $\vec{v} = v_x\hat{x} + v_y\hat{y}$, where $\hat{x}$ and $\hat{y}$ point one unit along the two axes. The numbers v_x and v_y are the arrow's **components**: they say how much of $\hat{x}$ and how much of $\hat{y}$ must be combined to build $\vec{v}$. Once axes are chosen, it is equivalent, and often more convenient, to drop the arrow notation entirely and just write the column of numbers, $\vec{v} \doteq (v_x, v_y)$.

The **dot product** of two vectors, $\vec{u}\cdot\vec{v} = u_xv_x + u_yv_y$, answers a specific geometric question: how much does one vector "overlap" with another? Two perpendicular vectors have zero overlap, and a vector dotted with itself gives the square of its own length, $\vec{v}\cdot\vec{v} = |\vec{v}|^2$. For a concrete check, $\vec{u} = (3, 4)$ and $\vec{v} = (0, 1)$ give $\vec{u}\cdot\vec{v} = 3(0) + 4(1) = 4$—exactly the y-component of $\vec{u}$, since $\vec{v} = \hat{y}$. Dotting a vector with a basis vector simply reads off the corresponding component.

Everything below repeats this pattern with two changes: the "components" a and b are complex numbers rather than real ones, and the dot product is replaced by an *inner product* built to handle that. Nothing about the underlying idea—build a vector from basis pieces, then use a product to read the pieces back off—is new.

The preparation called $z+$ is represented by the ket $|+z\rangle$, and the preparation $z-$ by $|-z\rangle$. These two state vectors form an orthonormal basis for the spin state space: every other spin-$\frac{1}{2}$ preparation can be built from them by superposition, in exact analogy to how two perpendicular unit vectors span a plane. Any pure spin state can therefore be written

$$|\psi\rangle = a|+z\rangle + b|-z\rangle, \tag{2.4}$$

where a and b are complex numbers. Think of a and b as *coordinates*: they fix how much of $|+z\rangle$ and how much of $|-z\rangle$ go into building $|\psi\rangle$, in the same way that the numbers 3 and 4 fix an arrow once we agree on x- and y-axes. They are not themselves probabilities—Section 2.3 will show exactly how probabilities are built from them.

It is often convenient to write $|\psi\rangle$ as a column of numbers. Choosing the representation

$$|+z\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |-z\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{2.5}$$

gives $|\psi\rangle \doteq (a, b)^{\mathsf{T}}$. The symbol $\doteq$ means "is represented, in the chosen basis, by." A ket is an abstract vector; a column matrix is only its representation once a basis has been chosen.

This distinction is the same one made for an ordinary geometric vector. The arrow in space is not identical to the list of three numbers used to represent it: rotating the coordinate axes changes the numbers but not the arrow itself. In exactly the same way, a spin ket does not change when we rewrite it in the x basis instead of the z basis—only its column of numbers changes.

Superposition is a physical principle

Because states are vectors, any nonzero linear combination of two allowed states can be normalized to represent an allowed state. This is the **superposition principle**, and it says something stronger than "the atom is secretly in one state or the other." In

$$|\psi\rangle = a|+z\rangle + b|-z\rangle, \tag{2.6}$$

the complex coefficients preserve phase information that can affect a later measurement. Compare this with a classical mixture, in which a fraction $|a|^2$ of atoms is separately prepared in $z+$ and the rest in $z-$. That mixture and the superposition above generally make *different* predictions, even though both give the same probabilities for a z measurement. The section on pure states and mixtures makes this comparison explicit.

2.2 Bras and inner products

A ket by itself cannot yet produce a number we could compare to laboratory data—it is a direction in an abstract space, not a prediction. To extract predictions, we pair every ket $|\psi\rangle$ with a companion object called a **bra**, written $\langle\psi|$. A bra combines with a ket to its right to produce a single complex number, called the inner product. This pairing is what lets us ask "how much of outcome ϕ is contained in state ψ?" and get back a definite amplitude. The bra corresponding to $|\psi\rangle$ is its Hermitian conjugate,

$$\langle\psi| = a^*\langle+z| + b^*\langle-z| \doteq \begin{pmatrix} a^* & b^* \end{pmatrix}. \tag{2.7}$$

The z basis is orthonormal, meaning each basis ket has unit length and the two are perpendicular in the sense of the inner product:

$$\langle+z|+z\rangle = \langle-z|-z\rangle = 1, \qquad \langle+z|-z\rangle = \langle-z|+z\rangle = 0. \tag{2.8}$$

Using these two facts, pairing $\langle\psi|$ with $|\psi\rangle$ gives

$$\langle\psi|\psi\rangle = |a|^2 + |b|^2. \tag{2.9}$$

A physical pure state is normalized, meaning $|a|^2 + |b|^2 = 1$. If a nonzero vector $|v\rangle$ is not already normalized, we can always rescale it into a valid physical state by dividing by its length:

$$|\psi\rangle = \frac{|v\rangle}{\sqrt{\langle v|v\rangle}}. \tag{2.10}$$

Example 2.1: normalization

Normalize $|v\rangle = 2|+z\rangle + (1+i)|-z\rangle$. Applying equation (2.9) term by term, the first coefficient contributes $|2|^2 = 4$ and the second contributes $|1+i|^2 = (1+i)^*(1+i) = (1-i)(1+i) = 1+1 = 2$, so the squared norm is $\langle v|v\rangle = 4 + 2 = 6$, hence

$$|\psi\rangle = \frac{1}{\sqrt{6}}\left(2|+z\rangle + (1+i)|-z\rangle\right). \tag{2.11}$$

Algebra check

If $|\chi\rangle = (|+z\rangle - i|-z\rangle)/\sqrt{2}$, what are $\langle\chi|$ and $\langle\chi|\chi\rangle$?

Answer

Conjugating the coefficients gives $\langle\chi| = (\langle +z| + i\langle -z|)/\sqrt{2}$. Orthonormality then gives $\langle\chi|\chi\rangle = 1/2 + 1/2 = 1$.

Inner-product habits

Three small algebraic rules prevent most early errors. First, swapping the order of an inner product conjugates it:

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*. \tag{2.12}$$

Second, a bra acts linearly on a ket to its right:

$$\langle\phi|(a|u\rangle + b|v\rangle) = a\langle\phi|u\rangle + b\langle\phi|v\rangle. \tag{2.13}$$

Third, the inner product is **conjugate-linear in its first slot**. If a ket is combined first and then converted into a bra, its coefficients are conjugated:

$$\langle au + bv|\psi\rangle = a^*\langle u|\psi\rangle + b^*\langle v|\psi\rangle. \tag{2.14}$$

For example, if $|\phi\rangle = (|+z\rangle + i|-z\rangle)/\sqrt{2}$, then $\langle\phi| = (\langle +z| - i\langle -z|)/\sqrt{2}$. Once a bra has already been written as a row, ordinary row–column multiplication is linear in the entries of that row. Keeping these two statements separate prevents the common error of forgetting conjugation when forming the bra.

Complex numbers as amplitude geometry

Complex numbers enter quantum mechanics because an amplitude needs to carry both a magnitude and a phase at once, and a single real number cannot do that. In polar form,

$$z = re^{i\phi} = r(\cos\phi + i\sin\phi), \qquad z^* = re^{-i\phi}, \qquad |z|^2 = z^* z = r^2. \tag{2.15}$$

Multiplying two amplitudes multiplies their magnitudes and adds their phases. Adding two amplitudes is ordinary geometric addition in the complex plane—and this is the key fact that makes interference possible. Because vectors in a plane can point in opposite directions, two *nonzero* amplitudes can add to *zero*. For example,

$$\frac{1}{2} + \left(-\frac{1}{2}\right) = 0, \tag{2.16}$$

even though the squared magnitude of each term separately is 1/4. This possibility—two nonzero contributions canceling exactly—is the algebraic core of destructive interference.

For arbitrary amplitudes A_1 and A_2, expanding $|A_1 + A_2|^2$ directly gives

$$|A_1 + A_2|^2 = |A_1|^2 + |A_2|^2 + 2\,\mathrm{Re}(A_1^* A_2). \tag{2.17}$$

The final term is the **interference term**. It depends on the relative phase between A_1 and A_2, and it can be positive, negative, or zero. If the two alternatives are instead recorded separately, or their relative phase is randomized by the apparatus, this cross term averages to zero and only $|A_1|^2 + |A_2|^2$ remains—the ordinary rule for adding probabilities of mutually exclusive events.

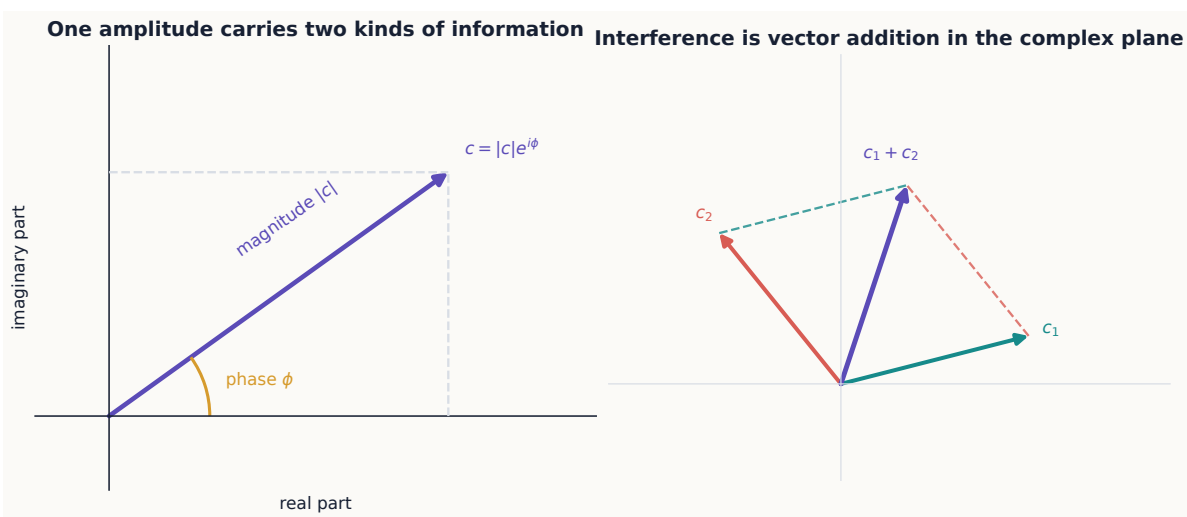

Figure 2.1: A complex amplitude is a two-dimensional arrow: its length is the magnitude and its direction is the phase. Coherent alternatives interfere because these arrows add before the squared magnitude is taken.

> **Algebra checkpoint**
> Do not replace $|A_1+A_2|^2$ by $|A_1|^2+|A_2|^2$ unless the physical alternatives are distinguishable or incoherent. Whether a cross term belongs in the calculation is decided by the apparatus, not by algebraic convenience.

2.3 The Born rule

We now have the machinery to turn a state into an actual prediction. The amplitude for obtaining $z+$ from $|\psi\rangle$ is $\langle +z|\psi\rangle = a$—simply the coefficient of $|+z\rangle$ in the expansion of $|\psi\rangle$. The probability of that outcome is the squared magnitude of that amplitude:

$$P(z+ \mid \psi) = |\langle +z|\psi\rangle|^2 = |a|^2, \qquad P(z- \mid \psi) = |b|^2. \tag{2.18}$$

This rule for turning an amplitude into a probability is the **Born rule**, and it is one of the central postulates of quantum mechanics. Notice that the order matters in the amplitude: the bra labels the possible measurement outcome, and the ket labels the prepared state.

Because the z basis spans the state space, the two overlaps recover the whole ket:

$$|\psi\rangle = \langle +z|\psi\rangle|+z\rangle + \langle -z|\psi\rangle|-z\rangle. \tag{2.19}$$

Comparing this with equation (2.4) shows that a and b are precisely the amplitudes $\langle +z|\psi\rangle$ and $\langle -z|\psi\rangle$. Chapter 3 will package this basis expansion into an identity operator after operators have been introduced.

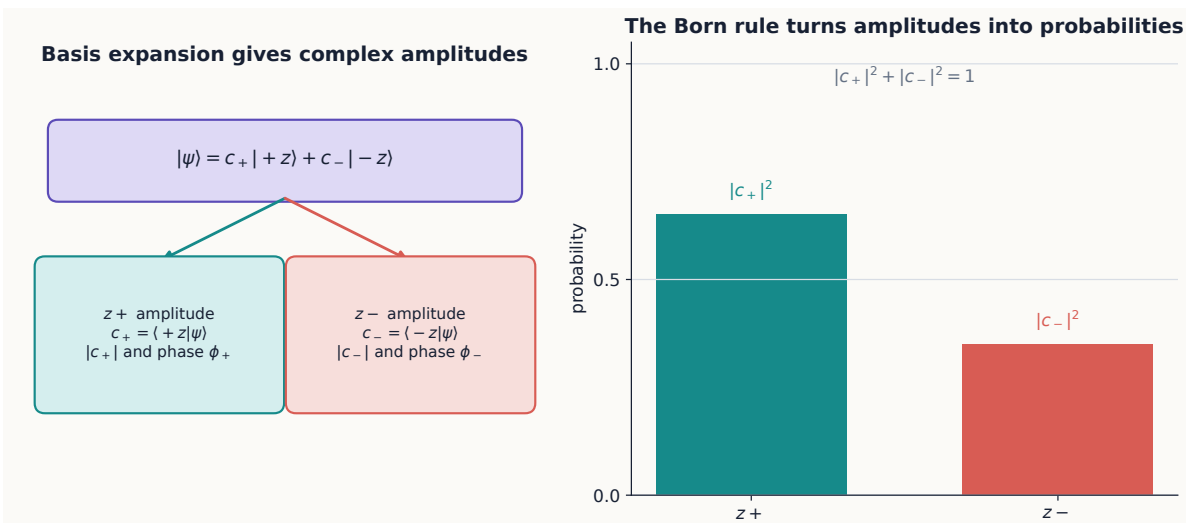

Figure 2.2: Expansion in a measurement basis produces complex amplitudes; the Born rule turns their squared magnitudes into normalized outcome probabilities. Each amplitude can carry a phase even though the probability bars cannot show it.

A probability calculation in four steps

For any ideal spin measurement, the same workflow applies:

1. Write the prepared ket and the outcome ket in a common basis.
2. Hermitian-conjugate the outcome ket to make the bra.
3. Form the amplitude $\langle \text{outcome}|\text{prepared}\rangle$.
4. Take its squared magnitude, and check that all exclusive outcome probabilities sum to one.

The amplitude computed in step 3 is generally complex, and there is no requirement that it lie between zero and one—only its squared magnitude in step 4 is a probability. Keeping the amplitude, rather than squaring it immediately, is essential whenever the alternative will later be recombined with another one.

Concept check 2.1

After computing $\langle\phi|\psi\rangle = -i/2$, a student reports a probability of $-i/2$. What should the student report instead?

Answer

The amplitude may be complex, but the probability is its squared magnitude: $|-i/2|^2 = 1/4$.

2.4 Changing the question means changing the basis

Every measurement so far has used the z basis, but nothing forces that choice—an analyzer built along x or y asks a different physical question, and it needs its own basis kets. The 50–50 result alone fixes the magnitudes of the two z-basis coefficients but not their relative phases. The formulas below therefore contain two ingredients: equal magnitudes demanded by the experiment, and a standard phase convention that ties the labels x and y to perpendicular directions in physical space. Chapter 3 will recover the same kets as eigenvectors of the spin-component matrices. With that convention, the x basis is

$$|+x\rangle = \frac{1}{\sqrt{2}}\left(|+z\rangle + |-z\rangle\right), \qquad |-x\rangle = \frac{1}{\sqrt{2}}\left(|+z\rangle - |-z\rangle\right). \tag{2.20}$$

And the y basis is

$$|+y\rangle = \frac{1}{\sqrt{2}}\left(|+z\rangle + i|-z\rangle\right), \qquad |-y\rangle = \frac{1}{\sqrt{2}}\left(|+z\rangle - i|-z\rangle\right). \tag{2.21}$$

Example 2.2: a relative phase becomes a definite result

Prepare

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|+z\rangle + i|-z\rangle\right) \tag{2.22}$$

and ask for the probability of $y+$. From equation (2.21),

$$\langle +y| = \frac{1}{\sqrt{2}}(\langle +z| - i\langle -z|). \tag{2.23}$$

Therefore

$$\langle +y|\psi\rangle = \frac{1}{2}\left(1 + (-i)i\right) = 1, \tag{2.24}$$

so $P(y+) = 1$ and $P(y-) = 0$. The state is exactly $|+y\rangle$, written in the z basis. This deliberately simple example exposes the important algebra: the i in the ket becomes $-i$ in the bra, and $(-i)i = 1$. A missed complex conjugation would destroy the certainty that defines the $y+$ state.

Let's check that this reproduces the known result. Applying the Born rule to a $z+$ input,

$$P(x+ \mid z+) = |\langle +x| + z\rangle|^2 = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}, \tag{2.25}$$

exactly as observed in Chapter 1. To express any state $|\psi\rangle$ in the x basis instead of the z basis, project it onto the two x-basis kets:

$$|\psi\rangle = |+x\rangle\langle +x|\psi\rangle + |-x\rangle\langle -x|\psi\rangle. \tag{2.26}$$

It is important to see what has and has not changed here. No physical state has changed—the atom's preparation is exactly what it was. Only the *coordinates* used to describe that state have changed, in the same way that rewriting an arrow's components in a rotated coordinate frame does not move the arrow.

Optional extension: basis-change matrices

The inner-product method above is the core skill. Readers who want a compact matrix version can organize all the overlaps into one basis-change matrix.

The same calculation can be organized as a single matrix multiplication, which is often more convenient for bookkeeping. If $[\psi]_z = (a, b)^\mathsf{T}$ is the column representing $|\psi\rangle$ in the z basis, its x-basis column is

$$[\psi]_x = \begin{pmatrix} \langle +x| + z\rangle & \langle +x| - z\rangle \\ \langle -x| + z\rangle & \langle -x| - z\rangle \end{pmatrix} [\psi]_z = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} [\psi]_z. \tag{2.27}$$

Each row of this matrix is one of the x-basis bras, written out in z coordinates. A matrix U is called **unitary** when $U^\dagger U = UU^\dagger = I$; this condition guarantees that lengths and inner products are preserved. The change-of-basis matrix is unitary, so a properly normalized state stays normalized after the change of basis. Chapter 3 will develop operators and unitary matrices systematically. Reading the entries here as inner products also makes clear that a basis change is *not* a physical Stern–Gerlach measurement—no outcome has been selected, and the abstract ket itself has not been disturbed in any way.

Quick calculation: carrying out the matrix multiplication

It is worth performing this multiplication by hand once, so that equation (2.27) reads as an ordinary matrix–vector product—row dotted into column, exactly as in the vector review of Section 2.1—rather than as an unfamiliar piece of bra–ket bookkeeping. Take $|\psi\rangle = (|+z\rangle + i|-z\rangle)/\sqrt{2}$, so $[\psi]_z = \frac{1}{\sqrt{2}}\binom{1}{i}$. Multiplying,

$$[\psi]_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1+i \\ 1-i \end{pmatrix}. \tag{2.28}$$

The top entry comes from dotting the row $(1, 1)$ into the column $(1, i)$, giving $1 + i$; the bottom entry comes from dotting $(1, -1)$ into $(1, i)$, giving $1 - i$; each is then scaled by the $\frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{2}$ collected from the two prefactors. As a check, normalization must survive the change of basis:

$$\left|\frac{1+i}{2}\right|^2 + \left|\frac{1-i}{2}\right|^2 = \frac{1}{2} + \frac{1}{2} = 1, \tag{2.29}$$

confirming directly, by arithmetic rather than by appeal to unitarity alone, that the x-basis column is a properly normalized state.

Example 2.3: probabilities in another basis

For

$$|\psi\rangle = \frac{\sqrt{3}}{2}|+z\rangle + \frac{i}{2}|-z\rangle, \tag{2.30}$$

the $x+$ amplitude is

$$\langle +x|\psi\rangle = \frac{1}{\sqrt{2}}\left(\frac{\sqrt{3}}{2} + \frac{i}{2}\right). \tag{2.31}$$

Therefore $P(x+) = \frac{1}{2}$ and $P(x-) = \frac{1}{2}$. The same state has unequal z probabilities, 3/4 and 1/4. Probabilities belong to a state–measurement *pair*, not to the state alone.

2.5 Sequential measurements with amplitudes

We can now redo the analyzer chains of Chapter 1 using the vector formalism, and see that it reproduces the same answers by a more systematic route. If $+a$ is prepared, $+b$ is selected, and $+c$ is finally detected, the probability for that selected route is

$$P(+c, +b \mid +a) = |\langle +c|+b\rangle|^2 |\langle +b|+a\rangle|^2. \tag{2.32}$$

The two factors multiply because the first measurement prepares $|+b\rangle$, and it is *that* state, not the original $|+a\rangle$, which enters the second stage. It would be incorrect to replace this product with $|\langle +c|+a\rangle|^2$: the intervening selected measurement genuinely changes the state, and the calculation must reflect that.

Example 2.4: a three-stage analyzer problem

A $z+$ beam enters SG_x. The $x+$ output is retained and sent into SG_y. What fraction of the original beam emerges as $y-$?

The route probability is

$$P(y-, x+ \mid z+) = |\langle -y|+x\rangle|^2 |\langle +x|+z\rangle|^2. \tag{2.33}$$

Using the basis kets,

$$\langle +x|+z\rangle = \frac{1}{\sqrt{2}}, \qquad \langle -y|+x\rangle = \frac{1+i}{2}. \tag{2.34}$$

Both squared magnitudes are 1/2, so one quarter of the original $z+$ beam reaches the $y-$ detector. Notice that the second amplitude is complex even though its probability is simple.

Selected paths versus coherent paths

Equation (2.32) applies when the intermediate result is recorded or filtered, so the state is updated. If intermediate alternatives remain coherent and are recombined, their **amplitudes** must be added before squaring. The apparatus determines which calculation is appropriate.

2.6 Relative and overall phase

Not every phase in a ket matters physically, and it is important to know which ones do. Multiplying an entire ket by an overall phase does not change any probability at all:

$$|\psi'\rangle = e^{i\gamma}|\psi\rangle, \qquad |\langle\phi|\psi'\rangle|^2 = |e^{i\gamma}|^2|\langle\phi|\psi\rangle|^2. \tag{2.35}$$

Since $|e^{i\gamma}|^2 = 1$, the kets $|\psi\rangle$ and $e^{i\gamma}|\psi\rangle$ represent exactly the same physical state, for any choice of γ.

A **relative phase**—one that appears on only one term of a superposition, not on the whole ket—is entirely different, and it does change interference. Compare

$$\frac{|+z\rangle + |-z\rangle}{\sqrt{2}} = |+x\rangle \quad \text{with} \quad \frac{|+z\rangle - |-z\rangle}{\sqrt{2}} = |-x\rangle. \tag{2.36}$$

Concept check 2.2

Do $|+y\rangle$ and $|-y\rangle$ differ only by an overall phase?

Answer

No. If $|-y\rangle = c|+y\rangle$, matching the $|+z\rangle$ coefficient requires $c = 1$, while matching the $|-z\rangle$ coefficient requires $c = -1$. They are orthogonal states, not two representations of the same state.

Both $|+x\rangle$ and $|-x\rangle$ give 50–50 probabilities in the z basis—the two states look identical if all we ever do is measure S_z. Yet a relative sign is all that separates them, and an S_x measurement distinguishes them perfectly. This is the essential fact about relative phase: it can be completely invisible to one choice of measurement and yet fully visible to another.

More generally, consider the family of states

$$|\psi_\phi\rangle = \frac{|+z\rangle + e^{i\phi}|-z\rangle}{\sqrt{2}}, \tag{2.37}$$

parameterized by a single relative phase ϕ. Its $x+$ probability is

$$P(x+) = \left|\frac{1+e^{i\phi}}{2}\right|^2 = \cos^2\frac{\phi}{2}. \tag{2.38}$$

Concept check 2.3

A normalized ket has $P(z+) = P(z-) = 1/2$. Is the state determined uniquely?

Answer

No. It has the form $(|+z\rangle + e^{i\phi}|-z\rangle)/\sqrt{2}$. Different relative phases give different predictions in other bases.

Every member of this family has exactly the same z statistics—a 50–50 split, regardless of ϕ. Rotating the final analyzer to x is what converts the otherwise hidden relative phase into a visible count rate. Interference, seen this way, is not some extra phenomenon bolted onto the vector formalism; it is simply what ordinary basis projection does to complex coefficients once more than one term is present.

2.7 A complete coherent path calculation

We can now close the experimental loop opened in Chapter 1. Consider an ideal device that splits $z+$ into coherent $x+$ and $x-$ paths and then implements the inverse routing operation, returning both paths to the same spatial and motional state. With no uncontrolled path-dependent phase, the amplitude for each final outcome is the sum of the two indistinguishable route amplitudes:

$$A(z+) = \langle +z|+x\rangle\langle +x|+z\rangle + \langle +z|-x\rangle\langle -x|+z\rangle. \tag{2.39}$$

The two route amplitudes are each 1/2, so

$$A(z+) = \frac{1}{2} + \frac{1}{2} = 1, \qquad P(z+) = 1. \tag{2.40}$$

For the other exit,

$$A(z-) = \langle -z|+x\rangle\langle +x|+z\rangle + \langle -z|-x\rangle\langle -x|+z\rangle = \frac{1}{2} - \frac{1}{2} = 0. \tag{2.41}$$

The mathematics describes the ideal apparatus specified above; merely writing a sum over intermediate states would not prove that an arbitrary physical beam splitter preserves coherence or recombines every other degree of freedom.

Concept check 2.4

Both route amplitudes in equation (2.41) have squared magnitude 1/4. Why is the final probability zero rather than 1/2?

Answer

The routes are indistinguishable and must be added as complex amplitudes. Their relative minus sign makes the sum zero. Adding the two route probabilities would describe a different apparatus—one that records or decoheres the intermediate alternatives.

The cancellation is valid only because the intermediate x result was never recorded. If it *is* recorded, the two complete routes become exclusive, and we add their probabilities:

$$P_{\text{recorded}}(z+) = \left|\frac{1}{2}\right|^2 + \left|\frac{1}{2}\right|^2 = \frac{1}{2}. \tag{2.42}$$

The same route magnitudes produce certainty or a 50–50 split depending on whether the intermediate alternatives remain coherent.

Concept check 2.5

Does changing the phase convention of $|-x\rangle$ change the interference prediction?

Answer

No. A phase assigned to the intermediate basis ket appears once in a ket overlap and once with the opposite sign in the corresponding bra overlap. It cancels from each complete route amplitude. Observable interference depends on physical relative phases introduced by transformations, not on basis notation.

A controllable phase

Suppose one path adds a phase $e^{i\delta}$ before recombination, perhaps because it travels a slightly longer distance. Repeating the calculation,

$$A(z+) = \frac{1+e^{i\delta}}{2}, \qquad A(z-) = \frac{1-e^{i\delta}}{2}, \tag{2.43}$$

so that

$$P(z+) = \cos^2\frac{\delta}{2}, \qquad P(z-) = \sin^2\frac{\delta}{2}. \tag{2.44}$$

At $\delta = 0$, the original $z+$ preparation is recovered. At $\delta = \pi$, the two exits exchange roles and $z-$ occurs with certainty. In between, a continuously adjustable phase controls discrete detector outcomes through continuously changing ensemble frequencies.

The calculation supplies a reusable rule:

1. identify each complete, indistinguishable route to one final record;
2. multiply amplitudes along each route;
3. include every transformation phase;
4. add the route amplitudes; and
5. square the magnitude only after the coherent sum is complete.

If a record distinguishes routes, group together only routes that remain indistinguishable to that record, and add probabilities between the resulting exclusive groups.

2.8 The Bloch-sphere coordinates

After normalizing a state and discarding its irrelevant overall phase, every pure spin-$\frac{1}{2}$ state can be written in the standard form

$$|\psi\rangle = \cos\frac{\theta}{2}|+z\rangle + e^{i\phi}\sin\frac{\theta}{2}|-z\rangle, \tag{2.45}$$

with $0 \le \theta \le \pi$ and $0 \le \phi < 2\pi$. Because only two real parameters remain once normalization and overall phase are fixed, the angles (θ, ϕ) locate a point on an ordinary sphere—the **Bloch sphere**—and every possible pure spin state corresponds to exactly one point on it. The half-angle is essential here, not optional. The angles (θ, ϕ) label a direction on the Bloch sphere; they should not be confused with a history of physical rotation undergone by the atom. Chapter 4 will show that when a representative spinor is physically rotated through 2π, its ket acquires a minus sign. The corresponding physical ray is unchanged because the sign is an overall phase, though it can be observed relative to another coherent amplitude.

For a measurement along a unit vector $\mathbf{n}$ making angle θ with $+z$, equation (2.45) gives

$$P(+\mathbf{n} \mid +z) = \cos^2\frac{\theta}{2}, \qquad P(-\mathbf{n} \mid +z) = \sin^2\frac{\theta}{2}. \tag{2.46}$$

This spin-$\frac{1}{2}$ rule resembles Malus's law for polarized light, but with one important difference: it uses *half* the geometric angle on the Bloch sphere, not the angle itself.

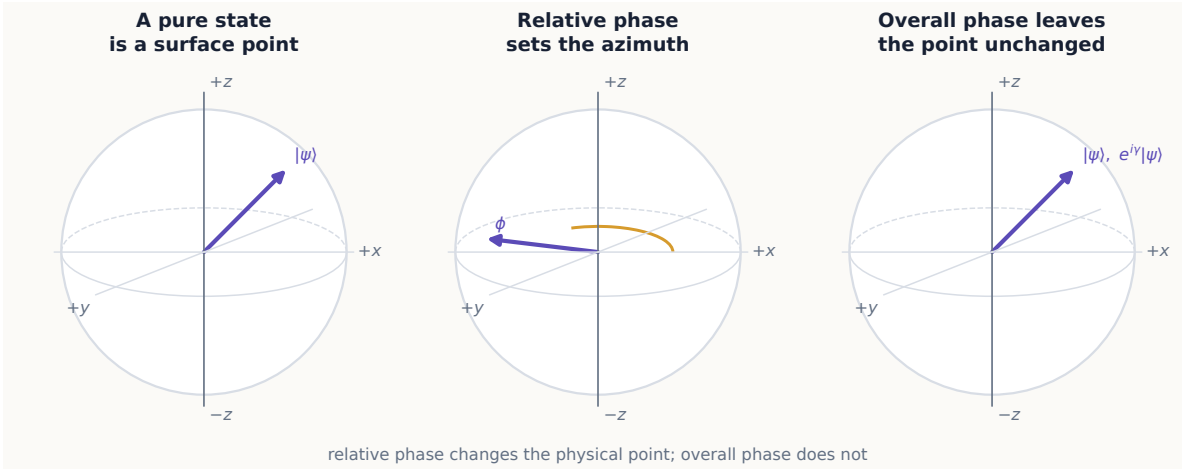

Figure 2.3: The Bloch sphere represents physical pure states rather than individual ket representatives. Polar angle fixes the relative magnitudes, azimuth fixes the relative phase, and overall phase has no coordinate because it leaves the physical state unchanged.

For two arbitrary pure states, let their Bloch angles be (θ_a, ϕ_a) and (θ_b, ϕ_b). Substituting equation (2.45) for both kets gives the amplitude

$$\langle +\mathbf{b}|+\mathbf{a}\rangle = \cos\frac{\theta_b}{2}\cos\frac{\theta_a}{2} + e^{i(\phi_a-\phi_b)}\sin\frac{\theta_b}{2}\sin\frac{\theta_a}{2}. \tag{2.47}$$

Taking its squared magnitude and using the half-angle identities produces

$$|\langle +\mathbf{b}|+\mathbf{a}\rangle|^2 = \frac{1}{2}\left[1+\cos\theta_a\cos\theta_b+\sin\theta_a\sin\theta_b\cos(\phi_a-\phi_b)\right]. \tag{2.48}$$

The expression in square brackets after the 1 is exactly $\mathbf{a}\cdot\mathbf{b} = \cos\alpha$, where α is the ordinary angle between the two analyzer directions. Therefore

$$|\langle +\mathbf{b}|+\mathbf{a}\rangle|^2 = \frac{1+\mathbf{a}\cdot\mathbf{b}}{2} = \cos^2\frac{\alpha}{2}. \tag{2.49}$$

This completes the derivation, promised in Chapter 1, of the empirical angle rule from state vectors. Notice a geometric subtlety worth pausing on: orthogonal directions in ordinary space ($\alpha = 90°$) do **not** correspond to orthogonal kets. They give an overlap probability of 1/2, not zero. The state that *is* orthogonal to $|+\mathbf{a}\rangle$ is $|-\mathbf{a}\rangle$, represented by the antipodal, oppositely pointing point on the Bloch sphere.

Example 2.5: recover Bloch angles from a ket

Consider

$$|\chi\rangle = \frac{\sqrt{3}}{2}|+z\rangle - \frac{i}{2}|-z\rangle. \tag{2.50}$$

Comparison with equation (2.45) gives $\cos(\theta/2) = \sqrt{3}/2$, so $\theta = \pi/3$. The relative phase is $e^{i\phi} = -i$, so one convenient choice is $\phi = 3\pi/2$. The state therefore lies 60° from $+z$ and points toward negative y.

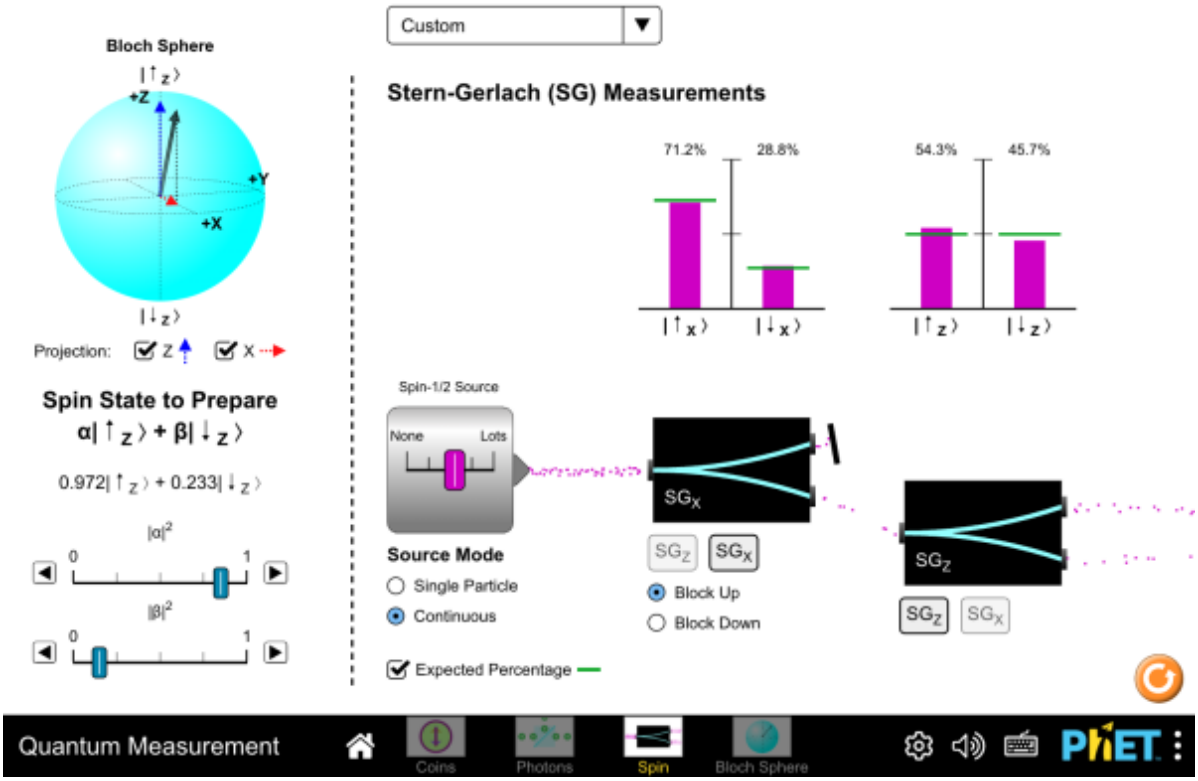

Figure 2.4: Prepare a spin-$\frac{1}{2}$ state and read its (θ, ϕ) coordinates directly off the sphere, or drag the state vector and watch the ket in equation (2.45) update to match. Reproduce Example 2.5 by dragging to $\theta = 60°$ and checking the phase the sphere reports. Interactive simulation: Quantum Measurement: Bloch Sphere

2.9 Pure states and mixtures

Section 2.1 promised that a coherent superposition and a classical mixture are not the same thing, even when they agree on some measurements. We can now make that comparison precise. A pure state assigns one ket to every member of an ideally prepared ensemble. Compare the coherent state $|+x\rangle$ with a different preparation: flip a fair classical coin, and produce $z+$ on heads and $z-$ on tails. Both preparations give identical z statistics,

$$P(z+) = P(z-) = \frac{1}{2}. \tag{2.51}$$

An x measurement, however, tells the two apart immediately. The coherent $|+x\rangle$ state gives $x+$ with certainty, by construction. For the coin-toss ensemble, half of the atoms began in $z+$ and half in $z-$; each of those subensembles independently gives 50–50 x results, so the combined statistics also remain 50–50.

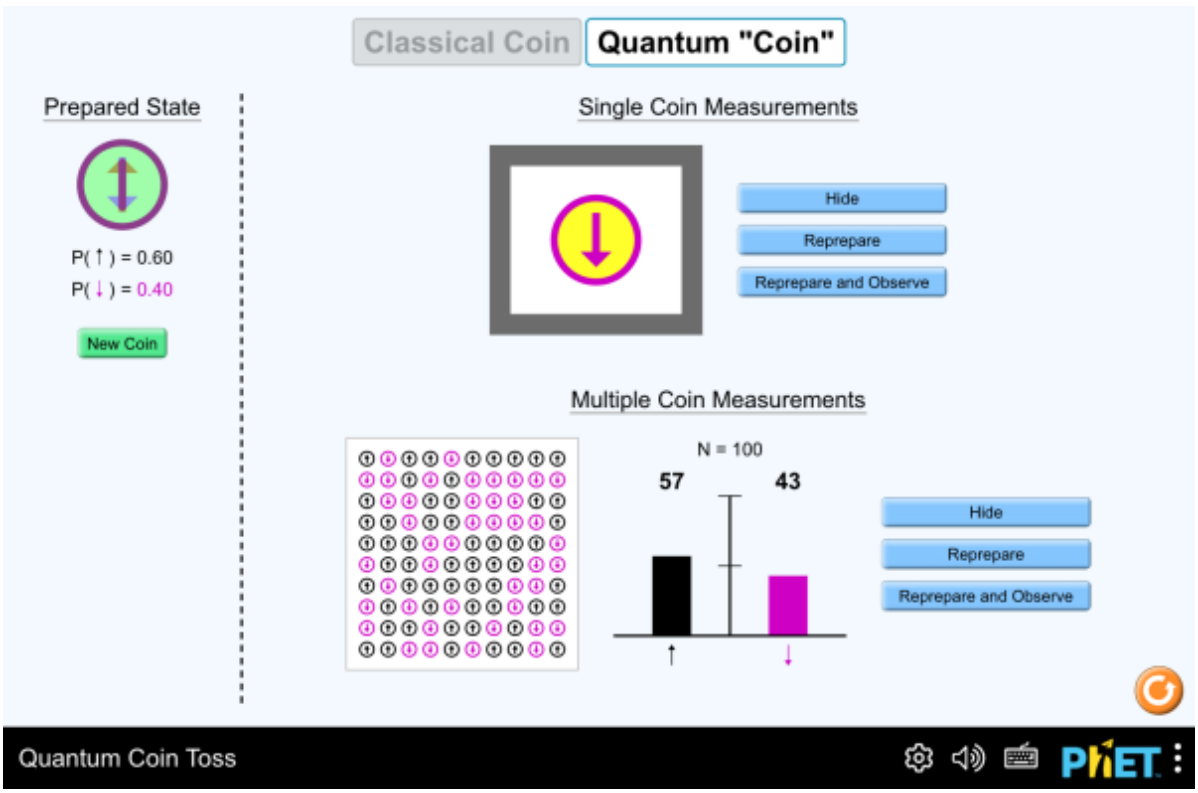

Figure 2.5: Compare a classical biased coin with a quantum "coin" prepared in superposition. Both can be tuned to land heads half the time, yet only the quantum coin has a measurement basis—an analog of the x analyzer above—in which the 50–50 outcome disappears entirely. Interactive simulation: Quantum Coin Toss

No single ket can represent that classical mixture—there is simply no linear combination of $|+z\rangle$ and $|-z\rangle$ that reproduces its statistics in *every* basis. The core operational lesson is therefore: identical probabilities for one measurement do not prove that two preparations are the same quantum state.

Optional preview: density operators

A later chapter introduces the density operator, which describes pure states and mixtures within one framework. In the z basis, the two preparations above will be represented by

$$\hat{\rho}_{\text{coh}} = |+x\rangle\langle +x| = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \tag{2.52}$$

whereas

$$\hat{\rho}_{\text{mix}} = \frac{1}{2}|+z\rangle\langle +z| + \frac{1}{2}|-z\rangle\langle -z| = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.53}$$

The off-diagonal entries present in $\hat{\rho}_{\text{coh}}$ but absent from $\hat{\rho}_{\text{mix}}$ are exactly the coherence that can influence measurements in another basis. This notation is a preview, not a prerequisite for the remaining sections or exercises.

2.10 Optional extension: spinors for an arbitrary direction

The core chapter is complete. This optional extension packages an arbitrary analyzer orientation into a pair of spinors and a basis-change matrix.

The Bloch-sphere coordinates do more than just label a state—they also give us the eigenstates of an analyzer built along any direction we like. For an analyzer oriented along

$$\mathbf{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \tag{2.54}$$

a convenient phase convention gives its two outcome states as

$$|+\mathbf{n}\rangle = \cos\frac{\theta}{2}|+z\rangle + e^{i\phi}\sin\frac{\theta}{2}|-z\rangle, \tag{2.55}$$

$$|-\mathbf{n}\rangle = -e^{-i\phi}\sin\frac{\theta}{2}|+z\rangle + \cos\frac{\theta}{2}|-z\rangle. \tag{2.56}$$

A direct calculation shows that each ket has unit norm and that $\langle +\mathbf{n}|-\mathbf{n}\rangle = 0$, so together they form a valid orthonormal basis, exactly as $|\pm z\rangle$, $|\pm x\rangle$, and $|\pm y\rangle$ did. The overall phase of either ket is conventional: multiplying $|+\mathbf{n}\rangle$ or $|-\mathbf{n}\rangle$ by its own overall phase changes its coordinates but no outcome probability. The *relative* phase written inside each ket, by contrast, is not a matter of convention once the z-basis convention has already been fixed.

Checking special cases first

Before applying equation (2.55) to an unfamiliar direction, it is a good habit to check that it reproduces the three directions already known, since a mistake in a general formula often shows up as soon as you specialize it. Setting $\theta = 0$ (pointing along $+z$) gives $|+\mathbf{n}\rangle = |+z\rangle$, exactly as it should, since the formula's ϕ-dependent term is multiplied by $\sin(\theta/2) = 0$. Setting $\theta = \pi/2,\ \phi = 0$ (pointing along $+x$) gives $|+\mathbf{n}\rangle = (|+z\rangle + |-z\rangle)/\sqrt{2} = |+x\rangle$, matching equation (2.20). Setting $\theta = \pi/2,\ \phi = \pi/2$ (pointing along $+y$) gives $|+\mathbf{n}\rangle = (|+z\rangle + i|-z\rangle)/\sqrt{2} = |+y\rangle$, matching equation (2.21). With these three checks passed, the formula can be trusted for a direction that is not simply one of the six coordinate poles.

Example 2.6: an analyzer not in a coordinate plane

Let $\theta = 2\pi/3$ and $\phi = \pi/2$. Then

$$|+\mathbf{n}\rangle = \frac{1}{2}|+z\rangle + i\frac{\sqrt{3}}{2}|-z\rangle. \tag{2.57}$$

A $z+$ input gives

$$P(+\mathbf{n} \mid z+) = |\langle +\mathbf{n}|+z\rangle|^2 = \frac{1}{4}. \tag{2.58}$$

The same result follows from the real-space angle rule: $\cos^2(\theta/2) = \cos^2(\pi/3) = 1/4$. Notice that the azimuth ϕ does not affect a $z+$ input, because rotating the analyzer around z leaves its angle from the preparation axis unchanged. It does affect inputs that carry their own transverse phase information.

The column matrix whose two columns are the two spinors above,

$$V_{\mathbf{n}} = \begin{pmatrix} \cos(\theta/2) & -e^{-i\phi}\sin(\theta/2) \\ e^{i\phi}\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}, \tag{2.59}$$

is unitary, and it converts coordinates between the analyzer basis and the z basis. Its columns describe how the *new* basis vectors look in *old* coordinates, while $V_{\mathbf{n}}^{\dagger}$ performs the reverse change, from old coordinates to new. Keeping those two directions of conversion straight prevents a common basis-change error.

Summary

- A ket is an abstract state vector; its column depends on the chosen basis.
- Inner products produce amplitudes, and squared magnitudes produce probabilities.
- An orthonormal basis supplies expansion amplitudes through inner products.
- Overall phase is unobservable, while relative phase controls interference.
- Every pure spin-$\frac{1}{2}$ state corresponds to a point on the Bloch sphere.
- A selected intermediate result resets the state used in subsequent amplitudes, while unobserved coherent alternatives must be added at the amplitude level.
- A coherent superposition and a classical mixture can agree in one basis and disagree in another.
- In the optional extension, arbitrary analyzer directions correspond to orthonormal spinors whose half-angle coordinates form a unitary basis-change matrix.
- Coherent routes are multiplied along each route and added across indistinguishable alternatives before the Born rule is applied.

Exercises

Exercises 1–7 establish the core notation and probability skills, 8–14 consolidate basis changes, phases, and mixtures, and 15–21 are extensions on arbitrary directions and coherent paths. Complete the first group before moving to the more algebraic problems.

1. Normalize $(3, 4i)^{\mathsf{T}}$ and find the probabilities of $z+$ and $z-$.

2. Verify directly that the two states in equation (2.20) are normalized and orthogonal.

3. Write $|+z\rangle$ and $|-z\rangle$ in the x basis.

4. Calculate all four quantities $|\langle \pm y| \pm x\rangle|^2$. Interpret the result as a sequence of analyzers.

5. For the state in Example 2.3, calculate $P(y+)$ and $P(y-)$. Explain why the relative phase matters here although it did not affect the x probabilities.

6. Show that equation (2.45) is normalized. Identify the Bloch-sphere angles for the six states $|\pm x\rangle$, $|\pm y\rangle$, and $|\pm z\rangle$.

7. Give two normalized kets that have identical z probabilities but are orthogonal to one another.

8. Prove that replacing both basis kets by phase-shifted basis kets changes the column coordinates but not any measurement probability.

9. Use equation (2.27) to express $(|+z\rangle + i|-z\rangle)/\sqrt{2}$ in the x basis. Check the norm before and after the transformation.

10. A $y+$ beam enters SG_z; the $z-$ output is retained and sent to SG_x. Calculate the probability for every complete detected route from the original beam.

11. Derive equation (2.38) and find $P(x-)$ without doing a second inner-product calculation.

12. Find the Bloch angles (θ, ϕ) for the normalized ket $(1, \sqrt{3}i)^{\mathsf{T}}/2$. What spin direction does it represent?

13. Without using density operators, show that a coherent $|+x\rangle$ preparation gives $x+$ with certainty while a 50–50 classical mixture of $z+$ and $z-$ gives $P(x+) = 1/2$. State where probabilities, rather than amplitudes, must be added in the mixture calculation.

14. Construct two physically different ensembles that give identical z statistics. Specify one additional analyzer orientation that distinguishes them, and calculate the predicted probabilities.

15. Verify the normalization and orthogonality of the two spinors in equation (2.55).

16. Write $|-\mathbf{n}\rangle$ for $\theta = \pi/2$ and $\phi = \pi/2$. Identify it as one of the six coordinate-axis states, up to overall phase.

17. Verify by direct matrix multiplication that $V_{\mathbf{n}}^{\dagger} V_{\mathbf{n}} = \hat{I}$ for equation (2.59).

18. Starting from equation (2.17), let $A_1 = a$ and $A_2 = be^{i\delta}$ for real nonnegative a and b. Find the largest and smallest possible total probabilities as δ varies.

19. Reproduce equations (2.40) and (2.41) with the y basis inserted instead of the x basis. Track every factor of i.

20. A phase shifter adds $e^{i\delta}$ to the $x-$ route of the recombiner. Predict both final probabilities for $\delta = \pi/2$, $2\pi/3$, and π.

21. A path marker leaves the $x+$ route unchanged but correlates the $x-$ route with a distinguishable marker state. Explain, without yet using composite state notation, why equations (2.44) no longer describe full-visibility fringes.

Selected exercise guidance

Use these answers only after making a complete attempt.

Exercise 1

The norm is 5, so the normalized ket is $(3/5, 4i/5)^{\mathsf{T}}$. The probabilities are 9/25 and 16/25.

Exercise 5

$$P(y+) = \frac{2+\sqrt{3}}{4}, \qquad P(y-) = \frac{2-\sqrt{3}}{4}. \tag{2.60}$$

The imaginary relative coefficient aligns constructively with $|+y\rangle$ and destructively with $|-y\rangle$.

Exercise 12

$\theta = 2\pi/3$ and $\phi = \pi/2$. The Bloch vector is $(0, \sqrt{3}/2, -1/2)$: it lies 120° from $+z$ toward $+y$.

Exercise 20

For $\delta = \pi/2$, $(P(z+), P(z-)) = (1/2, 1/2)$. For $2\pi/3$ the pair is $(1/4, 3/4)$, and for π it is $(0, 1)$.

Chapter 3

Operators and Measurement

3.1 Operators and measurement

Learning objectives

After this chapter, you should be able to:

- represent observables by Hermitian operators;
- calculate eigenvalues, eigenvectors, and spectral decompositions;
- apply the Born rule and the projection postulate;
- compute expectation values and uncertainties;
- use commutators to diagnose incompatible observables;
- reconstruct a spin state from component-measurement statistics;
- use Pauli-matrix identities to simplify spin calculations; and
- represent pure states, mixtures, and unread measurements with density operators.

Sections 3.8 and 3.9 are optional mathematical extensions. They prove the spectral structure used earlier and collect compact Pauli identities, but the measurement and density-operator sections do not depend on completing every derivation there.

3.1 Observables as operators

Chapter 2 treated S_z, S_x, and every other spin component separately, building a fresh pair of basis kets each time we wanted to ask a new question. That approach works, but it is repetitive: the same idea—"measure spin along some axis"—kept reappearing with different labels attached. It is both more efficient and more revealing to package that whole idea into a single mathematical object, called an **operator**. Its eigenvalues list the possible outcomes, and its eigenspaces identify the states that make those outcomes certain. Applying the operator to an arbitrary ket is an algebraic operation, not a literal simulation of one laboratory measurement.

Quick review: matrices, eigenvectors, and eigenvalues

Skip this box if diagonalizing a matrix is already routine for you.

A matrix acting on a column vector is a rule for turning one vector into another, in general rotating and rescaling it into some new direction. Remarkably, for almost any square matrix, a few special input vectors are *not* redirected at all: the matrix sends each of them straight back along its own original direction, merely stretched or shrunk by some

factor. Such a vector is called an **eigenvector** of the matrix, and the stretching factor is its **eigenvalue**.

For a concrete, purely mathematical example with no physics attached yet, take the real matrix

$$M = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}. \tag{3.1}$$

Acting on the column $(1,1)$ gives $(4,4) = 4\,(1,1)$: the same direction, scaled by 4. Acting on $(1,-1)$ gives $(2,-2) = 2\,(1,-1)$: the same direction again, this time scaled by 2. Try this on almost any other vector, say $(1,0)$, and the result, $(3,1)$, points in a genuinely new direction rather than merely rescaling $(1,0)$. So $(1,1)$ and $(1,-1)$ are eigenvectors of M, with eigenvalues 4 and 2 respectively; for this particular M, these two directions are also essentially the only ones with this special property.

This is exactly the mathematical structure a measurement needs. An eigenvector is a preparation for which "applying the operator" returns a definite, predictable multiple of the same state, rather than mixing it into something new; the multiple it returns is the number a measurement would report. Sections 3.1–3.2 make this identification precise for spin, where the vectors involved are complex kets rather than real columns like $(1,1)$.

A physical observable is represented by a **Hermitian** linear operator: one equal to its own conjugate transpose. Section 3.8 will show exactly why this property is what guarantees real, physically sensible measurement outcomes; for now, treat it as the defining requirement an operator must satisfy to represent something measurable. For spin $\frac{1}{2}$, it is convenient to introduce the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.2}$$

The spin operators are built directly from them:

$$\hat{S}_j = \frac{\hbar}{2}\sigma_j, \qquad j \in \{x, y, z\}. \tag{3.3}$$

Hermiticity, written $\hat{A}^\dagger = \hat{A}$, guarantees two things at once: real eigenvalues, and an orthonormal eigenbasis. These are exactly the two properties we need from an ideal measurement: real numbers on a meter, and a definite state associated with each possible reading.

For example, applying $\hat{S}_z$ to the two basis kets just returns them, scaled by the corresponding measured value:

$$\hat{S}_z|+z\rangle = +\frac{\hbar}{2}|+z\rangle, \qquad \hat{S}_z|-z\rangle = -\frac{\hbar}{2}|-z\rangle. \tag{3.4}$$

It is worth stressing what the possible measured values actually are: they are the **eigenvalues** of the operator, not just any numbers that happen to appear when we write out its matrix in some basis.

One apparatus, three mathematical objects

For an ideal measurement, it helps to keep three related objects distinct, since students often blur them together at first:

Laboratory meaning	Mathematical object	For S_z
A possible detector reading	eigenvalue	$+\hbar/2$ or $-\hbar/2$
A state that makes that reading certain	eigenket	\$
The test for membership in that outcome state	projector	\$

The observable operator packages all three pieces of information into one object. Solving its eigenvalue equation

$$\hat{A}|a_n\rangle = a_n|a_n\rangle \tag{3.5}$$

extracts the possible readings a_n and the corresponding definite-outcome states $|a_n\rangle$ in one stroke. The prepared state $|\psi\rangle$ need not itself be one of those eigenkets; in general, it will not be. Instead, its *projections* onto them are what determine the outcome probabilities, as the next section makes precise.

3.2 Constructing an operator from its measurement states

We can build an operator representing a measurement directly out of its outcome states, using a tool already implicit in Chapter 2's basis expansions: the **projector**. A projector is an operator that takes an arbitrary state, extracts its component along one particular direction in state space, and discards the rest. Define the projectors

$$\hat{P}_{z+} = |+z\rangle\langle+z|, \qquad \hat{P}_{z-} = |-z\rangle\langle-z|. \tag{3.6}$$

Direct from the definition, they satisfy three properties: $\hat{P}^2 = \hat{P}$ (projecting twice does nothing new), $\hat{P}^\dagger = \hat{P}$ (each is Hermitian), and $\hat{P}_{z+}\hat{P}_{z-} = 0$ (the two outcomes are mutually exclusive). Because the two basis states span the space, $\hat{P}_{z+} + \hat{P}_{z-} = \hat{I}$. This is the operator form of the basis expansion used in Chapter 2.

Combining the eigenvalues with their projectors reconstructs the full operator. The spectral decomposition of $\hat{S}_z$ is

$$\hat{S}_z = +\frac{\hbar}{2}\hat{P}_{z+} - \frac{\hbar}{2}\hat{P}_{z-}. \tag{3.7}$$

More generally, if a Hermitian observable $\hat{A}$ has nondegenerate eigenstates $|a_n\rangle$ and eigenvalues a_n, the same pattern holds for any number of outcomes:

$$\hat{A} = \sum_n a_n|a_n\rangle\langle a_n|. \tag{3.8}$$

This expression cleanly separates two pieces of information that are easy to conflate: the possible *values* a_n, and the state *projectors* associated with those values.

Example 3.1: construct $\hat{S}_x$

Using equation (2.20),

$$\hat{S}_x = \frac{\hbar}{2}\left(|+x\rangle\langle+x| - |-x\rangle\langle-x|\right) = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.9}$$

The last matrix is written in the z basis. In the x basis, the same operator is diagonal—a reminder that "the matrix of an operator" always means the matrix *in a particular basis*, never an absolute property of the operator alone.

This construction also runs in reverse: if the matrix of an observable is already known, its eigenvectors identify the corresponding analyzer states. This two-way connection deserves practice in both directions:

$$\text{measurement states and values} \quad \longleftrightarrow \quad \text{Hermitian operator.} \tag{3.10}$$

Keep in mind what the operator is *not*. It is not an additional physical substance carried around by the atom. It is simply the linear map that encodes one particular experimental question—nothing more, and nothing less.

Concept check: reading a spectral decomposition

In $\hat{A} = a_1|a_1\rangle\langle a_1| + a_2|a_2\rangle\langle a_2|$, which symbols denote detector readings, definite-outcome states, and tests for the two outcomes?

Answer

The numbers a_1, a_2 are detector readings; the kets $|a_1\rangle, |a_2\rangle$ are definite-outcome states; and the projectors $|a_1\rangle\langle a_1|, |a_2\rangle\langle a_2|$ test for those outcomes.

3.3 Measurement probabilities and state update

For a system prepared in $|\psi\rangle$, the probability of obtaining result a_n follows directly from the Born rule of Chapter 2, now written using the projector language just introduced:

$$P(a_n) = |\langle a_n|\psi\rangle|^2 = \langle\psi|\hat{P}_n|\psi\rangle, \qquad \hat{P}_n = |a_n\rangle\langle a_n|. \tag{3.11}$$

If a_n is obtained in an ideal nondegenerate measurement, the state immediately after the measurement is $|a_n\rangle$, up to an overall phase. This is the **projection postulate**: before the result is known, the theory supplies a probability distribution over outcomes; the instant a result is selected, that same theory supplies a new preparation to use going forward.

A reusable measurement workflow

Given a prepared state and an observable, the same five steps apply every time:

1. Diagonalize the observable if its eigenvalues and normalized eigenkets are not already known.
2. Build the projector $\hat{P}_n$ for each distinct outcome.
3. Compute $P(a_n) = \langle\psi|\hat{P}_n|\psi\rangle$ and check that the probabilities sum to one.
4. If an outcome is selected, apply its projector and normalize the result.
5. Use the resulting state—not the original state—for any later measurement.

This workflow is nothing more than the operator version of tracing a route through sequential Stern–Gerlach analyzers, which Chapter 1 did by hand and Chapter 2 did with explicit amplitudes.

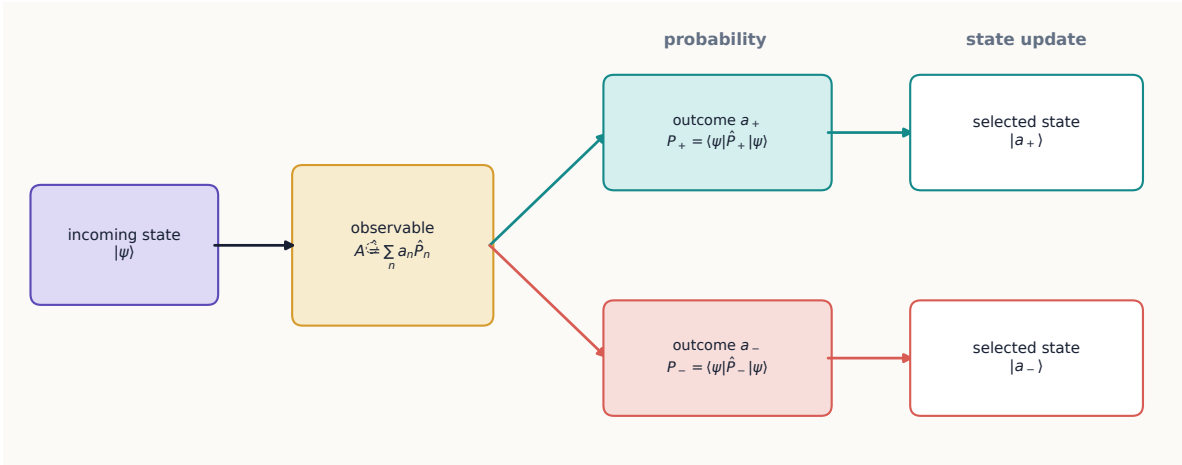

Figure 3.1: A projective measurement does two jobs. Its spectral projectors determine the probabilities of the possible records, and a selected record supplies the state used in every later step.

Example 3.2: measure along a tilted axis

Let $\mathbf{n} = (\sin\theta, 0, \cos\theta)$ lie in the xz plane. In the z basis,

$$\hat{S}_{\mathbf{n}} = \frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}. \tag{3.12}$$

Its normalized $+$ eigenket can be chosen as

$$|+\mathbf{n}\rangle = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}. \tag{3.13}$$

For a $z+$ input,

$$P(+\mathbf{n} \mid z+) = \langle +z|\hat{P}_{\mathbf{n}+}|+z\rangle = |\langle +\mathbf{n}|+z\rangle|^2 = \cos^2\frac{\theta}{2}. \tag{3.14}$$

The operator method has recovered the same half-angle rule found empirically in Chapter 1, while also identifying the exact state prepared once the $+\mathbf{n}$ result is selected, something the earlier probability-tree approach could not supply on its own.

Optional extension: degenerate outcomes

For a degenerate result, several orthogonal states share one eigenvalue. If $\hat{P}_n$ projects onto that entire eigenspace, the normalized post-measurement state generalizes to

$$|\psi'\rangle = \frac{\hat{P}_n|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_n|\psi\rangle}}. \tag{3.15}$$

The denominator here is just the square root of the probability for that outcome, ensuring $|\psi'\rangle$ comes out properly normalized.

This projector construction also makes repeatability, Chapter 1's first empirical rule, mathematically transparent. For a one-dimensional outcome space,

$$\hat{P}_n\left(\frac{\hat{P}_n|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_n|\psi\rangle}}\right) = \frac{\hat{P}_n|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_n|\psi\rangle}}, \tag{3.16}$$

because $\hat{P}_n^2 = \hat{P}_n$. Reading this equation in words: applying the same projector a second time changes nothing. A second immediate measurement therefore finds the same outcome with certainty. This is an important **consistency check** on the projection postulate: once ideal measurement has been modeled by a projector update, repeatability follows from idempotence. The algebra does not derive the update rule from the Born rule; the update itself remains part of the ideal-measurement postulate motivated by the experiments of Chapter 1.

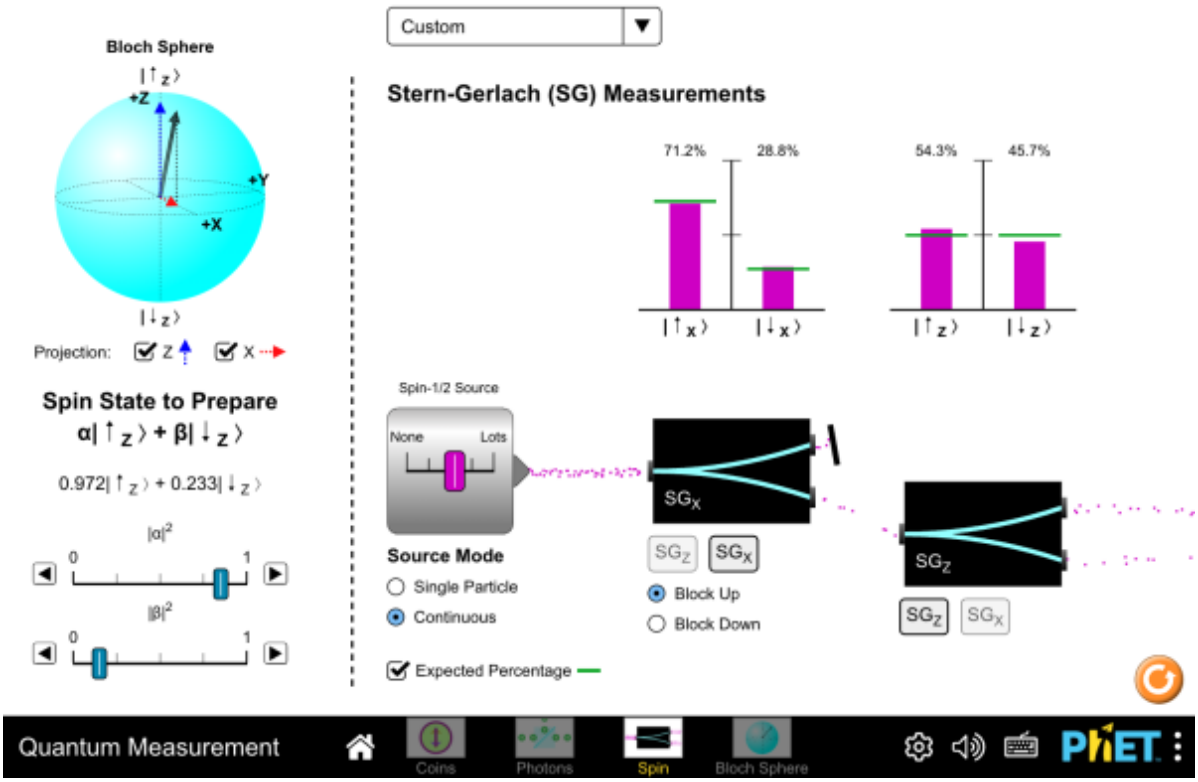

Figure 3.2: Prepare a spin-$\frac{1}{2}$ state, choose an analyzer orientation $\mathbf{n}$, and watch the projection postulate act: a Bloch-sphere readout shows the prepared state collapse to $|\pm\mathbf{n}\rangle$ the instant a result is recorded. Reproduce Example 3.2 by tilting the analyzer to angle θ from a $z+$ preparation and comparing the displayed probability with $\cos^2(\theta/2)$. Interactive simulation: Quantum Measurement: Spin

3.4 Expectation values

A single measurement returns one eigenvalue, chosen randomly according to the Born-rule probabilities. It is also useful to ask a different question: what would the *average* reading be, over many repetitions of the same preparation and measurement? This is a number a laboratory can compare directly to a meter that reports a running mean, rather than a single random outcome. We call it the **expectation value**, and it is the ensemble average predicted for many identically prepared systems:

$$\langle A \rangle = \sum_n a_n P(a_n) = \langle \psi | \hat{A} | \psi \rangle. \tag{3.17}$$

An important subtlety: the expectation value need not itself be a possible result of any single measurement. For spin $\frac{1}{2}$, $\langle S_z \rangle = 0$ is a perfectly good ensemble average even though no individual S_z measurement can ever return zero: every single measurement returns $+\hbar/2$ or $-\hbar/2$.

Example 3.3: mean spin

Let $|\psi\rangle = a|+z\rangle + b|-z\rangle$. Then

$$\langle S_z \rangle = \frac{\hbar}{2}\left(|a|^2 - |b|^2\right). \tag{3.18}$$

For the Bloch-sphere state of equation (2.45), this becomes

$$\langle S_z \rangle = \frac{\hbar}{2}\cos\theta. \tag{3.19}$$

Repeating the same calculation for $\langle S_x \rangle$ and $\langle S_y \rangle$ gives the complete result

$$\langle \mathbf{S} \rangle = \frac{\hbar}{2}(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta). \tag{3.20}$$

So the Bloch vector points along the expectation value of spin. Be careful not to overread this: it does not mean every component simultaneously *has* that value in any single measurement, only that this is the value each component averages to.

Expectation values as state tomography

For a spin-$\frac{1}{2}$ pure state, the three mean Pauli components together determine the Bloch vector completely:

$$\mathbf{r} = (\langle \sigma_x \rangle, \langle \sigma_y \rangle, \langle \sigma_z \rangle). \tag{3.21}$$

Each component is measured experimentally from counts, not read off directly. If N_{x+} and N_{x-} are the numbers measured on identically prepared members of an ensemble, then

$$\langle \sigma_x \rangle \approx \frac{N_{x+} - N_{x-}}{N_{x+} + N_{x-}}, \tag{3.22}$$

with analogous expressions for y and z. Three *separate* subensembles are required, one per axis, because the three spin components cannot all be measured sharply on the same atom: measuring one disturbs what can be known about the others. The general process of reconstructing a state from a collection of measurement statistics like this is called **quantum-state tomography**.

For a pure state, $|\mathbf{r}| = 1$ exactly. Values with $|\mathbf{r}| < 1$ describe mixed states, which Section 3.10 introduces properly; the completely unpolarized mixture sits at $\mathbf{r} = 0$. So "zero mean spin" does not, by itself, mean that every measurement returns zero—indeed, zero is not even an allowed result for a single spin-$\frac{1}{2}$ component measurement.

3.5 Variance and uncertainty

Just as the expectation value captures the *center* of a distribution of outcomes, we also want a number that captures its *spread*. The variance of A is defined, as in ordinary statistics, by

$$(\Delta A)^2 = \left\langle (\hat{A} - \langle A \rangle)^2 \right\rangle = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2. \tag{3.23}$$

For every spin component, $\hat{S}_j^2 = (\hbar^2/4)\hat{I}$, a fact that follows directly from $\sigma_j^2 = \hat{I}$. Applying this to a $z+$ state gives

$$\Delta S_z = 0, \qquad \Delta S_x = \Delta S_y = \frac{\hbar}{2}. \tag{3.24}$$

Zero uncertainty means exactly one thing: the state is an eigenstate of the observable in question. It does *not* mean that all observables are simultaneously sharp: here S_z is perfectly definite while S_x and S_y are maximally spread.

Concept check 3.1

If $\langle S_z \rangle = 0$, must the state have a definite value of S_x?

Answer

No. Every equatorial Bloch-sphere state has $\langle S_z \rangle = 0$, but only two of them, $|+x\rangle$ and $|-x\rangle$, have definite S_x.

Concept check 3.2

Can an expectation value lie outside the range of an observable's eigenvalues?

Answer

No. It is a probability-weighted average of the eigenvalues. It can lie between them, even when that intermediate value cannot occur in one measurement.

For an arbitrary component $S_{\mathbf{n}}$ in a pure state with Bloch vector $\mathbf{r}$, the two possible outcomes immediately give both the mean and the spread in closed form:

$$\langle S_{\mathbf{n}} \rangle = \frac{\hbar}{2}\mathbf{n} \cdot \mathbf{r}, \qquad \Delta S_{\mathbf{n}} = \frac{\hbar}{2}\sqrt{1 - (\mathbf{n} \cdot \mathbf{r})^2}. \tag{3.25}$$

The uncertainty vanishes precisely when the analyzer points along or opposite the Bloch vector, and it is maximal for every direction perpendicular to it. This is the cleanest way to see that uncertainty is a property of a *state relative to a specified measurement*—not some intrinsic fuzziness of the state considered alone.

Concept check 3.3

A student calculates $\Delta S_z = 0$ in a state $|\psi\rangle$. What can be concluded about $|\psi\rangle$?

Answer

It must be an eigenstate of $\hat{S}_z$, so it is $|+z\rangle$ or $|-z\rangle$ up to an overall phase. The calculation does not imply zero uncertainty for S_x or S_y; each of those uncertainties is $\hbar/2$.

3.6 Commutators and incompatibility

Chapter 1 showed experimentally that measuring S_x then S_z gives a different final distribution than measuring S_z then S_x: order matters for incompatible observables. We now want a single algebraic object that captures this order-dependence directly, without having to redo the Stern–Gerlach bookkeeping every time. The **commutator** of two operators is

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{3.26}$$

If $\hat{A}$ and $\hat{B}$ commute, the order in which we apply them makes no difference; if the commutator is nonzero, order matters. It is worth multiplying the Pauli matrices out by hand once, so that the result is not taken purely on faith. Using equation (3.2),

$$\sigma_x\sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = i\sigma_z, \tag{3.27}$$

while multiplying in the opposite order gives

$$\sigma_y\sigma_x = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = -i\sigma_z. \tag{3.28}$$

The two products are not equal—already a direct demonstration that matrix multiplication need not commute—and their difference is $\sigma_x\sigma_y - \sigma_y\sigma_x = 2i\sigma_z$. Inserting $\hat{S}_j = (\hbar/2)\sigma_j$ from equation (3.3) then gives

$$[\hat{S}_x, \hat{S}_y] = \left(\frac{\hbar}{2}\right)^2 (2i\sigma_z) = i\hbar\left(\frac{\hbar}{2}\sigma_z\right) = i\hbar\hat{S}_z, \tag{3.29}$$

with cyclic permutations for the other pairs, obtained by the same steps with $x \to y \to z \to x$ relabeled throughout. This nonzero commutator is the precise algebraic counterpart of the sequential-analyzer behavior observed in Chapter 1.

The commutator also controls how sharply two observables can be simultaneously known. For any state,

$$\Delta A\,\Delta B \geq \frac{1}{2}\left|\langle[\hat{A}, \hat{B}]\rangle\right|. \tag{3.30}$$

In a $z+$ state, equation (3.30) gives $\Delta S_x \Delta S_y \geq \hbar^2/4$; both uncertainties equal $\hbar/2$, so the bound is saturated exactly—there is no slack left.

The right-hand side is **state dependent**. A nonzero commutator operator does not guarantee a positive numerical lower bound in every state. For example, in $|+z\rangle$,

$$\langle[\hat{S}_x, \hat{S}_z]\rangle = -i\hbar\langle S_y\rangle = 0, \tag{3.31}$$

so Robertson gives only $\Delta S_x \Delta S_z \geq 0$. The observables are still incompatible; this particular bound is merely uninformative for this state. Indeed $\Delta S_z = 0$ and $\Delta S_x = \hbar/2$. Noncommutation is an operator statement, whereas the numerical lower bound also depends on the preparation.

A common misreading

An uncertainty relation describes the spread of outcomes for ensembles prepared in the same state. It is not merely a statement about poor instruments or a careless observer.

Noncommutation is closely related to analyzer order, but be careful not to equate an operator product directly with "perform one projective measurement and then the other." A selected sequence also includes a state update in between, which the bare commutator does not track by itself. For example, the route probability for $z+ \to x+ \to z-$ is

$$\|\hat{P}_{z-}\hat{P}_{x+}|+z\rangle\|^2 = \frac{1}{4}. \tag{3.32}$$

Reversing the two projectors changes the physical route being described: it answers a different experimental question, not just a reordered version of the same one. The projector product keeps both the ordering and the conditional selection explicit, which the commutator alone does not.

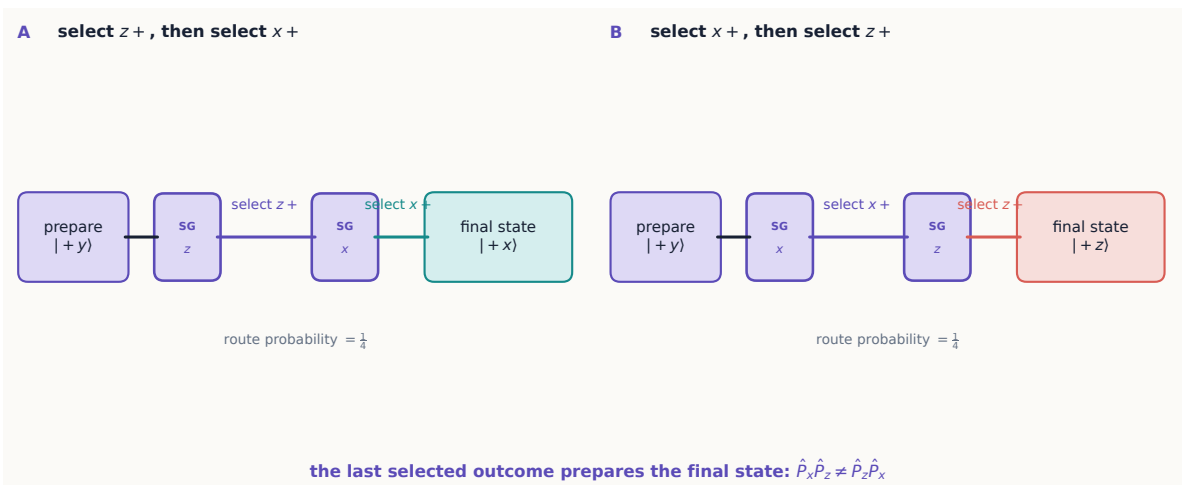

Figure 3.3: Reversing two selected measurements reverses which eigenstate is prepared last. Even when the displayed routes happen to have equal probabilities, their final states differ: $\hat{P}_x\hat{P}_z$ and $\hat{P}_z\hat{P}_x$ are not the same operation.

3.7 Operators along an arbitrary direction

For a unit vector $\mathbf{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, define the spin component along that direction as

$$\hat{S}_{\mathbf{n}} = \mathbf{n}\cdot\hat{\mathbf{S}} = \frac{\hbar}{2}\mathbf{n}\cdot\boldsymbol{\sigma}. \tag{3.33}$$

Its eigenvalues are $\pm\hbar/2$, for any choice of $\mathbf{n}$, a fact consistent with Chapter 1's observation that every analyzer, no matter how it is oriented, gives exactly two outcomes. Its $+$ eigenstate is the Bloch-sphere state in equation (2.45), up to an overall phase. The compact identity

$$(\mathbf{n}\cdot\boldsymbol{\sigma})^2 = \hat{I} \tag{3.34}$$

makes the eigenvalues immediate, without any further diagonalization.

The associated outcome projectors can be written down directly, without first solving for the eigenvectors at all:

$$\hat{P}_{\mathbf{n}\pm} = \frac{1}{2}(\hat{I} \pm \mathbf{n}\cdot\boldsymbol{\sigma}). \tag{3.35}$$

They sum to the identity, as any pair of complementary projectors must, and satisfy $(\mathbf{n} \cdot \boldsymbol{\sigma})\hat{P}_{\mathbf{n}\pm} = \pm\hat{P}_{\mathbf{n}\pm}$. For a state with Bloch vector $\mathbf{r}$, equation (3.35) yields a compact probability rule that is often faster to use than working with explicit eigenkets:

$$P(\pm\mathbf{n}) = \frac{1}{2}(1 \pm \mathbf{n} \cdot \mathbf{r}). \tag{3.36}$$

This is the operator version of equation (2.49), and it remains useful even for *mixed* states, whose Bloch vectors lie strictly inside the sphere rather than on its surface.

3.8 Optional extension: why Hermitian operators have the needed structure

Section 3.1 asserted that Hermiticity guarantees real eigenvalues and orthogonal eigenstates. It is worth proving both claims directly because they explain why Hermitian operators fit ideal observables in finite-dimensional quantum mechanics.

Suppose

$$\hat{A}|a\rangle = a|a\rangle \tag{3.37}$$

for a normalized eigenket. Taking the inner product with $\langle a|$ gives

$$a = \langle a|\hat{A}|a\rangle. \tag{3.38}$$

For a Hermitian operator, complex-conjugating this expression and using $\hat{A}^\dagger = \hat{A}$ gives

$$\langle a|\hat{A}|a\rangle^* = \langle a|\hat{A}^\dagger|a\rangle = \langle a|\hat{A}|a\rangle, \tag{3.39}$$

so $a = a^*$, meaning the eigenvalue is real. This is precisely what we need: detector readings can be real numbers with physical units, not complex quantities with no direct meter reading.

Now let $|a\rangle$ and $|b\rangle$ have distinct eigenvalues a and b. Hermiticity gives

$$\langle a|\hat{A}|b\rangle = b\langle a|b\rangle \tag{3.40}$$

but also, moving $\hat{A}$ to act on the bra instead,

$$\langle a|\hat{A}|b\rangle = a\langle a|b\rangle. \tag{3.41}$$

Subtracting these two expressions for the same quantity gives $(a-b)\langle a|b\rangle = 0$. If $a \neq b$, the only way this can hold is $\langle a|b\rangle = 0$: states belonging to distinct ideal outcomes are automatically orthogonal. In a degenerate eigenspace, where several states share one eigenvalue, orthonormal eigenvectors can still be chosen within that subspace, but the measurement outcome alone does not identify which particular vector within it was present.

Together these two facts give the finite-dimensional spectral theorem: every Hermitian operator has an orthonormal eigenbasis, and can be written

$$\hat{A} = \sum_n a_n \hat{P}_n, \tag{3.42}$$

where $\hat{P}_n$ projects onto the full eigenspace belonging to the distinct value a_n. The nondegenerate formula already used in equation (3.8) is just the special case $\hat{P}_n = |a_n\rangle\langle a_n|$, where each eigenspace happens to be one-dimensional.

Example 3.4: diagonalize a two-state observable

Consider

$$\hat{A} = \begin{pmatrix} 2 & 1-i \\ 1+i & 4 \end{pmatrix}. \tag{3.43}$$

The matrix is Hermitian because its off-diagonal entries are complex conjugates of each other. Its characteristic equation is

$$\det(\hat{A} - \lambda\hat{I}) = (2-\lambda)(4-\lambda) - 2 = \lambda^2 - 6\lambda + 6 = 0, \tag{3.44}$$

so the possible outcomes are the real values $\lambda_\pm = 3 \pm \sqrt{3}$—real, exactly as Hermiticity guaranteed in advance.

A quick check catches most sign or arithmetic errors before going any further: the sum of the eigenvalues must equal the trace of $\hat{A}$, and their product must equal its determinant. Here $\lambda_+ + \lambda_- = 6 = 2+4 = \operatorname{Tr}\hat{A}$, and $\lambda_+\lambda_- = (3+\sqrt{3})(3-\sqrt{3}) = 9-3 = 6$, which indeed matches the determinant computed above, $2 \cdot 4 - |1-i|^2 = 8 - 2 = 6$. Both checks pass, so it is safe to proceed.

To find the eigenvector for λ_+, write $|\lambda_+\rangle = (v_1, v_2)$ and use either row of $(\hat{A} - \lambda_+\hat{I})|\lambda_+\rangle = 0$: the two rows are not independent equations once λ_+ is an exact root, so either one alone determines the ratio v_2/v_1. The second row reads $(1+i)v_1 + (4-\lambda_+)v_2 = 0$, so

$$v_2 = -\frac{1+i}{4-\lambda_+}v_1 = -\frac{1+i}{1-\sqrt{3}}v_1. \tag{3.45}$$

Choosing $v_1 = 1$ gives a squared norm

$$1 + \left|\frac{1+i}{\sqrt{3}-1}\right|^2 = 3 + \sqrt{3}, \tag{3.46}$$

so one normalized eigenket is

$$|\lambda_+\rangle = \frac{1}{\sqrt{3+\sqrt{3}}}\begin{pmatrix} 1 \\ \dfrac{1+i}{\sqrt{3}-1} \end{pmatrix}. \tag{3.47}$$

For $\lambda_- = 3 - \sqrt{3}$, the second row instead gives $v_2 = -(1+i)/(1+\sqrt{3})$ when $v_1 = 1$. Its squared norm is $3 - \sqrt{3}$, so

$$|\lambda_-\rangle = \frac{1}{\sqrt{3-\sqrt{3}}}\begin{pmatrix} 1 \\ -\dfrac{1+i}{1+\sqrt{3}} \end{pmatrix}. \tag{3.48}$$

Before trusting the result, check the inner product. The product of the two lower components before normalization is

$$\left(\frac{1+i}{\sqrt{3}-1}\right)^* \left(-\frac{1+i}{1+\sqrt{3}}\right) = -1, \tag{3.49}$$

which cancels the product 1^*1 of the upper components. Thus $\langle\lambda_+|\lambda_-\rangle = 0$. We have now completed the full workflow: find the eigenvalues, solve for each component ratio, normalize, and check orthogonality. Hermiticity told us in advance that the roots would be real and that such an orthonormal eigenbasis must exist.

3.9 Optional extension: the Pauli algebra as a calculation tool

Many spin calculations become much shorter once a few Pauli-matrix identities are available, so they are collected here as tools rather than rederived from scratch each time. Direct multiplication of the matrices in equation (3.2) gives

$$\sigma_i \sigma_j = \delta_{ij} \hat{I} + i \sum_k \epsilon_{ijk} \sigma_k. \tag{3.50}$$

Here δ_{ij} is one when the indices agree and zero otherwise, while ϵ_{ijk} supplies the sign associated with a cyclic ordering of x, y, z. Packaging all nine possible products into a single vector identity gives

$$(\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) = (\mathbf{a} \cdot \mathbf{b})\hat{I} + i(\mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma}. \tag{3.51}$$

Setting $\mathbf{a} = \mathbf{b} = \mathbf{n}$ immediately recovers $(\mathbf{n} \cdot \boldsymbol{\sigma})^2 = \hat{I}$ for a unit vector, without any further calculation. Subtracting the same formula with $\mathbf{a}$ and $\mathbf{b}$ exchanged—and using $\mathbf{b} \times \mathbf{a} = -\mathbf{a} \times \mathbf{b}$—gives the commutator for components along two arbitrary directions:

$$[\hat{S}_{\mathbf{a}}, \hat{S}_{\mathbf{b}}] = i\hbar\, \hat{\mathbf{S}} \cdot (\mathbf{a} \times \mathbf{b}). \tag{3.52}$$

Parallel components commute, since their cross product vanishes; perpendicular components have the largest possible commutator magnitude. In this way, the vector geometry of analyzer directions is encoded directly in the operator algebra, rather than needing to be checked case by case.

Example 3.5: two projectors in succession

Using equation (3.35),

$$\hat{P}_{\mathbf{b}+} \hat{P}_{\mathbf{a}+} = \frac{1}{4} \left[(1 + \mathbf{a} \cdot \mathbf{b})\hat{I} + (\mathbf{a} + \mathbf{b}) \cdot \boldsymbol{\sigma} + i(\mathbf{b} \times \mathbf{a}) \cdot \boldsymbol{\sigma} \right]. \tag{3.53}$$

The imaginary cross-product term changes sign if the order is reversed, so the two operators are identical only for parallel or antiparallel axes; any other pair of directions gives an order-dependent result. Acting on $|+\mathbf{a}\rangle$ and taking the squared norm yields $(1 + \mathbf{a} \cdot \mathbf{b})/2$, recovering the analyzer angle rule while also retaining the post-selected output state, something the bare probability number alone would not do.

3.10 Density operators and unread measurements

A ket describes a *pure* preparation: every atom in the ensemble was prepared exactly the same way. But real experiments sometimes involve a classical mixture of different preparations—an ensemble in which state $|\psi_j\rangle$ occurs with classical probability w_j. This is represented by the **density operator**

$$\hat{\rho} = \sum_j w_j |\psi_j\rangle\langle\psi_j|, \qquad w_j \geq 0, \quad \sum_j w_j = 1. \tag{3.54}$$

Recall that the trace of an operator is the sum of its diagonal entries in any orthonormal basis; the result does not depend on which basis is chosen. (This is the same fact used as a check in Example 3.4: the trace, computed from the diagonal entries, equals the sum of the eigenvalues, computed from an entirely different calculation.) A density operator is always

Hermitian, has trace one, and has no negative eigenvalues. In this language, probabilities and expectation values take compact forms that work for pure states and mixtures alike:

$$P(a_n) = \mathrm{Tr}(\hat{\rho}\hat{P}_n), \qquad \langle A \rangle = \mathrm{Tr}(\hat{\rho}\hat{A}). \tag{3.55}$$

For a pure state, $\hat{\rho} = |\psi\rangle\langle\psi|$ and $\hat{\rho}^2 = \hat{\rho}$, so $\mathrm{Tr}(\hat{\rho}^2) = 1$. A genuine mixture instead has $\mathrm{Tr}(\hat{\rho}^2) < 1$. This **purity** test gives a single number that tells pure states and mixtures apart, and it does not depend on which basis we use to compute it.

Every spin-$\frac{1}{2}$ density operator, pure or mixed, can be written in Bloch-vector form:

$$\hat{\rho} = \frac{1}{2}(\hat{I} + \mathbf{r} \cdot \boldsymbol{\sigma}), \qquad |\mathbf{r}| \leq 1. \tag{3.56}$$

Pure states lie exactly on the surface of the Bloch sphere, where $|\mathbf{r}| = 1$; mixed states lie strictly inside it. The completely unpolarized state $\hat{\rho} = \hat{I}/2$ sits at the very center, $\mathbf{r} = 0$, and predicts equal probabilities for every possible analyzer orientation.

Example 3.6: tomography from three analyzer settings

Suppose three equal subensembles produce

$$P(x+) = 0.80, \qquad P(y+) = 0.10, \qquad P(z+) = 0.50. \tag{3.57}$$

Since $r_j = P(j+) - P(j-) = 2P(j+) - 1$ for each axis,

$$\mathbf{r} = (0.60, -0.80, 0). \tag{3.58}$$

Its length is exactly one, so this idealized data describes a pure equatorial state. A representative ket reproducing it is

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|+z\rangle + e^{i\phi}|-z\rangle\Big), \qquad e^{i\phi} = 0.60 - 0.80i. \tag{3.59}$$

In a real experiment, finite-sample noise means the reconstructed vector may come out slightly outside the unit sphere, with $|\mathbf{r}|$ marginally greater than one. A statistically sound tomography method then finds the nearest physically allowed density operator, rather than trying to interpret $|\mathbf{r}| > 1$ as if it described a valid quantum state.

An outcome ignored is not a measurement undone

Here is a question that trips up many students on first encounter: if an ideal measurement occurs but nobody looks at the result, has anything changed? The answer is yes—and understanding why is one of the more important lessons in this chapter. If an ideal projective measurement occurs but its result is not retained, the post-measurement ensemble is

$$\hat{\rho}' = \sum_n \hat{P}_n \hat{\rho} \hat{P}_n. \tag{3.60}$$

For an initial $|+x\rangle$ state followed by an unread S_z measurement,

$$\hat{\rho}' = \hat{P}_{z+}|+x\rangle\langle+x|\hat{P}_{z+} + \hat{P}_{z-}|+x\rangle\langle+x|\hat{P}_{z-} = \frac{1}{2}\hat{P}_{z+} + \frac{1}{2}\hat{P}_{z-} = \frac{1}{2}\hat{I}. \tag{3.61}$$

Notice what happened to the off-diagonal z-basis terms: they have disappeared entirely. As a direct consequence, a later S_x measurement is now 50–50, even though no observer ever

actually used the intermediate result. Discarding a record after the fact is not the same as preventing that record from being created in the first place. It is the physical correlation that made the alternatives distinguishable—not anyone's act of reading it—that destroyed the interference. This is exactly the distinction the quantum-coin-toss simulation of Chapter 2 makes tangible: a coin that has decohered into a fixed outcome behaves just like $\hat{\rho}'$ above, no matter whether that outcome has been read yet.

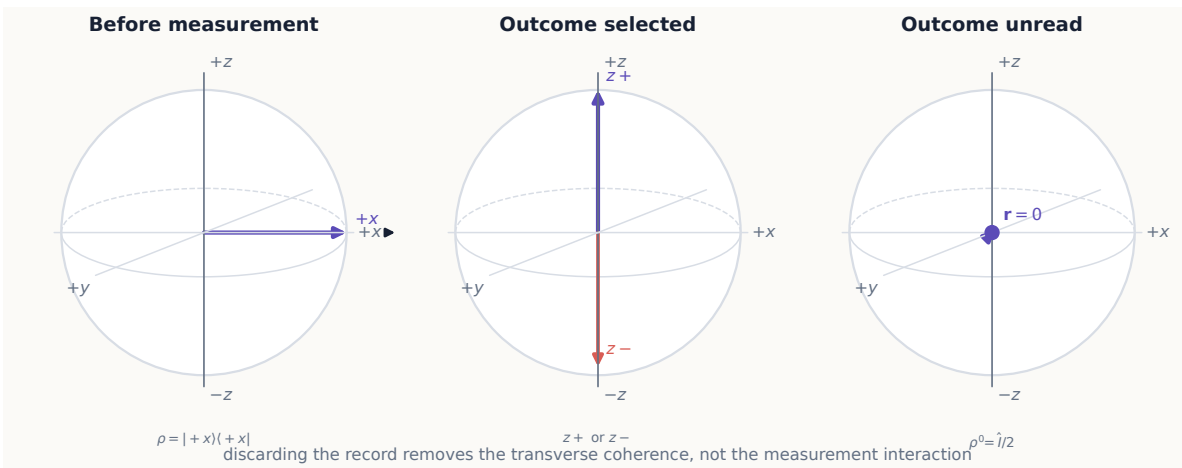

Figure 3.4: An S_z measurement sends each member of the ensemble to a pole. Keeping the outcome selects one pole; pooling the outcomes leaves no net Bloch vector and removes the transverse coherence of the initial $|+x\rangle$ state.

Concept check 3.4

Can the same density operator arise from different preparation recipes?

Answer

Yes. For example, equal mixtures of $z+$ and $z-$, of $x+$ and $x-$, or of $y+$ and $y-$ all give $\hat{I}/2$. No measurement on the spin alone can distinguish preparation recipes that produce the same density operator.

Concept check 3.5

An unread S_z measurement is made on a $z+$ state. Does its density operator change?

Answer

No. Equation (3.60) returns $\hat{P}_{z+}\hat{\rho}\hat{P}_{z+} = \hat{\rho}$ and a zero contribution from the other projector. The state was already definite for the measured observable.

By contrast, if the result a_n *is* retained, the conditional density operator is

$$\hat{\rho}'_n = \frac{\hat{P}_n\hat{\rho}\hat{P}_n}{\mathrm{Tr}(\hat{\rho}\hat{P}_n)}. \tag{3.62}$$

Equations (3.60) and (3.62) distinguish two common laboratory instructions that are easy to conflate in words but behave quite differently: "measure and forget the result" versus "measure and select one result."

Summary

- Hermitian operators represent observables; their eigenvalues are possible outcomes and their eigenvectors are definite-outcome states.
- Projectors give both outcome probabilities and post-measurement states.
- An expectation value is an ensemble mean, not generally a single outcome.
- Variance quantifies the spread of repeated measurements on identically prepared systems.
- Noncommuting observables encode order dependence and uncertainty relations.
- Projector products describe selected routes through sequential measurements.
- The three mean Pauli components reconstruct a spin state's Bloch vector.
- Hermiticity guarantees real outcomes and orthogonal eigenspaces for distinct outcomes.
- Density operators describe pure states and mixtures in one framework; unread measurements remove coherence in the measured basis.

Exercises

Exercises 1–7 build the essential operator, projector, and expectation-value skills. Exercises 8–14 consolidate sequential measurement and state reconstruction. Exercises 15–22 draw on the optional mathematical extensions or density-operator material. A first pass can therefore use 1–7, 10–12, 18, and 20–21 without treating the entire list as one assignment.

1. Find the eigenvalues and normalized eigenvectors of each Pauli matrix.

2. Verify the projector properties for $\hat{P}_{x+}$ and $\hat{P}_{x-}$ using explicit z-basis matrices.

3. For $|\psi\rangle = (|+z\rangle + i|-z\rangle)/\sqrt{2}$, calculate $\langle S_x\rangle$, $\langle S_y\rangle$, and $\langle S_z\rangle$.

4. Calculate $[\hat{S}_y, \hat{S}_z]$ by direct matrix multiplication.

5. Show that $(\mathbf{n}\cdot\boldsymbol{\sigma})^2 = \hat{I}$ for any unit vector $\mathbf{n}$.

6. Derive the projectors $\hat{P}_{\mathbf{n}\pm} = (\hat{I} \pm \mathbf{n}\cdot\boldsymbol{\sigma})/2$.

7. A state gives $P(x+) = 3/4$. Find $\langle S_x\rangle$ and ΔS_x without first finding a ket.

8. Give an example showing that $\langle \hat{A}\hat{B}\rangle$ and $\langle \hat{B}\hat{A}\rangle$ can differ. Explain how the commutator measures the difference.

9. For $\mathbf{n} = (\sqrt{3}/2, 0, 1/2)$, write the z-basis matrix of $\hat{S}_{\mathbf{n}}$ and verify the proposed $|+\mathbf{n}\rangle$ eigenvector in Example 3.2.

10. Starting with $|+z\rangle$, calculate the state and probability associated with the selected route $z+ \to y- \to x+ \to z-$ using a product of projectors.

11. An ensemble produces counts $(N_{x+}, N_{x-}) = (750, 250)$, $(N_{y+}, N_{y-}) = (500, 500)$, and $(N_{z+}, N_{z-}) = (933, 67)$ on three equal subensembles. Estimate its Bloch vector and decide whether the data are approximately consistent with a pure state.

12. Derive equation (3.25) from the two outcome probabilities in equation (3.36).

13. Prove all three projector properties for equation (3.35): Hermiticity, idempotence, and orthogonality of the $+$ and $-$ projectors.

14. Compare the two route probabilities $\|\hat{P}_{z+}\hat{P}_{x+}|+z\rangle\|^2$ and $\|\hat{P}_{x+}\hat{P}_{z+}|+z\rangle\|^2$. Explain why they answer different experimental questions even though they contain the same two projectors.

15. For Example 3.4, verify both eigenvalue equations with the normalized eigenkets given in the text, then reconstruct $\hat{A}$ from its spectral decomposition.

16. Derive equation (3.51) by expanding both dot products and using equation (3.50).

17. Use equation (3.52) to calculate the commutator of components along $\mathbf{a} = (1, 1, 0)/\sqrt{2}$ and $\mathbf{b} = (0, 1, 1)/\sqrt{2}$.

18. Find the density matrix and purity for a mixture containing $z+$ with probability 3/4 and $z-$ with probability 1/4. Draw or describe its Bloch vector.

19. Show from equation (3.56) that $\mathrm{Tr}(\hat{\rho}^2) = (1 + |\mathbf{r}|^2)/2$.

20. Verify every line of the unread-measurement calculation for an initial $|+x\rangle$ state using explicit 2×2 matrices.

21. An initial $|+y\rangle$ state undergoes an unread S_x measurement. Find the final density matrix and predict subsequent x, y, and z statistics.

22. A detector selects $+\mathbf{n}$ from a mixed input $\hat{\rho} = (\hat{I} + r\sigma_z)/2$. Use equation (3.62) to show that the transmitted ensemble is the pure state $|+\mathbf{n}\rangle$, provided the transmission probability is nonzero.

Selected exercise guidance

Use these answers only after making a complete attempt.

Exercise 3

The state is $|+y\rangle$, so $(\langle S_x\rangle, \langle S_y\rangle, \langle S_z\rangle) = (0, \hbar/2, 0)$.

Exercise 7

$\langle S_x\rangle = (\hbar/2)(3/4 - 1/4) = \hbar/4$. Since $\langle S_x^2\rangle = \hbar^2/4$, $\Delta S_x = \sqrt{3}\hbar/4$.

Exercise 11

The estimated Bloch vector is $\mathbf{r} = (0.500, 0, 0.866)$. Its length is approximately one, so the idealized counts are consistent with a nearly pure state.

Exercise 18

$$\hat{\rho} = \begin{pmatrix} 3/4 & 0 \\ 0 & 1/4 \end{pmatrix}, \qquad \mathrm{Tr}(\hat{\rho}^2) = \frac{5}{8}, \tag{3.63}$$

and its Bloch vector is $(0, 0, 1/2)$.

Chapter 4

Time Evolution

4.1 Time evolution

Learning objectives

After this chapter, you should be able to:

- use the Hamiltonian to evolve a quantum state in time;
- construct and interpret the unitary time-evolution operator;
- distinguish stationary states from general superpositions;
- predict spin precession in a uniform magnetic field;
- connect relative phase evolution to changing measurement probabilities;
- compose piecewise evolution and measurement steps in a complete protocol;
- translate controlled field pulses into Bloch-sphere rotations; and
- predict a Ramsey interference signal.

Sections 4.1–4.7 form the core path through unitary evolution and spin precession. Sections 4.8–4.9 are mathematical extensions, while Sections 4.10–4.11 form an applications track on laboratory pulses and Ramsey interference. Those tracks may be assigned independently after the core.

4.1 The dynamical postulate

Every chapter so far has treated the spin state as something that only changes at the instant of a measurement. Between measurements, we have simply left the state alone. That gap needs to be filled: real experiments involve waiting time, traveling atoms, and applied fields, and the state must be doing *something* during all of that. This chapter supplies the missing piece. Between measurements, the state of an isolated system evolves according to the **time-dependent Schrödinger equation**,

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle. \tag{4.1}$$

The Hamiltonian $\hat{H}$ is the observable associated with energy, and it plays a second role here as the *generator* of time evolution: the object that tells the state how to change from one instant to the next. Equation (4.1) is a first-order differential equation for a vector, which means specifying the initial state $|\psi(0)\rangle$ and the Hamiltonian $\hat{H}(t)$ is enough to determine the state at every later time.

Quick review: what does it mean to exponentiate an operator?
Skip this box if exponentiating a matrix is already routine for you.

For an ordinary number x, the exponential function is defined by its power series,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots . \tag{4.2}$$

The very same series makes sense if x is replaced by a square matrix X, since matrices can be multiplied by themselves and added together just like numbers:

$$e^X = \hat{I} + X + \frac{X^2}{2!} + \frac{X^3}{3!} + \cdots . \tag{4.3}$$

This infinite sum looks intimidating, but it collapses to something simple whenever X is diagonal, say $X = \mathrm{diag}(x_1, x_2)$. Every power of a diagonal matrix is just the same diagonal matrix with each entry raised to that power, so the whole series acts independently, entry by entry:

$$e^X = \mathrm{diag}(e^{x_1}, e^{x_2}). \tag{4.4}$$

Exponentiating a diagonal matrix, in other words, is nothing more than exponentiating each diagonal number separately, then reassembling the result. This is precisely the strategy used throughout this chapter: diagonalize $\hat{H}$ first—its eigenvalues are the possible energies— exponentiate the resulting simple diagonal matrix, and transform back to whichever basis is convenient. You have already met a 1×1 special case of this idea, under a different name, in Chapter 2's Euler formula $e^{i\phi} = \cos\phi + i \sin\phi$.

When the Hamiltonian does not depend on time, this differential equation has an explicit solution:

$$|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle, \qquad \hat{U}(t) = e^{-i\hat{H}t/\hbar}. \tag{4.5}$$

Check directly that this solves equation (4.1), rather than accepting it purely on authority. Differentiating the defining power series of $\hat{U}(t)$ term by term with respect to t,

$$\frac{d\hat{U}(t)}{dt} = \frac{d}{dt}\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{-i\hat{H}t}{\hbar}\right)^n = -\frac{i\hat{H}}{\hbar}\hat{U}(t), \tag{4.6}$$

because $\hat{H}$ commutes with every power of itself, so each term of the series differentiates exactly as an ordinary power of t would. Multiplying both sides by $i\hbar$ and applying the result to $|\psi(0)\rangle$ reproduces equation (4.1) exactly. This calculation used only that $\hat{H}$ does not depend on time; nothing about spin specifically entered.

Because $\hat{H}$ is Hermitian, the operator $\hat{U}$ built from it is automatically **unitary**: $\hat{U}^\dagger\hat{U} = \hat{I}$. This single fact has an important consequence. It guarantees that a state's normalization never drifts as it evolves:

$$\langle\psi(t)|\psi(t)\rangle = \langle\psi(0)|\hat{U}^\dagger\hat{U}|\psi(0)\rangle = 1. \tag{4.7}$$

So unitary evolution preserves normalization and, more generally, every inner product between two evolving states. Something here can feel paradoxical at first: within the ideal

isolated-system model, this evolution is completely *deterministic*, even though the outcomes of a later projective measurement are probabilistic. The Schrödinger equation itself introduces no random step; the measurement postulate supplies probabilities when a measurement is made.

What an evolution problem is asking

Most two-state dynamics problems, however different they look on the surface, can be organized into the same five-step sequence:

1. Write the Hamiltonian and choose a basis in which to represent it.
2. Construct $\hat{U}(t)$ by using energy eigenstates, diagonalizing the matrix, or exploiting Pauli-matrix identities.
3. Apply $\hat{U}(t)$ to the **initial** ket.
4. Project the evolved ket onto the states of the measurement performed at time t.
5. Check normalization, the $t = 0$ limit, and any conserved probabilities.

The distinction between steps 3 and 4 is the single most common source of error in this material, so it is worth stating plainly: the ket evolves continuously under $\hat{U}(t)$, but a *probability* is only ever defined once a measurement question has been specified. Two different choices of final analyzer can reveal very different time dependence from the very same evolving state, so step 4 cannot be skipped or guessed.

For a time-independent Hamiltonian, the exponential in step 2 can always be understood through its power series, which is most useful once $\hat{H}$ has been diagonalized. If $\hat{H} = \hat{V}\hat{D}\hat{V}^\dagger$, with $\hat{D}$ diagonal, then

$$e^{-i\hat{H}t/\hbar} = \hat{V}e^{-i\hat{D}t/\hbar}\hat{V}^\dagger. \tag{4.8}$$

In words: diagonalize the Hamiltonian, let each diagonal component evolve by its own simple phase, and transform back to the original basis. This is a completely general method for solving the Schrödinger equation, not a trick that happens to work only for spin.

4.2 Energy eigenstates and phases

Let $|E_n\rangle$ satisfy the time-independent eigenvalue equation

$$\hat{H}|E_n\rangle = E_n|E_n\rangle. \tag{4.9}$$

Because the evolution operator is built from the exponential of $\hat{H}$, it acts especially simply on one of $\hat{H}$'s own eigenstates:

$$|E_n, t\rangle = e^{-iE_nt/\hbar}|E_n\rangle. \tag{4.10}$$

Only an overall phase changes here, and Chapter 2 established that an overall phase never affects any measurement probability. For this reason, an energy eigenstate is called a **stationary state**: every time-independent measurement probability computed from it stays exactly constant, even though the ket itself is visibly changing.

Concept check 4.1

Can a stationary state have a time-dependent ket?

Answer

Yes. Its ket acquires the phase $e^{-iEt/\hbar}$. The state is called stationary because this overall phase leaves all time-independent measurement probabilities unchanged.

A general initial state need not be a single energy eigenstate; it can be any superposition of them,

$$|\psi(0)\rangle = \sum_n c_n |E_n\rangle, \tag{4.11}$$

and by linearity of the Schrödinger equation, each term simply evolves on its own:

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_n t/\hbar} |E_n\rangle. \tag{4.12}$$

Here is the key qualitative difference from a single stationary state: the energy probabilities $|c_n|^2$ stay constant, but the *relative* phases between different terms evolve, at angular frequencies $(E_n - E_m)/\hbar$ set by the energy differences. Any observable that mixes different energy eigenstates can therefore vary in time, even though no individual energy probability ever does.

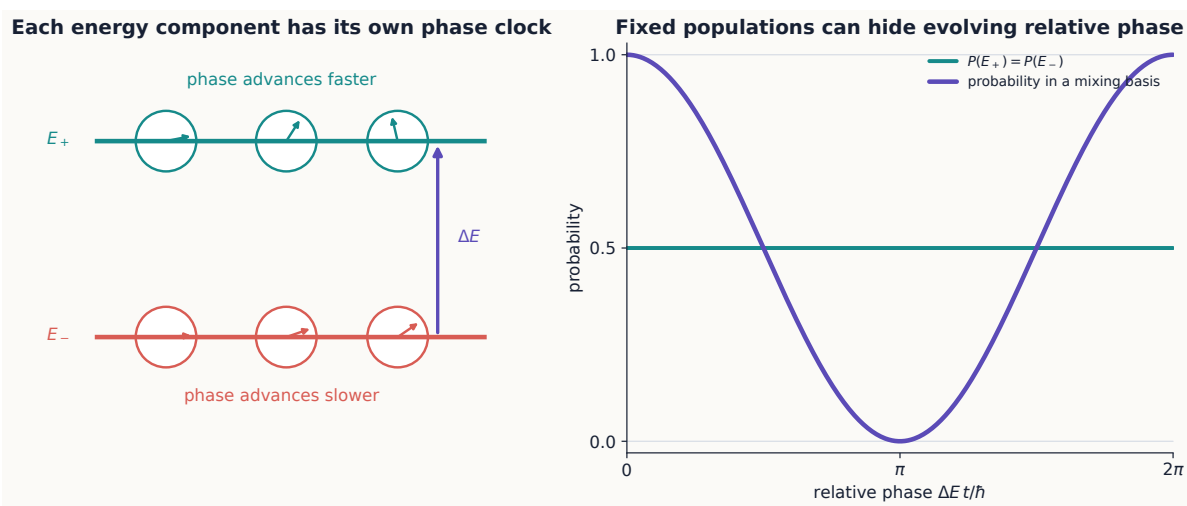

Figure 4.1: Energy-basis populations stay fixed while the phase clocks advance at rates set by their energies. Only the relative clock angle is observable, and a measurement in a mixing basis converts it into an oscillating probability.

Example 4.1: a two-level superposition

Suppose $E_+ > E_-$ and

$$|\psi(0)\rangle = \frac{|E_+\rangle + |E_-\rangle}{\sqrt{2}}. \tag{4.13}$$

Removing the physically irrelevant common phase $e^{-iE_- t/\hbar}$ gives the equivalent representative

$$|\psi(t)\rangle \sim \frac{e^{-i\omega t}|E_+\rangle + |E_-\rangle}{\sqrt{2}}, \qquad \omega = \frac{E_+ - E_-}{\hbar}. \tag{4.14}$$

The energy probabilities remain one half at all times. But in a basis built from coherent sums of $|E_+\rangle$ and $|E_-\rangle$, the corresponding probabilities oscillate with frequency ω. If these two energy states are chosen as the z basis of a spin-$\frac{1}{2}$ system, the x and y bases are familiar examples of such mixing bases.

Only energy differences drive observable phase changes

For two energy components, the cross term in any measurement probability contains a phase of the form

$$e^{-i(E_n - E_m)t/\hbar}. \tag{4.15}$$

This predicts an experimentally visible angular frequency $|E_n - E_m|/\hbar$, a useful fact to remember, since it also gives a quick dimensional check: energy divided by $\hbar$ has units of inverse time, as an angular frequency should. A common phase built from, say, the *average* energy of two levels drops out of every isolated-system probability, exactly like the overall phase discussed in Section 4.1.

Example 4.2: evolution from a nondiagonal matrix

In the z basis, consider

$$\hat{H} = E_0\hat{I} + \frac{\hbar\Omega}{2}\sigma_x = \begin{pmatrix} E_0 & \hbar\Omega/2 \\ \hbar\Omega/2 & E_0 \end{pmatrix}. \tag{4.16}$$

The energy eigenstates are $|+x\rangle$ and $|-x\rangle$, with energies $E_0 \pm \hbar\Omega/2$. Because the identity commutes with σ_x,

$$\hat{U}(t) = e^{-iE_0 t/\hbar}\left(\cos\frac{\Omega t}{2}\,\hat{I} - i\sin\frac{\Omega t}{2}\,\sigma_x\right). \tag{4.17}$$

Applied to $|+z\rangle$, and suppressing the common phase,

$$|\psi(t)\rangle \sim \cos\frac{\Omega t}{2}|+z\rangle - i\sin\frac{\Omega t}{2}|-z\rangle. \tag{4.18}$$

Therefore

$$P(z+,t) = \cos^2\frac{\Omega t}{2}, \qquad P(z-,t) = \sin^2\frac{\Omega t}{2}. \tag{4.19}$$

Notice what is happening here: the initial state was definite along z, but the Hamiltonian was *not* diagonal in the z basis. It is exactly this mismatch—the off-diagonal entries connecting $|+z\rangle$ and $|-z\rangle$—that drives coherent transfer between the two z outcomes over time.

4.3 Spin in a uniform magnetic field

We now specialize the general machinery above to the physical situation that will occupy the rest of this chapter: a spin in a magnetic field. A magnetic moment in a uniform field has Hamiltonian

$$\hat{H} = -\hat{\boldsymbol{\mu}} \cdot \mathbf{B}, \tag{4.20}$$

the same energy expression introduced in Chapter 1, now promoted to an operator. For a spin-$\frac{1}{2}$ particle whose magnetic moment is proportional to its spin, write $\hat{\boldsymbol{\mu}} = \gamma \hat{\mathbf{S}}$, where γ is the gyromagnetic ratio. For $\mathbf{B} = B_0 \hat{\mathbf{z}}$, this becomes

$$\hat{H} = -\gamma B_0 \hat{S}_z = -\frac{\hbar \omega_0}{2} \sigma_z, \qquad \omega_0 = \gamma B_0. \tag{4.21}$$

The sign of ω_0 depends on the sign convention chosen for γ. Keeping track of that sign carefully, rather than dropping it, avoids silently reversing the direction of precession later on.

Because $\sigma_z^2 = \hat{I}$, its exponential can be evaluated directly from the power series, without any diagonalization step. Every even power of σ_z collapses back to $\hat{I}$, and every odd power collapses back to σ_z itself: $\sigma_z^3 = \sigma_z(\sigma_z^2) = \sigma_z$, $\sigma_z^4 = (\sigma_z^2)^2 = \hat{I}$, and so on. Writing $\theta = \omega_0 t/2$ for brevity and separating the exponential's power series into its even and odd terms,

$$e^{i\theta\sigma_z} = \sum_{n=0}^{\infty} \frac{(i\theta\sigma_z)^n}{n!} = \underbrace{\left(\sum_{n \text{ even}} \frac{(i\theta)^n}{n!} \right)}_{=\cos\theta} \hat{I} + \underbrace{\left(\sum_{n \text{ odd}} \frac{(i\theta)^n}{n!} \right)}_{=i\sin\theta} \sigma_z, \tag{4.22}$$

using $\sigma_z^n = \hat{I}$ for even n and $\sigma_z^n = \sigma_z$ for odd n to pull σ_z out of every odd term. The two bracketed sums are exactly the ordinary Taylor series whose sum is $\cos\theta$ and $i\sin\theta$ respectively: the same series reviewed above for $e^{i\phi}$, now split into its even and odd parts instead of read off directly. This gives

$$e^{i\omega_0 t\sigma_z/2} = \cos\frac{\omega_0 t}{2}\, \hat{I} + i \sin\frac{\omega_0 t}{2}\, \sigma_z. \tag{4.23}$$

The identical argument applies to $e^{i\alpha\, \mathbf{n}\cdot\boldsymbol{\sigma}/2}$ for any unit vector $\mathbf{n}$, since $(\mathbf{n}\cdot\boldsymbol{\sigma})^2 = \hat{I}$ as well; this is exactly the fact reused in Section 4.5 for a field pointing in an arbitrary direction and in Section 4.10 for a general control pulse.

In the z basis, where $\hat{H}$ is already diagonal, this is simply

$$\hat{U}(t) = \begin{pmatrix} e^{i\omega_0 t/2} & 0 \\ 0 & e^{-i\omega_0 t/2} \end{pmatrix}. \tag{4.24}$$

Concept check 4.2

A spin begins in $|+z\rangle$ in a uniform z-directed field. Does it precess into $|-z\rangle$?

Answer

No. It is an energy eigenstate and gains only an overall phase. A state with a component transverse to the field has a Bloch vector that visibly precesses.

4.4 Larmor precession

The matrix above shows that a $z+$ or $z-$ state, being an eigenstate of $\hat{H}$, only picks up a phase and never changes. To see the field actually *do* something visible, we need a state with some component transverse to z. Prepare $|+x\rangle = (|+z\rangle + |-z\rangle)/\sqrt{2}$ and let it evolve in the z-directed field. Then

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}\left(e^{i\omega_0 t/2}|+z\rangle + e^{-i\omega_0 t/2}|-z\rangle\right). \tag{4.25}$$

After removing an overall phase, the relative phase between the two terms is $e^{-i\omega_0 t}$, and Chapter 2 already showed exactly what a relative phase like this does to measurement probabilities in a different basis. The probabilities for a later S_x measurement are

$$P(x+,t) = \cos^2\frac{\omega_0 t}{2}, \qquad P(x-,t) = \sin^2\frac{\omega_0 t}{2}. \tag{4.26}$$

Meanwhile $P(z+) = P(z-) = 1/2$ at every time t: the z-basis populations never change at all. Geometrically, the Bloch vector simply rotates around the z axis at the Larmor angular frequency $|\omega_0|$. This is the general pattern to remember: dynamics changes a *relative phase*, and it is the choice of final analyzer that converts an otherwise hidden phase into a measurable probability.

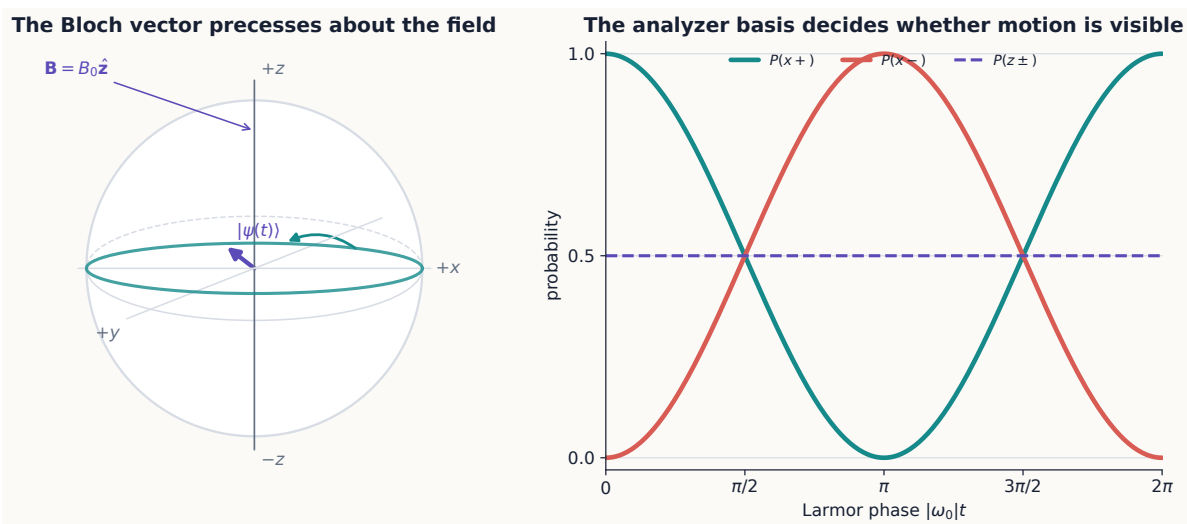

Figure 4.2: Larmor precession leaves the z-basis populations unchanged but moves the Bloch vector around the equator. An x analyzer reveals this motion as complementary oscillations in its two output probabilities.

Concept check 4.3

A spin in a z-directed field always has $P(z+) = P(z-) = 1/2$. Must its state be constant in time?

Answer

No. Its relative phase may evolve even while its z probabilities remain fixed. An x or y analyzer can reveal that phase evolution through changing counts. A single measurement basis does not provide complete state information.

This experiment is a two-state interferometer, in the same spirit as the recombination experiments of Chapters 1 and 2. Preparing $|+x\rangle$ creates two equal z-basis amplitudes, the field gives them different phases as time passes, and the final x analyzer recombines those amplitudes into an interference pattern. The observed oscillation is the spin-space analog of moving a phase plate through one arm of an ordinary path interferometer.

Two different roles for magnetic fields
A field gradient in a Stern–Gerlach analyzer separates outcomes and enables a measurement. A uniform field, ideally, does not separate the beam; it changes the spin state coherently. The former correlates spin with a path, while the latter implements unitary evolution.

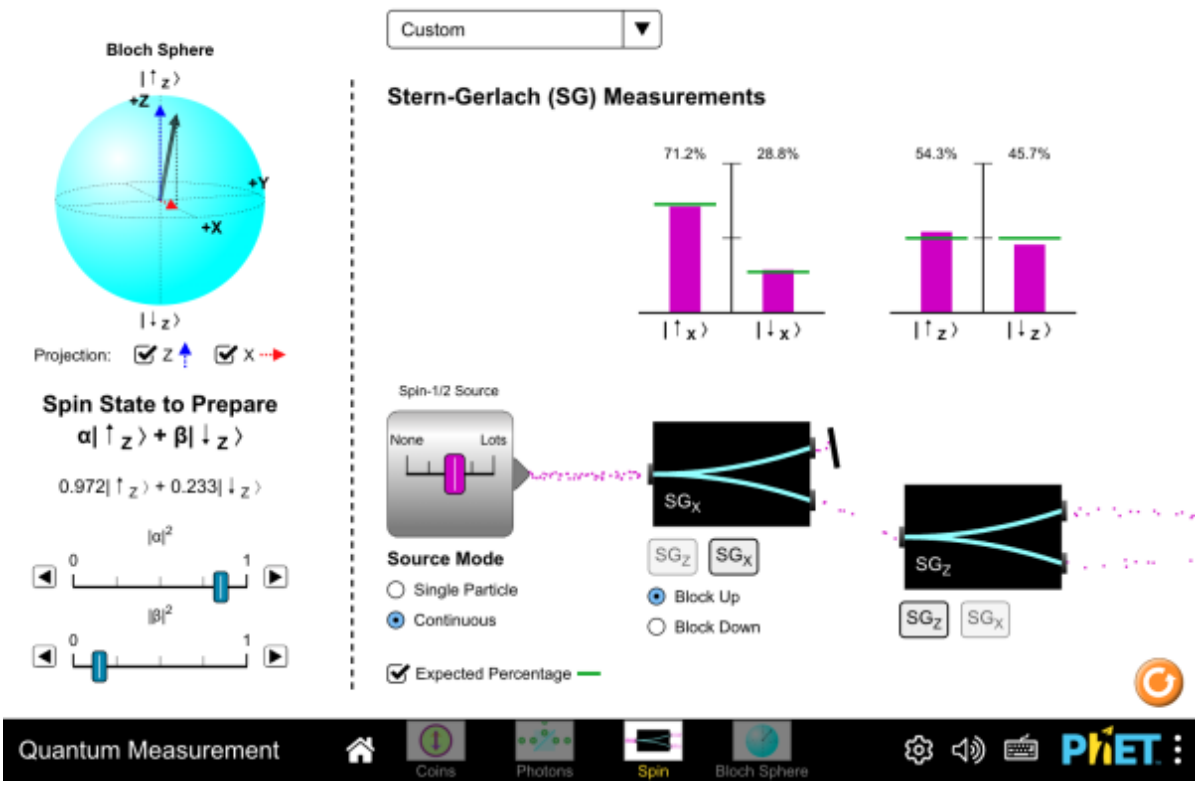

Figure 4.3: Prepare $|+x\rangle$, switch on a z-directed field, and watch the Bloch vector sweep around the equator at the Larmor frequency exactly as equation (4.25) predicts, while a live readout of $P(x\pm, t)$ confirms equation (4.26). Interactive simulation: Quantum Measurement: Bloch Sphere

The classical system with the same equation of motion is a gyroscope: a torque perpendicular to an angular momentum does not tip it over but walks it around a cone at fixed polar angle, exactly as a transverse field walks the Bloch vector around the equator without changing $P(z\pm)$.

Example 4.3: a spin-flip time

Starting from $|+x\rangle$, the first time at which an S_x measurement gives $x-$ with certainty satisfies $|\omega_0|t = \pi$. Thus

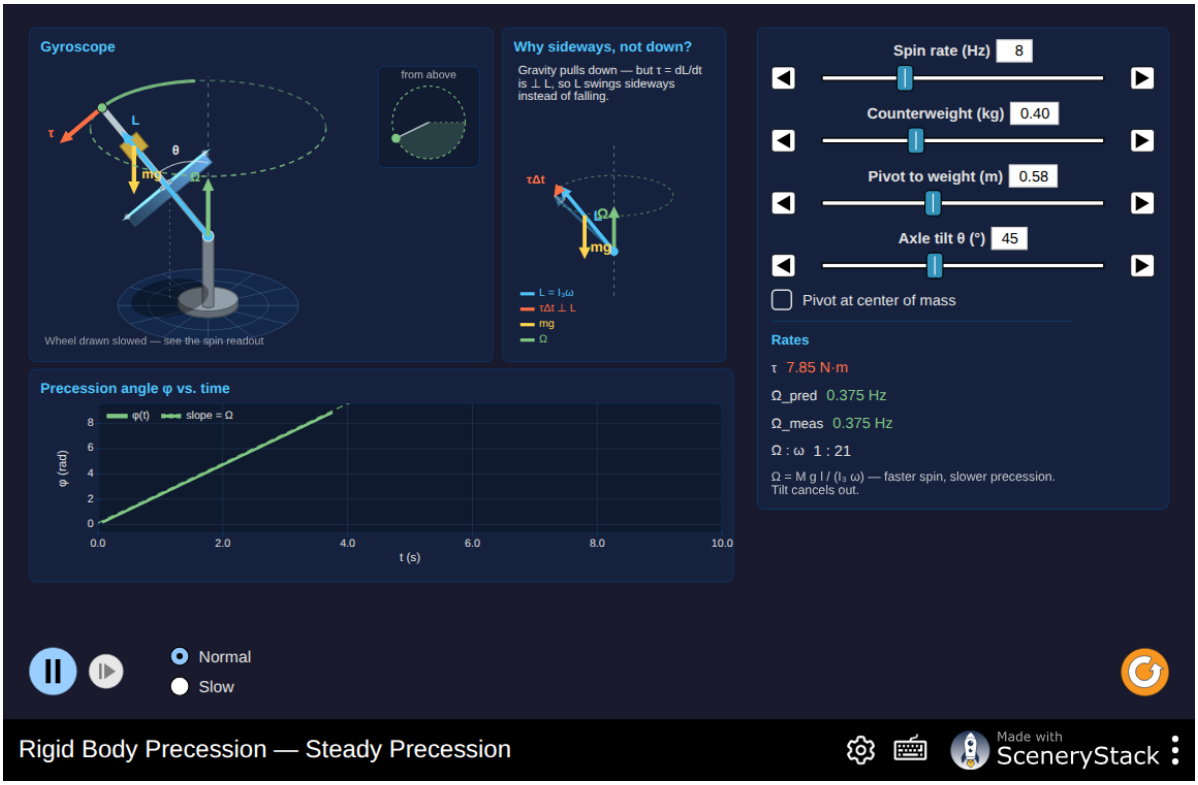

Figure 4.4: A spinning top under gravity, precessing steadily about the vertical. $|\vec{L}|$ and L_z stay fixed while L_x and L_y trace out a circle—the classical motion behind the Bloch vector's rotation above, and behind the orbital-angular-momentum vector model of Chapter 9. Interactive simulation: Precession

$$t_{\text{flip}} = \frac{\pi}{|\omega_0|}. \tag{4.27}$$

Be precise about what this is *not*: it is not a transition between the S_z energy levels, since their populations remain fixed throughout. It is a rotation of the state relative to the x measurement basis—the energy eigenstates are untouched; only the relative phase between them has advanced by π.

Example 4.4: infer a field from counts

A $+x$ state evolves for $2.0\,\mu$s in a z-directed field. A later S_x measurement gives $P(x+) = 3/4$. Equation (4.26) implies

$$\cos^2 \frac{\omega_0 t}{2} = \frac{3}{4}. \tag{4.28}$$

The smallest positive phase is $|\omega_0|t = \pi/3$, giving

$$|\omega_0| = \frac{\pi}{3t} \approx 5.24 \times 10^5 \text{ rad s}^{-1}. \tag{4.29}$$

Two ambiguities remain, and neither should be overlooked. The probability alone does not determine the *sign* of ω_0, since $\cos^2$ cannot distinguish clockwise from counterclockwise rotation. And because the cosine is periodic, this same data is also consistent with larger frequencies differing by whole cycles. Additional measurements at other times, or a y analyzer sensitive to the direction of rotation, remove these ambiguities.

4.5 A field in an arbitrary direction

Everything so far assumed the field points along z. For a constant field along an arbitrary unit vector $\mathbf{n}$,

$$\hat{H} = -\frac{\hbar\omega}{2}\mathbf{n}\cdot\boldsymbol{\sigma}. \tag{4.30}$$

Since $(\mathbf{n}\cdot\boldsymbol{\sigma})^2 = \hat{I}$—the same identity used in Chapter 3—the same power-series trick used for σ_z generalizes immediately:

$$\hat{U}(t) = \cos\frac{\omega t}{2}\,\hat{I} + i\sin\frac{\omega t}{2}\,\mathbf{n}\cdot\boldsymbol{\sigma}. \tag{4.31}$$

This operator rotates the Bloch vector around $\mathbf{n}$, and it also reveals why the exponent naturally involves a *half*-angle, echoing the half-angle rule from Chapter 1. A 2π spatial rotation gives $\hat{U} = -\hat{I}$, and only a 4π rotation returns the ket to exactly itself. That minus sign after 2π is an overall phase for one isolated evolving path, so it has no effect by itself, but it becomes observable through interference with a second, unrotated path.

Three geometric cases deserve separate attention. If the initial Bloch vector is parallel to $\mathbf{n}$, the state is an energy eigenstate and gains only an overall phase, never changing direction. If it starts perpendicular to $\mathbf{n}$, it sweeps out a full circle around $\mathbf{n}$. At any intermediate angle, it traces a cone around $\mathbf{n}$, with its component along $\mathbf{n}$ staying fixed throughout.

Example 4.5: rotation about x

Take the Hamiltonian $\hat{H} = -(\hbar\omega/2)\sigma_x$ and initial state $|+z\rangle$. Equation (4.31) gives

$$|\psi(t)\rangle = \cos\frac{\omega t}{2}|+z\rangle + i\sin\frac{\omega t}{2}|-z\rangle. \tag{4.32}$$

The probability of a $z-$ result is $\sin^2(\omega t/2)$. At $|\omega|t = \pi/2$, the state has equal z probabilities but a definite transverse direction; at $|\omega|t = \pi$, it is $|-z\rangle$ up to phase. Equal z populations halfway through the rotation do not mean coherence has been lost, however: the state is still a perfectly definite pure state, just one pointing along a different axis.

4.6 Measurement interrupts unitary evolution

So far we have evolved a state continuously and only asked about measurement at the very end. Real experimental protocols often measure, select an outcome, and then let the resulting state evolve further, and it is important to know exactly how to handle that. Suppose the spin evolves until time t_1, when S_x is measured. Just before the measurement,

$$|\psi(t_1)\rangle = \hat{U}(t_1)|\psi(0)\rangle. \tag{4.33}$$

The outcome probabilities follow from projecting this state onto $|\pm x\rangle$, exactly as in Chapter 3. If the result is $x+$, the state used for all subsequent evolution is $|+x\rangle$ itself, not the pre-measurement superposition:

$$|\psi(t > t_1)\rangle = \hat{U}(t, t_1)|+x\rangle. \tag{4.34}$$

Unitary evolution and ideal projective measurement are two genuinely distinct parts of the model, governed by different rules, and confusing them leads to a common mistake: continuing to evolve the *pre*-measurement superposition after a definite outcome has already been selected, as if the measurement had never happened.

Example 4.6: a complete prepare–evolve–measure protocol

A spin is prepared in $|+x\rangle$, evolves in a z-directed field for a time $t_1 = \pi/(2|\omega_0|)$, and is measured along x. Suppose the $x-$ outcome is selected. The same field then acts for another time $t_2 = \pi/(2|\omega_0|)$ before a final S_x measurement.

Just before the intermediate measurement, $P(x+) = P(x-) = 1/2$. Selecting $x-$ prepares $|-x\rangle$; it does not leave the earlier phase-evolved ket in place, per the rule just stated. Starting fresh from $|-x\rangle$, another quarter-period of evolution produces equal final x probabilities. So the probability for either final result, *conditional* on having selected the $x-$ branch, is 1/2. The probability for the *complete* route starting from the original preparation must also include the probability of the intermediate selection itself, giving 1/4 overall.

4.7 Constants in the Hamiltonian

Adding a constant energy E_0 to every state changes the Hamiltonian to $\hat{H}' = \hat{H} + E_0\hat{I}$. Because $E_0\hat{I}$ commutes with anything, the corresponding evolution operator factors cleanly:

$$\hat{U}'(t) = e^{-iE_0t/\hbar}\hat{U}(t). \tag{4.35}$$

Every state acquires exactly the same overall phase from this shift, so isolated-system probabilities are completely unchanged. This confirms a pattern seen throughout the chapter: it is energy *differences*, not an arbitrary choice for the zero of energy, that control observable relative phases.

4.8 Optional mathematical extension: time-dependent Hamiltonians

Every result so far assumed a time-independent Hamiltonian, which is what let us write $\hat{U}(t) = e^{-i\hat{H}t/\hbar}$ so simply. If $\hat{H}$ instead changes with time, that simple exponential formula no longer applies, and the evolution operator must instead be defined through its own differential equation:

$$i\hbar \frac{d\hat{U}(t,t_0)}{dt} = \hat{H}(t)\hat{U}(t,t_0), \qquad \hat{U}(t_0,t_0) = \hat{I}. \tag{4.36}$$

If the Hamiltonians at different times all commute with each other, the solution still takes a simple exponential form, just with an integral in place of a product:

$$\hat{U}(t,t_0) = \exp\left[-\frac{i}{\hbar}\int_{t_0}^{t} \hat{H}(t')\,dt'\right]. \tag{4.37}$$

But if $[\hat{H}(t_1), \hat{H}(t_2)] \neq 0$ for some pair of times, this shortcut fails, and the ordering of successive infinitesimal rotations genuinely matters; a time-ordered exponential or a piecewise product must be used instead. For example, a pulse about x followed by a pulse about y generally does not produce the same final state as the reversed sequence. This is the dynamical counterpart of the noncommuting analyzer questions from Chapter 3.

Example 4.7: a piecewise-constant Hamiltonian

Suppose a control system applies

$$\hat{H}(t) = \begin{cases} (\hbar\Omega/2)\sigma_x, & 0 \le t < \tau, \\ (\hbar\Omega/2)\sigma_z, & \tau \le t < 2\tau, \end{cases} \qquad \Omega\tau = \frac{\pi}{2}. \tag{4.38}$$

Within each interval the Hamiltonian is constant. The complete evolution is therefore the ordered product

$$\hat{U}(2\tau,0) = \underbrace{\hat{R}_z(\pi/2)}_{\text{second interval}}\ \underbrace{\hat{R}_x(\pi/2)}_{\text{first interval}}. \tag{4.39}$$

The earliest operation stands nearest the initial ket. Starting from $|+z\rangle$, the first interval gives $\hat{R}_x(\pi/2)|+z\rangle = |-y\rangle$. Applying the z rotation then gives

$$\hat{R}_z(\pi/2)|-y\rangle = e^{-i\pi/4}|+x\rangle. \tag{4.40}$$

The displayed phase is overall, so the final Bloch direction is $+x$. If the two intervals are reversed, the z pulse first changes $|+z\rangle$ only by an overall phase, after which the x pulse leaves the Bloch vector at $-y$. The two schedules use the same Hamiltonians for the same durations but end in different physical states because the interval propagators do not commute. This piecewise construction is the practical finite-step version of time ordering.

4.9 Optional mathematical extension: motion of expectation values

The Schrödinger equation for the ket also implies a direct equation of motion for the expectation value of any observable, which is often a faster route to physical insight than tracking the full ket. For an operator with no explicit time dependence of its own,

$$\frac{d}{dt}\langle A\rangle = \frac{i}{\hbar}\langle[\hat{H},\hat{A}]\rangle. \tag{4.41}$$

Specializing to $\hat{H} = -\gamma\mathbf{B}\cdot\hat{\mathbf{S}}$ and using the spin commutators from Chapter 3, this becomes

$$\frac{d}{dt}\langle\mathbf{S}\rangle = \gamma\langle\mathbf{S}\rangle\times\mathbf{B}. \tag{4.42}$$

Equation (4.42) has exactly the form of classical magnetic precession, even though each individual component measurement remains fully quantized, with only two discrete outcomes. It offers a useful bridge between two pictures developed so far: the geometric Bloch-sphere picture of Chapter 2, and the explicit matrix evolution of a ket developed in this chapter. The expectation vector traces out a smooth, classical-looking trajectory, even though the individual Stern–Gerlach outcomes behind it remain discrete.

4.10 Applications track: controlled rotations and pulse language

Laboratory control of a spin is usually described not in terms of a continuously running field, but in terms of discrete **pulses**—a field switched on for a controlled duration to produce a specific rotation. A constant control Hamiltonian applied for a chosen duration implements exactly such a rotation. Define

$$\hat{R}_{\mathbf{n}}(\alpha) = \exp\left(-i\frac{\alpha}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right) = \cos\frac{\alpha}{2}\hat{I} - i\sin\frac{\alpha}{2}\mathbf{n}\cdot\boldsymbol{\sigma}. \tag{4.43}$$

For $\hat{H} = (\hbar\Omega/2)\mathbf{n}\cdot\boldsymbol{\sigma}$ applied for time τ, the rotation angle is $\alpha = \Omega\tau$. This control Hamiltonian is defined with the opposite sign from the field Hamiltonian $\hat{H} = -(\hbar\omega/2)\mathbf{n}\cdot\boldsymbol{\sigma}$ of Sections 4.3–4.5; matching the two conventions, $\Omega = -\omega$, recovers $\hat{R}_{\mathbf{n}}(\Omega\tau) = \hat{U}(\tau)$ from equation (4.31). Two particular pulses come up often enough to deserve names: a π **pulse** carries a Bloch vector to the opposite side of its rotation circle, while a $\pi/2$ **pulse** can turn a definite z state into an equal-amplitude superposition. For example,

$$\hat{R}_y\left(\frac{\pi}{2}\right)|+z\rangle = |+x\rangle, \qquad \hat{R}_x(\pi)|+z\rangle = -i|-z\rangle. \tag{4.44}$$

The factor $-i$ may be dropped here because it multiplies the entire state; subsequent unitary operations cannot turn a truly global phase into a relative one. In multibranch calculations, however, a phase belonging to only one branch is *not* global and must be retained until the branches are recombined.

If an x pulse acts first and a y pulse second, the combined operator is

$$|\psi_{\mathrm{f}}\rangle = \hat{R}_y(\beta)\hat{R}_x(\alpha)|\psi_{\mathrm{i}}\rangle. \tag{4.45}$$

Read this expression from right to left: the earliest operation in time stands closest to the initial ket. Rotations about different axes generally do not commute, so reversing the order of these two matrices describes a genuinely different laboratory sequence, not merely a different way of writing the same one.

Example 4.8: pulse order on a prepared spin

Begin in $|+z\rangle$. An x-axis $\pi/2$ pulse produces

$$\hat{R}_x\left(\frac{\pi}{2}\right)|+z\rangle = \frac{|+z\rangle - i|-z\rangle}{\sqrt{2}} = |-y\rangle. \tag{4.46}$$

A subsequent y-axis pulse changes this S_y eigenstate only by an overall phase, since $|-y\rangle$ is already an eigenstate of the rotation axis. In the reverse order, the first y pulse produces $|+x\rangle$, which the x pulse then changes only by an overall phase, for the same reason. The two final Bloch vectors, $-y$ and $+x$, are clearly different states. This geometric shortcut reaches the same conclusion as direct matrix multiplication, with considerably less arithmetic.

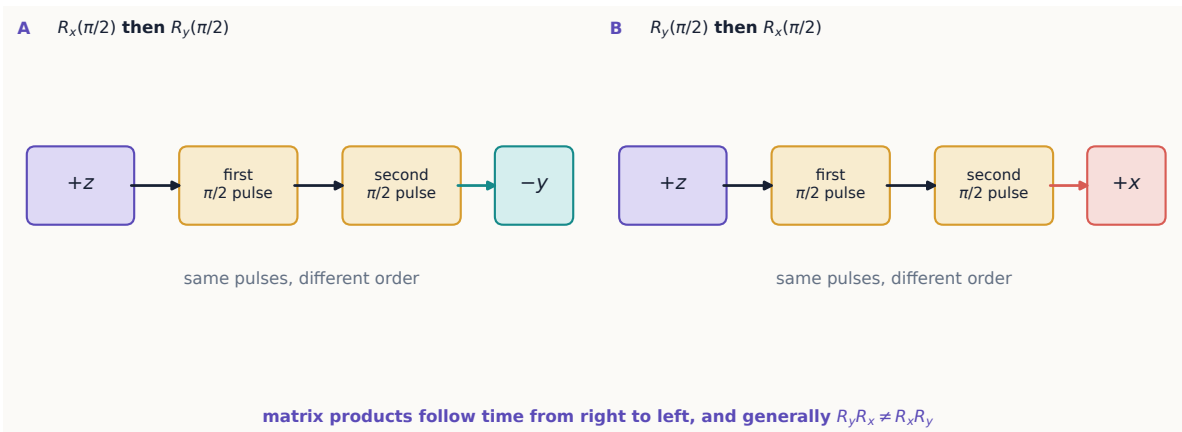

Figure 4.5: The same two control pulses lead to different final states when their order is reversed. The laboratory sequence runs left to right, while the corresponding operator product acts on the ket from right to left.

4.11 Applications track: Ramsey interference

We can now assemble everything in this chapter into one of the most important protocols in atomic physics: a Ramsey sequence, which is really an interferometer built entirely in a two-state space, using pulses instead of beam splitters. The protocol has five steps:

1. prepare $|+z\rangle$;
2. apply $\hat{R}_y(\pi/2)$ to create $|+x\rangle$;
3. allow free phase evolution for time T;
4. apply $\hat{R}_y(-\pi/2)$; and
5. measure S_z.

Represent the free evolution in step 3 by

$$\hat{U}_{\text{free}}(T) = \hat{R}_z(\delta T) = \begin{pmatrix} e^{-i\delta T/2} & 0 \\ 0 & e^{i\delta T/2} \end{pmatrix}, \tag{4.47}$$

where δ is the relative angular frequency in the chosen reference frame. Combining all five steps, the final state is

$$|\psi_{\mathrm{f}}\rangle = \hat{R}_y\left(-\frac{\pi}{2}\right)\hat{R}_z(\delta T)\hat{R}_y\left(\frac{\pi}{2}\right)|+z\rangle. \tag{4.48}$$

Projecting onto $|\pm z\rangle$ gives the observed fringes

$$P(z+) = \cos^2\frac{\delta T}{2} = \frac{1+\cos(\delta T)}{2}, \qquad P(z-) = \sin^2\frac{\delta T}{2}. \tag{4.49}$$

It is worth naming what each step contributes: the first pulse creates two coherent amplitudes out of a single definite state, free evolution changes their relative phase by an amount set by δ and T, and the second pulse recombines those two amplitudes so that the final measurement can read out the accumulated phase as a probability. This is precisely the coherent path logic first introduced in Chapter 2, now implemented with controlled rotations rather than physically separated paths.

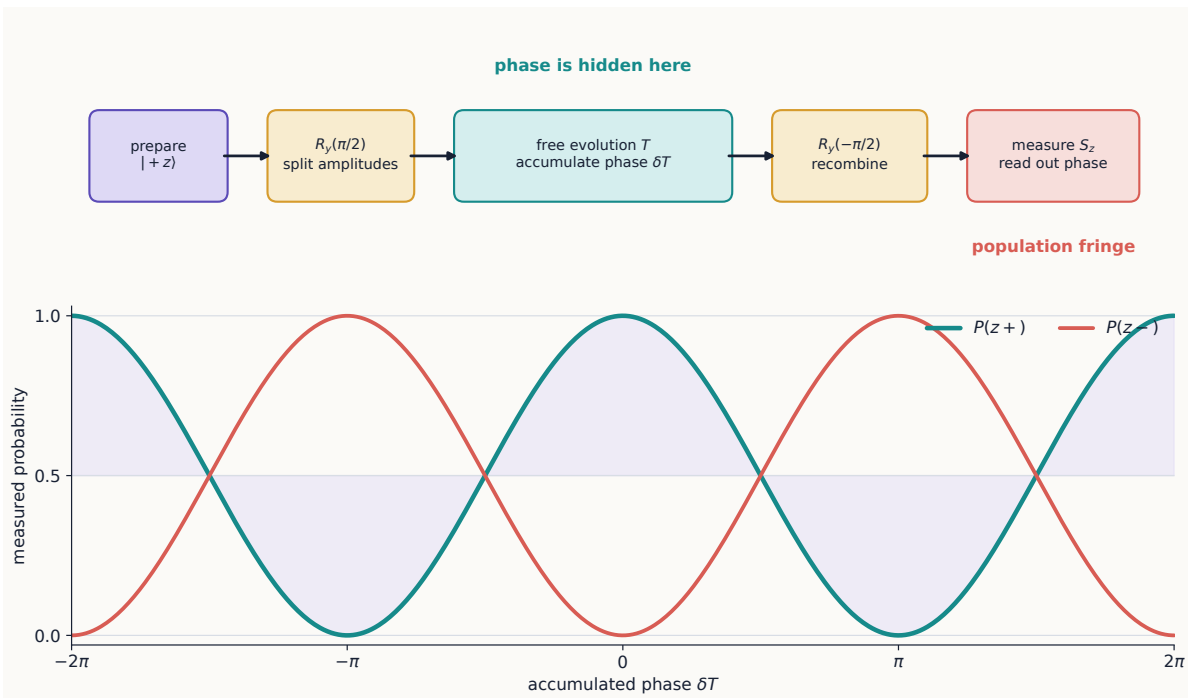

Figure 4.6: A Ramsey experiment is a two-state interferometer: the first pulse splits an amplitude, free evolution shifts the relative phase, and the second pulse recombines the alternatives. The final S_z counts trace the accumulated phase as complementary fringes.

Example 4.9: the first fringe minimum

If $|\delta| = 2\pi(25\,\mathrm{kHz})$, the first $z+$ minimum occurs at $|\delta|T = \pi$:

$$T_{\min} = \frac{\pi}{|\delta|} = 20\,\mu\mathrm{s}. \tag{4.50}$$

Scanning T at fixed detuning reveals these oscillations directly. Scanning the control frequency instead, at fixed T, also produces fringes, now as a function of detuning rather than time. This sensitivity of a simple probability to a small phase is the operating principle behind atomic clocks and many modern quantum sensors.

If the accumulated phase varies unpredictably between different trials of the experiment, averaging over those trials reduces the fringe contrast. Each individual run may still be perfectly unitary; it is the *pooled* ensemble that becomes mixed, because its relative phase is no longer controlled from run to run. Density operators, introduced in Chapter 3, are exactly the tool needed to distinguish this loss of ensemble coherence from ordinary, coherent precession.

Concept check 4.4

Why not measure S_z after the first Ramsey pulse to verify both alternatives?

Answer

That measurement would make the alternatives distinguishable and remove their coherence. If its outcomes were pooled, the second pulse would yield equal final z probabilities rather than a full-contrast fringe.

Summary

- The Hamiltonian generates deterministic, unitary evolution between measurements.
- Energy eigenstates acquire only overall phases and are stationary.
- Superpositions acquire changing relative phases set by energy differences.
- A uniform magnetic field rotates a spin state; a field gradient can analyze it.
- Larmor precession converts phase evolution into oscillating measurement probabilities.
- Adding a constant to every energy changes only an unobservable overall phase.
- A complete dynamics calculation specifies a preparation, an evolution operator, and a final measurement basis.
- Time-dependent Hamiltonians require attention to the order of rotations when the Hamiltonians at different times do not commute.
- Spin expectation values precess smoothly even though individual component measurements have only two discrete outcomes.
- Controlled pulses implement rotations whose matrix order follows the laboratory sequence from right to left.
- Ramsey sequences turn accumulated relative phase into measurable population fringes.

Exercises

Exercises 1–9 establish the core evolution and measurement workflow. Exercises 10–13 consolidate spin precession and interrupted evolution. Exercises 14–20 support the optional mathematical and applications tracks; a focused first assignment need not include all three groups.

1. Starting from equation (4.1), show directly that the norm of a state is constant when $\hat{H}$ is Hermitian.

2. A two-level system has energies 0 and ϵ. Evolve the initial state $(|E_0\rangle + i|E_1\rangle)/\sqrt{2}$ and find the probability of returning to the initial state at time t.

3. Derive equation (4.23) by separating the exponential power series into even and odd powers.

4. Derive equation (4.26) from equation (4.25) using inner products.

5. For the same evolving state, calculate $P(y+, t)$ and identify the direction of precession for positive ω_0 under the conventions used here.

6. Find the earliest time at which a spin initially in $|+z\rangle$ becomes $|-z\rangle$ under a uniform field in the $+x$ direction.

7. Prove that equation (4.31) is unitary.

8. An S_x measurement is made halfway through a full Larmor period and its $x+$ result is selected. Describe the state immediately before and after the measurement, then find $P(x+)$ one quarter-period later.

9. Show that adding $E_0\hat{I}$ to a Hamiltonian cannot affect the expectation value of any time-independent observable in an isolated system.

10. For the Hamiltonian in Example 4.2, calculate $P(y+,t)$ for an initial $|+z\rangle$ state. At what earliest time is the state $|+y\rangle$ up to overall phase?

11. A $+x$ state precesses in a z-directed field. Find both $P(x+,t)$ and $P(y+,t)$, and explain how their relative timing reveals the sign of ω_0.

12. Derive the result of Example 4.5 by expanding $|+z\rangle$ in the energy basis $|\pm x\rangle$, evolving the two components, and transforming back.

13. Work through every branch of Example 4.6 rather than selecting only $x-$. Show that pooling the intermediate outcomes gives final statistics different from a protocol with no intermediate measurement.

14. A Hamiltonian equals $\hbar\Omega\sigma_x/2$ for a duration τ and then $\hbar\Omega\sigma_y/2$ for the same duration. Write the total evolution operator in the correct order. Repeat for the reversed pulse sequence and show that the results differ unless $\Omega\tau$ takes a special value.

15. Derive equation (4.41) from the Schrödinger equation and its Hermitian conjugate.

16. Use equation (4.42) to identify a conserved component of $\langle\mathbf{S}\rangle$ in a constant magnetic field and connect it to energy conservation.

17. Verify both pulse actions following equation (4.43) using explicit Pauli matrices.

18. Calculate both pulse orders in Example 4.8 by matrix multiplication.

19. Derive equation (4.49) from equation (4.48) and check $\delta T = 0, \pi, 2\pi$.

20. Insert an unread S_z measurement after the first Ramsey pulse and use density operators to show that both final outcomes have probability 1/2.

Selected exercise guidance

Use these answers only after making a complete attempt.

Exercise 2

The return amplitude is $(1 + e^{-i\epsilon t/\hbar})/2$, so the return probability is $\cos^2(\epsilon t/2\hbar)$.

Exercise 5

For positive ω_0 under the chapter's sign convention, the Bloch vector is $(\cos\omega_0 t, -\sin\omega_0 t, 0)$ and

$$P(y+,t) = \frac{1-\sin(\omega_0 t)}{2}. \tag{4.51}$$

It initially moves from $+x$ toward $-y$.

Exercise 6

$P(z-,t) = \sin^2(|\omega|t/2)$, so the earliest certain flip occurs at $t = \pi/|\omega|$.

Exercise 19

At $\delta T = 0, \pi, 2\pi$, the $z+$ probabilities are respectively $1, 0, 1$; the $z-$ probabilities are the complements.

Part II

Composite and continuous systems

Chapter 5

Composite Systems and Entanglement

5.1 Composite systems and entanglement

Learning objectives

After this chapter, you should be able to:

- construct product states and operators with tensor products;
- calculate joint, marginal, and conditional probabilities;
- distinguish separable states from entangled states, including in mixtures;
- use reduced density operators to describe a subsystem;
- predict correlations in Bell states and quantify partial entanglement with the Schmidt decomposition;
- explain why entanglement does not enable faster-than-light signaling;
- state the Tsirelson bound and what it does and does not imply about nature; and
- analyze simple two-qubit gates and measurement protocols.

Sections 5.1–5.6 provide the core route: build joint states, calculate local and joint predictions, interpret Bell correlations, and create entanglement. Section 5.7 is an optional quantitative extension using the Schmidt decomposition. Section 5.8 begins with the essential distinction between correlation and entanglement, then marks its Werner-state test as an advanced extension.

5.1 Building a joint state space

Chapters 2 through 4 developed a complete quantum description of a single spin: a two-dimensional complex state space, Hermitian operators for measurable quantities, and unitary time evolution. But that formalism, on its own, does not say how to describe *two* spins that were prepared together and are later measured separately, perhaps in two different laboratories. A first guess might be that each spin simply carries its own independent ket, and a joint experiment just reports the two kets' separate predictions side by side. That guess is sometimes right and sometimes badly wrong, and sorting out the difference is the subject of this chapter.

The correct joint state space is built from the two one-spin spaces by a specific mathematical construction called the **tensor product**. It is not merely "the two spaces considered together": it is a larger space, big enough to hold every possible correlation between the two spins, including correlations with no classical counterpart at all. Learning to build this space, and

learning to recognize when a joint state does or does not reduce to "two independent spins," is the first step toward composite systems of any size—two electrons, a spin coupled to a field mode, or the many-particle systems of Chapter 11.

Quick review: why do the dimensions multiply?

Skip this box if it is already clear why a joint system of two qubits needs four basis states rather than, say, two or three.

Suppose you are choosing an outfit: a shirt, from 2 options, paired with a pair of pants, also from 2 options. How many complete outfits are possible? Every shirt can be paired with every pair of pants independently, so the count is $2 \times 2 = 4$, not $2 + 2$: (shirt 1, pants 1), (shirt 1, pants 2), (shirt 2, pants 1), (shirt 2, pants 2). Two independent choices, each with two options, combine multiplicatively.

A joint quantum state space is built the same way. Spin A has two basis states and spin B has two basis states, so the *joint* system needs one basis state for every combination of one A-label and one B-label: four in total, listed just below. This is why joint dimensions multiply rather than add, and it generalizes immediately: three spins need $2 \times 2 \times 2 = 8$ basis states, and an A with m outcomes joined to a B with n outcomes needs mn.

It is worth flagging now exactly where this analogy will later break down, since that break is the entire point of the chapter. A person can always describe a chosen outfit by separately naming one shirt and one pair of pants. Section 5.2 shows that a general joint quantum state *cannot* always be described by separately naming one state for A and one state for B. That extra possibility—a joint description with no corresponding pair of separate local descriptions—is what makes composite quantum systems genuinely different from combinatorics with ordinary independent choices.

One spin-$\frac{1}{2}$ system has a two-dimensional state space. Two distinguishable spins require four independent basis states:

$$|+z\rangle_A|+z\rangle_B, \quad |+z\rangle_A|-z\rangle_B, \quad |-z\rangle_A|+z\rangle_B, \quad |-z\rangle_A|-z\rangle_B. \tag{5.1}$$

The joint space is the tensor product $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. We abbreviate $|+z\rangle_A|-z\rangle_B$ as $|+z, -z\rangle$ or, for qubits, $|01\rangle$. In the ordered z basis,

$$|00\rangle \doteq \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad |01\rangle \doteq \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad |10\rangle \doteq \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad |11\rangle \doteq \begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{5.2}$$

If $|\psi\rangle_A = a|0\rangle + b|1\rangle$ and $|\phi\rangle_B = c|0\rangle + d|1\rangle$, then

$$|\psi\rangle_A \otimes |\phi\rangle_B = ac|00\rangle + ad|01\rangle + bc|10\rangle + bd|11\rangle. \tag{5.3}$$

Notice that the coefficient array factors into one list for A and one for B. This factorization is the defining feature of a pure **product state**, and it is exactly what will fail, in Section 5.2, for a state that cannot be built this way.

Quick calculation: multiplying out a concrete product state

Take $|\psi\rangle_A = |+x\rangle_A = (|0\rangle_A + |1\rangle_A)/\sqrt{2}$ and $|\phi\rangle_B = |0\rangle_B$, so in the notation of equation (5.3), $a = b = 1/\sqrt{2}$ while $c = 1$ and $d = 0$. Multiplying out term by term,

$$|+x\rangle_A|0\rangle_B = \frac{1}{\sqrt{2}}(1)|00\rangle + \frac{1}{\sqrt{2}}(0)|01\rangle + \frac{1}{\sqrt{2}}(1)|10\rangle + \frac{1}{\sqrt{2}}(0)|11\rangle = \frac{|00\rangle + |10\rangle}{\sqrt{2}}. \tag{5.4}$$

Only the terms ending in B-label 0 survive, because $|\phi\rangle_B$ has no $|1\rangle_B$ component at all: the amplitude for any joint outcome with $B = 1$ is forced to zero regardless of what A does. Contrast this with the entangled state introduced next: here both surviving terms share the same B label while A's label is free to be either 0 or 1, exactly the signature of A and B having been prepared independently, with no correlation between their outcomes.

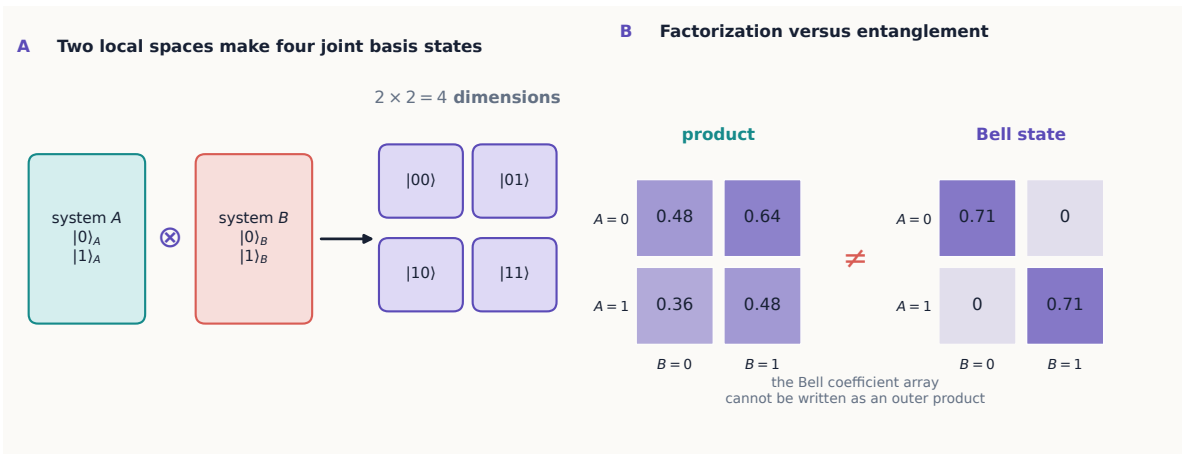

Figure 5.1: The tensor product expands two local two-state spaces into one four-state joint space. Product-state coefficients form an outer product of two local lists; the Bell-state pattern cannot be factored that way.

Operators on one subsystem

An operation on A alone is represented by $\hat{A}\otimes\hat{I}_B$; an operation on B alone is $\hat{I}_A\otimes\hat{B}$. Operators acting on different subsystems always commute:

$$[\hat{A}\otimes\hat{I}, \hat{I}\otimes\hat{B}] = 0. \tag{5.5}$$

To construct these matrices rather than merely name them, replace every entry of the first matrix by that entry times the entire second matrix. For example, in the ordered basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$,

$$\sigma_x \otimes \hat{I} = \begin{pmatrix} 0\hat{I} & 1\hat{I} \\ 1\hat{I} & 0\hat{I} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad \hat{I} \otimes \sigma_z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{5.6}$$

The first matrix flips A's label while leaving B's unchanged: $|00\rangle \mapsto |10\rangle$ and $|01\rangle \mapsto |11\rangle$. The second multiplies a basis ket by the eigenvalue associated with B's label while leaving A alone. Multiplying these two explicit matrices in either order gives the same result, verifying the commutator above concretely. This block-matrix recipe is the **Kronecker product** used whenever subsystem operators must be written as a matrix on the joint space.

Because these operators commute, measurements made on the two distinct systems can always be assigned one common joint probability distribution, even when the individual components being measured on either spin, taken alone, would be incompatible. This is a stronger and more precise statement than just "the two measurements don't interfere." It says that a single joint probability distribution can be written down for the two experiments at all, which is exactly the kind of object that incompatible single-spin observables (Chapter 3) cannot generally be assigned to begin with.

Example 5.1: independent preparations

For $|+x\rangle_A|-z\rangle_B$, measurements of S_z on both spins give $P(0,1) = P(1,1) = 1/2$ and zero for the other outcomes. Since the preparation is a product, its joint probabilities factor:

$$P(a,b) = P_A(a)P_B(b). \tag{5.7}$$

Learning B's result supplies no new information about A. This factorization is the operational signature of a product state: every joint question about A and B can be answered by combining two separate one-spin calculations, and nothing is lost by treating the two laboratories as running independent experiments on a shared clock.

5.2 States that do not factor

Now consider the normalized state

$$|\Phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}. \tag{5.8}$$

Try to write it in the product form of equation (5.3). If $ad = bc = 0$ while both ac and bd are nonzero, the required single-system coefficients contradict one another: there is no consistent choice of a, b, c, d that works. A state like this, which cannot be factored, is called **entangled**.

A joint z measurement on $|\Phi^+\rangle$ produces 00 or 11, each with probability 1/2. Each individual outcome is random, yet the two records always agree. Rewriting the state in the x basis gives

$$|\Phi^+\rangle = \frac{|+x,+x\rangle + |-x,-x\rangle}{\sqrt{2}}, \tag{5.9}$$

so the x records also always agree. This is the key point: entanglement is not merely "both spins were prepared the same way." No assignment of one definite pure state to each spin separately can reproduce all of these correlations at once.

The four Bell states form an orthonormal basis:

$$|\Phi^\pm\rangle = \frac{|00\rangle \pm |11\rangle}{\sqrt{2}}, \qquad |\Psi^\pm\rangle = \frac{|01\rangle \pm |10\rangle}{\sqrt{2}}. \tag{5.10}$$

Relative signs that were invisible in a z measurement become visible once we switch to another basis, exactly as relative phase did for a single spin in Chapter 2.

Being precise about what "entangled" means matters here, since the word invites an all-or-nothing reading it does not deserve. Entanglement is not a label attached loosely to "the two particles are correlated somehow"; it is a precise property of the joint state, and it comes in degrees. A state can fail to factor into a product yet still be much "closer" to a product than $|\Phi^+\rangle$ is. Section 5.7 makes this quantitative with the Schmidt decomposition. For now, treat

the Bell states as one convenient orthonormal family of *maximally* entangled states, not as the only entangled states that exist.

Foundation check: factor or not?

Can $(|00\rangle + |10\rangle)/\sqrt{2}$ be written as a product? What about $(|00\rangle + |11\rangle)/\sqrt{2}$?

Answer

The first state factors as $|+x\rangle_A|0\rangle_B$. The second does not: making the $|00\rangle$ and $|11\rangle$ coefficients nonzero while making both cross coefficients zero is impossible for a product of two one-qubit kets.

5.3 Joint, marginal, and conditional probabilities

For outcome kets $|a\rangle_A$ and $|b\rangle_B$, the joint Born rule is

$$P(a,b) = |\langle a,b|\Psi\rangle|^2. \tag{5.11}$$

The marginal probability for A ignores B's result:

$$P_A(a) = \sum_b P(a,b). \tag{5.12}$$

If $B = b$ is recorded, the conditional probability is $P(a|b) = P(a,b)/P_B(b)$ when $P_B(b) \neq 0$.

Example 5.2: joint, marginal, and conditional records

Begin with $|\Phi^+\rangle$ and measure both spins in the z basis. Reading the four amplitudes directly gives

$$P(0,0) = \frac{1}{2}, \qquad P(1,1) = \frac{1}{2}, \qquad P(0,1) = P(1,0) = 0. \tag{5.13}$$

To ignore B's record, sum across the two possibilities for B:

$$P_A(0) = P(0,0) + P(0,1) = \frac{1}{2}, \qquad P_A(1) = \frac{1}{2}. \tag{5.14}$$

Thus A alone looks random. If $B = 0$ is known, however,

$$P(A = 0 \mid B = 0) = \frac{P(0,0)}{P_B(0)} = 1, \tag{5.15}$$

so the same data are perfectly predictable after conditioning. Keep the three questions distinct: a **joint** probability concerns both records, a **marginal** deliberately ignores one record, and a **conditional** probability uses a known record to sort the trials.

The arbitrary-angle singlet formulas used in the Bell discussion are stated next. Their derivation is optional on a first pass.

Optional derivation: singlet analyzers at arbitrary angles

For the singlet

$$|\Psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{5.16}$$

Measure A along $\mathbf{a}$ and B along $\mathbf{b}$, and assign numerical results $r, s = \pm 1$ to the two outcomes. Chapter 3 gives the corresponding projectors as

$$\hat{P}_{\mathbf{a},r} = \frac{1}{2}(\hat{I} + r\,\mathbf{a}\cdot\boldsymbol{\sigma}), \qquad \hat{P}_{\mathbf{b},s} = \frac{1}{2}(\hat{I} + s\,\mathbf{b}\cdot\boldsymbol{\sigma}). \tag{5.17}$$

The singlet density operator has the useful Pauli form

$$\hat{\rho}_{\Psi^-} = \frac{1}{4}\left(\hat{I}\otimes\hat{I} - \sum_{j=x,y,z}\sigma_j\otimes\sigma_j\right). \tag{5.18}$$

This identity can be checked by inserting the Pauli matrices and comparing the resulting 4×4 matrix with $|\Psi^-\rangle\langle\Psi^-|$. Now use the trace rules $\mathrm{Tr}(\sigma_i) = 0$ and $\mathrm{Tr}(\sigma_i\sigma_j) = 2\delta_{ij}$. The joint Born rule gives

$$\begin{aligned} P(r,s|\mathbf{a},\mathbf{b}) &= \mathrm{Tr}\left[\hat{\rho}_{\Psi^-}(\hat{P}_{\mathbf{a},r}\otimes\hat{P}_{\mathbf{b},s})\right] \\ &= \frac{1}{4}\left(1 - rs\sum_j a_j b_j\right) \\ &= \frac{1}{4}(1 - rs\,\mathbf{a}\cdot\mathbf{b}). \end{aligned} \tag{5.19}$$

This calculation combines projectors, tensor products, density operators, and trace identities. It is an efficient derivation once those ingredients are familiar, but it should not be the first example used to learn the meaning of a marginal probability.

The resulting joint distribution is

$$P(r,s|\mathbf{a},\mathbf{b}) = \frac{1}{4}(1 - rs\,\mathbf{a}\cdot\mathbf{b}), \tag{5.20}$$

and the correlation is

$$E(\mathbf{a},\mathbf{b}) = \sum_{r,s} rsP(r,s|\mathbf{a},\mathbf{b}) = -\mathbf{a}\cdot\mathbf{b}. \tag{5.21}$$

Aligned analyzers always disagree, perpendicular analyzers show no correlation, and oppositely aligned analyzers always agree, under these numerical labels. One more feature of equation (5.20) deserves a direct check: summing it over s gives $P_A(r) = 1/2$ for every r, independent of $\mathbf{a}$ or $\mathbf{b}$. Each side's own marginal statistics are, on their own, featureless coin flips. All the physical content lives in how the two records *compare*, not in either record alone—a point Section 5.4 develops further.

5.4 Reduced states

Suppose an experimenter has access only to spin A, and never touches spin B at all—perhaps B has been carried off to a distant laboratory. What can that experimenter predict, given only the full joint state $\hat{\rho}_{AB}$? The answer is captured by the **reduced density operator**,

$$\hat{\rho}_A = \mathrm{Tr}_B(\hat{\rho}_{AB}), \tag{5.22}$$

where the partial trace sums over any orthonormal basis of B. Concretely, for any orthonormal basis $\{|k\rangle_B\}$,

$$\mathrm{Tr}_B(\hat{\rho}_{AB}) = \sum_k {}_B\langle k|\hat{\rho}_{AB}|k\rangle_B. \tag{5.23}$$

Each term ${}_B\langle k|\hat{\rho}_{AB}|k\rangle_B$ is itself still an operator, but now acting on A alone: B's basis ket has been contracted away on both sides, and the sum over k adds up the resulting A-only pieces.

Let's carry this out explicitly for $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, using the z basis for B. Writing out the full operator,

$$\hat{\rho}_{AB} = |\Phi^+\rangle\langle\Phi^+| = \frac{1}{2}\Big(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|\Big). \tag{5.24}$$

Tracing over B means sandwiching this between ${}_B\langle 0|\cdots|0\rangle_B$ and separately between ${}_B\langle 1|\cdots|1\rangle_B$, then adding the two results. Because ${}_B\langle 0|0\rangle_B = {}_B\langle 1|1\rangle_B = 1$ while ${}_B\langle 0|1\rangle_B = {}_B\langle 1|0\rangle_B = 0$, any term whose two B labels disagree vanishes immediately, including both off-diagonal terms, $|00\rangle\langle 11|$ and $|11\rangle\langle 00|$, which are exactly the terms carrying $|\Phi^+\rangle$'s coherence. What survives is only

$$\hat{\rho}_A = {}_B\langle 0|\hat{\rho}_{AB}|0\rangle_B + {}_B\langle 1|\hat{\rho}_{AB}|1\rangle_B = \frac{1}{2}|0\rangle_A\langle 0|_A + \frac{1}{2}|1\rangle_A\langle 1|_A = \frac{1}{2}\hat{I}, \tag{5.25}$$

and the identical calculation, tracing over A instead, gives $\hat{\rho}_B = \hat{I}/2$ as well.

Notice exactly which piece of information the partial trace discarded. It kept the diagonal terms, which correctly predict that a z measurement of A alone is 50–50. It discarded the off-diagonal $|00\rangle\langle 11|$ coherence—the very feature that made $|\Phi^+\rangle$ a *pure* joint state rather than the classical mixture of Section 5.8. That coherence is a property of the joint operator alone; it has no counterpart living inside either local factor, which is exactly why no measurement confined to A by itself can ever recover it.

The joint state is pure, but each subsystem, considered alone, is maximally mixed. This is not ordinary ignorance about which pre-existing Bell component actually occurred; the coherence that makes $|\Phi^+\rangle$ a pure state resides in joint off-diagonal terms such as $|00\rangle\langle 11|$, which the reduced state at A alone cannot see.

If B is measured in the z basis and 0 is selected, A is conditionally prepared in $|0\rangle$. If B instead measures in the x basis and selects $+x$, A is conditionally prepared in $|+x\rangle$. Yet before learning which setting was used, or what its outcome was, an observer at A still assigns exactly $\hat{I}/2$, no matter what choice is made at B.

Correlation is not a signal

Changing the analyzer at B changes how joint data are sorted after the two records are compared. It does not change A's local marginal probabilities. Entanglement therefore creates correlations stronger than classical factorization but cannot transmit a controllable message without an ordinary communication channel.

This is a general feature, not a special property of $|\Phi^+\rangle$: for *any* joint state, the reduced density operator $\hat{\rho}_A = \mathrm{Tr}_B(\hat{\rho}_{AB})$ used to predict A's local statistics does not depend on which measurement, if any, is later performed at B. Choosing a different analyzer at B—or none at all—only changes how an already-fixed set of joint records can be sorted into subsets once the two sides compare notes; it cannot change the unconditional distribution of A's own outcomes.

Here is the algebra behind that statement. If B performs a complete projective measurement and the outcome is not supplied to A, the joint state becomes

$$\hat{\rho}'_{AB} = \sum_b (\hat{I} \otimes \hat{P}_b)\hat{\rho}_{AB}(\hat{I} \otimes \hat{P}_b). \tag{5.26}$$

Taking the partial trace and cycling operators that act only on B within that trace,

$$\begin{aligned}\hat{\rho}'_A &= \sum_b \mathrm{Tr}_B\Big[\hat{\rho}_{AB}(\hat{I}\otimes\hat{P}_b^2)\Big] \\ &= \mathrm{Tr}_B\left[\hat{\rho}_{AB}(\hat{I}\otimes\sum_b\hat{P}_b)\right] \\ &= \mathrm{Tr}_B(\hat{\rho}_{AB}) = \hat{\rho}_A.\end{aligned} \tag{5.27}$$

Completeness, $\sum_b \hat{P}_b = \hat{I}$, is the decisive step. The same result holds for any local operation at B whose outcomes are pooled and whose probabilities sum to one. If one outcome is selected and communicated, however, A uses a *conditional* state; that sorting requires the ordinary communication channel and cannot produce a signal before the message arrives.

Midpoint mastery check

Before moving to Bell inequalities, make sure you can do all four of the following without consulting the preceding pages:

1. expand a product ket in the four-state basis;
2. decide whether a pure two-qubit ket factors;
3. compute a joint probability and sum it to obtain a marginal; and
4. take a partial trace of a Bell-state density operator.

Diagnostic answers

A product expansion has coefficients ac, ad, bc, bd. A pure coefficient array factors exactly when its 2×2 coefficient matrix has determinant zero. Marginalizing means summing the joint probabilities over the ignored outcome. Tracing either half of any Bell state gives $\hat{I}/2$. If any of these steps is not yet routine, review Sections 5.1–5.4 before proceeding: the Bell and Schmidt calculations reuse all four skills.

Concept check 5.1

If $\hat{\rho}_A = \hat{I}/2$, does that mean the spin at A is definitely unpolarized and uncorrelated with B?

Answer

No. The reduced state $\hat{\rho}_A$ fixes only A's own unconditional statistics; it says nothing about whether A's results agree, disagree, or vary independently with B's. All four Bell states, the Werner state of Section 5.8 at every mixing parameter, and the classical mixture $\hat{\rho}_{\text{class}}$ of Section 5.8 all give $\hat{\rho}_A = \hat{I}/2$ despite having very different—or no—correlations with B. Correlation is a joint-state property; the reduced state has already thrown that information away.

5.5 Bell's test of local hidden variables

A local hidden-variable model supposes that a measurement's result is determined by information carried along from the shared source, and that a setting freely chosen at one station cannot affect the result recorded at a distant one. One experimentally useful constraint on any such model is the CHSH inequality. For two settings $\mathbf{a}, \mathbf{a}'$ at A and $\mathbf{b}, \mathbf{b}'$ at B, define

$$S = E(\mathbf{a}, \mathbf{b}) + E(\mathbf{a}, \mathbf{b}') + E(\mathbf{a}', \mathbf{b}) - E(\mathbf{a}', \mathbf{b}'). \tag{5.28}$$

Every local hidden-variable model of this type satisfies $|S| \leq 2$.

Where does the classical bound of 2 come from?
Here is the one-line idea behind that bound; Exercise 18 works it out in full. A local hidden-variable model says that each of the four numbers A, A', B, B' is fixed in advance to +1 or -1 before any measurement choice is made, independent of which pair actually gets measured together. Then a bit of algebra regroups the CHSH combination as $AB + AB' + A'B - A'B' = A(B + B') + A'(B - B')$. Since B and B' are each ± 1, exactly one of $B + B'$ and $B - B'$ is always 0 while the other is always ± 2; they can never both be large at the same time. So this expression can never exceed 2 in magnitude, no matter how the four hidden values were assigned. Quantum mechanics evades this argument not by breaking the algebra, but because no single joint assignment of definite, pre-existing values to all four settings needs to exist in the first place.

The short argument makes several assumptions that should remain visible. The setting choices are statistically independent of the source variables (often called **measurement independence**); the result at either station does not depend on the distant setting (**locality**); and the recorded sample fairly represents the emitted pairs. Allowing a hidden-variable model to be stochastic rather than deterministic does not evade the bound: any local response probability can be treated as an average over additional local randomness, and averaging numbers that each lie between -2 and 2 cannot leave that interval. Real Bell experiments are designed to close setting, locality, and detection loopholes closely enough that these assumptions can be tested rather than silently presumed.

For the singlet, suitable coplanar settings separated by 45° instead give $|S| = 2\sqrt{2}$. Quantum theory violates the classical bound while still preserving the no-signaling marginal probabilities from Section 5.4.

The conclusion this supports is precise, and it is worth stating carefully: the observed correlations cannot be reproduced by local models in which every result is fixed by shared pre-existing variables, independent of the later settings. Bell tests do not show that experimenters can send information backward in time or faster than light; Section 5.4 already ruled that out on its own.

The Tsirelson bound

Equation (5.28) shows that quantum mechanics *can* exceed the classical bound of 2, but it cannot exceed it arbitrarily. One can show that for the singlet, and more generally for any quantum state and any choice of ± 1-valued observables, $|S| \leq 2\sqrt{2} \approx 2.828$. This number is the **Tsirelson bound**. The coplanar, 45°-separated setting above is not a lucky combination that happens to violate the classical bound; it is the specific configuration that *saturates* the strongest violation quantum mechanics allows.

Why quantum mechanics stops at $2\sqrt{2}$

Let $\hat{A}, \hat{A}', \hat{B}, \hat{B}'$ be the four ± 1-valued observables and define the CHSH operator

$$\hat{C} = \hat{A} \otimes (\hat{B} + \hat{B}') + \hat{A}' \otimes (\hat{B} - \hat{B}'). \tag{5.29}$$

Because each observable squares to the identity, direct multiplication gives

$$\hat{C}^2 = 4\hat{I} - [\hat{A}, \hat{A}'] \otimes [\hat{B}, \hat{B}']. \tag{5.30}$$

A ± 1 observable cannot stretch a normalized vector by more than one, so each commutator has maximum stretch at most 2. Consequently $\hat{C}^2$ has maximum stretch at most $4 + 2 \times 2 = 8$, and $\hat{C}$ at most $\sqrt{8} = 2\sqrt{2}$. Every expectation value therefore obeys $|\langle \hat{C} \rangle| \leq 2\sqrt{2}$. Suitable singlet analyzer directions reach the bound, so the inequality is tight rather than merely an estimate.

Enrichment: post-quantum correlations

That quantum correlations stop at $2\sqrt{2}$, rather than climbing all the way to the logically largest conceivable value of 4, is itself a substantive fact about nature, not a bookkeeping accident. Hypothetical "post-quantum" correlations reaching all the way to $|S| = 4$, called Popescu–Rohrlich boxes, are mathematically consistent with the no-signaling constraint of Section 5.4—each individual marginal can still come out exactly 1/2—yet no quantum state realizes them. Quantum mechanics is more correlated than any classical local theory can be, but less correlated than logical consistency alone would permit. This comparison is useful context but is not required for the remaining sections.

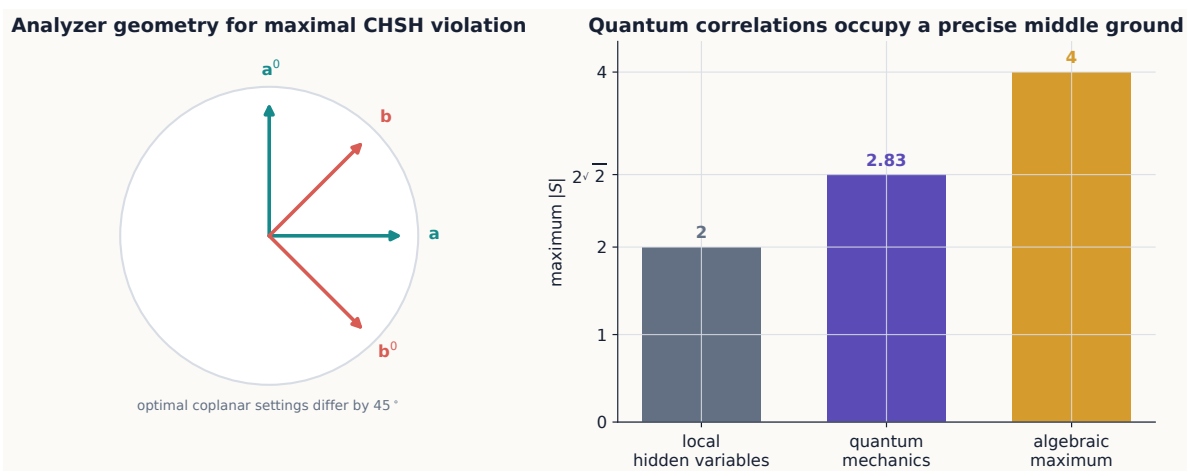

Figure 5.2: The optimal CHSH geometry places the two settings at each station between those at the other station. Quantum mechanics exceeds the local bound but stops well below the largest value that an arbitrary correlation table could produce.

Example 5.3: a setting that does not reveal a violation

A $2\sqrt{2}$ violation requires a specific relationship between the four analyzer directions; it does not happen automatically for any four directions we might pick. Suppose instead the four directions are chosen 22.5° apart in sequence: $\theta_{\mathbf{a}} = 0°$, $\theta_{\mathbf{b}} = 22.5°$, $\theta_{\mathbf{a}'} = 45°$, $\theta_{\mathbf{b}'} = 67.5°$. Using $E(\mathbf{a}, \mathbf{b}) = -\cos(\theta_{\mathbf{a}} - \theta_{\mathbf{b}})$ from equation (5.21),

$$S = -\cos(22.5°) - \cos(67.5°) - \cos(22.5°) + \cos(22.5°) = -\cos(22.5°) - \cos(67.5°) \approx -1.31. \tag{5.31}$$

Here $|S| \approx 1.31$, well below even the classical bound of 2. The very same singlet, analyzed with a different—equally legitimate—choice of four directions, shows no sign of any nonclassical correlation at all. Demonstrating a Bell violation is therefore an experiment-design problem as much as a fact about the state: the directions must be chosen, as in the 45° configuration above, to make the classical and quantum predictions disagree as sharply as possible. This is the same lesson Chapter 1 drew from sequential Stern–Gerlach analyzers: a single measurement rarely distinguishes rival explanations, and a well-chosen combination is needed to force them apart.

Concept check 5.2

Could a future theory find correlations stronger than quantum mechanics's $2\sqrt{2}$, while still forbidding faster-than-light signaling?

Answer

Logically, yes: no-signaling alone only forbids S from being used to send a controllable message, and mathematical correlations (Popescu–Rohrlich boxes) exist that reach the algebraic maximum $|S| = 4$ while keeping every local marginal exactly $1/2$. Quantum mechanics happens to stop at $2\sqrt{2}$. That intermediate value—more correlated than any classical local theory, less correlated than logical consistency alone would allow—is a specific, falsifiable feature of quantum theory, not a generic consequence of forbidding signaling.

5.6 Creating entanglement with gates

In quantum-information notation, the Hadamard gate is

$$\hat{H}_{\mathrm{q}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{5.32}$$

and the controlled-NOT gate acts as $|a, b\rangle \mapsto |a, b \oplus a\rangle$. Starting from $|00\rangle$,

$$|00\rangle \xrightarrow{\hat{H}_{\mathrm{q}} \otimes \hat{I}} \frac{|00\rangle + |10\rangle}{\sqrt{2}} \xrightarrow{\mathrm{CNOT}} \frac{|00\rangle + |11\rangle}{\sqrt{2}}. \tag{5.33}$$

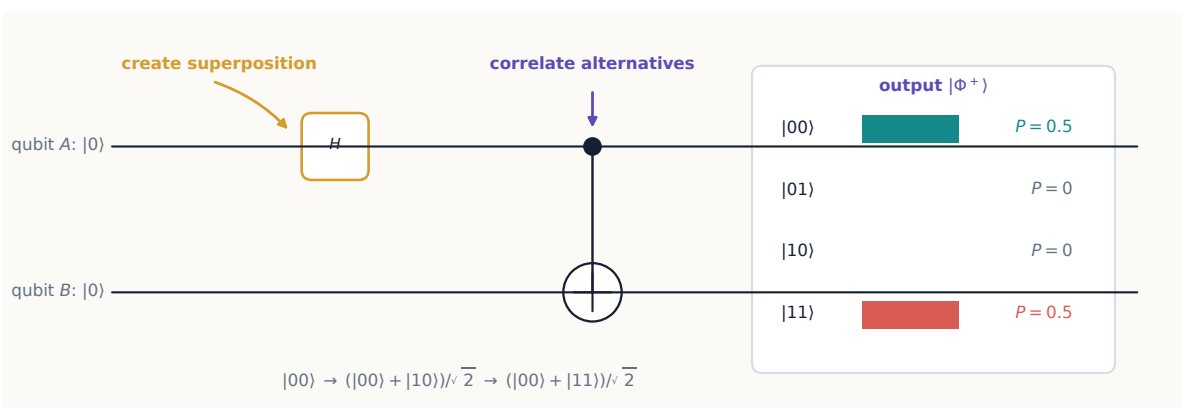

Figure 5.3: The Hadamard creates two coherent alternatives; CNOT correlates the second qubit with them. The output is still a pure state, but it is no longer a product of separate states for A and B.

The Hadamard creates a superposition in A; the controlled operation then correlates each alternative with a distinguishable state of B. A reversible unitary process has produced a pure entangled state, with no measurement involved anywhere.

This last point deserves emphasis: entanglement is created by ordinary unitary evolution acting on a product input, just as a beam splitter followed by a phase shift can entangle two photons, or a spin–spin interaction term in a Hamiltonian can entangle two initially independent spins. Chapter 4's rotations and this section's entangling gates are the same kind of mathematical object—unitary operators—differing only in how many subsystems they act on jointly.

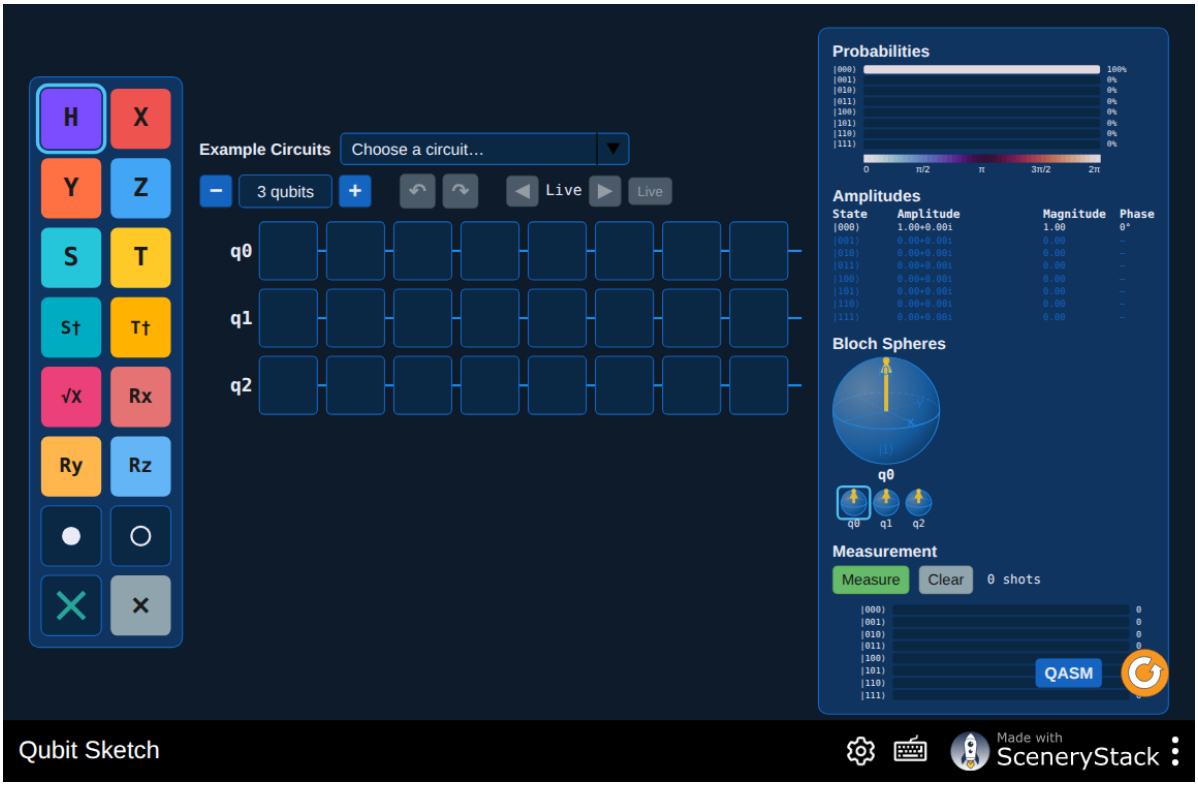

Figure 5.4: Drag a Hadamard gate onto qubit A and a controlled-NOT between A and B to build the circuit above. Watch the two Bloch spheres and the joint-state amplitudes update live, and confirm that conditioning on A's recorded result makes a definite prediction for a later measurement of B, while B's unconditional statistics remain unchanged. Interactive simulation: Qubit Sketch

5.7 Optional extension: Schmidt decomposition and amount of entanglement

Section 5.2 promised a precise, quantitative way to say "how entangled" a state is. Every pure state of two finite systems can be expressed in suitable local orthonormal bases as

$$|\Psi\rangle = \sum_{k=1}^{r} \sqrt{\lambda_k}\, |u_k\rangle_A |v_k\rangle_B, \qquad \lambda_k \geq 0, \quad \sum_k \lambda_k = 1. \tag{5.34}$$

This is the **Schmidt decomposition**. The number r of nonzero coefficients is the Schmidt rank. A pure state is a product exactly when $r = 1$; it is entangled precisely when $r > 1$. The two reduced density operators share the same nonzero eigenvalues λ_k:

$$\hat{\rho}_A = \sum_k \lambda_k |u_k\rangle\langle u_k|, \qquad \hat{\rho}_B = \sum_k \lambda_k |v_k\rangle\langle v_k|. \tag{5.35}$$

Worked method: finding a Schmidt decomposition

When a state is not already in Schmidt form, arrange its amplitudes as a coefficient matrix C, with A labeling rows and B labeling columns:

$$|\Psi\rangle = \sum_{a,b} C_{ab}|a,b\rangle. \tag{5.36}$$

Then $\hat{\rho}_A = CC^\dagger$. Its eigenvalues are the Schmidt weights λ_k, and its eigenvectors are the $|u_k\rangle_A$. The corresponding B vectors can be recovered from

$$|v_k\rangle_B = \frac{1}{\sqrt{\lambda_k}}(\langle u_k| \otimes \hat{I})|\Psi\rangle. \tag{5.37}$$

For example, consider a normalized state that is not visibly in Schmidt form,

$$|\chi\rangle = \frac{2|00\rangle + |01\rangle + |10\rangle + 2|11\rangle}{\sqrt{10}}, \qquad C = \frac{1}{\sqrt{10}}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \tag{5.38}$$

The reduced state is

$$\hat{\rho}_A = CC^\dagger = \frac{1}{10}\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}. \tag{5.39}$$

Its eigenvectors are $|+x\rangle$ and $|-x\rangle$, with eigenvalues 9/10 and 1/10. Applying the recovery formula gives the same two vectors on B, so

$$|\chi\rangle = \frac{3}{\sqrt{10}}|+x,+x\rangle + \frac{1}{\sqrt{10}}|-x,-x\rangle. \tag{5.40}$$

This is the Schmidt decomposition. It immediately shows Schmidt rank two and therefore entanglement. In matrix language, the same procedure is the singular value decomposition of C; its singular values are $\sqrt{\lambda_k}$.

For two qubits, the state

$$|\Psi_\eta\rangle = \cos\eta\,|00\rangle + \sin\eta\,|11\rangle, \qquad 0 \le \eta \le \frac{\pi}{4}, \tag{5.41}$$

is already written in Schmidt form. At $\eta = 0$ it is a product. At $\eta = \pi/4$ its coefficients are equal and each subsystem is maximally mixed: the fully entangled case.

A single number summarizing pure-state entanglement is the entropy of either reduced state:

$$S(\rho_A) = -\operatorname{Tr}(\hat{\rho}_A \log_2 \hat{\rho}_A) = -\sum_k \lambda_k \log_2 \lambda_k. \tag{5.42}$$

It is zero for a product state and exactly one bit for a maximally entangled pair of qubits. For a pure joint state, this entropy measures entanglement by quantifying the mixedness of either reduced state. It is not, by itself, a general measure of all information or correlation in an arbitrary mixed state.

Example 5.4: a partially entangled state

Not every entangled state is maximally entangled. Take

$$|\psi\rangle = \sqrt{\frac{3}{4}}\,|00\rangle + \sqrt{\frac{1}{4}}\,|11\rangle, \tag{5.43}$$

already in Schmidt form with $\lambda_1 = 3/4$, $\lambda_2 = 1/4$ (compare equation (5.34); this is the $\eta = 30°$ member of the $|\Psi_\eta\rangle$ family above). It does not factor: no choice of single-system amplitudes reproduces both nonzero coefficients while leaving the $|01\rangle$ and $|10\rangle$ coefficients zero, so the state is entangled. Its reduced state is

$$\hat{\rho}_A = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|, \tag{5.44}$$

which has $\mathrm{Tr}(\hat{\rho}_A^2) = 9/16 + 1/16 = 5/8 < 1$: A alone is mixed, confirming the entanglement directly. Its entropy is

$$S(\hat{\rho}_A) = -\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4} \approx 0.811 \text{ bit}, \tag{5.45}$$

less than the 1 bit of a Bell state. So this state is entangled, verifiably so, and yet quantifiably "less entangled" than $|\Phi^+\rangle$— exactly the kind of statement the all-or-nothing language of Section 5.2 could not make precise, but the Schmidt decomposition now can.

Example 5.5: conditional states in a rotated basis

Write $|\Phi^+\rangle$ as

$$|\Phi^+\rangle = \frac{|+y,-y\rangle + |-y,+y\rangle}{\sqrt{2}}. \tag{5.46}$$

The y records always disagree, even though the z and x records always agree. Suppose B obtains $+y$. Applying $\hat{I}\otimes|+y\rangle\langle +y|$ and normalizing leaves A in $|-y\rangle$. The joint state determines this conditional preparation exactly, while the outcome itself is still random. Changing basis reveals a different correlation pattern without changing the entangled state one bit.

5.8 Entanglement, mixtures, and correlation

Correlation alone does not prove entanglement; here is a clean example where it does not. The mixed state

$$\hat{\rho}_{\text{class}} = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11| \tag{5.47}$$

has perfect agreement in z, just like $|\Phi^+\rangle$. But it lacks the off-diagonal $|00\rangle\langle 11|$ coherence that made $|\Phi^+\rangle$ special. Rewriting $\hat{\rho}_{\text{class}}$ in the x basis shows that its x outcomes are completely uncorrelated, whereas the Bell state remains perfectly correlated there too.

A mixed state is called **separable** if it can be written

$$\hat{\rho}_{\text{sep}} = \sum_j w_j \hat{\rho}_A^{(j)} \otimes \hat{\rho}_B^{(j)}. \tag{5.48}$$

Such states can still contain shared classical randomness, but nothing more. Entangled mixed states cannot be written this way. Bell-inequality violation certifies entanglement when

it occurs, but some entangled mixed states do not violate any particular Bell test: the choice of measurement matters.

Optional advanced extension: the Werner-state boundary

A cleaner illustration of the boundary between separable and entangled is the **Werner state**,

$$\hat{\rho}_W = p\,|\Phi^+\rangle\langle\Phi^+| + (1-p)\frac{\hat{I}}{4}, \qquad 0 \le p \le 1, \tag{5.49}$$

a mixture of the maximally entangled $|\Phi^+\rangle$ with the maximally mixed two-qubit background $\hat{I}/4$, which carries no correlation at all. At $p = 0$ this is two independent, fully random spins; at $p = 1$ it is the pure Bell state. To locate the entanglement boundary rather than merely quote it, use the **partial-transpose test**. Transposing only B changes a matrix element by

$$\langle a, b|\hat{\rho}|a', b'\rangle \longmapsto \langle a, b'|\hat{\rho}^{T_B}|a', b\rangle. \tag{5.50}$$

Every separable state remains positive under this operation. For two qubits, the converse also holds: a positive partial transpose is sufficient for separability. The partial transpose of $|\Phi^+\rangle\langle\Phi^+|$ has three eigenvalues $+1/2$ and one eigenvalue $-1/2$. Adding the maximally mixed part therefore gives

$$\mathrm{eig}(\hat{\rho}_W^{T_B}) = \left\{\frac{1+p}{4}, \frac{1+p}{4}, \frac{1+p}{4}, \frac{1-3p}{4}\right\}. \tag{5.51}$$

The final eigenvalue becomes negative precisely when $p > 1/3$, certifying entanglement. For $p \le 1/3$, all four eigenvalues are nonnegative and the two-qubit sufficiency result guarantees separability. Mixing in only a little of the entangled state therefore leaves a classically explicable mixture; only past a specific threshold does the joint state require entanglement. "How entangled" a mixed state is does not scale simply with how much of an entangled state was mixed in.

Curiously, the reduced state $\hat{\rho}_A = \mathrm{Tr}_B(\hat{\rho}_W) = \hat{I}/2$ for every value of p: local measurements at A alone cannot distinguish any Werner state from any other, entangled or not. Whether $\hat{\rho}_W$ is entangled is a fact about the joint state that no amount of data collected at A alone, or at B alone, can ever reveal.

Concept check 5.3

The reduced states of $\hat{\rho}_{\text{class}}$, $|\Phi^+\rangle$, and every Werner state $\hat{\rho}_W$ are all $\hat{I}/2$. Does this mean "entangled" is not a meaningful physical distinction, since it cannot be measured locally?

Answer

It means entanglement is a property of the *joint* state, not of either local marginal, so it cannot be certified from A's data or B's data considered separately. It remains an operationally meaningful and experimentally testable distinction: comparing the two players' records after the fact—as in the CHSH statistic of Section 5.5, or the raw joint probabilities of Section 5.3—distinguishes them perfectly well. "Not locally measurable" is not the same as "not physically real"; it simply identifies which experiments are capable of detecting it.

Local unitary operations can rotate the Schmidt bases, but they cannot change the Schmidt coefficients, the reduced-state purity, or the entanglement entropy. Creating or destroying entanglement genuinely requires an interaction, a joint measurement, or discarding information into an external system; no amount of purely local fiddling will do it.

Enrichment: beyond two subsystems

Three or more entangled qubits raise questions this two-qubit chapter deliberately sets aside. The state $(|000\rangle + |111\rangle)/\sqrt{2}$, for example, does not factor, yet—as you can check by tracing out any one qubit—every *pair* of its three qubits is left in exactly the unentangled mixture $\hat{\rho}_{\text{class}}$ of this section. The three-qubit correlations are entirely genuinely tripartite; no two of the three qubits are entangled with each other at all. This is an instance of **monogamy of entanglement**: a qubit maximally entangled with one partner has none left over to share with a third. Multipartite entanglement is a rich subject in its own right and is not developed further here.

Concept check 5.4

Is $(|00\rangle + |01\rangle)/\sqrt{2}$ entangled?

Answer

No. It factors as $|0\rangle_A(|0\rangle_B + |1\rangle_B)/\sqrt{2} = |0\rangle_A|+x\rangle_B$.

Concept check 5.5

For $|\Phi^+\rangle$, does measuring B create a random result at A?

Answer

The joint state already predicts correlated records. Conditioning on B's record changes the state used for predictions at A, but without that record A's local results remain 50–50. No controllable local change identifies whether or how B was measured.

Concept check 5.6

Can two different pure joint states have the same reduced state for A?

Answer

Yes. All four Bell states give $\hat{\rho}_A = \hat{I}/2$. Local statistics do not contain all information in a joint state.

Summary

- Composite state spaces are tensor products, and their dimensions multiply.
- Product states factor into subsystem states; entangled states do not.
- Joint probabilities can contain perfect correlations while each marginal distribution remains random.
- A partial trace gives the state relevant to local measurements.
- Bell correlations violate a bound obeyed by local hidden-variable models but do not permit faster-than-light signaling, and quantum mechanics itself never exceeds the Tsirelson bound $2\sqrt{2}$.
- Superposition followed by a controlled gate can create entanglement.
- Entanglement is quantitative, not binary: Schmidt coefficients and entanglement entropy measure how far a state is from a product, and mixed states such as the Werner state are entangled only above a threshold mixing parameter.

Exercises

Exercises 1–12 are the foundational set. Exercises 13–18 consolidate the core ideas and begin using the optional Schmidt machinery. Exercises 19–32 are extensions and synthesis problems; assigning a representative subset is more useful than treating all fourteen as routine end-of-chapter practice.

Core practice

1. Expand $|+x\rangle_A|+y\rangle_B$ in the joint z basis.
2. Determine which states factor: $(|00\rangle - |10\rangle)/\sqrt{2}$, $(|00\rangle + |11\rangle)/\sqrt{2}$, and $(|00\rangle + |01\rangle + |10\rangle + |11\rangle)/2$.
3. Calculate all joint z probabilities for each Bell state.
4. Rewrite $|\Phi^-\rangle$ in the x basis and describe its correlations.
5. For the singlet, use equation (5.20) with an angle of $60°$ to find all four joint probabilities.
6. Compute $\hat{\rho}_A$ and the entanglement entropy for $\sqrt{2/5}|00\rangle + \sqrt{3/5}|11\rangle$. Is A pure?
7. Show that a product pure state has a pure reduced density operator.
8. Prove from equation (5.20) that both local marginal distributions are $1/2$ for every pair of settings.
9. Use singlet correlations to evaluate the CHSH expression for directions with planar angles $0°, 90°, 45°, -45°$.
10. Write the 4×4 matrix of CNOT in the ordered computational basis and verify its action on all four basis kets.
11. Start with each computational basis state, apply Hadamard to A and then CNOT, and identify the resulting Bell state.
12. Explain why the phrase "each entangled spin has a definite but unknown state" fails for $|\Phi^+\rangle$ when both z and x correlations are considered.
13. Find the reduced density operators and entanglement entropy of $\cos\eta|00\rangle + \sin\eta|11\rangle$.
14. Verify the y-basis expansion in Example 5.5 and its anticorrelations.
15. Compare all joint x probabilities of $\hat{\rho}_{\text{class}}$ and $|\Phi^+\rangle$.
16. Prove that applying $\hat{U}_A \otimes \hat{U}_B$ does not change the eigenvalues of either reduced density operator.
17. For the singlet, evaluate $E(\mathbf{a}, \mathbf{b})$ at relative angles $30°$ and $150°$, and explain the sign of each result in terms of typical agreement or disagreement between the two records.
18. Derive the elementary classical CHSH bound directly: for any four numbers $A, A', B, B' \in \{+1, -1\}$, show that $AB + AB' + A'B - A'B' = A(B + B') + A'(B - B')$, and that this expression cannot exceed 2 in magnitude because $B + B'$ and $B - B'$ cannot both be nonzero.

Extensions and synthesis

19. Verify the arithmetic of Example 5.3 by direct substitution into equation (5.21), and find one nearby setting—a small deviation from the 45°-spaced configuration of Section 5.5—that still violates the classical bound of 2.

20. Show that the reduced state $\hat{\rho}_A$ of the Werner state (5.49) equals $\hat{I}/2$ for every value of p, by computing the partial trace of both terms separately.

21. For $p = 1/2$, decide whether the Werner state is separable or entangled and justify your answer using the threshold stated in Section 5.8.

22. Verify that the two-qubit reduced state of the GHZ state $(|000\rangle + |111\rangle)/\sqrt{2}$, obtained by tracing out any one qubit, equals $\hat{\rho}_{\text{class}}$ from Section 5.8. Explain why this shows the pairwise correlations contain no entanglement even though the three-qubit state does not factor.

23. Two spins are prepared independently, each in $|+z\rangle$, so the joint state is a product. If A is measured along $\mathbf{a}$ and B along $\mathbf{b}$, find $E(\mathbf{a}, \mathbf{b})$ and show that the resulting CHSH combination cannot exceed the classical bound of 2 for any four settings.

24. A student claims that because $\hat{\rho}_A = \hat{I}/2$ for $|\Phi^+\rangle$, the two spins "must actually be independently random, like a coin flip on each side." Using the CHSH violation of Section 5.5, explain precisely what is wrong with this claim.

25. Show that applying a local phase gate $\hat{S} \otimes \hat{I}$, with $\hat{S} = \text{diag}(1, i)$, to $|\Phi^+\rangle$ changes the state but not its Schmidt coefficients or entanglement entropy.

26. Construct the two-qubit state produced by applying CNOT (control A, target B) to $|+x\rangle_A|0\rangle_B$, and show it is a Bell state.

27. Compute the entanglement entropy for Schmidt coefficients $\lambda_1 = 0.9, \lambda_2 = 0.1$ and compare it with the 0.811-bit result of Example 5.4. Which state is closer to a product state, and does the entropy order match your intuition?

28. For a general two-qubit pure state written in Schmidt form with $\lambda_1 = \cos^2\eta$, $\lambda_2 = \sin^2\eta$, argue from the shape of $-\lambda\log_2\lambda - (1-\lambda)\log_2(1-\lambda)$ as a function of $\lambda \in [0, 1]$ that the entanglement entropy is maximized at $\eta = \pi/4$ and vanishes at $\eta = 0$ or $\pi/2$.

29. Two experimentalists share 10,000 singlet pairs and measure CHSH correlations using the optimal 45°-spaced settings of Section 5.5. Each correlation estimate is subject to the same $1/\sqrt{N}$ statistical fluctuation discussed in Chapter 1. Estimate the order of magnitude of the statistical uncertainty in S, and explain why a modest violation just above 2, rather than any positive excess at all, is needed before ruling out statistical fluctuation as an explanation.

30. Explain, in your own words, why "the two spins are entangled" and "the two spins are correlated" are not synonyms, using $\hat{\rho}_{\text{class}}$ and $|\Phi^+\rangle$ as your two contrasting examples.

31. A colleague proposes building a faster-than-light telegraph by having A choose to measure in the z basis to send a "0" and the x basis to send a "1," using a shared $|\Phi^+\rangle$ pair with B far away. Using the results of Sections 5.3 and 5.4, explain specifically what B would need to do to detect which basis A used, and why the attempt fails.

32. Using the definition of a separable mixed state in Section 5.8, verify explicitly that $\hat{\rho}_{\text{class}} = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$ has the separable form $\sum_j w_j \hat{\rho}_A^{(j)} \otimes \hat{\rho}_B^{(j)}$, and identify the weights w_j and single-qubit states explicitly.

Selected exercise guidance

Use these answers to check method and conventions after making a complete attempt. They are intentionally selective rather than a substitute for working the full set.

Exercise 1

$$|+x\rangle_A|+y\rangle_B = \frac{1}{2}\left(|00\rangle + i|01\rangle + |10\rangle + i|11\rangle\right). \tag{5.52}$$

Exercise 2

The first state factors as $|-x\rangle_A|0\rangle_B$; the second is the Bell state $|\Phi^+\rangle$ and is entangled; the third factors as $|+x\rangle_A|+x\rangle_B$.

Exercise 5

At $60°$, $\mathbf{a}\cdot\mathbf{b} = 1/2$. Each equal-result probability is $1/8$, and each opposite-result probability is $3/8$. Their sum is one and the correlation is $-1/2$.

Exercise 9

With $(\theta_a, \theta_{a'}, \theta_b, \theta_{b'}) = (0°, 90°, 45°, -45°)$, the four terms in the stated CHSH convention combine to $S = -2\sqrt{2}$, so $|S| = 2\sqrt{2}$.

Exercise 21

For $p = 1/2$, the final partial-transpose eigenvalue is $(1-3p)/4 = -1/8$. Its negativity certifies that the state is entangled.

Chapter 6

Position and Momentum Representations

6.1 Position and momentum representations

Learning objectives

After this chapter, you should be able to:

- interpret a wavefunction as the coordinates of a state in a continuous basis;
- normalize wavefunctions and calculate position probabilities;
- represent position and momentum operators in coordinate space;
- transform between position and momentum representations;
- use commutators to derive translations and uncertainty bounds; and
- calculate expectation values for continuous observables.

Sections 6.1–6.7 form the conceptual core. Sections 6.8–6.9 are applications to packet spreading and expectation-value dynamics; they may be taken as a second pass before the chapter's final representation check.

6.1 From discrete labels to a continuum

Every chapter so far has worked with a spin's *discrete* state space: two basis kets, a handful of amplitudes, sums instead of integrals. A particle's position, by contrast, can take a continuum of values, and this chapter extends the same logic that worked for spin to that continuous setting. Spin states already taught us the key idea: a ket is abstract, and a column of numbers is only its representation in some chosen basis. Position simply uses a basis with a continuous label instead of a discrete one. The generalized ket $|x\rangle$ represents an ideal position outcome, and the position-space **wavefunction** is defined exactly as an amplitude was for spin:

$$\psi(x) = \langle x|\psi\rangle. \tag{6.1}$$

Quick review: from a discrete sum to a continuous integral
Skip this box if replacing a sum with an integral already feels natural.

For spin, completeness read $\sum_n |n\rangle\langle n| = \hat{I}$, a *finite* sum over two labels. Position has a continuous label x in place of a discrete index n, and turning a sum into its continuum

analog always follows the same two-step recipe: chop the line into narrow bins of width Δx, write a Riemann sum over those bins, and then let $\Delta x \to 0$ so the sum becomes an integral. If $|n;\Delta x\rangle$ is a normalized state localized in the nth bin, its approximate completeness relation is an ordinary sum, $\sum_n |n;\Delta x\rangle\langle n;\Delta x| \approx \hat{I}$. Define sampled generalized kets by $|x_n\rangle = |n;\Delta x\rangle/\sqrt{\Delta x}$. Then

$$\sum_n |x_n\rangle\langle x_n|\,\Delta x \xrightarrow[\Delta x\to 0]{} \int_{-\infty}^{\infty} |x\rangle\langle x|\,dx. \tag{6.2}$$

The factor of Δx accompanies the *generalized* sampled kets; it is what turns into dx under the integral sign. A normalized finite-bin ket, by contrast, belongs in the sum without that factor. This scaling is why an orthonormality condition that used to be a simple 1 or 0—the Kronecker delta δ_{mn}—must turn into something that blows up at coincidence and vanishes elsewhere just fast enough to cancel a factor of dx: the Dirac delta function introduced next. Everything below follows from taking this limit seriously, rather than treating $|x\rangle$ as though it obeyed the same finite normalization already familiar from $|+z\rangle$.

The completeness relation, which for spin was a finite sum over two basis kets, becomes an integral over the continuum of possible positions:

$$\int_{-\infty}^{\infty} |x\rangle\langle x|\,dx = \hat{I}, \tag{6.3}$$

so that

$$|\psi\rangle = \int_{-\infty}^{\infty} \psi(x)|x\rangle\,dx. \tag{6.4}$$

The basis is **delta-normalized**:

$$\langle x|x'\rangle = \delta(x-x'). \tag{6.5}$$

The Dirac delta is not an ordinary finite function; it is defined entirely by its effect inside an integral:

$$\int f(x')\delta(x'-x)\,dx' = f(x). \tag{6.6}$$

One consequence follows immediately: ideal position kets are not normalizable physical states at all. Real preparations are wave packets of finite width, as Section 6.5 develops in detail.

Optional mathematical detail: why the delta is singular

It helps to see directly why an ordinary finite function cannot play the role of $\langle x|x'\rangle$, rather than simply asserting it. Suppose instead that $\langle x|x'\rangle = g(x-x')$ for some bounded function g. Completeness would then require

$$\langle x'|\psi\rangle = \int g(x'-x)\psi(x)\,dx \tag{6.7}$$

to equal $\psi(x')$ for *every* state $|\psi\rangle$. But a bounded kernel convolves and smooths a function as it integrates against it; it cannot reproduce the function's exact value at a single point. The only way it could would be for the kernel to be concentrated entirely at $x = x'$ with unit weight

there, and no ordinary function is simultaneously concentrated at one point and normalized to integrate to one. Only the singular limit of an infinitely narrow, infinitely tall spike—the Dirac delta—has the sifting property we actually need. This is the same logic that made a discrete basis's Kronecker delta δ_{mn} work for spin: orthonormality is exactly what lets us read off a coefficient by taking an inner product. The continuum version of that same requirement forces delta-function normalization in place of unit inner products, and it is exactly what pushes the generalized kets $|x\rangle$ just outside the space of normalizable states.

A generalized basis is a bookkeeping device

No apparatus prepares $|x\rangle$ itself: doing so would require infinite momentum uncertainty and, for the usual nonrelativistic kinetic Hamiltonian $\hat{p}^2/(2m)$, infinite mean kinetic energy. What is physically prepared are normalizable wave packets, expressed as superpositions $\int \psi(x)|x\rangle\, dx$ over the generalized basis. The basis kets earn their keep by making the expansion coefficients—the wavefunction—easy to extract via (6.1), exactly as $|+z\rangle, |-z\rangle$ did for spin.

6.2 The continuous Born rule

The probability of finding the particle between a and b is

$$P(a \le x \le b) = \int_a^b |\psi(x)|^2\, dx. \tag{6.8}$$

Normalization requires

$$\int_{-\infty}^{\infty} |\psi(x)|^2\, dx = 1. \tag{6.9}$$

Unlike a discrete probability, $|\psi(x)|^2$ is a probability **density**, with units of inverse length. This has a striking consequence: the probability of finding the particle at any one exact point is zero, even at a point where the density itself is nonzero.

This is a genuine conceptual shift from the spin chapters, not merely a change of notation. There, $|\langle +z|\psi\rangle|^2$ was already, by itself, a probability. Here one more step—multiplying by an interval dx, or integrating over a finite range—is required before a number between 0 and 1 appears. Forgetting this step is the most common bookkeeping error when moving from discrete to continuous quantum mechanics: a "large" value of $|\psi(x)|^2$ at some point says only that a narrow interval there carries appreciable probability, not that the point itself is likely.

Example 6.1: normalize an exponential packet

Let $\psi(x) = Ae^{-\kappa|x|}$ with $\kappa > 0$. Because $|x|$ makes the integrand symmetric about the origin, the integral over the whole line splits into twice the integral over $x > 0$, where $|x| = x$ and the exponential is easy to integrate directly:

$$1 = \int_{-\infty}^{\infty} |A|^2 e^{-2\kappa|x|}\, dx = 2|A|^2 \int_0^{\infty} e^{-2\kappa x}\, dx = 2|A|^2 \left[\frac{-e^{-2\kappa x}}{2\kappa}\right]_0^{\infty} = \frac{|A|^2}{\kappa}, \tag{6.10}$$

so one convenient choice is $A = \sqrt{\kappa}$. For $\langle x\rangle$, the integrand $x|\psi(x)|^2 = |A|^2 x e^{-2\kappa|x|}$ is an odd function times an even one, hence odd overall, and an odd function integrated

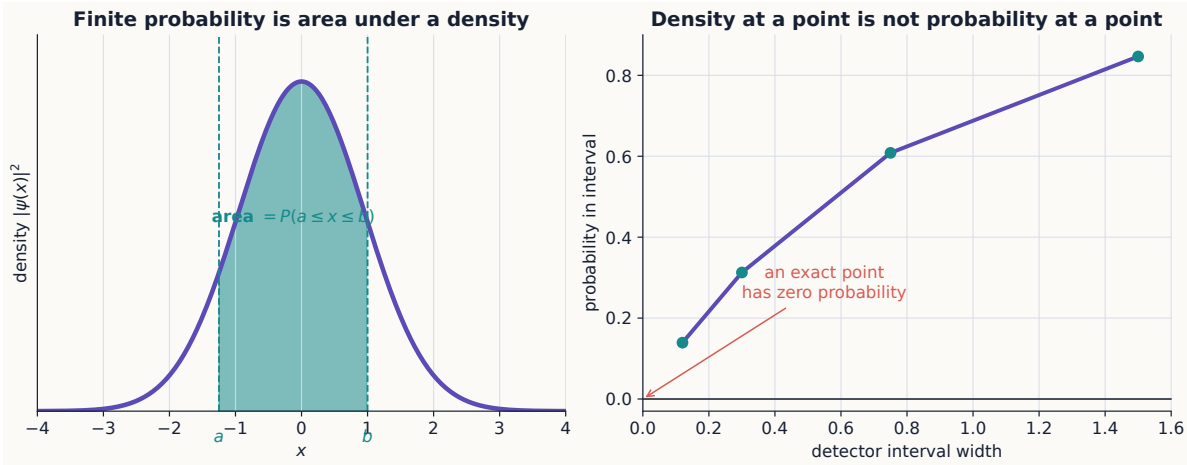

Figure 6.1: In a continuous measurement, height is a density and area is probability. Shrinking the detector interval to zero removes its area even when the density at the selected point remains nonzero.

over a range symmetric about the origin always integrates to zero; so $\langle x \rangle = 0$ without any further calculation. For $\langle x^2 \rangle$, the same even/odd split as before applies, and the standard integral $\int_0^\infty x^2 e^{-2\kappa x}\, dx = 2/(2\kappa)^3 = 1/(4\kappa^3)$ (obtained by integrating by parts twice, or by differentiating $\int_0^\infty e^{-sx} dx = 1/s$ twice with respect to s) gives

$$\langle x^2 \rangle = 2\kappa \int_0^\infty x^2 e^{-2\kappa x}\, dx = 2\kappa \cdot \frac{1}{4\kappa^3} = \frac{1}{2\kappa^2}. \tag{6.11}$$

Thus $\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} = 1/(\sqrt{2}\kappa)$.

Concept check 6.1

Is $|\psi(x_0)|^2$ the probability that a position measurement gives exactly x_0?

Answer

No. It is a density. A finite probability requires integration over a finite interval, and the density carries units of inverse length.

6.3 Operators in the position representation

Position acts by multiplication:

$$\langle x|\hat{x}|\psi\rangle = x\psi(x). \tag{6.12}$$

Momentum acts by differentiation:

$$\langle x|\hat{p}|\psi\rangle = -i\hbar \frac{d\psi}{dx}. \tag{6.13}$$

Apply both operator orders explicitly to an arbitrary function $\psi(x)$, rather than taking the commutator below on faith. The term $\hat{x}\hat{p}\,\psi$ first differentiates, then multiplies by x:

$$\hat{x}\hat{p}\,\psi(x) = x\left(-i\hbar \frac{d\psi}{dx}\right) = -i\hbar\, x \frac{d\psi}{dx}. \tag{6.14}$$

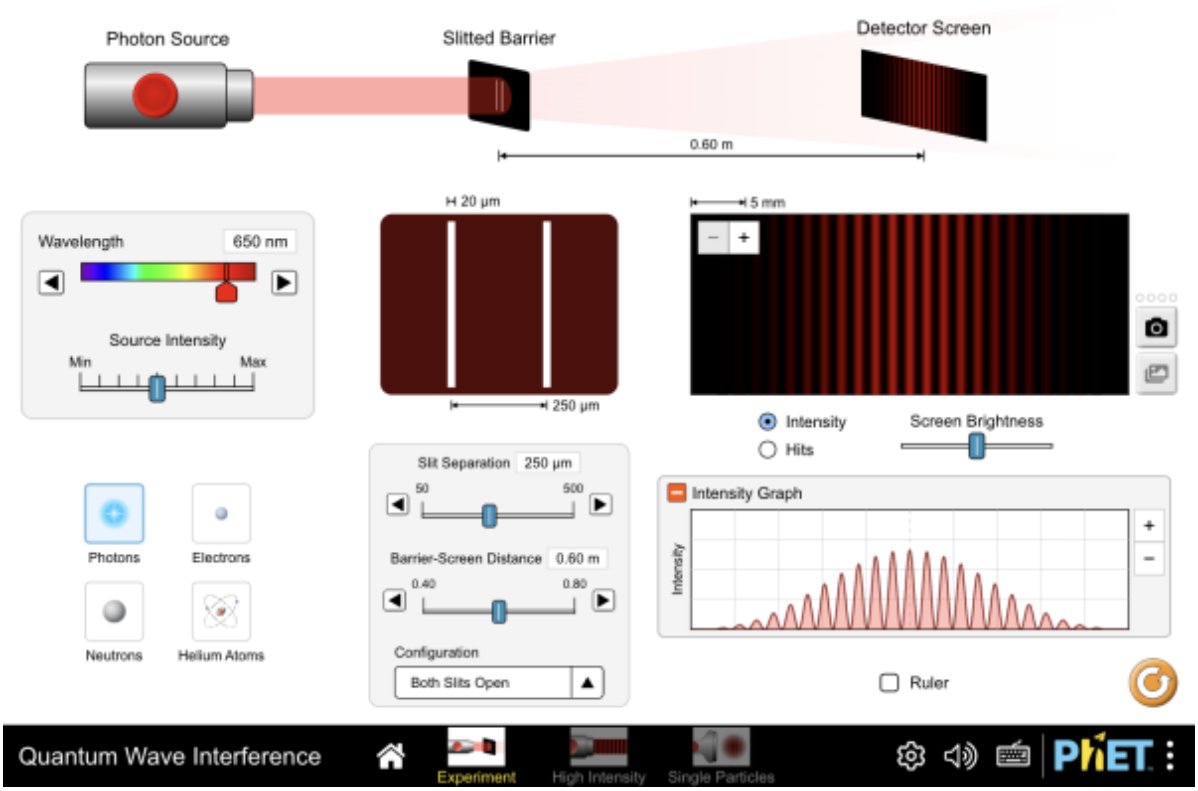

Figure 6.2: Send electrons, photons, or atoms one at a time through a double slit and watch each arrive as a single dot, yet build up the interference pattern predicted by $|\psi(x)|^2$ over many trials. Cover one slit with a detector and the pattern collapses to two bands, the continuous-variable version of Chapter 1's Rule 3. Interactive simulation: Quantum Wave Interference

The term $\hat{p}\hat{x}\,\psi$ first multiplies by x, then differentiates the product $x\psi(x)$ using the ordinary product rule:

$$\hat{p}\hat{x}\,\psi(x) = -i\hbar\frac{d}{dx}\big(x\psi(x)\big) = -i\hbar\,\psi(x) - i\hbar\,x\frac{d\psi}{dx}. \tag{6.15}$$

Subtracting, the term proportional to $d\psi/dx$—the piece generated by the derivative landing on ψ itself—is identical in both orderings and cancels exactly, leaving only the extra term the product rule generates when the derivative instead lands on the factor of x:

$$[\hat{x},\hat{p}]\,\psi(x) = \hat{x}\hat{p}\,\psi - \hat{p}\hat{x}\,\psi = -i\hbar\,x\frac{d\psi}{dx} - \left(-i\hbar\,\psi - i\hbar\,x\frac{d\psi}{dx}\right) = i\hbar\,\psi(x). \tag{6.16}$$

Since $\psi(x)$ was an arbitrary function, this is an operator identity, true independent of any particular state:

$$[\hat{x},\hat{p}]\psi(x) = i\hbar\psi(x), \qquad [\hat{x},\hat{p}] = i\hbar\hat{I}. \tag{6.17}$$

The expectation values are

$$\langle x\rangle = \int \psi^*(x)x\psi(x)\,dx, \qquad \langle p\rangle = \int \psi^*(x)\left(-i\hbar\frac{d}{dx}\right)\psi(x)\,dx. \tag{6.18}$$

Boundary behavior matters here in a way it did not for spin.

Quick calculation: integration by parts and the boundary term

Hermiticity of $\hat{p}$ would require $\langle\phi|\hat{p}\psi\rangle = \langle\hat{p}\phi|\psi\rangle$ for every pair of states ϕ,ψ. Writing the left side as an integral and integrating by parts once,

$$\int_{-\infty}^{\infty}\phi^*(x)\left(-i\hbar\frac{d\psi}{dx}\right)dx = -i\hbar\Big[\phi^*(x)\psi(x)\Big]_{-\infty}^{\infty} + i\hbar\int_{-\infty}^{\infty}\frac{d\phi^*}{dx}\psi(x)\,dx. \tag{6.19}$$

The remaining integral on the right is exactly $\langle\hat{p}\phi|\psi\rangle$, the quantity symmetry of $\hat{p}$ requires the left side to equal. Thus the two sides agree when the boundary form $\big[\phi^*\psi\big]_{-\infty}^{\infty}$ vanishes for every pair in the operator's domain. Square integrability alone does not guarantee a pointwise limit at infinity; one also needs suitable regularity and boundary behavior. The standard momentum domain on the whole line is chosen so these conditions hold.

An unbounded operator is therefore not fully defined by a differential expression alone: its domain is part of its definition. A further subtlety appears on a finite interval. Requiring wavefunctions to vanish at two hard walls makes the boundary form vanish, so $-i\hbar d/dx$ is symmetric on that domain, but it is not self-adjoint there and does not define the well's momentum observable. The operator $-\hbar^2 d^2/dx^2/(2m)$ with the hard-wall boundary conditions *does* define a self-adjoint Hamiltonian. Chapter 7 uses that Hamiltonian directly.

Concept check 6.2

If $\psi(x)$ is multiplied by the constant phase $e^{i\gamma}$, do $\langle x\rangle$ or $\langle p\rangle$ change?

Answer

No. The phase cancels from all expectation values. A position-dependent phase, such as e^{ik_0x}, can change momentum statistics.

Example 6.2: phase carries mean momentum

That last remark is worth demonstrating explicitly, since it is easy to mistake "the phase doesn't matter" for a blanket statement rather than one about *constant* phase specifically. Let $\psi(x) = f(x)e^{ik_0x}$, where f is real, normalized, and vanishes at infinity. Then

$$\langle p\rangle = \int f e^{-ik_0x}\left[-i\hbar\frac{d}{dx}(fe^{ik_0x})\right]dx = \hbar k_0, \tag{6.20}$$

because $\int ff'\,dx = 0$. The density $|\psi|^2 = f^2$ does not reveal k_0 at all; only the spatial phase carries that information. As in the spin interferometer of Chapter 4, phase information becomes observable only through a suitable measurement.

6.4 Momentum states and Fourier transforms

Quick review: what a Fourier transform does

Skip this box if decomposing a function into a superposition of waves e^{ikx} already feels natural.

Just as a spin state can be built out of two special basis states, $|+z\rangle$ and $|-z\rangle$, an ordinary function $f(x)$ can be built out of a continuum of complex "wave" functions e^{ikx}, one for every real wave number k, each carrying its own complex weight $c(k)$:

$$f(x) = \int_{-\infty}^{\infty} c(k)\, e^{ikx}\, dk. \tag{6.21}$$

The **Fourier transform** is simply the rule that reads off the weight $c(k)$ belonging to each wave number, given $f(x)$, playing exactly the role that an inner product $\langle +z|\psi\rangle$ played in extracting a coefficient from a ket. A function that oscillates rapidly in x draws mostly on large-$|k|$ waves; a function that varies slowly draws mostly on small-$|k|$ waves; and a function sharply localized in x, like a narrow spike, turns out to require comparable contributions from every wave number, spread very broadly in k. This trade-off, between how localized a function is in x and how localized its transform is in k, is a purely mathematical fact about waves that has nothing to do with quantum mechanics yet: it is the same reason a sharply plucked guitar string rings with many overtones, while a slow, pure tone contains almost only one frequency.

Equations (6.23) and (6.24) below are exactly this relationship, with the wave number relabeled $p/\hbar$ and the weight relabeled as the momentum-space wavefunction $\phi(p)$. What is new, and genuinely quantum, is the physical claim made just after these equations: that $\phi(p)$ is not merely a bookkeeping coefficient, but a probability amplitude for an independently measurable physical quantity.

Momentum eigenstates satisfy $\hat{p}|p\rangle = p|p\rangle$. In position space,

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}. \tag{6.22}$$

The momentum-space wavefunction is

$$\phi(p) = \langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} e^{-ipx/\hbar}\psi(x)\,dx, \tag{6.23}$$

with inverse

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{ipx/\hbar} \phi(p)\, dp. \tag{6.24}$$

These are basis changes, not dynamical evolution: exactly the same distinction Chapter 2 drew between rewriting a spin ket in a new basis and performing an actual measurement. Normalization is preserved either way: $\int |\psi(x)|^2 dx = \int |\phi(p)|^2 dp = 1$.

Coordinates are not outcomes

Calculating $\phi(p)$ from $\psi(x)$ only rewrites the same state. It neither moves the particle nor measures its momentum. A momentum measurement is a physical interaction that produces one outcome and changes the conditional state used afterward. Section 6.10 returns to this distinction after the applications.

A broad packet in position generally requires a narrow range of wave numbers; a narrow position packet requires many. This Fourier fact underlies the position–momentum uncertainty relation, and separating what is and is not quantum here matters. As a mathematical statement about pairs of Fourier partners, it has no quantum content of its own: it is true of classical wave pulses too. What quantum mechanics adds is the physical claim that $\phi(p)$, not merely $\psi(x)$, is a genuine probability amplitude for an equally real, independently measurable observable. A classical wave packet's Fourier width is a statement about its mathematical shape; a quantum wavefunction's Fourier width is simultaneously a statement about the distribution of a different physical quantity's measured values.

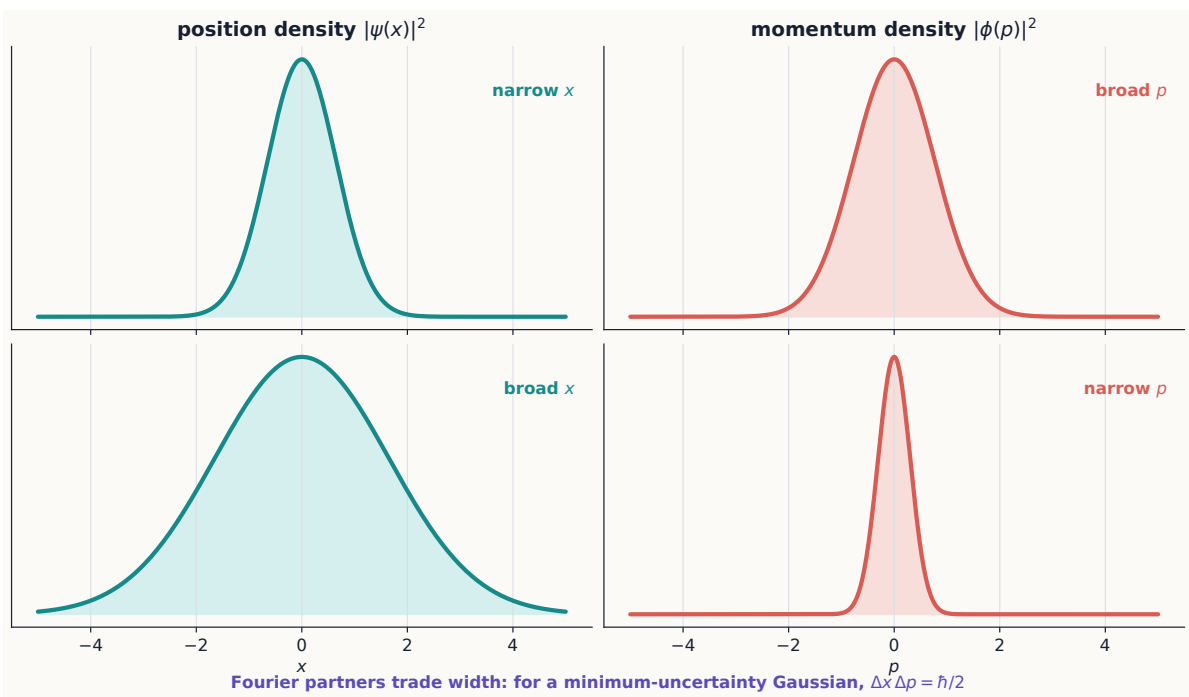

Figure 6.3: Position and momentum descriptions are Fourier partners. Localizing the same state more tightly in one representation necessarily spreads it across more basis states in the other.

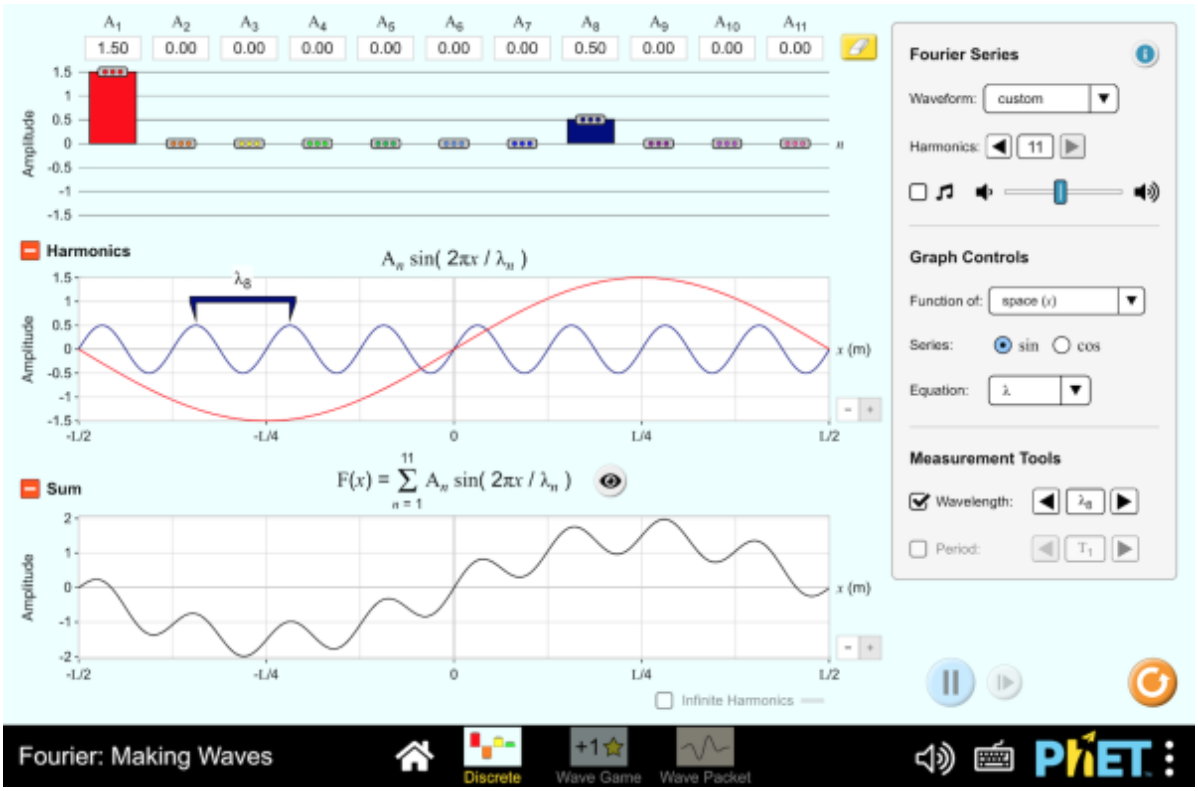

Figure 6.4: Add harmonics one at a time and watch a broad range of wave numbers synthesize a narrow, localized packet in position, or remove harmonics and watch the reverse. The spread displayed for each domain is exactly the trade-off behind equation (6.27) below. Interactive simulation: Fourier: Making Waves

Concept check 6.3

Does a momentum eigenfunction describe a localized physical particle?

Answer

No. Its magnitude is constant over all space and it cannot be normalized in the ordinary sense. It is a generalized basis state; physical beams are modeled by normalizable superpositions with a range of momenta.

6.5 Gaussian wave packets

A normalized Gaussian centered at x_0 with mean wave number k_0 can be written

$$\psi(x) = \frac{1}{(2\pi\sigma_x^2)^{1/4}} \exp\left[-\frac{(x-x_0)^2}{4\sigma_x^2} + ik_0 x\right]. \tag{6.25}$$

It has

$$\langle x\rangle = x_0, \quad \Delta x = \sigma_x, \quad \langle p\rangle = \hbar k_0, \quad \Delta p = \frac{\hbar}{2\sigma_x}. \tag{6.26}$$

Thus

$$\Delta x\, \Delta p = \frac{\hbar}{2}. \tag{6.27}$$

The Gaussian saturates the Robertson bound associated with equation (6.17). This uncertainty is not, however, caused by a badly calibrated position detector kicking the particle around. It describes the spreads found in two genuinely different ensembles, each prepared identically in the same state, one measured for position and the other for momentum.

Optional mathematical extension: why the Gaussian saturates the bound

For the canonical pair, equality in the Cauchy–Schwarz argument behind the uncertainty relation can occur only when the two centered state vectors are related by

$$(\hat{A} - \langle A\rangle)|\psi\rangle = i\lambda(\hat{B} - \langle B\rangle)|\psi\rangle \tag{6.28}$$

for some real number λ. It is worth solving this condition explicitly for $\hat{A} = \hat{x}$, $\hat{B} = \hat{p}$, to see directly why the minimizing states turn out to be Gaussian and not some other shape. In the position representation, with $\langle p\rangle$ momentarily set to zero for simplicity,

$$(x - \langle x\rangle)\psi(x) = i\lambda\left(-i\hbar\frac{d\psi}{dx}\right) = \lambda\hbar\frac{d\psi}{dx}, \tag{6.29}$$

a first-order differential equation with solution

$$\psi(x) \propto \exp\left[\frac{(x-\langle x\rangle)^2}{2\lambda\hbar}\right]. \tag{6.30}$$

Normalizability requires the exponent to decay, so λ must be negative; writing $\lambda = -2\sigma_x^2/\hbar$ reproduces exactly the Gaussian profile of equation (6.25) (restoring $\langle p\rangle = \hbar k_0$ simply reinserts the phase e^{ik_0x}, which affects neither Δx nor Δp). No other line shape solves equation (6.28): any state that is not Gaussian must have $\Delta x\, \Delta p > \hbar/2$ strictly. So Gaussian wave packets, far from being merely a convenient textbook example, are *the* states of least joint position–momentum uncertainty: there is no better one to be found.

Example 6.3: overlap of two wave packets

How distinguishable are two Gaussian packets prepared with the same width σ_x and mean wave number k_0, but different centers x_1 and x_2? Their overlap is

$$\langle\psi_2|\psi_1\rangle = \frac{1}{\sqrt{2\pi\sigma_x^2}}\int_{-\infty}^{\infty}\exp\left[-\frac{(x-x_1)^2}{4\sigma_x^2}-\frac{(x-x_2)^2}{4\sigma_x^2}\right]dx, \tag{6.31}$$

where the common phase e^{ik_0x} has canceled between ψ_2^* and ψ_1. Writing $\bar{x}=(x_1+x_2)/2$, the identity $(x-x_1)^2+(x-x_2)^2 = 2(x-\bar{x})^2+\frac{1}{2}(x_1-x_2)^2$ separates the exponent into a piece that integrates to the normalization and a piece that depends only on the separation:

$$\langle\psi_2|\psi_1\rangle = \exp\left[-\frac{(x_1-x_2)^2}{8\sigma_x^2}\right]. \tag{6.32}$$

Packets separated by many widths ($|x_1-x_2|\gg\sigma_x$) are nearly orthogonal and behave as reliably distinguishable preparations; packets separated by less than a width overlap substantially, and no measurement can sort individual particles between them with certainty. This is the continuous counterpart of the Stern–Gerlach statement that only exactly orthogonal spin states are perfectly distinguishable: here "exactly orthogonal" is replaced by "infinitely far apart," and any finite separation always leaves some residual overlap.

Concept check 6.4

Two Gaussian packets of width σ_x are centered $6\sigma_x$ apart. Are they reliably distinguishable by an ideal position measurement?

Answer

Yes, to excellent approximation. Equation (6.32) gives an overlap of $\exp(-36/8) = \exp(-4.5)\approx 0.011$, so the two states are nearly orthogonal and a position measurement almost always lands cleanly in the region associated with the correct packet.

6.6 Momentum generates translations

Define the translation operator

$$\hat{T}(a) = e^{-ia\hat{p}/\hbar}. \tag{6.33}$$

Using $\hat{p}=-i\hbar\,d/dx$, the exponent becomes $-ia\hat{p}/\hbar = -a\,d/dx$, so in the position representation the exponential's defining power series—the same construction reviewed in Chapter 4 for a matrix exponential, now applied to the differentiation operator d/dx instead—gives

$$\langle x|\hat{T}(a)|\psi\rangle = e^{-a\,d/dx}\psi(x) = \sum_{n=0}^{\infty}\frac{(-a)^n}{n!}\frac{d^n\psi}{dx^n} = \psi(x) - a\,\psi'(x) + \frac{a^2}{2}\,\psi''(x) - \cdots. \tag{6.34}$$

This is exactly the Taylor series of $\psi(x-a)$ expanded about the point x, so

$$\langle x|\hat{T}(a)|\psi\rangle = \psi(x-a). \tag{6.35}$$

The translated packet is centered a distance a to the right: its value at x equals the old packet's value at $x-a$. Infinitesimally,

$$\hat{T}(a) \approx \hat{I} - \frac{ia}{\hbar}\hat{p}. \tag{6.36}$$

Momentum is therefore the **generator** of spatial translations, just as the Hamiltonian is the generator of time translations and spin components generate rotations. The pattern is always the same: a Hermitian observable exponentiates into a unitary operator that implements a continuous symmetry, and the observable's expectation value is conserved exactly when that symmetry leaves the Hamiltonian unchanged. Chapter 9 will show that orbital and spin angular momentum generate rotations in precisely this sense, so that the rotation operators already used for spin in Chapter 4 are a special case of a single unifying structure that also produces (6.33) here.

Conjugating position by the translation gives

$$\hat{T}^\dagger(a)\hat{x}\hat{T}(a) = \hat{x} + a\hat{I}. \tag{6.37}$$

This operator statement makes the geometry independent of any particular wavefunction: it is true for every state at once, not just the ones we happen to write down explicitly.

Concept check 6.5

Does translating a state by a change the *magnitude* of its momentum-space wavefunction?

Answer

No. Since $\hat{T}(a) = e^{-ia\hat{p}/\hbar}$ commutes with $\hat{p}$, translation multiplies $\phi(p)$ by the pure phase $e^{-iap/\hbar}$ and leaves $|\phi(p)|^2$ unchanged. Sliding a packet in space redistributes nothing among its momentum components; only *where* it is found changes, not *how fast* it is found to be moving.

6.7 Measurement resolution and wave packets

An ideal sharp position measurement uses projectors $|x\rangle\langle x|$ only as a mathematical limit. A detector pixel covering a region R corresponds instead to

$$\hat{P}_R = \int_R |x\rangle\langle x|\, dx, \qquad P(R) = \langle\psi|\hat{P}_R|\psi\rangle. \tag{6.38}$$

Selecting that detector outcome produces a state localized to the region, after normalization. A narrower region generally broadens the momentum distribution. This is the continuous version of an analyzer preparing an outcome state, with finite detector resolution replacing an ideal one-dimensional projector.

An idealized limit is not a recipe

Letting R shrink to a point in $\hat{P}_R = \int_R |x\rangle\langle x|\, dx$ sends the normalized post-measurement state toward the non-normalizable $|x\rangle\langle x|/\langle x|x\rangle$, which does not exist as an ordinary quantum state. Real position measurements always have finite resolution, and it is the finite-resolution projector, not its idealized limit, that correctly describes a laboratory detector pixel or a particle track in a bubble chamber.

6.8 Application: free-particle evolution and packet spreading

For $V = 0$, each momentum component is an energy eigenstate with $E(p) = p^2/(2m)$. The momentum wavefunction evolves as

$$\phi(p,t) = e^{-ip^2t/(2m\hbar)}\phi(p,0). \tag{6.39}$$

Its magnitude is unchanged, so the momentum distribution stays fixed for all time. The phase, however, is nonlinear in p, and transforming back to position shows that this nonlinear phase causes a packet to spread. For the minimum-uncertainty Gaussian of equation (6.25),

$$\Delta x(t) = \sigma_x\sqrt{1 + \left(\frac{\hbar t}{2m\sigma_x^2}\right)^2}. \tag{6.40}$$

The center, meanwhile, moves classically:

$$\langle x\rangle_t = x_0 + \frac{\hbar k_0}{m}t. \tag{6.41}$$

A narrow initial packet spreads faster than a broad one, because a narrow packet in position necessarily contains a broader range of momenta—and therefore a broader range of velocities—than a broad one does.

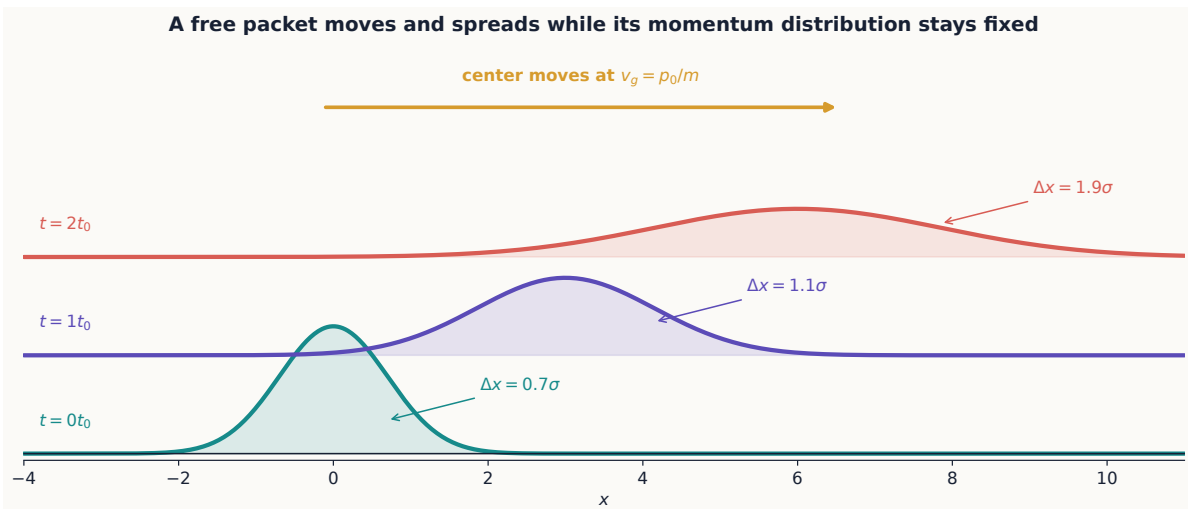

Figure 6.5: Free evolution advances the packet center at the group velocity while its constituent momentum components separate. The fixed momentum spread therefore appears as a growing position spread.

Group and phase velocity

A plane-wave component has phase velocity $v_{\text{ph}} = \omega/k = \hbar k/(2m)$, while the packet envelope itself travels at the different speed

$$v_{\text{g}} = \frac{d\omega}{dk} = \frac{\hbar k}{m} = \frac{p}{m}. \tag{6.42}$$

It is the group velocity that matches the particle's classical velocity. Confusing the two gives a factor-of-two error for the nonrelativistic dispersion relation $\omega = \hbar k^2/(2m)$. The phase

velocity is not meaningless—it governs how the fine ripples inside the envelope move—but it is not what a particle detector downstream registers as "how fast the particle arrived." For a narrow-band packet, the carrier oscillations do not appear separately in $|\psi|^2$, whereas the envelope locates the probability packet. Its group velocity is therefore the speed to compare with the packet's classical motion.

6.9 Application: Ehrenfest's theorem

For

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}), \tag{6.43}$$

the expectation-value equation from Chapter 4 gives

$$\frac{d\langle x\rangle}{dt} = \frac{\langle p\rangle}{m}, \qquad \frac{d\langle p\rangle}{dt} = -\left\langle \frac{dV}{dx} \right\rangle. \tag{6.44}$$

These resemble Hamilton's equations, and it is tempting to conclude that quantum expectation values always follow classical trajectories. They do not, in general: $\langle V'(\hat{x})\rangle \neq V'(\langle x\rangle)$ except in special cases. The center of a narrow packet in a slowly varying potential only *approximately* follows a classical trajectory.

There is one important exception, and the reason it arises deserves a closer look. For quadratic potentials, the equations close exactly for the means. The reason is purely algebraic: the discrepancy $\langle V'(\hat{x})\rangle - V'(\langle x\rangle)$ is controlled by the higher moments of $\hat{x} - \langle x\rangle$, through a Taylor expansion of V' about $\langle x\rangle$. Every term in that expansion beyond the first involves at least $\langle(\hat{x} - \langle x\rangle)^2\rangle$, or a higher power still. If V' is linear in x—equivalently, if V is at most quadratic—that Taylor expansion has no higher terms at all, and the mean-value equations close exactly, regardless of how wide the packet is. For any other potential, this gap is generally nonzero and grows with the packet's spread.

Ehrenfest's theorem does not turn a quantum state into a classical point. The packet retains its uncertainties, may spread, and may even split into reflected and transmitted pieces. Classical behavior emerges only for suitable states, observables, potentials, and time scales, not as a general guarantee.

Example 6.4: constant force

For $V(x) = -Fx$, equation (6.44) gives

$$\langle p\rangle_t = \langle p\rangle_0 + Ft, \qquad \langle x\rangle_t = \langle x\rangle_0 + \frac{\langle p\rangle_0}{m}t + \frac{F}{2m}t^2. \tag{6.45}$$

The mean follows the classical constant-acceleration trajectory exactly, even though the position distribution need not remain narrow.

Example 6.5: the harmonic oscillator

For $V(x) = \frac{1}{2}m\omega^2 x^2$, studied in full in Chapter 7, $V'(x) = m\omega^2 x$ is linear, so equation (6.44) closes exactly:

$$\frac{d\langle x\rangle}{dt} = \frac{\langle p\rangle}{m}, \qquad \frac{d\langle p\rangle}{dt} = -m\omega^2\langle x\rangle. \tag{6.46}$$

Differentiating the first equation and substituting the second gives

$$\frac{d^2\langle x\rangle}{dt^2} = -\omega^2\langle x\rangle, \tag{6.47}$$

the classical simple-harmonic equation of motion, satisfied exactly by $\langle x\rangle_t$ and $\langle p\rangle_t$ for *any* initial state—a narrow packet, a broad one, or a superposition of many energy eigenstates. Together with Example 6.4, this shows that exactly classical mean motion is not a generic feature of quantum dynamics; it is a special consequence of having a potential no steeper than quadratic. The width of the packet is a separate question entirely: for the free particle it grows without bound (equation (6.40)), while a harmonic packet prepared in the right shape returns periodically to its initial width, oscillating in step with $\langle x\rangle_t$.

Concept check 6.6

For the harmonic oscillator, does Ehrenfest's theorem guarantee that a widely spread-out state behaves like a classical particle in every respect?

Answer

No. It guarantees only that the means $\langle x\rangle_t$ and $\langle p\rangle_t$ trace the exact classical trajectory. The state can remain highly uncertain, and observables other than these first moments—such as $\Delta x(t)$ itself—need not follow any classical law.

6.10 A representation is not a physical location

In momentum space the same abstract operators act differently:

$$\langle p|\hat{p}|\psi\rangle = p\phi(p), \qquad \langle p|\hat{x}|\psi\rangle = i\hbar\frac{d\phi}{dp}. \tag{6.48}$$

Nothing has happened to the particle when one changes from $\psi(x)$ to $\phi(p)$. One has simply changed the basis used to describe the same ket. A physical position measurement, by contrast, changes conditional predictions through state update. This distinction is the continuous analog of rewriting a spin ket in the x basis versus actually sending it through an x analyzer.

Representation change versus measurement

Computing $\phi(p) = \langle p|\psi\rangle$ from a known $\psi(x)$ is arithmetic: it uses one fixed state and asks what its coefficients look like in a different, already-chosen basis. Performing a momentum measurement is physical: it couples the system to an apparatus, produces one random outcome, and updates the state. The two operations can produce numerically identical-looking distributions $|\phi(p)|^2$—one as a computed property of the original state, one as a set of relative frequencies from repeated trials—but only the second involves selecting an outcome and preparing a new state for whatever comes next.

Summary

- A wavefunction is a state's coordinate function in the position basis.
- Squared wavefunction magnitude is a probability density.
- Position multiplies by x, while momentum differentiates in the position representation.
- Position and momentum wavefunctions are Fourier transforms of one another.

- Gaussian packets attain the minimum uncertainty product $\hbar/2$ and, apart from their center, mean momentum, and overall phase, are the unique states that do so.
- Momentum generates translations, connecting symmetry with observables; the same pattern reappears for time evolution and rotations.
- Ehrenfest's theorem closes exactly only when the potential is at most quadratic; otherwise it describes an approximate classical trajectory for the mean.
- Finite-resolution measurements project onto ranges rather than exact points.

Exercises

Exercises 1–12 build the core representation, normalization, Fourier, and translation skills. Exercises 13–17 accompany the dynamics applications. Exercises 18–34 are deeper derivations and synthesis; a first course should select from them rather than assign the entire bank at once.

1. Normalize $\psi(x) = A$ on $0 < x < L$ and zero elsewhere. Find $\langle x \rangle$ and Δx.
2. For the exponential packet in Example 6.1, find the probability that $|x| < 1/\kappa$.
3. Verify equation (6.17) by applying both operator orders to an arbitrary differentiable function.
4. Show by integration by parts that $\langle p \rangle$ is real when $\psi(x)$ vanishes at infinity.
5. Normalize $xe^{-\kappa x}$ on $x > 0$ and zero elsewhere.
6. Find $\langle p \rangle$ for a real, normalized wavefunction that vanishes at the boundaries of its domain.
7. Verify the four moments stated for equation (6.25).
8. Fourier transform a Gaussian by completing the square and confirm $\Delta p = \hbar/(2\sigma_x)$.
9. Show that $\hat{T}(a)\hat{T}(b) = \hat{T}(a+b)$ and that $\hat{T}^\dagger(a) = \hat{T}(-a)$.
10. Derive $\hat{T}^\dagger \hat{x} \hat{T} = \hat{x} + a\hat{I}$ to first order in a using the commutator.
11. A detector reports only whether $x < 0$ or $x \geq 0$. Write its two projectors and the conditional wavefunction after the $x \geq 0$ result.
12. Explain how two wavefunctions can have identical position densities but different momentum distributions. Give an explicit pair.

Dynamics applications

13. Derive equation (6.40) by evolving the Gaussian in momentum space, or verify its short- and long-time limits.
14. Find the time at which the width of a free Gaussian has doubled.
15. Derive both Ehrenfest equations using commutators.
16. Give a potential for which $\langle V'(\hat{x}) \rangle = V'(\langle x \rangle)$ for every state and one for which it is generally false.
17. Verify the momentum-space representation of $\hat{x}$ by differentiating the Fourier kernel.

Further extensions and synthesis

18. Suppose $\langle x|x'\rangle = g(x - x')$ for a bounded function g with $\int g(x)\,dx = 1$. Show that $\int g(x' - x)\psi(x)\,dx$ generally differs from $\psi(x')$ for a wavefunction that varies on the scale over which g is spread, and explain why only the Dirac-delta limit removes this discrepancy for every ψ.

19. Using equation (6.32), find the separation $|x_1 - x_2|$, in units of σ_x, at which $|\langle\psi_2|\psi_1\rangle|^2 = \frac{1}{2}$.

20. Repeat the overlap calculation of Example 6.3 for two Gaussian packets with the same center but different mean wave numbers $k_1 \neq k_2$, and show that the result depends on $k_1 - k_2$ and σ_x in a way that parallels equation (6.32).

21. Solve equation (6.28) for $\hat{A} = \hat{x}$, $\hat{B} = \hat{p}$ keeping $\langle p\rangle$ general from the start, and confirm that the phase e^{ik_0x} appears with $\langle p\rangle = \hbar k_0$.

22. Verify by direct substitution that the harmonic-oscillator Ehrenfest equations of Example 6.5 combine to $d^2\langle x\rangle/dt^2 = -\omega^2\langle x\rangle$.

23. Identify the mathematical feature shared by the free particle and the harmonic oscillator that allows Ehrenfest's theorem to close exactly for both, and explain why a cubic potential $V(x) = cx^3$ does not share it.

24. A particle is confined with certainty to a finite interval $[0, L]$. Explain using the Fourier relation why its momentum-space wavefunction cannot be concentrated at a single value of p.

25. Verify $\langle x^2\rangle$ and $\langle p^2\rangle$ for the Gaussian packet of equation (6.25) by direct integration, and confirm the stated Δx and Δp.

26. Two Gaussian packets share a center x_0 but have different widths σ_1 and σ_2. Compute their overlap and show that it depends only on the ratio σ_1/σ_2.

27. Show that a momentum eigenstate is unchanged in magnitude, up to an overall phase, under any translation $\hat{T}(a)$, and interpret this as the statement that momentum eigenstates carry no distinguishing spatial location.

28. Show directly from its integral definition that $\hat{P}_R$ in Section 6.7 is Hermitian and idempotent ($\hat{P}_R^2 = \hat{P}_R$).

29. Model a detector pixel of width ε centered at x_0 by $\hat{P}_\varepsilon = \int_{x_0-\varepsilon/2}^{x_0+\varepsilon/2} |x\rangle\langle x|\,dx$. Show that $P(R)/\varepsilon \to |\psi(x_0)|^2$ as $\varepsilon \to 0$, recovering the probability density from a finite-resolution measurement.

30. Find the time at which $\Delta x(t) = \sqrt{2}\,\sigma_x$ for a free Gaussian packet, in terms of m, σ_x, and $\hbar$.

31. Compute the group velocity for the relativistic dispersion relation $\omega = \sqrt{c^2k^2 + (mc^2/\hbar)^2}$, and show that it reduces to p/m when $\hbar k \ll mc$.

32. A student measures momentum, obtains a result, and then asks for the "position-space picture" of the resulting state by Fourier transforming. Explain what has and has not been physically established about position by this point, contrasting it with an actual position measurement.

33. A real, even wavefunction satisfies $\psi(x) = \psi(-x)$. Using the Fourier relation between ψ and ϕ, show that $\phi(p)$ is real and even, and use that fact to show $\langle p \rangle = 0$ without evaluating the momentum integral directly.

34. Prove that $\hat{T}(a)$ is unitary directly from its exponential definition and the Hermiticity of $\hat{p}$.

Chapter 7

One-Dimensional Wave Mechanics

7.1 One-dimensional wave mechanics

Learning objectives

After this chapter, you should be able to:

- solve the time-independent Schrödinger equation in regions of constant potential;
- impose boundary and matching conditions;
- distinguish bound states from scattering states;
- calculate spectra and wavefunctions for the infinite square well;
- use probability current to interpret reflection and transmission;
- explain tunneling through a classically forbidden region; and
- describe the harmonic oscillator with ladder operators.

Sections 7.1–7.6 form a core progression from boundary-value problems to probability flow and tunneling. The above-barrier resonance subsection is an optional extension. Section 7.7 introduces a second solution strategy—the oscillator's algebraic ladder method—and can be taught after the well and current sections without requiring the resonance extension.

7.1 Schrödinger's equation in position space

Chapter 6 developed the position representation in the abstract: operators become differential expressions, and a ket becomes a wavefunction. This chapter puts that machinery to work on the simplest possible setting, a single spatial dimension, where boundary conditions and matching conditions turn into concrete, solvable calculations. For one nonrelativistic particle of mass m in a potential $V(x,t)$,

$$i\hbar\frac{\partial\psi}{\partial t} = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x,t)\right]\psi(x,t). \tag{7.1}$$

If V is time independent, a separated solution $\psi(x,t) = u(x)e^{-iEt/\hbar}$ obeys

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} + V(x)u(x) = Eu(x). \tag{7.2}$$

This is an eigenvalue equation for the Hamiltonian. A normalizable energy eigenfunction changes only by an overall phase, so its probability density is stationary.

The local behavior of a solution depends entirely on the sign of $E - V$:

- if $E > V$, solutions oscillate with $k = \sqrt{2m(E-V)}/\hbar$;
- if $E < V$, solutions grow or decay exponentially with $\kappa = \sqrt{2m(V-E)}/\hbar$; and
- at abrupt finite changes in V, both u and du/dx are continuous.

The last condition follows by integrating the differential equation across an arbitrarily narrow interval: a finite jump in V cannot produce a δ-function in u'', so u' cannot jump either. Infinite potentials are the one exception, and it is worth being precise about why. An infinite wall forces $u = 0$ at the wall, but u' *is* allowed to jump there, because the argument just given assumed a *finite* jump in V. An infinitely deep well is really the limit of an ever-taller finite well, not a literal potential one can integrate across.

Physically, equation (7.2) is nothing more than the same eigenvalue problem introduced abstractly in Chapter 3, specialized to the position representation of Chapter 6. $\hat{H} = \hat{p}^2/2m + V(\hat{x})$ acting on $|\psi\rangle$ becomes a second-order differential operator acting on $u(x) = \langle x|\psi\rangle$. Nothing new is being postulated here—only the concrete form that an old postulate takes once a continuous basis is chosen.

7.2 The infinite square well

This is the simplest bound-state problem in quantum mechanics, and it is worth solving completely: every later matching calculation in this chapter repeats the same basic steps on a harder potential. Let $V = 0$ for $0 < x < L$ and $V = \infty$ elsewhere. The wavefunction vanishes at $x = 0$ and $x = L$. Inside,

$$u(x) = A \sin kx + B \cos kx. \tag{7.3}$$

The first boundary gives $B = 0$; the second requires $\sin kL = 0$. Thus

$$k_n = \frac{n\pi}{L}, \qquad u_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}, \qquad n = 1, 2, 3, \ldots \tag{7.4}$$

and

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2}. \tag{7.5}$$

Boundary conditions have quantized the energy. There is no $n = 0$ state: it would be the zero function and could not be normalized. The nonzero ground-state energy is required by confinement and the uncertainty principle—a state squeezed into a region of size L cannot also have $p = 0$, so it cannot have $E = 0$.

The infinite well is the simplest model of confinement, and its qualitative lessons reappear whenever a particle is trapped: quantum dots, particles in optical lattices, and nucleons in a schematic nuclear well are all discussed using the same vocabulary of quantized levels and a nonzero ground-state energy. One more pattern is worth noting now, because it will recur for every bound-state problem in this chapter: the integer $n - 1$ counts the number of interior nodes of u_n. More nodes means more curvature, and more curvature means higher energy, since curvature is exactly what converts potential energy into kinetic energy in equation (7.2).

Example 7.1: a superposition in the well

Suppose

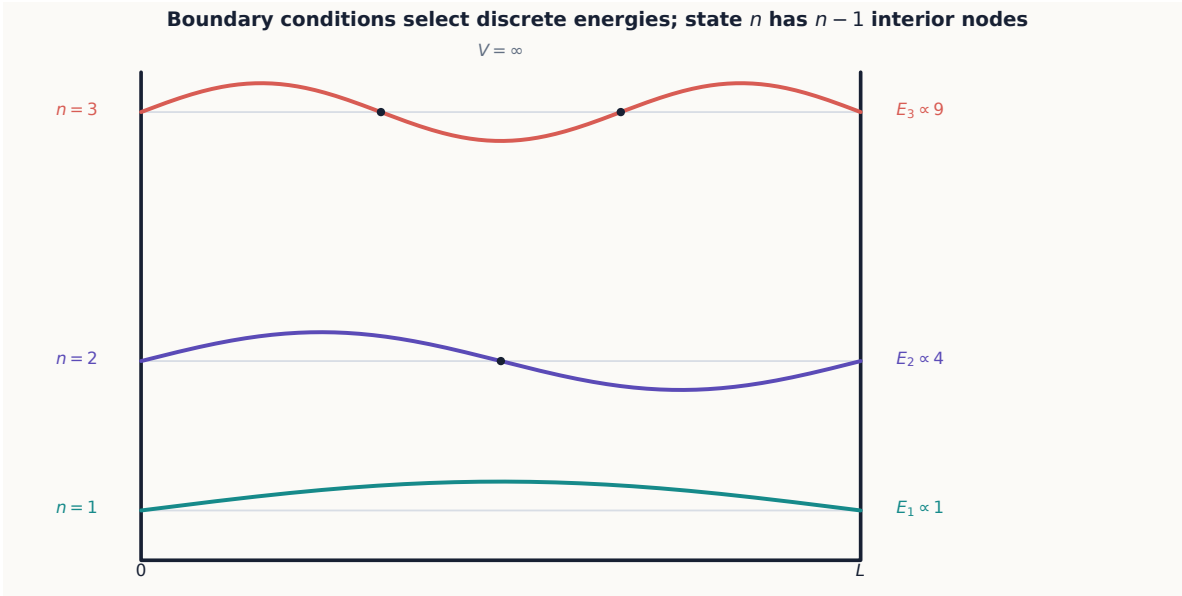

Figure 7.1: Hard-wall boundary conditions admit only wavelengths that fit the box. Higher n brings both more nodes and more curvature, so the kinetic energy rises as n^2 rather than in equal steps.

$$\psi(x,0) = \frac{u_1(x) + u_2(x)}{\sqrt{2}}. \tag{7.6}$$

Then

$$\psi(x,t) = \frac{1}{\sqrt{2}}\left[u_1(x)e^{-iE_1 t/\hbar} + u_2(x)e^{-iE_2 t/\hbar}\right]. \tag{7.7}$$

The energy probabilities stay 1/2, but the position density contains an interference term proportional to $\cos[(E_2 - E_1)t/\hbar]$. A superposition of stationary states need not itself have a stationary density.

Concept check 7.1

Why is $\psi(x,t) = u_n(x)e^{-iE_n t/\hbar}$ called stationary if it depends on time?

Answer

Its time dependence is an overall phase, so $|\psi|^2$ and every time-independent observable distribution are constant.

7.3 General bound states

A bound state is normalizable and localized. For potentials approaching constants at infinity, bound-state energies lie below the asymptotic potential, so the wavefunction decays exponentially outside the classically allowed region.

Several qualitative facts are useful before solving anything further:

- one-dimensional bound-state energies are nondegenerate;
- the ground state has no interior nodes;
- higher states have successively more nodes; and
- for an even potential, eigenstates can be chosen with even or odd parity.

The nondegeneracy claim deserves a moment's justification, and it is special to one dimension specifically. Two independent solutions of a second-order ordinary differential equation at the same E would be related by their Wronskian. Normalizability at both infinities forces that Wronskian to vanish, which in turn forces the two solutions to be proportional to one another—so there is really only one independent bound state at that energy. Chapter 8 will show that this argument fails in three dimensions, where rotational symmetry routinely produces several independent states at one energy.

Quantization is a boundary-value result
The differential equation alone permits solutions at many energies. Discrete energies arise when normalizability and boundary conditions can be satisfied only for selected values.

The finite square well

For a finite square well, oscillatory interior solutions must be matched to decaying exterior solutions. The matching conditions produce transcendental equations rather than the simple integer rule of the infinite well. The finite well has only finitely many bound states, and each leaks into the classically forbidden region.

Let $V(x) = -V_0$ for $|x| < L/2$ and $V(x) = 0$ for $|x| > L/2$, with $-V_0 < E < 0$. Inside, the wavefunction oscillates with $k = \sqrt{2m(E+V_0)}/\hbar$; outside, it decays with $\kappa = \sqrt{-2mE}/\hbar$. Because the well is symmetric about $x = 0$, every bound state can be chosen even or odd, and it is enough to match at $x = L/2$ and use the symmetry to fix the other boundary automatically. Matching u and u' gives

$$k \tan \frac{kL}{2} = \kappa \qquad \text{(even states)}, \tag{7.8}$$

$$k \cot \frac{kL}{2} = -\kappa \qquad \text{(odd states)}. \tag{7.9}$$

Neither equation can be solved in closed form, but both are constrained by energy conservation, $k^2 + \kappa^2 = 2mV_0/\hbar^2$, which is independent of E. In terms of the dimensionless variables $\xi = kL/2$ and $\eta = \kappa L/2$, this constraint is the quarter circle

$$\xi^2 + \eta^2 = z_0^2, \qquad z_0 \equiv \frac{L}{2}\frac{\sqrt{2mV_0}}{\hbar}, \tag{7.10}$$

and the bound states are exactly the intersections of this circle with the curves $\eta = \xi \tan \xi$ (even) and $\eta = -\xi \cot \xi$ (odd) in the first quadrant. The dimensionless number z_0 measures the well's "strength": how many oscillations the interior wavefunction can complete before decay must take over.

This picture, without any further algebra, already answers the two questions that matter most:

- **A one-dimensional well always binds at least one state.** As $\xi \to 0$, $\xi \tan \xi \to 0$, so the even branch always starts at the origin and rises to meet the circle of any radius $z_0 > 0$. However shallow or narrow the well, there is always exactly one even, nodeless bound state. This is a genuinely one-dimensional fact; shallow three-dimensional wells can fail to bind anything at all.

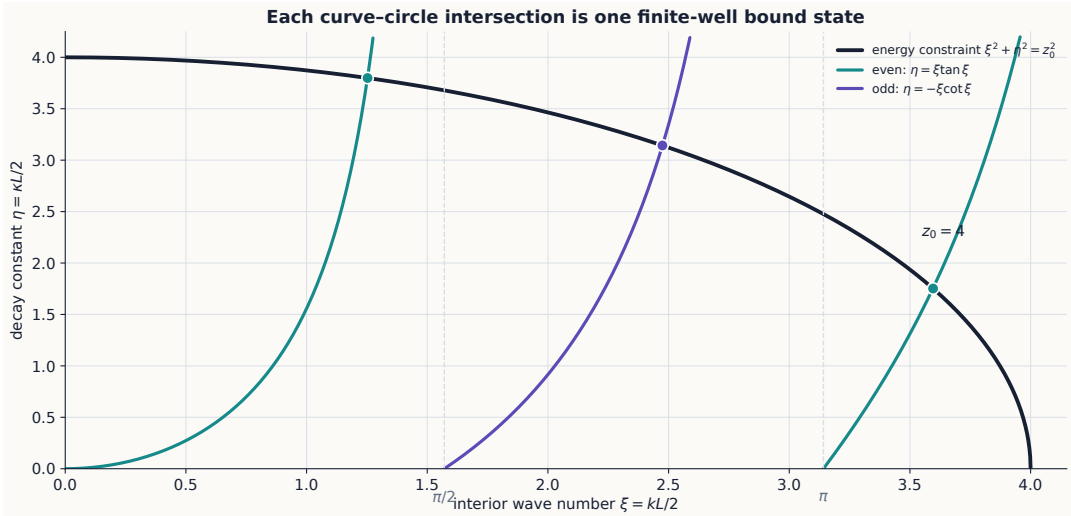

Figure 7.2: For $z_0 = 4$, the constraint circle meets two even branches and one odd branch. Enlarging the circle admits another state only after it crosses the next branch threshold.

- **New states appear at regular intervals of z_0.** The tangent branch diverges at $\xi = \pi/2, 3\pi/2, \ldots$ and the cotangent branch changes sign there, so a new branch of the even curve enters the first quadrant every time z_0 passes a multiple of π, and a new branch of the odd curve enters every time z_0 passes an odd multiple of $\pi/2$. The total bound-state count, for $z_0 > 0$, is therefore

$$N = \left\lceil \frac{2z_0}{\pi} \right\rceil . \tag{7.11}$$

At an exact threshold the prospective new solution has $E = 0$ and is not normalizable, so it does not yet count as a bound state; this is why the ceiling formula, rather than "floor plus one," handles threshold values correctly. Away from exact thresholds the two forms agree. The states alternate even, odd, even, odd, … as z_0 grows. Only once $z_0 \gg 1$ do the lowest levels approach the n^2-spaced infinite-well energies of equation (7.5); the topmost bound state, barely below $E = 0$, always looks qualitatively different—loosely bound with a long decaying tail.

Example 7.2: counting bound states

Take a well with $z_0 = 4$. Since $\pi/2 \approx 1.57$, $\pi \approx 3.14$, and $3\pi/2 \approx 4.71$, the value $z_0 = 4$ has passed the first two thresholds but not the third: one even branch exists from $z_0 > 0$, a second even branch appears once $z_0 > \pi$, and one odd branch appears once $z_0 > \pi/2$, while the second odd branch (threshold $3\pi/2$) has not yet appeared. Equation (7.11) confirms this directly: $N = \lceil 8/\pi \rceil = \lceil 2.55 \rceil = 3$. This well supports exactly three bound states: two even and one odd.

Concept check 7.2

Does a shallow, narrow finite well always fail to bind a particle if V_0 is small enough?

Answer

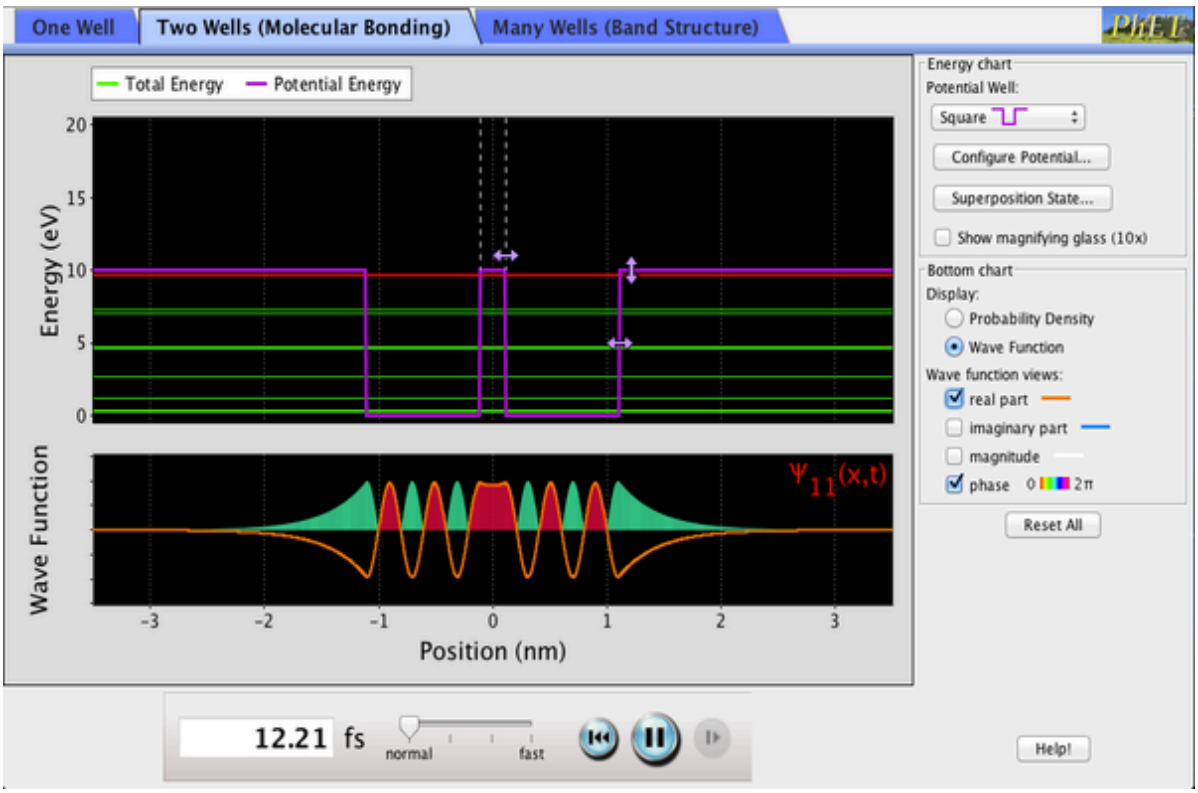

Figure 7.3: Set up a finite well and drag its width and depth to change z_0, watching bound states appear at the thresholds equation (7.11) predicts. The same sim's other potential shapes carry forward to the harmonic oscillator later in this chapter and to the variational estimates of Chapter 10. Interactive simulation: Bound States

No. Equation (7.11) gives $N \geq 1$ for every $z_0 > 0$: a one-dimensional attractive well of any strength binds at least one even, nodeless state. Only in higher dimensions can an arbitrarily weak well fail to bind anything.

7.4 Probability current

Scattering problems ask how probability flows, not merely how large a wave-amplitude coefficient is. Equation (7.1) implies the continuity equation

$$\frac{\partial |\psi|^2}{\partial t} + \frac{\partial j}{\partial x} = 0, \tag{7.12}$$

where

$$j = \frac{\hbar}{2mi}\left(\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x}\right). \tag{7.13}$$

For Ae^{ikx}, $j = (\hbar k/m)|A|^2$. Reflection and transmission probabilities are therefore ratios of outgoing to incoming current. This velocity weighting is essential when two regions have different wave numbers, as the step problem shows next.

7.5 Scattering from a step

Bound states are only half the story: a particle with enough energy to escape to infinity produces a **scattering state** instead, and the relevant questions become "how much reflects?" and "how much transmits?" rather than "what are the allowed energies?" Consider a particle incident from the left on

$$V(x) = \begin{cases} 0, & x < 0, \\ V_0, & x > 0. \end{cases} \tag{7.14}$$

For $E > V_0$, write

$$u_I = e^{ik_1 x} + re^{-ik_1 x}, \qquad u_{II} = te^{ik_2 x}, \tag{7.15}$$

where $k_1 = \sqrt{2mE}/\hbar$ and $k_2 = \sqrt{2m(E - V_0)}/\hbar$. Matching u and u' at $x = 0$ gives

$$r = \frac{k_1 - k_2}{k_1 + k_2}, \qquad t = \frac{2k_1}{k_1 + k_2}. \tag{7.16}$$

Even though $E > V_0$, reflection occurs because the wavelength changes abruptly. The amplitude $|t|^2$ is not by itself the transmission probability when the speeds differ. Using the current from Section 7.4 gives

$$R = |r|^2, \qquad T = \frac{k_2}{k_1}|t|^2, \qquad R + T = 1. \tag{7.17}$$

Probability conservation plays the role here that particle-number bookkeeping played in Stern–Gerlach chains.

Concept check 7.3

Can a particle reflect from an upward step when $E > V_0$?

Answer

Yes. Matching waves of different wavelengths produces a reflected amplitude. Classical certainty of transmission is not the quantum prediction.

Total reflection below the step

Now suppose $E < V_0$. Region II no longer oscillates. Boundedness as $x \to \infty$, together with the absence of an exponentially growing incoming term from the right, leaves the purely decaying solution $u_{II} = Ce^{-\kappa x}$, with $\kappa = \sqrt{2m(V_0 - E)}/\hbar$. Matching u and u' at $x = 0$ as before gives

$$r = \frac{ik_1 + \kappa}{ik_1 - \kappa}. \tag{7.18}$$

Both numerator and denominator have the same magnitude $\sqrt{k_1^2 + \kappa^2}$, so $|r| = 1$ exactly: every particle reflects, exactly as a classical particle would. What is *not* classical is that u_{II} is nonzero. The wavefunction penetrates into the classically forbidden region with an exponentially decaying tail, even though the probability current carried by a purely real exponential vanishes identically,

$$j_{II} = \frac{\hbar}{2mi}\left(u_{II}^* u_{II}' - u_{II} u_{II}^{*\prime}\right) = 0, \tag{7.19}$$

since u_{II} and u_{II}' differ only by the real factor $-\kappa$. The forbidden region has nonzero probability density but carries no net probability flow toward $x = +\infty$, consistent with $R = 1$. This statement does not assign a trajectory in which an individual particle travels in and then returns.

Writing r in polar form exposes a reflection *phase shift* with no classical analog. If $\kappa \to \infty$ (an infinitely high step, i.e. an infinite wall), $r \to -1$: the familiar node-forcing hard-wall reflection. For finite κ, the phase interpolates continuously between this hard-wall value and $r \to +1$ as $E \to V_0^-$—which matches the $E > V_0$ formula for r evaluated at $k_2 \to 0$, so the two regimes join smoothly at threshold even though their functional forms look unrelated.

Example 7.3: how a soft wall differs from a hard one

For $\kappa = k_1$ (a step exactly twice the particle's kinetic energy), equation (7.18) gives $r = (ik_1 + k_1)/(ik_1 - k_1) = (i + 1)/(i - 1) = -i$. The reflection probability is still $|r|^2 = 1$, but $r = -i = e^{-i\pi/2}$ differs by a quarter cycle from the $\kappa \to \infty$ hard-wall answer $r = -1 = e^{i\pi}$. A detector counting reflected atoms cannot see this phase directly; an interference experiment that recombines the reflected wave with another coherent path can.

Concept check 7.4

For $E < V_0$ at a step, is the probability density zero in the classically forbidden region?

Answer

No. $|u_{II}|^2 = |C|^2 e^{-2\kappa x}$ is nonzero and decays smoothly. What vanishes is the probability *current* there, so no probability permanently escapes to $x \to +\infty$; $R = 1$ is still exact.

7.6 Tunneling through a barrier

Section 7.5 showed total reflection when $E < V_0$ for a step that stays high forever. What changes if the classically forbidden region has finite width instead? Let a rectangular barrier have height $V_0 > E$ and width a. Inside the barrier, solutions are exponential rather than oscillatory, yet matching at both boundaries leaves a nonzero transmitted amplitude. The exact transmission probability is

$$T = \left[1 + \frac{V_0^2 \sinh^2(\kappa a)}{4E(V_0 - E)}\right]^{-1}, \qquad \kappa = \frac{\sqrt{2m(V_0 - E)}}{\hbar}. \tag{7.20}$$

For a thick or high barrier, $\kappa a \gg 1$,

$$T \approx \frac{16E(V_0 - E)}{V_0^2} e^{-2\kappa a}. \tag{7.21}$$

The exponential dependence explains why tunneling is highly sensitive to barrier width and particle mass. Tunneling does not mean the particle borrows energy and violates conservation. An energy measurement in the stationary scattering state still gives E; the spatial wavefunction simply extends through a region where a classical particle of that energy could not travel.

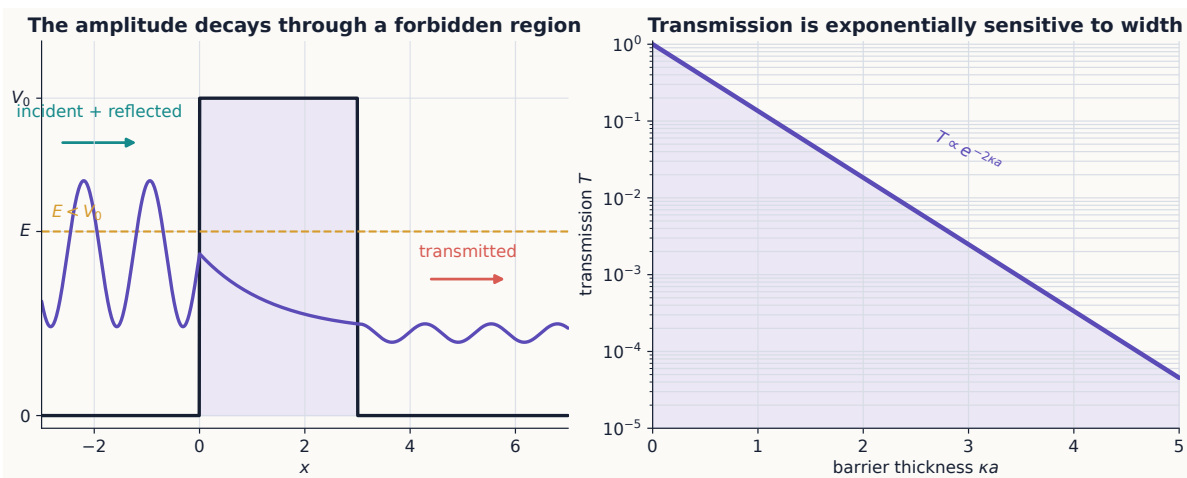

Figure 7.4: Inside a classically forbidden barrier the stationary wavefunction is exponential, not absent. Matching it at the second boundary leaves a transmitted amplitude whose probability is extraordinarily sensitive to barrier width.

Example 7.4: compare two barrier widths

In the thick-barrier regime, increasing a by Δa multiplies T by $e^{-2\kappa\Delta a}$. If $\kappa\Delta a = 1$, the transmission falls by $e^{-2} \approx 0.135$. This sensitivity underlies scanning tunneling microscopy and many decay processes.

Optional extension: above-barrier resonant transmission

Equation (7.20) was derived for $E < V_0$, but the same matching calculation applies for $E > V_0$, where the interior solution oscillates instead of decaying, with $k_2 = \sqrt{2m(E - V_0)}/\hbar$. Formally replacing $\kappa \to ik_2$—so that $\sinh(\kappa a) \to i\sin(k_2 a)$—turns equation (7.20) into

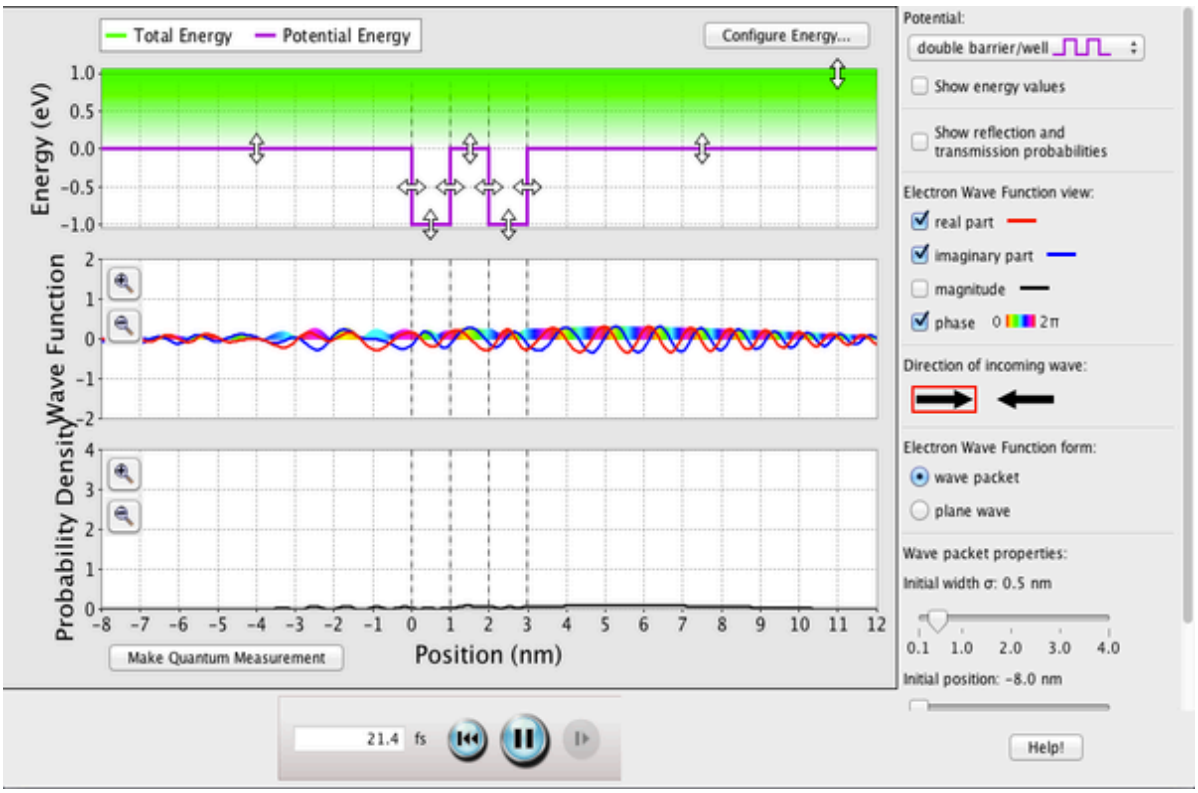

Figure 7.5: Send a wave packet at a rectangular barrier and adjust its width and height to see the exponential sensitivity of equation (7.20) directly: a reflected packet, a smaller transmitted one, and no borrowed energy anywhere. A CheerpJ simulation takes tens of seconds to load on a cold cache and is mouse-only. Interactive simulation: Quantum Tunneling

$$T = \left[1 + \frac{V_0^2 \sin^2(k_2 a)}{4E(E - V_0)}\right]^{-1}, \qquad E > V_0. \tag{7.22}$$

Unlike the tunneling regime, this transmission is not monotonic in the barrier width. Whenever $k_2 a = n\pi$ for an integer n—the barrier width equals a whole number of half wavelengths of the interior wave—$\sin(k_2 a) = 0$ and $T = 1$: the barrier becomes completely transparent, exactly as an antireflection coating of the right thickness eliminates reflected light. Away from these resonances, some reflection persists even though $E > V_0$ classically guarantees transmission. The two formulas meet consistently at $E = V_0$, where both κ and k_2 vanish and a careful limit of either expression gives the finite value $T = [1 + mV_0 a^2/(2\hbar^2)]^{-1}$ rather than an ambiguous 0/0.

One formula, two regimes

Equations (7.20) and (7.22) are the same analytic function of energy, continued across $E = V_0$. Tunneling and resonant transmission are not two different phenomena; they are the same matching calculation evaluated on either side of one threshold.

Concept check 7.5

If a barrier is thick enough that $\kappa a \gg 1$, can it ever be perfectly transparent?

Answer

Not for $E < V_0$: $\sinh^2(\kappa a)$ grows without bound, so T only shrinks with increasing width. Perfect transmission ($T = 1$) is a feature of the $E > V_0$ resonant regime in equation (7.22), where the interior wave oscillates rather than decays.

7.7 The harmonic oscillator

Every bound state so far was found by directly solving a differential equation and matching boundary conditions—a fresh calculation for each new energy level. The harmonic oscillator admits a much slicker, purely algebraic route that sidesteps this repetition entirely. For

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2} m\omega^2 \hat{x}^2, \tag{7.23}$$

define

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + \frac{i}{\sqrt{2m\hbar\omega}}\hat{p}, \qquad \hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} - \frac{i}{\sqrt{2m\hbar\omega}}\hat{p}. \tag{7.24}$$

They satisfy $[\hat{a}, \hat{a}^\dagger] = 1$, and

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger \hat{a} + \frac{1}{2}\right). \tag{7.25}$$

The spectrum does not follow merely by naming these operators, so here is the essential argument. Define $\hat{N} = \hat{a}^\dagger \hat{a}$. For every normalized state,

$$\langle\psi|\hat{N}|\psi\rangle = \|\hat{a}|\psi\rangle\|^2 \geq 0, \tag{7.26}$$

so its eigenvalues cannot be negative. The commutator above gives

$$[\hat{H}, \hat{a}^\dagger] = \hbar\omega\hat{a}^\dagger, \qquad [\hat{H}, \hat{a}] = -\hbar\omega\hat{a}. \tag{7.27}$$

Therefore $\hat{a}^\dagger$ raises any energy eigenvalue by $\hbar\omega$, while $\hat{a}$ lowers it by the same amount. Starting from a number eigenstate with eigenvalue ν, repeated lowering would produce $\nu - 1, \nu - 2, \ldots$. Because number eigenvalues cannot be negative, this sequence must terminate at a state annihilated by $\hat{a}$. For that state, $\|\hat{a}|0\rangle\|^2 = \langle 0|\hat{N}|0\rangle = 0$, so its number eigenvalue is zero. The original ν must therefore have been a nonnegative integer. Repeated raising gives every $n = 0, 1, 2, \ldots$, and the norms of the raised and lowered states fix the ladder factors:

$$\hat{a}|n\rangle = \sqrt{n}\,|n-1\rangle, \qquad \hat{a}^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle, \tag{7.28}$$

with

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right), \qquad n = 0, 1, 2, \ldots \tag{7.29}$$

The ground state is defined by $\hat{a}|0\rangle = 0$. In position space this first-order equation gives a Gaussian, as Example 7.5 works out explicitly. The ladder method obtains the complete spectrum without solving a new differential equation for every level: every excited state is reached from $|0\rangle$ by repeated algebra with $\hat{a}^\dagger$, rather than by solving a new second-order boundary-value problem for each n as the earlier well problems required.

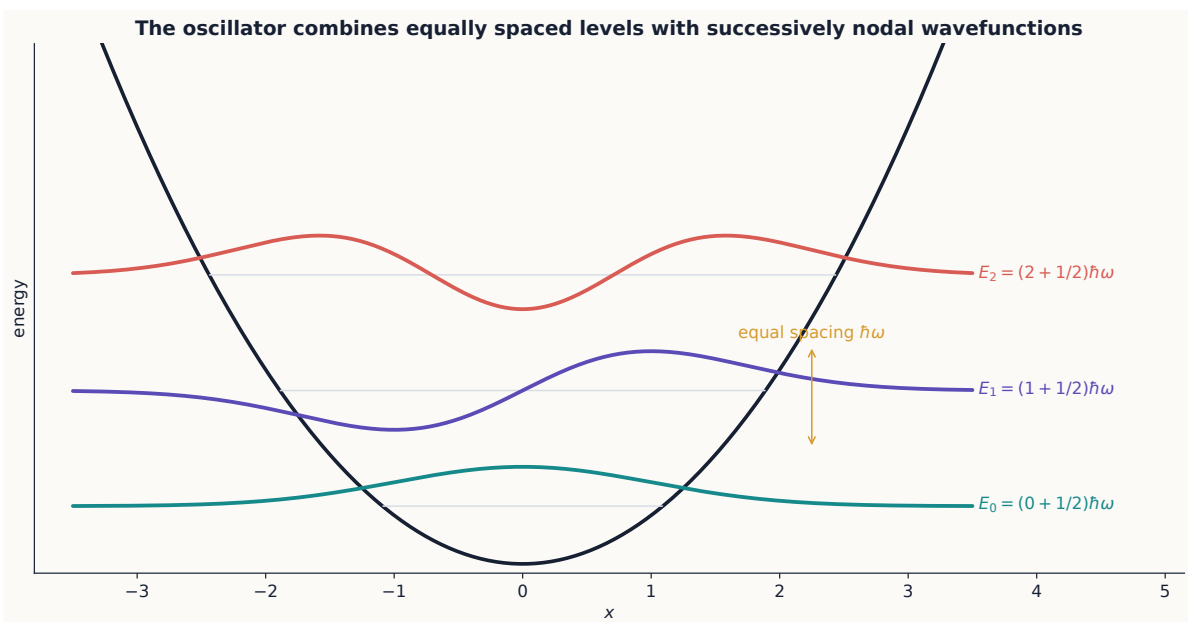

Figure 7.6: Unlike the square well, the oscillator has equal energy spacing. The ladder operator raises both the energy by $\hbar\omega$ and the node count by one, building every state from the Gaussian ground state.

Example 7.5: the ground state in position space

Writing $\hat{p} = -i\hbar\, d/dx$, the condition $\hat{a}\,\langle x|0\rangle = 0$ becomes the first-order equation

$$\sqrt{\frac{m\omega}{2\hbar}}\, x\, u_0(x) + \sqrt{\frac{\hbar}{2m\omega}}\,\frac{du_0}{dx} = 0, \qquad \text{so} \qquad \frac{du_0}{u_0} = -\frac{m\omega}{\hbar}x\, dx. \tag{7.30}$$

Integrating and normalizing gives

$$u_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/(2\hbar)}. \tag{7.31}$$

This is a Gaussian of exactly the form studied in Chapter 6 as a minimum-uncertainty wave packet, with width $\sigma_x = \sqrt{\hbar/(2m\omega)}$ and zero mean momentum. It therefore saturates $\Delta x\,\Delta p = \hbar/2$, just like the minimum-uncertainty packets of that chapter. But this particular Gaussian does *not* spread with time the way a free-particle packet does: it is an exact energy eigenstate, so $|\psi(x,t)|^2 = |u_0(x)|^2$ is stationary. Confining harmonic forces, unlike free propagation, can hold a minimum-uncertainty packet in a permanent balance between spreading and restoring force.

Concept check 7.6

Why does the oscillator not have $E = 0$?

Answer

The ground state has $E_0 = \hbar\omega/2$. A state with both zero kinetic and zero potential energy would require exact $p = 0$ and $x = 0$, contradicting $[\hat{x}, \hat{p}] = i\hbar$.

Summary

- Stationary wavefunctions solve a Hamiltonian boundary-value problem.
- Confinement produces discrete energies and a nonzero ground-state energy.
- A finite well always binds at least one state, and gains new states at regular intervals of its dimensionless strength z_0.
- Scattering probabilities are ratios of probability currents; total reflection below a step can still carry a nonclassical phase shift and a nonzero forbidden-region density.
- Wavefunctions can penetrate forbidden regions and tunnel through finite barriers without violating energy conservation, while the same matching calculation above a barrier predicts resonant, perfectly transparent energies.
- Ladder operators reveal the equally spaced harmonic-oscillator spectrum and generate every excited state algebraically from one Gaussian ground state.

Exercises

Exercises 1–12 form the core set on wells, probability flow, tunneling, and the oscillator algebra. Exercises 13–19 deepen boundary matching and below-step reflection. Exercises 20–24 are synthesis or optional-extension problems.

1. Derive equations (7.4) and (7.5) from both boundary conditions.

2. Find $\langle x \rangle$ for every infinite-well energy eigenstate by symmetry.

3. For $(u_1 + u_2)/\sqrt{2}$, calculate $|\psi(x,t)|^2$ explicitly and identify its oscillation period.

4. Show that no two distinct normalizable bound states in one dimension can have the same energy by analyzing their Wronskian.

5. Derive r and t for the potential step and verify equation (7.17).

6. Calculate the current for $Ae^{ikx} + Be^{-ikx}$.

7. Derive the continuity equation from the Schrödinger equation and its complex conjugate.

8. Estimate the change in thick-barrier transmission when the particle mass is multiplied by four.

9. Verify $[\hat{a}, \hat{a}^\dagger] = 1$ from $[\hat{x}, \hat{p}] = i\hbar$.

10. Derive equation (7.25).

11. Show that $\hat{a}^\dagger|n\rangle$ has energy $E_n + \hbar\omega$.

12. Example 7.5 derives the ground state $u_0(x)$. Apply $\hat{a}^\dagger$ to $u_0(x)$ in position space to find $u_1(x)$ explicitly, and verify by direct substitution that it solves equation (7.2) with $E_1 = \frac{3}{2}\hbar\omega$.

Boundary-matching consolidation

13. Starting from continuity of u and u' at $x = L/2$, and using the symmetry of the well to relate $u(-L/2)$ to $u(L/2)$, derive the even- parity condition in equation (7.8).

14. Derive the odd-parity condition in equation (7.9) following the same steps as Exercise 13.

15. For a well with $z_0 = 7$, find the total number of bound states and how many are even and how many are odd.

16. Explain, using the $\xi \to 0$ limit of $\xi \tan \xi$, why a one-dimensional finite well can never have zero bound states, no matter how small V_0 or L is.

17. Derive equation (7.18) by matching u and u' at $x = 0$ for $E < V_0$.

18. Show directly from equation (7.18) that $|r| = 1$ for every $\kappa > 0$, and confirm that $r \to -1$ as $\kappa \to \infty$.

19. Show that the probability current j_{II} vanishes identically for $u_{II} = Ce^{-\kappa x}$, for any constant C, and explain why this is consistent with $R = 1$.

Extensions and synthesis

20. Derive equation (7.22) from equation (7.20) by the substitution $\kappa \to ik_2$, and find the smallest barrier width a giving $T = 1$ for a given k_2.

21. Take the limit $E \to V_0$ of equation (7.20) carefully (expanding $\sinh(\kappa a) \approx \kappa a$) to show that $T(E{=}V_0) = [1 + mV_0a^2/(2\hbar^2)]^{-1}$, and check that the same value follows from equation (7.22) in the limit $k_2 \to 0$.

22. Using the Gaussian ground state of equation (7.31), calculate Δx and Δp explicitly and confirm $\Delta x\, \Delta p = \hbar/2$.

23. Explain, in terms of Ehrenfest's theorem from Chapter 6, why the mean position of an oscillator energy eigenstate is time independent even though a classical particle of the same energy oscillates back and forth.

24. Compare the ground-state energy $\hbar\omega/2$ of the harmonic oscillator with the ground-state energy of an infinite well of width $L = 2\sqrt{\hbar/(m\omega)}$ (a rough width match). Which is larger, and why might that be expected from the shape of the two potentials near their minima?

Selected exercise guidance

Use these checks only after making a complete attempt.

Exercise 5

Use $j = (\hbar k/m)|A|^2$ separately for the incident, reflected, and transmitted traveling waves. The reflected current points left, so R uses its magnitude. Substituting the matched amplitudes should give $R + T = 1$.

Exercise 8

Because $\kappa \propto \sqrt{m}$, multiplying the mass by four doubles κ. In the thick-barrier approximation the exponent therefore changes from $-2\kappa a$ to $-4\kappa a$, apart from the algebraic prefactor.

Exercise 15

$N = \lceil 14/\pi \rceil = 5$. The parity alternates from the even ground state, so there are three even states and two odd states.

Chapter 8

Three-Dimensional Wave Mechanics

8.1 Three-dimensional wave mechanics

Learning objectives

After this chapter, you should be able to:

- formulate normalization, probability, and Schrödinger equations in three dimensions;
- separate Cartesian problems into one-dimensional factors;
- interpret degeneracy in symmetric potentials;
- separate central-force problems into radial and angular equations;
- use radial probability densities correctly;
- explain why Coulomb degeneracy exceeds what rotational symmetry alone requires; and
- describe the quantum numbers and wavefunctions of hydrogen.

8.1 States in three-dimensional space

Chapters 1 through 7 built quantum mechanics from a small set of postulates — states as vectors, probabilities from the Born rule, unitary time evolution — first for a two-dimensional spin space and then for a particle confined to a line. None of those postulates mentioned how many dimensions space has. This chapter applies the same postulates to a particle free to move in three dimensions, which is where the theory finally meets real atoms, real scattering experiments, and real detectors. Nothing conceptually new is being added; the position label x is simply replaced by the vector label $\mathbf{r}$, and a single integral becomes a triple integral.

The position basis is labeled by $\mathbf{r} = (x, y, z)$, with

$$\psi(\mathbf{r}) = \langle \mathbf{r}|\psi\rangle, \qquad \int |\psi(\mathbf{r})|^2 \, d^3r = 1. \tag{8.1}$$

Quick review: the gradient and the Laplacian

Skip this box if multivariable calculus is already familiar.

In one dimension, the kinetic-energy operator involves a single second derivative, d^2/dx^2. In three dimensions a wavefunction depends on all three coordinates at once, $\psi(x, y, z)$, so "how fast ψ curves" must now account for curvature along every direction, not just one. The **gradient**, written $\boldsymbol{\nabla}\psi$, collects the three ordinary partial derivatives into a single vector,

$$\boldsymbol{\nabla}\psi = \left(\frac{\partial\psi}{\partial x}, \frac{\partial\psi}{\partial y}, \frac{\partial\psi}{\partial z}\right), \tag{8.2}$$

the direct three-dimensional generalization of an ordinary slope. The **Laplacian**, written $\nabla^2\psi$, then adds up the *second* derivative along each axis separately,

$$\nabla^2\psi = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}. \tag{8.3}$$

It measures total curvature: how much the value of ψ at a point differs from the average of its immediate neighbors, summed over all three directions at once. Restricting ψ to depend on x alone recovers the one-dimensional kinetic-energy operator $-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}$ from Chapter 7 exactly, so nothing here is conceptually new: ∇^2 is simply what d^2/dx^2 becomes once a wavefunction is allowed to vary in more than one direction.

For a time-independent potential,

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{8.4}$$

Expectation values use the volume element:

$$\langle A\rangle = \int \psi^*(\mathbf{r})\hat{A}\psi(\mathbf{r})\, d^3r. \tag{8.5}$$

The vector probability current is

$$\mathbf{j} = \frac{\hbar}{2mi}(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*), \tag{8.6}$$

and obeys $\partial|\psi|^2/\partial t + \boldsymbol{\nabla}\cdot\mathbf{j} = 0$. This is the same probability current introduced in one dimension in Chapter 7, now carrying a direction as well as a magnitude: $\mathbf{j}$ measures how much probability flows through a small oriented patch of space per unit time, exactly as an ordinary mass or charge flux would, and the continuity equation says that probability can shift from place to place but can never be created or destroyed.

8.2 Separation in a rectangular box

If

$$V(x, y, z) = V_x(x) + V_y(y) + V_z(z), \tag{8.7}$$

a product $\psi = X(x)Y(y)Z(z)$ separates equation (8.4) into three one-dimensional eigenproblems, with

$$E = E_x + E_y + E_z. \tag{8.8}$$

Quick calculation: carrying out the separation

It is worth watching this separation happen explicitly rather than taking it on faith, since the identical technique reappears in Section 8.3 and throughout later chapters. Substituting $\psi = X(x)Y(y)Z(z)$ into equation (8.4) and using $V(x, y, z) = V_x(x) + V_y(y) + V_z(z)$ gives

$$-\frac{\hbar^2}{2m}\Big(YZ\,X'' + XZ\,Y'' + XY\,Z''\Big) + \big(V_x + V_y + V_z\big)XYZ = E\,XYZ. \tag{8.9}$$

Dividing every term by $\psi = XYZ$ separates the variables completely:

$$\underbrace{\left[-\frac{\hbar^2}{2m}\frac{X''}{X} + V_x(x)\right]}_{\text{function of } x \text{ alone}} + \underbrace{\left[-\frac{\hbar^2}{2m}\frac{Y''}{Y} + V_y(y)\right]}_{\text{function of } y \text{ alone}} + \underbrace{\left[-\frac{\hbar^2}{2m}\frac{Z''}{Z} + V_z(z)\right]}_{\text{function of } z \text{ alone}} = E. \tag{8.10}$$

The first bracket depends only on x, the second only on y, the third only on z—yet their sum is the single constant E, the same number regardless of x, y, or z. The only way three functions of independent variables can add up to one fixed constant for *every* choice of x, y, and z is if each bracket is separately equal to its own constant:

$$-\frac{\hbar^2}{2m}\frac{d^2X}{dx^2} + V_x(x)X = E_xX, \tag{8.11}$$

and likewise for Y with constant E_y and Z with constant E_z, where $E = E_x + E_y + E_z$. One three-dimensional problem has become three ordinary one-dimensional eigenvalue problems, each of exactly the form already solved in Chapter 7.

For an infinite rectangular box with sides L_x, L_y, L_z,

$$\psi_{n_xn_yn_z} = \sqrt{\frac{8}{L_xL_yL_z}}\sin\frac{n_x\pi x}{L_x}\sin\frac{n_y\pi y}{L_y}\sin\frac{n_z\pi z}{L_z}, \tag{8.12}$$

$$E_{n_xn_yn_z} = \frac{\pi^2\hbar^2}{2m}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2}\right). \tag{8.13}$$

In a cube, permutations such as $(1,1,2)$, $(1,2,1)$, and $(2,1,1)$ have the same energy. This **degeneracy** reflects the box's symmetry. Changing one side length breaks the symmetry and usually splits those energies.

Concept check 8.1

In a cubic box, are $(1,1,2)$ and $(1,2,1)$ the same state?

Answer

No. They are orthogonal spatial functions with the same energy. Degeneracy means equal eigenvalues, not identical eigenstates.

The cube's degeneracy comes from a *geometric* symmetry: permuting x, y, and z leaves the Hamiltonian unchanged, so states related by that permutation must share an energy. Section 8.5 will meet a second kind of degeneracy in the hydrogen atom that cannot be explained this way. Keeping these two origins distinct now will make the hydrogen case less mysterious later.

8.3 Central potentials and spherical coordinates

Quick review: spherical coordinates

Skip this box if (r, θ, ϕ) is already a familiar coordinate system.

Ordinary Cartesian coordinates (x, y, z) locate a point by three signed distances along fixed axes. **Spherical coordinates** instead locate the same point using one distance and

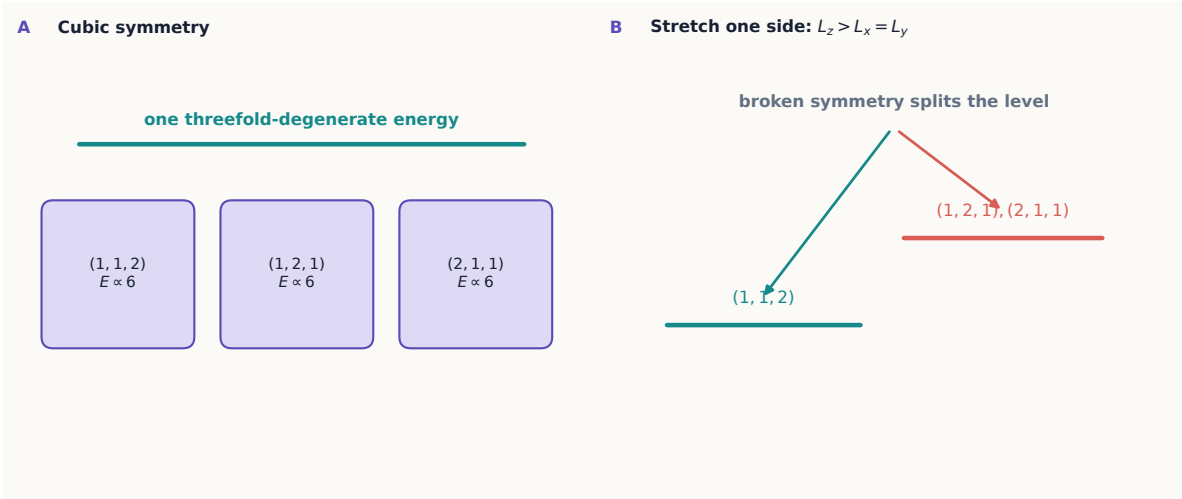

Figure 8.1: Permuting the quantum numbers of a cubic box does not change the energy. Stretching one side distinguishes that direction, partially lifting the degeneracy and exposing its geometric origin.

two angles: $r \geq 0$ is the distance from the origin, $\theta \in [0, \pi]$ is the polar angle measured down from the $+z$ axis, and $\phi \in [0, 2\pi)$ is the azimuthal angle measured around the z axis—the same way longitude is measured around the Earth, with θ playing the role of colatitude. The conversion to Cartesian coordinates is

$$x = r \sin\theta \cos\phi, \qquad y = r \sin\theta \sin\phi, \qquad z = r \cos\theta. \tag{8.14}$$

This system earns its keep whenever the physics singles out one special point, the origin, and treats every direction away from it on equal footing—exactly the situation for a central potential $V(r)$, which depends on distance from the origin alone and not on direction at all. Rewriting the Laplacian in these coordinates is what allows a problem with spherical symmetry to be split into a *radial* equation depending on r alone, plus angular equations depending only on θ and ϕ—precisely the split carried out in equation (8.15) next, by the same separation-of-variables logic just used for the rectangular box.

When $V(\mathbf{r}) = V(r)$, spherical coordinates respect the rotational symmetry. Write

$$\psi(r, \theta, \phi) = R(r)Y(\theta, \phi). \tag{8.15}$$

Separation produces angular eigenfunctions Y_ℓ^m, introduced fully in Chapter 9, and a radial equation.

The centrifugal term below is not an extra assumption; it is what remains of the Laplacian's angular part. In spherical coordinates the Laplacian splits into a radial piece and an angular piece,

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) - \frac{\hat{L}^2}{\hbar^2 r^2}, \tag{8.16}$$

where $\hat{L}^2$ is built entirely from angular derivatives — the same orbital-angular-momentum operator whose eigenstates and eigenvalues Chapter 9 derives from first principles. Its eigenfunctions are exactly the spherical harmonics,

$$\hat{L}^2 Y_\ell^m = \hbar^2 \ell(\ell+1) Y_\ell^m, \tag{8.17}$$

so substituting $\psi = R(r) Y_\ell^m(\theta, \phi)$ into (8.4) replaces $\hat{L}^2$ by its eigenvalue and leaves an ordinary differential equation in r alone, with $\ell(\ell+1)$ appearing as a fixed number rather than an operator. Writing $u(r) = rR(r)$ removes the first-derivative term from that equation and gives the compact form used throughout this chapter:

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dr^2} + \left[V(r) + \frac{\hbar^2 \ell(\ell+1)}{2mr^2}\right] u = Eu. \tag{8.18}$$

The term

$$V_{\text{cent}}(r) = \frac{\hbar^2 \ell(\ell+1)}{2mr^2} \tag{8.19}$$

is the centrifugal contribution. It repels states with $\ell > 0$ from the origin even when the physical potential is attractive.

The spherical volume element is $d^3r = r^2 \sin\theta \, dr \, d\theta \, d\phi$. If the spherical harmonic is normalized, the probability of finding the particle between r and $r + dr$ is

$$P(r)\, dr = r^2 |R(r)|^2 dr = |u(r)|^2 dr. \tag{8.20}$$

The radial probability is not $|R(r)|^2$ alone. The factor r^2 counts the growing volume of spherical shells.

Example 8.1: where is a $1s$ electron most likely?

For hydrogen,

$$\psi_{100}(r) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0}. \tag{8.21}$$

The spatial density is largest at $r = 0$, but the radial density is

$$P(r) = \frac{4r^2}{a_0^3} e^{-2r/a_0}. \tag{8.22}$$

Differentiating with the product rule,

$$\frac{dP}{dr} = \frac{4}{a_0^3}\left(2re^{-2r/a_0} - \frac{2r^2}{a_0} e^{-2r/a_0}\right) = \frac{8r}{a_0^3} e^{-2r/a_0}\left(1 - \frac{r}{a_0}\right), \tag{8.23}$$

which vanishes at $r = 0$ (a minimum, since $P(0) = 0$ there) and at $r = a_0$. The second root is the most probable radius, matching the length scale already built into the wavefunction itself. “Most likely point” and “most likely radius” answer different questions because their volume elements differ. Example 8.2 revisits this distinction for a state with nonzero ℓ, where the growing shell volume and the decaying exponential compete at a larger radius.

Concept check 8.2

Why can $|\psi(0)|^2$ be largest at the origin while the probability of finding the particle at exactly the origin is zero?

Answer

$|\psi|^2$ is a density. A single point has zero volume. Radial probabilities also include the shell factor $4\pi r^2$, which vanishes at the origin.

8.4 The hydrogen atom

For an electron and proton, the center-of-mass motion separates from their relative coordinate. The relative particle has reduced mass

$$\mu = \frac{m_e m_p}{m_e + m_p} \tag{8.24}$$

Quick review: why a reduced mass appears

Skip this box if the two-body-to-one-body reduction is already familiar from classical mechanics.

The hydrogen atom is really a two-body problem: an electron and a proton, each free to move, pulling on each other. Attacking the Schrödinger equation directly for two independent three-dimensional positions would mean six coordinates at once. Exactly as in classical mechanics, this splits cleanly into two separate and much simpler problems: the motion of the *center of mass*, which drifts freely like a single particle of total mass $m_e + m_p$ and carries no information about the atom's internal structure, and the motion of the *relative coordinate* $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_p$, which behaves exactly like one fictitious particle orbiting a fixed center— provided that fictitious particle is assigned the reduced mass $\mu = m_e m_p/(m_e + m_p)$ rather than either individual mass. Because $m_p \approx 1836\, m_e$, μ turns out extremely close to m_e itself, close enough that using m_e in its place is an excellent approximation for most purposes—though not for the highest-precision spectroscopy, and not when comparing hydrogen to deuterium, where the small difference in μ is exactly the quantity being measured.

and Coulomb potential

$$V(r) = -\frac{e^2}{4\pi\epsilon_0 r}. \tag{8.25}$$

Normalizable solutions are labeled by

$$n = 1, 2, \ldots, \qquad \ell = 0, 1, \ldots, n-1, \qquad m = -\ell, \ldots, \ell. \tag{8.26}$$

Ignoring spin and small corrections, the energy depends only on n:

$$E_n = -\frac{\mu e^4}{2(4\pi\epsilon_0)^2\hbar^2}\frac{1}{n^2} = -\frac{13.6\text{ eV}}{n^2} \tag{8.27}$$

when the reduced-mass correction is rounded to the usual value. The wavefunction factors as

$$\psi_{n\ell m}(r, \theta, \phi) = R_{n\ell}(r) Y_\ell^m(\theta, \phi). \tag{8.28}$$

The principal quantum number n sets the energy in the ideal Coulomb problem; ℓ sets the orbital angular-momentum magnitude; and m sets its component along the chosen z axis.

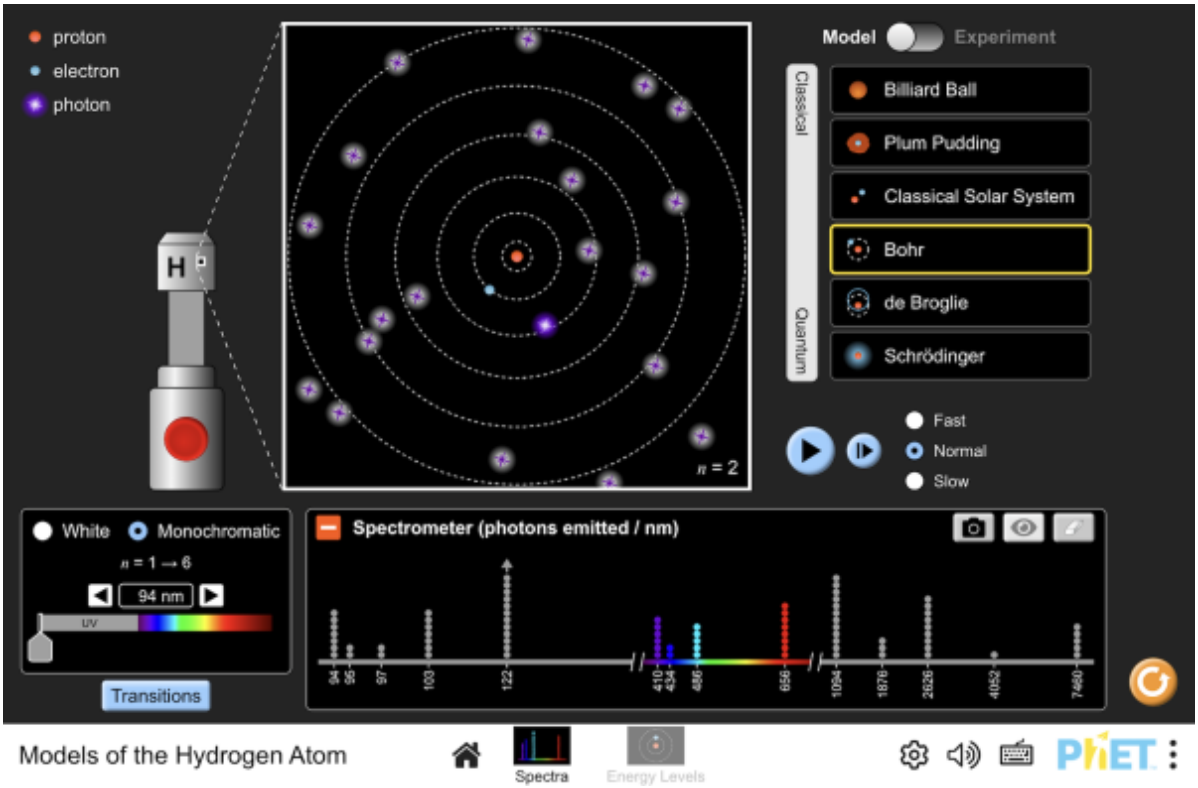

Figure 8.2: Compare the Bohr, de Broglie, and Schrödinger pictures of hydrogen side by side, and fire simulated photons at each model to see which one's predictions actually match the observed emission and absorption spectrum. Interactive simulation: Models of the Hydrogen Atom

Concept check 8.3

Does the quantum number m specify a three-dimensional orbital plane?

Answer

No. It specifies the eigenvalue $m\hbar$ of one chosen angular-momentum component. A stationary orbital is a probability amplitude, not a classical trajectory in a plane.

Nodes and scales

The angular function has $\ell - |m|$ angular nodes in polar angle, together with azimuthal phase structure. The radial function has $n - \ell - 1$ radial nodes. The total number of radial plus angular nodes is $n - 1$.

The natural length is the Bohr radius

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{\mu e^2}. \tag{8.29}$$

Hydrogen orbitals are stationary probability amplitudes, not classical paths traced by an electron. Names such as s, p, d refer to $\ell = 0, 1, 2$, not to little planetary orbits.

A checkable energy split: the Virial theorem

For any power-law potential $V(r) \propto r^k$, the Virial theorem relates the average kinetic and potential energies in a bound stationary state: $2\langle T\rangle = k\langle V\rangle$. The Coulomb potential has $k = -1$, so $2\langle T\rangle = -\langle V\rangle$. Combined with $E = \langle T\rangle + \langle V\rangle$, this fixes both pieces separately:

$$\langle T\rangle = -E_n, \qquad \langle V\rangle = 2E_n. \tag{8.30}$$

For $n = 1$, $E_1 = -13.6$ eV gives $\langle T\rangle = +13.6$ eV and $\langle V\rangle = -27.2$ eV — a compact, numerically checkable statement about how a hydrogen electron's energy divides between motion and attraction, without integrating the kinetic and potential operators separately.

Example 8.2: size and shape of a $2p$ orbital

The normalized $n = 2, \ell = 1$ radial function is

$$R_{21}(r) = \frac{1}{2\sqrt{6}}\, a_0^{-3/2}\, \frac{r}{a_0}\, e^{-r/(2a_0)}. \tag{8.31}$$

Its radial probability density is

$$P(r) = r^2 R_{21}^2(r) = \frac{1}{24a_0^5}\, r^4 e^{-r/a_0}. \tag{8.32}$$

Differentiating with the product rule, treating $1/(24a_0^5)$ as a constant prefactor,

$$\frac{dP}{dr} = \frac{1}{24a_0^5}\left(4r^3 e^{-r/a_0} - \frac{r^4}{a_0} e^{-r/a_0}\right) = \frac{e^{-r/a_0}}{24a_0^5}\left(4r^3 - \frac{r^4}{a_0}\right). \tag{8.33}$$

Setting $dP/dr = 0$ gives $4r^3 - r^4/a_0 = 0$ (the exponential factor is never zero, so only the polynomial part can vanish), so the most probable radius is $r_{\max} = 4a_0 = n^2 a_0$ — the same radius Bohr's original circular-orbit picture assigns to $n = 2$, even though the orbital has no orbit.

Quick review: the integral $\int_0^\infty x^n e^{-x/a}\, dx = n!\, a^{n+1}$

This integral—a rescaled Gamma-function integral—appears every time a hydrogen radial probability, always a polynomial in r times e^{-r/a_0}, is integrated to get a normalization constant or a mean value, so it is worth keeping on hand as a standing tool rather than re-deriving it by repeated integration by parts each time it is needed. A quick check at small n builds confidence in it: for $n = 0$, $\int_0^\infty e^{-x/a}\, dx = a = 0!\, a^1$; integrating once by parts for $n = 1$ gives $\int_0^\infty x e^{-x/a}\, dx = a^2 = 1!\, a^2$. Each additional power of x contributes one more integration by parts, and hence one more factor of a together with one more factor in the falling product that becomes $n!$—exactly the pattern the general formula packages into one closed form.

The mean radius, by contrast, is

$$\langle r\rangle = \int_0^\infty r P(r)\, dr = \frac{1}{24a_0^5}\int_0^\infty r^5 e^{-r/a_0}\, dr = \frac{5!\, a_0^6}{24a_0^5} = 5a_0, \tag{8.34}$$

somewhat larger than $r_{\max}$ because the exponential tail pulls the average outward. The same pattern held for the $1s$ state in Example 8.1 ($r_{\max} = a_0$, $\langle r \rangle = \frac{3}{2}a_0$, Exercise 6): a skewed distribution's most probable value and its mean are different numbers answering different questions, and neither should be read as "the orbital's radius" without saying which one is meant. In general,

$$\langle r \rangle_{n\ell} = \frac{a_0}{2}\left[3n^2 - \ell(\ell+1)\right], \tag{8.35}$$

which reproduces both results above and lets Exercises 15–16 extend the comparison to other orbitals without repeating the integral from scratch.

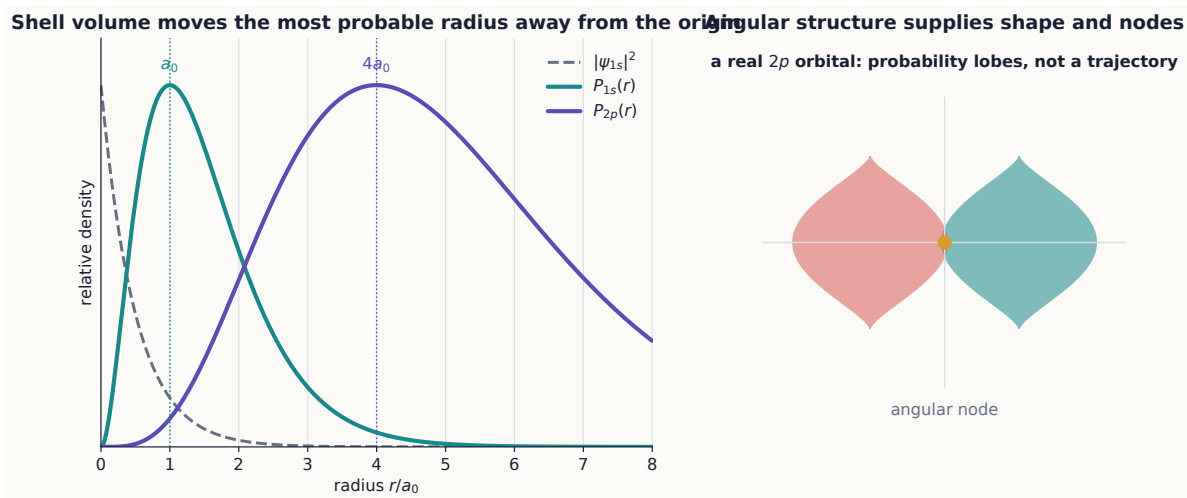

Figure 8.3: The shell-volume factor moves radial-probability peaks away from the point where the spatial density is largest. Angular factors add lobes, phases, and nodes; none of these structures represents a classical electron orbit.

Concept check 8.4

For the $2p$ orbital of Example 8.2, why is $\langle r \rangle = 5a_0$ larger than the most probable radius $r_{\max} = 4a_0$?

Answer

The radial probability density $P(r)$ is not symmetric: it rises from zero, peaks at $r_{\max}$, and then falls off with a long exponential tail. That tail contributes disproportionately to the mean, pulling $\langle r \rangle$ above the peak, exactly as for an everyday right-skewed distribution.

8.5 Degeneracy and symmetry

For fixed n, the allowed orbital states number

$$\sum_{\ell=0}^{n-1}(2\ell+1) = n^2. \tag{8.36}$$

This is worth checking directly for the first few values of n before trusting it in general. For $n = 1$, only $\ell = 0$ contributes: $1 = 1^2$. For $n = 2$, $\ell = 0$ and $\ell = 1$ contribute $1 + 3 = 4 = 2^2$. For $n = 3$, $\ell = 0, 1, 2$ contribute $1 + 3 + 5 = 9 = 3^2$. The pattern is just the familiar fact

that the first n odd numbers always sum to n^2, now dressed up in the $(2\ell+1)$-fold degeneracy that rotational symmetry assigns to each value of ℓ. Including two spin states doubles this count. The Coulomb degeneracy is larger than rotational symmetry alone requires: states with different ℓ also share the same energy. This is a genuinely different phenomenon from the box degeneracy of Section 8.2. Rotational symmetry by itself only guarantees the $(2\ell+1)$-fold degeneracy in m at fixed ℓ; it says nothing about why the $2s$ and $2p$ levels should coincide. The extra, exact $1/r$ dependence of the Coulomb potential hides a larger symmetry beyond ordinary rotations — associated with a conserved Runge–Lenz vector and often described as an SO(4) symmetry of the classical and quantum Kepler problem. Working out that symmetry is beyond this book's scope, but it is worth knowing by name: "accidental" degeneracy in hydrogen is not an accident at all, and it is fragile for exactly that reason. Relativistic corrections, spin–orbit coupling, external electric or magnetic fields, and finite nuclear effects break the exact $1/r$ form and lift parts of this degeneracy.

Concept check 8.5

Is the degeneracy between the cubic-box states $(1,1,2)$ and $(2,1,1)$ the same kind of phenomenon as the degeneracy between hydrogen's $2s$ and $2p$ states?

Answer

No. The box degeneracy comes from a geometric symmetry — permuting the three Cartesian directions of a cube. The $2s$–$2p$ degeneracy comes from the extra dynamical symmetry of the exact $1/r$ potential and disappears once the potential departs from a pure Coulomb form, for example in alkali atoms.

Degenerate states may be combined into other stationary states of the same energy. For example, real linear combinations of $m = \pm 1$ orbitals can be oriented along different spatial axes. Choosing a basis within a degenerate subspace is often guided by the additional interaction or measurement being considered.

8.6 Spectra and selection ideas

If an atom changes from energy E_i to E_f while emitting or absorbing one photon, energy conservation gives

$$\hbar\omega = |E_f - E_i|. \tag{8.37}$$

Not every energy difference necessarily appears with equal strength. The coupling operator and the angular symmetries determine transition amplitudes. For electric-dipole coupling, the most common orbital rules are $\Delta\ell = \pm 1$ and $\Delta m = 0, \pm 1$. These are not extra prohibitions placed on energy conservation; they identify matrix elements that vanish by symmetry.

Why $\Delta\ell = \pm 1$: a parity argument

Quick review: even and odd functions, and parity

Skip this box if "parity" already means something concrete to you.

A function $f(x)$ is called **even** if $f(-x) = f(x)$—flipping the sign of its argument leaves it unchanged, like $\cos x$ or x^2—and **odd** if $f(-x) = -f(x)$, like $\sin x$ or x itself. The key fact used repeatedly below concerns integrals: the integral of an *odd* function over a range symmetric about the origin, such as $\int_{-a}^{a} f(x)\,dx$ with f odd, is always exactly zero, since

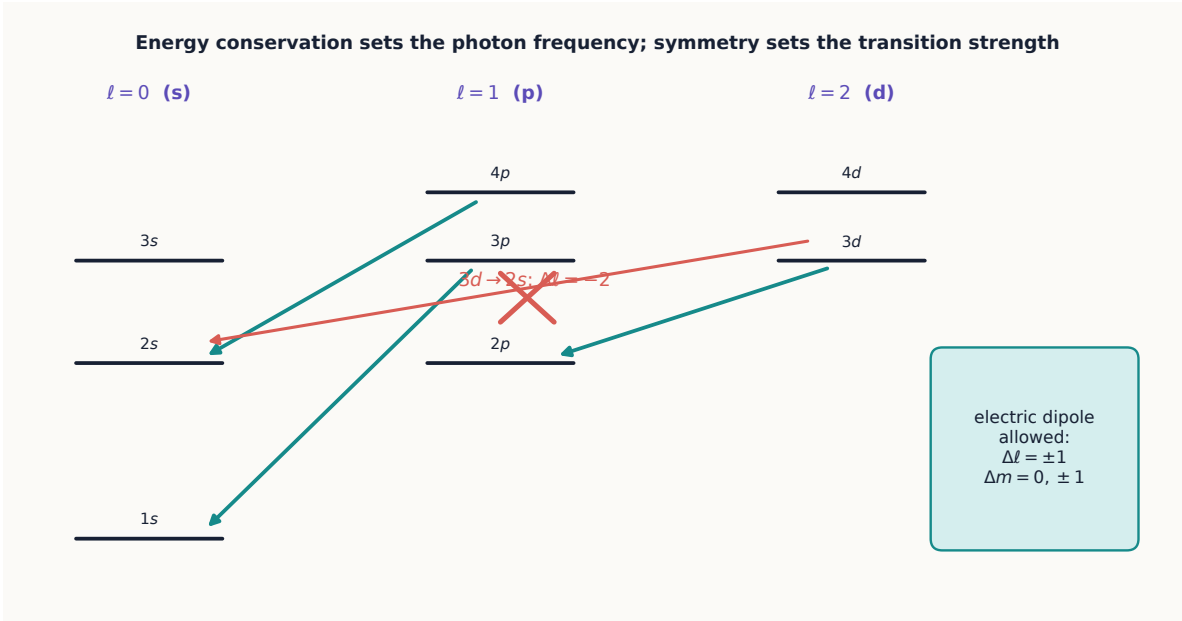

Figure 8.4: Energy differences determine possible photon frequencies, while symmetry determines which transitions have nonzero electric-dipole amplitudes. An energetically possible line may therefore be absent or extremely weak.

whatever the function contributes on the positive side is exactly canceled by its mirror image on the negative side. An even function has no such automatic cancellation.

Parity is this same even/odd idea promoted to three dimensions, where "flipping the sign of the argument" means reflecting every point through the origin, $\mathbf{r} \to -\mathbf{r}$. A three-dimensional function has **even parity** if it is unchanged by this reflection and **odd parity** if it flips sign, and the identical cancellation argument applies: an integral over all of space of a function with overall odd parity vanishes automatically, with no need to know any further details of the function. This is exactly the tool the argument below applies to the electric-dipole matrix element.

The rate for an electric-dipole transition is controlled by a matrix element of the position operator, $\langle n'\ell'm'|\mathbf{r}|n\ell m\rangle$. Under spatial inversion $\mathbf{r} \to -\mathbf{r}$, a spherical harmonic picks up a definite sign, $Y_\ell^m \to (-1)^\ell Y_\ell^m$, while $\mathbf{r}$ itself is odd. The full integrand therefore transforms with an overall factor $(-1)^{\ell'}(-1)(-1)^\ell = (-1)^{\ell+\ell'+1}$. Since the integral runs over all directions symmetrically, it can be nonzero only if this factor is $+1$, which requires $\ell+\ell'$ to be **odd**. This already forbids $\Delta\ell = 0, \pm 2, \ldots$ on symmetry grounds alone, independent of any detailed radial integral.

Parity alone would still allow $\Delta\ell = \pm 1, \pm 3, \ldots$. The tighter rule comes from angular momentum: the dipole operator behaves like an object carrying one unit of angular momentum, and Chapter 9 shows that combining angular momentum ℓ with one added unit can only produce a total in the range $|\ell - 1|$ to $\ell + 1$. Intersecting that with the odd-$\ell + \ell'$ requirement leaves exactly $\Delta\ell = \pm 1$. The $\Delta m = 0, \pm 1$ rule has an analogous origin: the three dipole components behave as objects with definite angular-momentum projection $q = 0, \pm 1$ along z, and the ϕ integral vanishes unless $m' = m + q$.

A selection rule is not a conservation law
$\Delta\ell = \pm 1$ says a particular coupling's matrix element vanishes by symmetry, not that angular momentum or parity is violated for other couplings. Magnetic-dipole and higher-multipole transitions have different, weaker selection rules and become important precisely when the electric-dipole matrix element is forbidden.

Concept check 8.6

A student argues that a $3d \to 2s$ transition should be electric-dipole allowed because energy is released and $\Delta n \neq 0$. Is the student correct?

Answer

No. Energy release only guarantees that a photon of the right frequency could in principle be emitted; it says nothing about the transition rate. Here $\ell = 2 \to \ell = 0$ gives $\Delta\ell = -2$, which is even and therefore forbidden by the parity argument in Section 8.6 for electric-dipole coupling. The transition can still occur through a weaker higher-multipole process.

Summary

- Three-dimensional wave mechanics uses the Laplacian and the physical volume element.
- Separable potentials reduce to lower-dimensional eigenproblems.
- Degeneracy reflects symmetry and allows many eigenstates to share one energy.
- Central potentials separate into radial functions and spherical harmonics.
- Radial probabilities include a factor r^2; the most probable radius and the mean radius generally differ.
- The Coulomb potential's centrifugal term traces back to the eigenvalue of $\hat{L}^2$ inside the Laplacian.
- The Virial theorem fixes $\langle T \rangle = -E_n$ and $\langle V \rangle = 2E_n$ for hydrogen.
- Hydrogen states are labeled by n, ℓ, m, with ideal Coulomb energies depending only on n; this extra degeneracy reflects a dynamical symmetry beyond ordinary rotations, distinct from the geometric degeneracy of a symmetric box.
- Electric-dipole selection rules ($\Delta\ell = \pm 1$, $\Delta m = 0, \pm 1$) follow from parity and angular-momentum addition, not from energy conservation.

Exercises

1. Normalize the rectangular-box state in equation (8.12).
2. List the degeneracies of the first four distinct energy values in a cubic box.
3. Show how a small change $L_z \neq L_x = L_y$ splits the $(1, 1, 2)$ family.
4. Derive equation (8.18) from the spherical-coordinate Laplacian, using $u = rR$.
5. Normalize the hydrogen $1s$ radial probability and verify that its maximum occurs at a_0.
6. Calculate $\langle r \rangle$ for the $1s$ state.
7. List all (ℓ, m) pairs for $n = 3$ and verify that there are nine orbital states.
8. Count the radial and angular nodes for $3s$, $3p$, and $3d$ states.

9. Find the wavelength of a photon emitted in the ideal hydrogen transition $n = 3$ to $n = 2$.

10. Explain why an arbitrary superposition inside one degenerate energy subspace is stationary up to an overall phase.

11. Give the electric-dipole-allowed possibilities for $(\ell, m) = (1, 0)$.

12. Compare position density, radial density, and probability in a thin spherical shell; state the units of each.

13. Substitute $\psi = R(r)Y_\ell^m(\theta, \phi)$ into equation (8.4), use equation (8.16) and $\hat{L}^2 Y_\ell^m = \hbar^2 \ell(\ell+1) Y_\ell^m$, and reduce the result to equation (8.18) with $u = rR$.

14. Verify directly that $R_{21}(r)$ in Example 8.2 satisfies $\int_0^\infty R_{21}^2(r)\, r^2\, dr = 1$.

15. The normalized $2s$ radial function is $R_{20}(r) = \frac{1}{2\sqrt{2}} a_0^{-3/2} \left(2 - \dfrac{r}{a_0}\right) e^{-r/(2a_0)}$. Find its radial node, and use equation (8.35) to find $\langle r \rangle$ for $2s$. Explain why it differs from the $2p$ result of Example 8.2 even though both belong to $n = 2$.

16. Use equation (8.30) to predict $\langle T \rangle$ and $\langle V \rangle$ for $n = 2$ and $n = 3$.

17. Using $\langle 1/r \rangle_{1s} = 1/a_0$, compute $\langle V \rangle = -\dfrac{e^2}{4\pi\epsilon_0} \langle 1/r \rangle$ for the $1s$ state directly and confirm it equals $2E_1$, as required by equation (8.30).

18. Using the parity argument of Section 8.6, decide which of these transitions are electric-dipole allowed: $3d \to 2p$, $3d \to 2s$, $3p \to 1s$, $3s \to 2s$.

19. Explain, using the angular-momentum addition idea previewed in Section 8.6, why $\Delta\ell = \pm 3$ is forbidden even though $\ell + \ell'$ is odd for such a transition.

20. A rectangular box has $L_x = L$, $L_y = 2L$, $L_z = 3L$. List its five lowest distinct energies in units of $\pi^2\hbar^2/(2mL^2)$ and state whether any of them are degenerate.

21. A weak external electric field is turned on in a hydrogen atom (the Stark effect). Using the distinction drawn in Section 8.5, explain qualitatively why this field can split the four degenerate $n = 2$ levels even though it is not a rotation. (Chapter 10 develops the quantitative tool for this calculation.)

22. Using R_{20} from Exercise 15, find the most probable radius for the $2s$ state. Does it equal $n^2 a_0 = 4a_0$ as it did for $2p$ in Example 8.2? What feature of the $2p$ state made $r_{\max} = n^2 a_0$ work there?

23. Considering only decays to $n = 2$, list every electric-dipole-allowed final state reachable from $3d,\ m = 0$ and from $3p,\ m = 0$, applying both the $\Delta\ell = \pm 1$ and $\Delta m = 0, \pm 1$ rules.

24. Explain in your own words why the phrase "hydrogen's degeneracy is accidental" is misleading, using the vocabulary introduced in Section 8.5.

Part III

Structure and applications

Chapter 9

Angular Momentum

9.1 Angular momentum

Learning objectives

After this chapter, you should be able to:

- use the angular-momentum commutation relations and ladder operators;
- identify allowed j, m values and measurement probabilities;
- distinguish orbital and intrinsic angular momentum;
- connect orbital eigenstates with spherical harmonics;
- add two angular momenta using coupled and uncoupled bases;
- construct the singlet and triplet states of two spin-$\frac{1}{2}$ systems;
- recover the Pauli matrices of Chapter 3 as the spin-$\frac{1}{2}$ case of the general ladder construction; and
- build explicit Clebsch–Gordan coefficients for coupling orbital and spin angular momentum.

9.1 The angular-momentum algebra

Chapters 1 through 4 introduced spin operationally, through Stern–Gerlach outcomes, and represented a spin-$\frac{1}{2}$ system's observables by $\hat{\mathbf{S}} = \frac{\hbar}{2}\boldsymbol{\sigma}$. That construction was built directly from the two-outcome apparatus; nothing in it referred to orbital motion. This chapter asks a more general question: what mathematical structure makes something "an angular momentum" in quantum mechanics at all? The answer is not a formula like $\mathbf{r}\times\mathbf{p}$—that expression reappears in Section 9.3, but only as one particular example—but a set of commutation relations. Anything obeying them, whatever its physical origin, inherits the entire apparatus developed below: discrete eigenvalues, ladder operators, and the role of generating rotations. Spin qualifies; so does orbital motion; so, as Section 9.4 shows, does any sum of angular momenta.

An angular momentum $\hat{\mathbf{J}}$ is, by definition, any vector operator whose Cartesian components satisfy

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z \tag{9.1}$$

and cyclic permutations. This is exactly the same kind of relation already proved directly from the Pauli matrices in Chapter 3, equation (3.29): $[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z$. What is new here is the direction of the logic. Chapter 3 started from a concrete 2×2 matrix representation and *derived*

the commutator. This chapter runs the argument in reverse: it starts from the commutator alone, treated as the defining property of "an angular momentum," and derives everything else—the allowed eigenvalues, the ladder construction, even the existence of representations other than spin-$\frac{1}{2}$—from that one algebraic fact. Nothing below assumes any particular matrix size in advance.

Before deriving anything from this relation, it is worth checking that it is physically plausible. Take an ordinary object—a book, a phone—and rotate it 90° about x and then 90° about y. Now reset it, and perform the same two rotations in the opposite order. The final orientations differ. Finite rotations about different axes simply do not commute, and equation (9.1) is the quantum statement of the corresponding infinitesimal fact: the generators of rotations about different axes cannot commute either. Section 9.3 makes this connection exact.

Its squared magnitude,

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2, \tag{9.2}$$

commutes with every component. This is not an additional assumption; it follows from equation (9.1) alone. For example,

$$[\hat{J}^2, \hat{J}_z] = [\hat{J}_x^2, \hat{J}_z] + [\hat{J}_y^2, \hat{J}_z] = \hat{J}_x[\hat{J}_x, \hat{J}_z] + [\hat{J}_x, \hat{J}_z]\hat{J}_x + \hat{J}_y[\hat{J}_y, \hat{J}_z] + [\hat{J}_y, \hat{J}_z]\hat{J}_y, \tag{9.3}$$

and substituting $[\hat{J}_x, \hat{J}_z] = -i\hbar\hat{J}_y$ and $[\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x$ makes every term cancel in pairs. The same pattern holds for $\hat{J}_x$ and $\hat{J}_y$ by the cyclic symmetry of the algebra. Because $\hat{J}^2$ commutes with each component, we may choose simultaneous eigenstates of $\hat{J}^2$ and any one component—conventionally $\hat{J}_z$:

$$\hat{J}^2|j,m\rangle = \hbar^2 j(j+1)|j,m\rangle, \qquad \hat{J}_z|j,m\rangle = \hbar m|j,m\rangle. \tag{9.4}$$

The allowed values are

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots, \qquad m = -j, -j+1, \ldots, j. \tag{9.5}$$

Section 9.2 derives this restriction rather than simply asserting it. For fixed j, the space has dimension $2j+1$. Measuring J_z in an arbitrary state in that space returns one of the $2j+1$ values $\hbar m$.

Concept check 9.1

If $J^2 = 2\hbar^2$, what values can a measurement of J_z return?

Answer

$j(j+1) = 2$ gives $j = 1$, so $m = -1, 0, 1$ and the possible values are $-\hbar, 0, \hbar$.

Concept check 9.2

Does $m = 0$ imply zero angular momentum?

Answer

No. It means only that the selected z component is zero. For $j > 0$, the magnitude remains $\hbar\sqrt{j(j+1)}$.

9.2 Ladder operators

Define

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y. \tag{9.6}$$

These are not Hermitian—$\hat{J}_+^\dagger = \hat{J}_-$—so they do not represent directly measurable quantities. Their role is purely constructive: acting with $\hat{J}_+$ or $\hat{J}_-$ on an eigenstate produces another eigenstate with m shifted by exactly one unit, which is what lets us build an entire j-multiplet from a single state. The commutators

$$[\hat{J}_z, \hat{J}_\pm] = \pm\hbar\hat{J}_\pm \tag{9.7}$$

show that $\hat{J}_\pm|j,m\rangle$ has m shifted by one: applying $\hat{J}_z$ to $\hat{J}_\pm|j,m\rangle$ and using the commutator gives eigenvalue $\hbar(m \pm 1)$.

To fix the normalization—and, along the way, to see *why* the ladder must terminate—compute the squared norm of $\hat{J}_+|j,m\rangle$ directly. Start by multiplying out the product $\hat{J}_-\hat{J}_+$ term by term, exactly as you would expand $(a - ib)(a + ib)$ for ordinary numbers:

$$\hat{J}_-\hat{J}_+ = (\hat{J}_x - i\hat{J}_y)(\hat{J}_x + i\hat{J}_y) = \hat{J}_x^2 + i\hat{J}_x\hat{J}_y - i\hat{J}_y\hat{J}_x - i^2\hat{J}_y^2 = \hat{J}_x^2 + \hat{J}_y^2 + i[\hat{J}_x, \hat{J}_y]. \tag{9.8}$$

The only difference from ordinary numbers is that $\hat{J}_x\hat{J}_y$ and $\hat{J}_y\hat{J}_x$ do not cancel, since $\hat{J}_x$ and $\hat{J}_y$ do not commute—their difference is exactly the commutator $[\hat{J}_x, \hat{J}_y]$ that survives in the last term. Substituting $[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$ from equation (9.1) and $\hat{J}_x^2 + \hat{J}_y^2 = \hat{J}^2 - \hat{J}_z^2$ then gives

$$\hat{J}_-\hat{J}_+ = \hat{J}^2 - \hat{J}_z^2 + i(i\hbar\hat{J}_z) = \hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z. \tag{9.9}$$

Therefore

$$\left\|\hat{J}_+|j,m\rangle\right\|^2 = \langle j,m|\hat{J}_-\hat{J}_+|j,m\rangle = \hbar^2\big[j(j+1) - m(m+1)\big]. \tag{9.10}$$

A squared norm can never be negative, so $j(j+1)-m(m+1) \geq 0$ for every physically realized m. Equality—a state that $\hat{J}_+$ annihilates—occurs exactly when $m = j$, not before and not after. This is the actual reason the ladder "must terminate": it is forced by the non-negativity of a norm, the same kind of argument used throughout this book to rule out states that a formula might otherwise seem to allow. The analogous calculation with $\hat{J}_+\hat{J}_- = \hat{J}^2 - \hat{J}_z^2 + \hbar\hat{J}_z$ shows that $\hat{J}_-$ annihilates a state exactly at $m = -j$. Together these fix both ends of the ladder, and since each application of $\hat{J}_-$ changes m by exactly one, they force the number of steps from $m = j$ down to $m = -j$—namely $2j$—to be a non-negative integer. That is precisely the statement that j itself must be an integer or a half-integer. Normalizing the resulting states gives

$$\hat{J}_\pm|j,m\rangle = \hbar\sqrt{j(j+1) - m(m\pm 1)}\,|j,m\pm 1\rangle. \tag{9.11}$$

The transverse components follow from

$$\hat{J}_x = \frac{\hat{J}_+ + \hat{J}_-}{2}, \qquad \hat{J}_y = \frac{\hat{J}_+ - \hat{J}_-}{2i}. \tag{9.12}$$

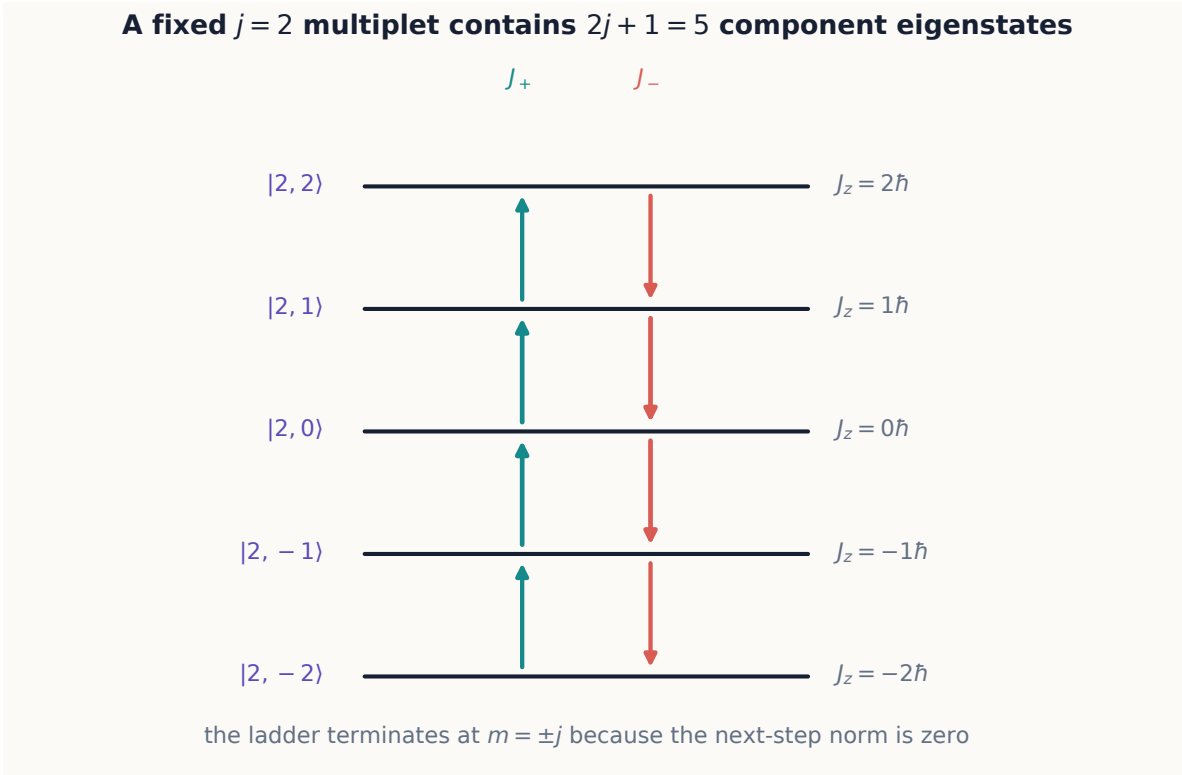

Figure 9.1: For fixed j, the ladder operators move only among the $2j + 1$ allowed component values. Their action vanishes at $m = \pm j$, so the ladder cannot continue beyond its two physical endpoints.

Example 9.1: a $j = 1$ measurement

Let

$$|\psi\rangle = \frac{1}{2}|1, 1\rangle + \frac{i}{\sqrt{2}}|1, 0\rangle + \frac{1}{2}|1, -1\rangle. \tag{9.13}$$

The J_z probabilities are $1/4, 1/2, 1/4$, so $\langle J_z\rangle = 0$ and $\langle J_z^2\rangle = \hbar^2/2$. Hence $\Delta J_z = \hbar/\sqrt{2}$. Every term has the same j, so a measurement of J^2 gives $2\hbar^2$ with certainty.

Example 9.2: recovering the Pauli matrices

For $j = \frac{1}{2}$, the ladder relations should reproduce the Pauli-matrix spin operators of Chapter 3—after all, spin-$\frac{1}{2}$ is nothing but the $j = \frac{1}{2}$ representation of the same algebra. Order the two basis states as $|\frac{1}{2}, \frac{1}{2}\rangle$ then $|\frac{1}{2}, -\frac{1}{2}\rangle$. Equation (9.11) gives

$$\hat{J}_+\big|\tfrac{1}{2}, -\tfrac{1}{2}\big\rangle = \hbar\sqrt{\tfrac{3}{4} - \big(-\tfrac{1}{2}\big)\big(\tfrac{1}{2}\big)}\,\big|\tfrac{1}{2}, \tfrac{1}{2}\big\rangle = \hbar\big|\tfrac{1}{2}, \tfrac{1}{2}\big\rangle, \qquad \hat{J}_+\big|\tfrac{1}{2}, \tfrac{1}{2}\big\rangle = 0, \tag{9.14}$$

and, by the same formula with the lower sign, $\hat{J}_-|\frac{1}{2}, \frac{1}{2}\rangle = \hbar|\frac{1}{2}, -\frac{1}{2}\rangle$ and $\hat{J}_-|\frac{1}{2}, -\frac{1}{2}\rangle = 0$. In matrix form,

$$\hat{J}_+ \doteq \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \hat{J}_- \doteq \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{9.15}$$

Combining these through $\hat{J}_x = (\hat{J}_+ + \hat{J}_-)/2$, $\hat{J}_y = (\hat{J}_+ - \hat{J}_-)/2i$, and $\hat{J}_z|\frac{1}{2}, \pm\frac{1}{2}\rangle = \pm(\hbar/2)|\frac{1}{2}, \pm\frac{1}{2}\rangle$ gives

$$\hat{J}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{J}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \hat{J}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{9.16}$$

which is exactly $\hat{\mathbf{J}} = (\hbar/2)\boldsymbol{\sigma}$ with the Pauli matrices defined in Chapter 3. Nothing about spin was assumed beyond $j = \frac{1}{2}$: the entire matrix structure of spin-$\frac{1}{2}$ observables, including the noncommutativity that drove Chapters 2 through 4, is a special case of the abstract ladder construction.

Concept check 9.3

A calculation of $\langle j, m|\hat{J}_-\hat{J}_+|j, m\rangle$ for some candidate m in the range $-j \le m \le j$ comes out negative. What does this signal?

Answer

An error. The quantity equals $\hbar^2[j(j+1) - m(m+1)]$, which is a squared norm and cannot be negative for any m in the physical range; it reaches zero only at the top of the ladder, $m = j$. A negative result means an arithmetic slip, not a new physical possibility.

9.3 Orbital angular momentum

For motion in space,

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}. \tag{9.17}$$

This is the historical origin of the name "angular momentum," and it is one concrete realization of the abstract algebra of Section 9.1: substituting the canonical commutator $[\hat{x}, \hat{p}] = i\hbar$ (Chapter 6) into the components of $\hat{\mathbf{r}} \times \hat{\mathbf{p}}$ shows directly that they satisfy equation (9.1). In spherical coordinates,

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}, \tag{9.18}$$

and the simultaneous eigenfunctions of $\hat{L}^2$ and $\hat{L}_z$ are the spherical harmonics:

$$\hat{L}^2 Y_\ell^m = \hbar^2\ell(\ell+1)Y_\ell^m, \qquad \hat{L}_z Y_\ell^m = \hbar m Y_\ell^m. \tag{9.19}$$

For orbital motion, single-valued spatial wavefunctions require integer $\ell = 0, 1, 2, \ldots$. The reason is direct: a wavefunction must return to the same value after $\phi \to \phi + 2\pi$, and $e^{im\phi}$ does that only for integer m, hence integer ℓ. The intrinsic spin label s may also be half-integer, precisely because spin is never built from a spatial wavefunction's dependence on ϕ in the first place—it is defined directly by the abstract algebra of Section 9.1, with no analog of "returning to the same value" to enforce. Both $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ obey the same operator algebra, but spin is not generated by a particle orbiting around an internal path.

A state with sharp L and L_z but indefinite L_x, L_y is often drawn as a static cone, which makes that indefiniteness look like mere ignorance of a vector sitting still somewhere on it. The classical system that gets the geometry right is not static at all.

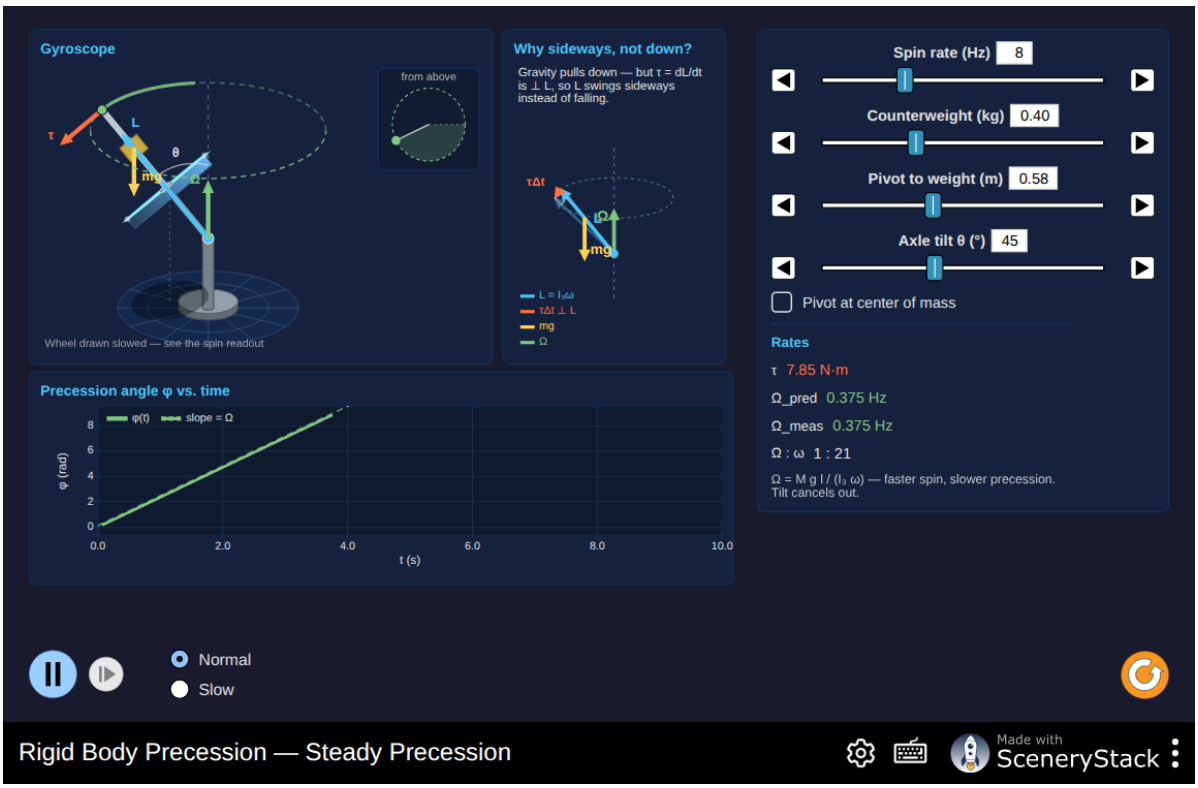

Figure 9.2: A spinning top under gravity: the torque stays perpendicular to $\vec{L}$, so steady precession changes the direction of the angular momentum without changing either its magnitude or its vertical component, walking the vector around the cone at fixed polar angle exactly as L_x and L_y remain indefinite while L and L_z stay sharp. This is the same equation of motion behind the Larmor precession of Chapter 4—a magnetic moment in a field in place of a top in gravity. Interactive simulation: Precession

Rotation generators

A rotation through angle α about $\mathbf{n}$ is

$$\hat{R}_{\mathbf{n}}(\alpha) = e^{-i\alpha\mathbf{n}\cdot\hat{\mathbf{J}}/\hbar}. \tag{9.20}$$

This is the precise sense in which angular momentum "generates" rotations, foreshadowed by the noncommuting-rotations argument in Section 9.1: expanding the exponential to first order in α gives $\hat{R}_{\mathbf{n}}(\alpha) \approx \hat{I} - i\alpha\,\mathbf{n}\cdot\hat{\mathbf{J}}/\hbar$, so $\hat{\mathbf{J}}$ is literally the infinitesimal generator, and composing two small rotations about different axes reproduces $[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$ to leading order. For spin $\frac{1}{2}$, $\hat{\mathbf{J}} = \hbar\boldsymbol{\sigma}/2$ and this reduces to the rotation operator of Chapter 4. Angular momentum is the generator of rotations for every quantum system.

A striking consequence follows immediately from equation (9.20). A full 2π rotation about any axis $\mathbf{n}$ gives $\hat{R}_{\mathbf{n}}(2\pi) = e^{-2\pi i\hat{J}_z/\hbar}$ once evaluated on an eigenbasis, and acting on $|j, m\rangle$ this is simply the phase $e^{-2\pi i m}$. For integer j (hence integer m) this equals $+1$: a full turn returns the state unchanged, matching classical intuition. For half-integer j, m is also half-integer and $e^{-2\pi i m} = -1$: the state acquires an overall minus sign after one full turn and needs a second full turn to return to itself. This is not an experimental defect. It is a genuine prediction of treating spin as a $j = \frac{1}{2}$ angular momentum, and it has been confirmed directly with neutron interferometry.

Concept check 9.4

Orbital angular momentum requires integer ℓ. Does this mean every total angular momentum encountered in this chapter must be an integer?

Answer

No. The integer restriction in this section applies specifically to a spatial wavefunction's single-valuedness under $\phi \to \phi + 2\pi$. Spin is not built from such a wavefunction, so it can be half-integer, and any system built by adding an odd number of half-integer spins—Sections 9.4 through 9.6—can have half-integer total j as well.

9.4 Adding two angular momenta

For subsystems with $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$, define

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2. \tag{9.21}$$

Each subsystem might be two different particles' spins, or one particle's orbital and spin angular momenta, or the total angular momenta of two larger subsystems already built from more elementary pieces; the algebra does not care which. The **uncoupled basis** diagonalizes $\hat{J}_1^2, \hat{J}_{1z}, \hat{J}_2^2, \hat{J}_{2z}$ and is written $|j_1, m_1; j_2, m_2\rangle$. It is the natural basis for describing the two subsystems separately. The **coupled basis** diagonalizes $\hat{J}_1^2, \hat{J}_2^2, \hat{J}^2, \hat{J}_z$ and is written $|j_1, j_2; j, m\rangle$. It is the natural basis whenever a physical interaction or a measurement cares about the *total* angular momentum rather than the two parts individually—a spin–orbit coupling being the leading example, taken up in Section 9.6.

The possible totals obey the triangle rule:

$$j = |j_1 - j_2|, |j_1 - j_2| + 1, \ldots, j_1 + j_2. \tag{9.22}$$

Why a range of totals, and why these particular endpoints?
If $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ behaved like two ordinary classical arrows of fixed lengths j_1 and j_2, their vector sum could range anywhere from $j_1 + j_2$ (the two arrows pointing the same way) down to $|j_1 - j_2|$ (the two arrows pointing as oppositely as their lengths allow), depending on the angle between them. The quantum triangle rule reproduces exactly this classical range of possible total lengths, even though no single state has both $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ pointing in a sharp direction at once. What quantum mechanics adds is that only the values in equation (9.22)—spaced one unit apart, never a continuum—are actually realized, and that a *given* pair of subsystem states does not have one fixed total j; it is generally a superposition over several allowed values, as Example 9.4 makes concrete below.

In every term of a coupled-state expansion, $m = m_1 + m_2$, since $\hat{J}_z = \hat{J}_{1z} + \hat{J}_{2z}$. The coefficients connecting the two bases are **Clebsch–Gordan coefficients**. They are not arbitrary: each coupled state is built, as in the example below, by starting from the unique state of maximum m and applying $\hat{J}_- = \hat{J}_{1-} + \hat{J}_{2-}$ repeatedly, exactly as one builds any j-multiplet from its top state.

A useful check on any such construction is dimension counting: the uncoupled basis has $(2j_1 + 1)(2j_2 + 1)$ states, and the coupled basis, summing the multiplicities $(2j + 1)$ over the triangle-rule range of j, must total the same number. This bookkeeping catches many errors before a single coefficient is computed.

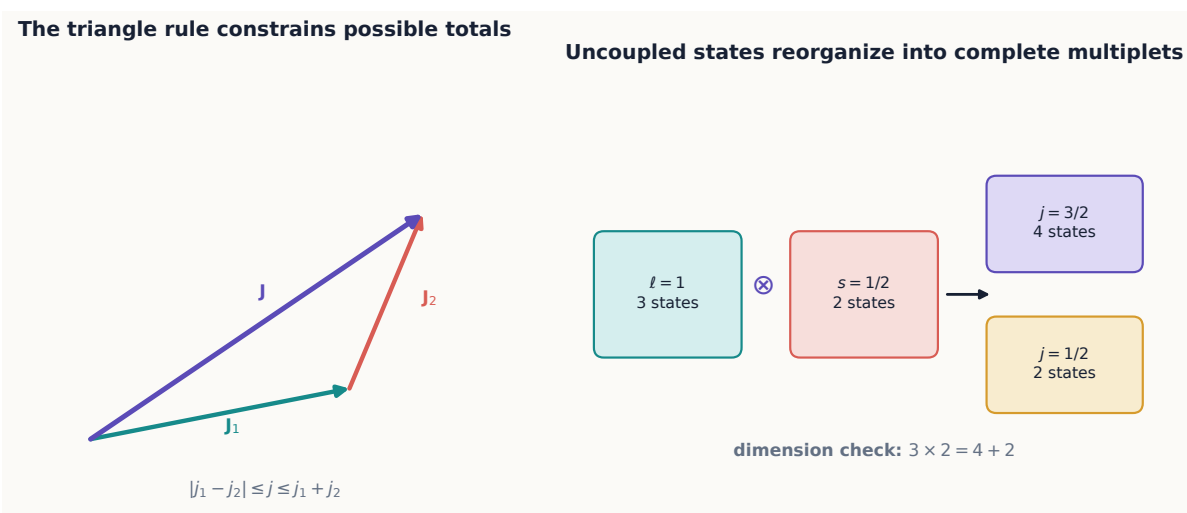

Figure 9.3: The triangle rule lists the allowed total magnitudes, while dimension counting checks that the coupled multiplets account for every uncoupled product state. For $\ell = 1$ and $s = 1/2$, six states become a quadruplet and a doublet.

Example 9.3: coupling orbital motion and spin

Couple $j_1 = \ell = 1$ (a p electron's orbital angular momentum) with $j_2 = s = \frac{1}{2}$ (its spin). The triangle rule allows $j = \frac{3}{2}$ and $j = \frac{1}{2}$, and the dimension check confirms $(2 \cdot 1 + 1)(2 \cdot \frac{1}{2} + 1) = 6 = (2 \cdot \frac{3}{2} + 1) + (2 \cdot \frac{1}{2} + 1) = 4 + 2$. Abbreviate the spin states $|\frac{1}{2}, \pm\frac{1}{2}\rangle$ as $|\uparrow\rangle, |\downarrow\rangle$.

The top state is unique, since only one uncoupled combination reaches the maximum $m = \frac{3}{2}$:

$$\left|\tfrac{3}{2},\tfrac{3}{2}\right\rangle = |1,1\rangle|\uparrow\rangle. \tag{9.23}$$

Lowering with $\hat{J}_- = \hat{L}_- + \hat{S}_-$ and equation (9.11) gives $\hat{J}_-|\frac{3}{2},\frac{3}{2}\rangle = \hbar\sqrt{3}\,|\frac{3}{2},\frac{1}{2}\rangle$ on the left, while on the right $\hat{L}_-|1,1\rangle = \hbar\sqrt{2}\,|1,0\rangle$ and $\hat{S}_-|\uparrow\rangle = \hbar|\downarrow\rangle$, so

$$\hbar\sqrt{3}\,\left|\tfrac{3}{2},\tfrac{1}{2}\right\rangle = \hbar\sqrt{2}\,|1,0\rangle|\uparrow\rangle + \hbar\,|1,1\rangle|\downarrow\rangle, \tag{9.24}$$

Dividing both sides by the common factor $\hbar\sqrt{3}$ isolates the ket on the left and turns the two right-hand coefficients into the Clebsch–Gordan numbers themselves:

$$\left|\tfrac{3}{2},\tfrac{1}{2}\right\rangle = \frac{\sqrt{2}}{\sqrt{3}}\,|1,0\rangle|\uparrow\rangle + \frac{1}{\sqrt{3}}\,|1,1\rangle|\downarrow\rangle, \tag{9.25}$$

giving

$$\left|\tfrac{3}{2},\tfrac{1}{2}\right\rangle = \sqrt{\tfrac{2}{3}}\,|1,0\rangle|\uparrow\rangle + \sqrt{\tfrac{1}{3}}\,|1,1\rangle|\downarrow\rangle. \tag{9.26}$$

Repeating the same lowering step (Exercise 18 asks you to check this instead by lowering from the bottom of the ladder) produces

$$\left|\tfrac{3}{2},-\tfrac{1}{2}\right\rangle = \sqrt{\tfrac{2}{3}}\,|1,0\rangle|\downarrow\rangle + \sqrt{\tfrac{1}{3}}\,|1,-1\rangle|\uparrow\rangle, \qquad \left|\tfrac{3}{2},-\tfrac{3}{2}\right\rangle = |1,-1\rangle|\downarrow\rangle. \tag{9.27}$$

At $m = \frac{1}{2}$, the uncoupled basis has two states, $|1,0\rangle|\uparrow\rangle$ and $|1,1\rangle|\downarrow\rangle$; equation (9.26) uses one combination, and the state orthogonal to it must be the $j = \frac{1}{2}$ state at the same m, since $j = \frac{3}{2}$ and $j = \frac{1}{2}$ are the only totals the triangle rule allows. Choosing the conventional phase (the coefficient of the state with the larger m_1 taken positive),

$$\left|\tfrac{1}{2},\tfrac{1}{2}\right\rangle = \sqrt{\tfrac{2}{3}}\,|1,1\rangle|\downarrow\rangle - \sqrt{\tfrac{1}{3}}\,|1,0\rangle|\uparrow\rangle, \qquad \left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle = \sqrt{\tfrac{1}{3}}\,|1,0\rangle|\downarrow\rangle - \sqrt{\tfrac{2}{3}}\,|1,-1\rangle|\uparrow\rangle. \tag{9.28}$$

Equations (9.26) and (9.28) give the Clebsch–Gordan coefficients for $\ell = 1$ coupled with $s = \frac{1}{2}$; the same six coefficients reappear throughout atomic fine structure (Section 9.6) and in any p-orbital spin–orbit problem.

Concept check 9.5

Why does coupling $\ell = 1$ with $s = \frac{1}{2}$ produce one quadruplet ($j = \frac{3}{2}$) and one doublet ($j = \frac{1}{2}$), rather than, say, two doublets with a state left over?

Answer

Dimension counting fixes it: the six uncoupled product states must reorganize into complete j-multiplets. The triangle rule permits only $j = \frac{3}{2}$ and $j = \frac{1}{2}$, and $(2\cdot\frac{3}{2}+1)+(2\cdot\frac{1}{2}+1) = 4+2 = 6$ exactly accounts for all six states, with no other combination of allowed j values doing so.

9.5 Two spin-one-half systems

The highest state is uniquely

$$|1,1\rangle = |+z,+z\rangle. \tag{9.29}$$

Applying the total lowering operator $\hat{J}_- = \hat{S}_{1-} + \hat{S}_{2-}$ gives

$$|1,0\rangle = \frac{|+z,-z\rangle + |-z,+z\rangle}{\sqrt{2}}. \tag{9.30}$$

A second lowering gives $|1,-1\rangle = |-z,-z\rangle$. These three symmetric states form the $j=1$ **triplet**. The remaining orthogonal state is

$$|0,0\rangle = \frac{|+z,-z\rangle - |-z,+z\rangle}{\sqrt{2}}, \tag{9.31}$$

the antisymmetric $j=0$ **singlet**.

It is worth pausing to recognize these states: equation (9.31) is exactly the entangled state $|\Psi^-\rangle$ from Chapter 5, and $|1,0\rangle$ above is exactly $|\Psi^+\rangle$. Chapter 5 introduced them as the states with the strongest possible nonclassical correlations between two separated spins, without yet having a name for *why* nature singles out precisely this pair of combinations. This section supplies that reason: the singlet and the $m=0$ triplet state are not an arbitrary choice of entangled states—they are exactly the states of definite *total* angular momentum, $j=0$ and $j=1$ respectively, that a rotationally symmetric interaction between two spins naturally produces and conserves. Entanglement and the addition of angular momentum turn out to be two views of the same construction for a two-spin system.

Example 9.4: measuring total spin

The product state $|+z,-z\rangle$ is not an eigenstate of $\hat{J}^2$. Inverting the coupled-basis relations,

$$|+z,-z\rangle = \frac{|1,0\rangle + |0,0\rangle}{\sqrt{2}}. \tag{9.32}$$

A total-spin measurement therefore gives $j=1$ or $j=0$, each with probability $1/2$, while $J_z = 0$ is certain.

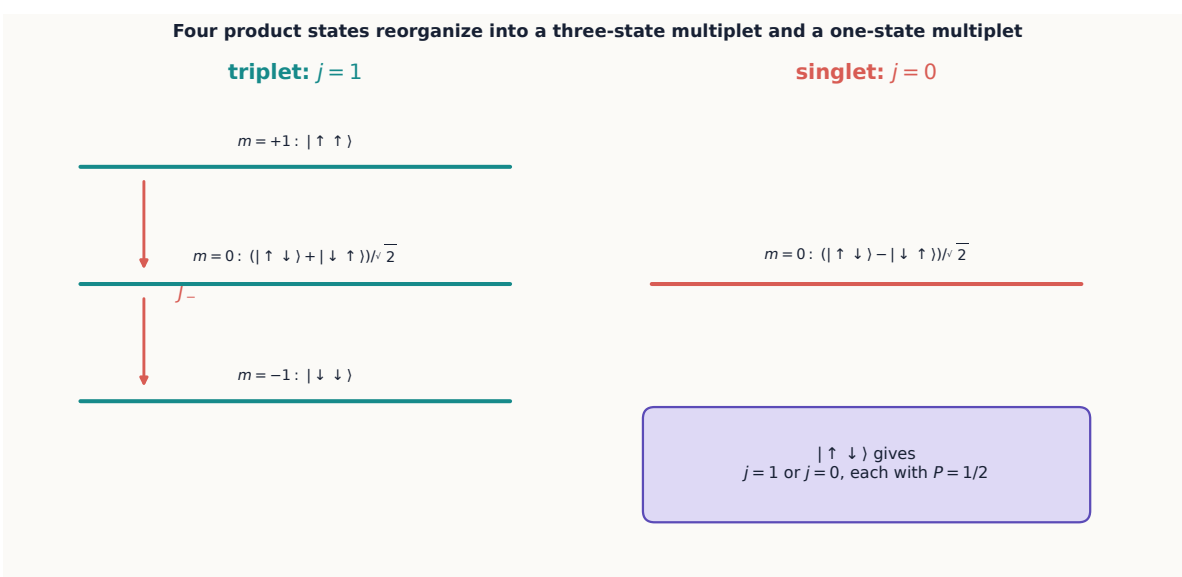

Figure 9.4: Coupling two spin-1/2 systems produces one symmetric triplet and one antisymmetric singlet. A product state with opposite component values has a definite J_z but not a definite total J^2.

Concept check 9.6

Why is the singlet unchanged by applying the same spatial rotation to both spins?

Answer

It has total $j = 0$, so every component of total angular momentum annihilates it. Equation (9.20) therefore acts as the identity on the singlet.

9.6 Coupling in atoms

An electron has orbital angular momentum $\hat{\mathbf{L}}$ and spin $\hat{\mathbf{S}}$. Their total is $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$, with

$$j = \ell \pm \frac{1}{2} \tag{9.33}$$

except that $\ell = 0$ permits only $j = 1/2$—consistent with the triangle rule of Section 9.4 applied to $j_1 = \ell$ and $j_2 = \frac{1}{2}$. Example 9.3 constructed the explicit coupled states for $\ell = 1$; the same lowering construction works for any ℓ. A spin–orbit interaction is proportional to $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$. Since

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \frac{1}{2}(\hat{J}^2 - \hat{L}^2 - \hat{S}^2), \tag{9.34}$$

its energy shifts are simplest in the coupled basis: on a state $|j, m\rangle$ built from fixed ℓ and $s = \frac{1}{2}$,

$$\langle \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \rangle = \frac{\hbar^2}{2}\Big[j(j+1) - \ell(\ell+1) - \tfrac{3}{4}\Big], \qquad j = \ell \pm \frac{1}{2}. \tag{9.35}$$

This is a general lesson: choose a basis that diagonalizes the dominant terms of the Hamiltonian and the measurement being asked about.

Example 9.5: fine-structure splitting of a p level

Apply equation (9.35) to $\ell = 1$. For $j = \frac{3}{2}$,

$$\langle \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \rangle = \frac{\hbar^2}{2}\Big[\frac{15}{4} - 2 - \frac{3}{4}\Big] = \frac{\hbar^2}{2}, \tag{9.36}$$

while for $j = \frac{1}{2}$,

$$\langle \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \rangle = \frac{\hbar^2}{2}\Big[\frac{3}{4} - 2 - \frac{3}{4}\Big] = -\hbar^2. \tag{9.37}$$

For a spin–orbit coupling of the usual (positive) sign, the energy shift is proportional to $\langle \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \rangle$, so the $j = \frac{3}{2}$ level is pushed up and the $j = \frac{1}{2}$ level is pushed down: a single unperturbed p level splits into two, with $j = \frac{3}{2}$ higher. This is the same splitting pattern responsible for the sodium D-line doublet, where the $3p$ level's $j = \frac{3}{2}$ and $j = \frac{1}{2}$ components acquire slightly different energies and hence emit at two closely spaced wavelengths.

Summary

- Angular momentum is defined by its commutation algebra, not by any single formula; $\mathbf{r} \times \mathbf{p}$ and $\frac{\hbar}{2}\boldsymbol{\sigma}$ are two realizations of the same structure.
- Angular momentum has simultaneous eigenstates of J^2 and one component.

- A non-negative-norm argument, not a separate postulate, forces the ladder to terminate and fixes the allowed j, m values.
- Ladder operators generate all m states for a fixed j, and for $j = \frac{1}{2}$ they exactly reproduce the Pauli matrices.
- Orbital and spin angular momentum share an algebra but have different physical origins; only orbital motion is restricted to integer values.
- Quantum rotations are generated by angular momentum, and half-integer j states pick up a sign under a full 2π rotation.
- Coupled and uncoupled bases organize the addition of angular momenta, with Clebsch–Gordan coefficients connecting them.
- Two spin-one-half systems combine into a symmetric triplet and an antisymmetric singlet; orbital and spin angular momentum combine into fine-structure multiplets with a predictable splitting pattern.

Exercises

1. Use equation (9.11) to build the matrix of $\hat{J}_x$ for $j = 1$.
2. Calculate $\langle J_x \rangle$ in the state of Example 9.1.
3. Show that $[\hat{J}^2, \hat{J}_z] = 0$ from the component commutators.
4. Verify that $Y_\ell^m \propto e^{im\phi}$ follows from the L_z equation.
5. List the allowed total j values for $j_1 = 1$ and $j_2 = 3/2$ and verify the dimensions on both sides.
6. Derive all triplet states by lowering from $|1, 1\rangle$.
7. Verify directly that the singlet has zero total J_z, J_+, and J_-.
8. Find probabilities for total j and m in $|-z, +z\rangle$.
9. For $\ell = 2$ and $s = 1/2$, find the eigenvalues of $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ for both allowed j values.
10. Explain why measuring S_{1z} on the singlet destroys certainty in a subsequent total-spin measurement.
11. Show that a 2π rotation acts differently on integer-j and half-integer-j kets.
12. Construct projectors onto the singlet and triplet subspaces using $\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2$.
13. Derive $\hat{J}_+\hat{J}_- = \hat{J}^2 - \hat{J}_z^2 + \hbar\hat{J}_z$ and use it to find the norm of $\hat{J}_-|j, m\rangle$, confirming it matches the lowering case of equation (9.11).
14. Using the 2×2 matrices from Example 9.2, compute $\hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$ directly and verify it equals $\frac{3}{4}\hbar^2\hat{I}$, consistent with $j = \frac{1}{2}$.
15. Using the $j = 1$ matrix for $\hat{J}_x$ found in Exercise 1, construct $\hat{J}_y$ and $\hat{J}_z$ in the same basis and verify $[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$ explicitly.
16. Verify that the four $j = \frac{3}{2}$ states and two $j = \frac{1}{2}$ states of Example 9.3 are normalized and mutually orthogonal.

17. A p electron is prepared in $|j = \frac{3}{2}, m = \frac{1}{2}\rangle$. Find the probability that a measurement of L_z returns 0.

18. Starting instead from $|j = \frac{3}{2}, -\frac{3}{2}\rangle = |1, -1\rangle|\downarrow\rangle$ in Example 9.3 and raising with $\hat{J}_+$, rederive $|j = \frac{3}{2}, -\frac{1}{2}\rangle$ and confirm it agrees with the result found by lowering from the top.

19. For an f electron ($\ell = 3$), list the allowed j values and use equation (9.35) to find $\langle \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \rangle$ for each.

20. Show that the two values found in Example 9.5 both follow from equation (9.35) evaluated at $j = \ell \pm \frac{1}{2}$ with $\ell = 1$.

21. Explain why the non-negative-norm argument behind Concept check 9.3 constrains j itself, and not merely the range of m available once j is fixed.

22. A 2π rotation about an arbitrary axis is applied to a $j = \frac{3}{2}$ state. State the result and compare it with the $j = 1$ and $j = \frac{1}{2}$ cases.

23. Two electrons occupy a p subshell, so $\ell_1 = \ell_2 = 1$. Use the triangle rule to list the allowed values of their total orbital angular momentum L.

24. A beam of atoms with $j = 2$ has all five m outcomes equally likely. Find $\langle J_z \rangle$ and $\langle J_z^2 \rangle$.

Chapter 10

Approximation Methods

10.1 Approximation methods

Learning objectives

After this chapter, you should be able to:

- calculate first- and second-order stationary perturbation corrections;
- handle a perturbation inside a degenerate subspace;
- apply the variational principle to estimate a ground-state energy;
- identify the semiclassical regime and use WKB connection ideas;
- calculate first-order transition amplitudes from a time-dependent perturbation; and
- judge an approximation using dimensionless parameters and limiting cases.

10.1 Why approximations are part of the theory

Exactly solvable Hamiltonians are rare. Real atoms interact with fields and other particles; real wells are not perfectly square; and most coupled differential equations do not have closed-form solutions. Approximation methods are controlled ways to retain the most important physics.

Before calculating, identify a dimensionless small or large parameter. An answer labeled "first order" is useful only when omitted terms are expected to be smaller. After calculating, check normalization, units, symmetry, and the limit in which the perturbation disappears.

This chapter develops five complementary tools. Stationary perturbation theory (Sections 10.2–10.3) corrects the energies and states of a Hamiltonian that is close to one already solved. The variational principle (Section 10.4) needs no such reference Hamiltonian at all, only a plausible guess for the shape of a ground state. WKB (Section 10.5) exploits a different smallness—a slowly varying potential rather than a weak one—and applies even when no nearby solvable problem exists. Time-dependent perturbation theory (Section 10.6) answers a different question altogether: not "what is the corrected energy" but "what is the probability of a transition." Section 10.7 collects these into a single decision procedure. None of the five methods is a substitute for solving the Schrödinger equation exactly; each is a controlled bookkeeping scheme for what happens when an exact solution is out of reach.

10.2 Nondegenerate stationary perturbation theory

Let

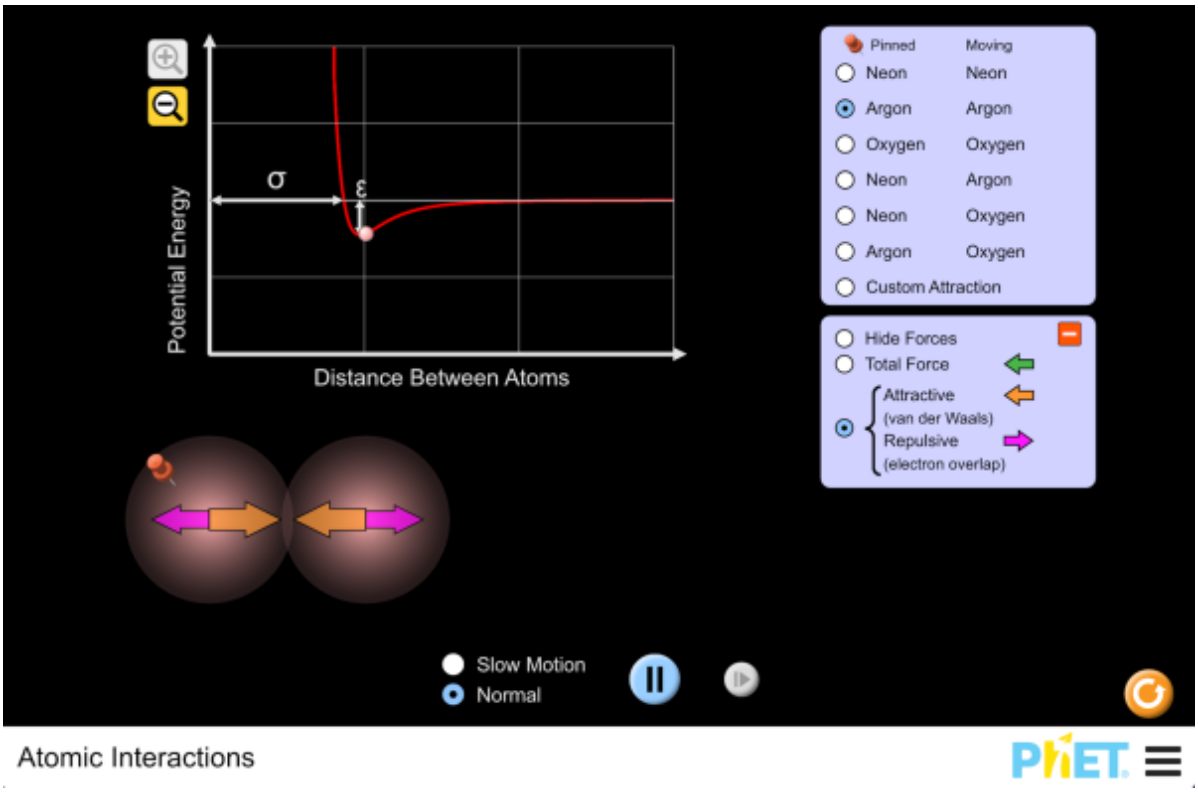

Figure 10.1: The interaction between two real atoms is not a square well or a harmonic oscillator; it is closer to the Lennard-Jones potential this simulation plots. Adjust the atom types and separation and watch the force and potential energy curves change shape—exactly the kind of realistic potential the methods below are built to handle without an exact solution. Interactive simulation: Atomic Interactions

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}, \tag{10.1}$$

where $\hat{H}_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle$ is solved and λ tracks order. Expand

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \tag{10.2}$$

and similarly for the ket. Matching powers of λ gives

$$E_n^{(1)} = \langle n^{(0)}|\hat{V}|n^{(0)}\rangle, \tag{10.3}$$

$$|n^{(1)}\rangle = \sum_{m\neq n} |m^{(0)}\rangle \frac{\langle m^{(0)}|\hat{V}|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}, \tag{10.4}$$

and

$$E_n^{(2)} = \sum_{m\neq n} \frac{|\langle m^{(0)}|\hat{V}|n^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \tag{10.5}$$

The first energy correction is the perturbing potential averaged in the unperturbed state. State mixing is large when a coupling matrix element is large or an energy denominator is small.

Example 10.1: a two-level perturbation

Take

$$\hat{H}_0 = \frac{\Delta}{2}\sigma_z, \qquad \hat{V} = g\sigma_x, \qquad |g| \ll |\Delta|. \tag{10.6}$$

The diagonal matrix elements of $\hat{V}$ vanish, so both first-order energy shifts vanish. For the upper level,

$$E_+^{(2)} = \frac{g^2}{\Delta}, \tag{10.7}$$

while $E_-^{(2)} = -g^2/\Delta$. Expanding the exact eigenvalues $\pm\sqrt{\Delta^2/4 + g^2}$ confirms these results. Comparison with an exact answer is an excellent way to understand an approximation's domain.

Selection rules save work

If symmetry forces $\langle m|\hat{V}|n\rangle = 0$, that state does not appear in the corresponding sum. For example, an odd perturbation has zero diagonal matrix element in a parity eigenstate, so its first-order energy shift vanishes. A zero correction often expresses symmetry, not absence of a physical effect at every order.

Concept check 10.1

Can a first-order correction vanish while the exact energy still changes?

Answer

Yes. Symmetry may make the diagonal matrix element zero while second- and higher-order corrections remain.

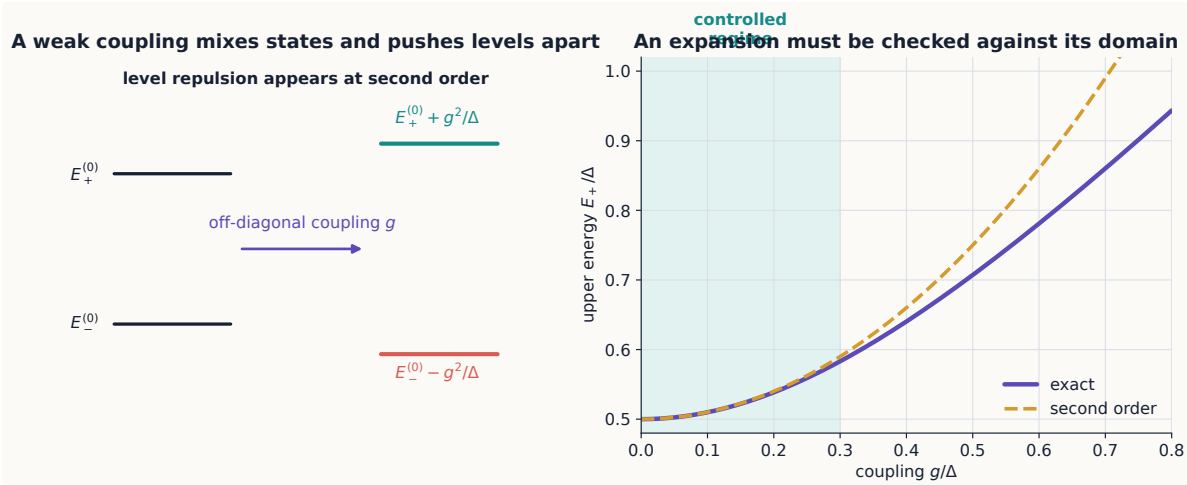

Figure 10.2: Off-diagonal coupling produces level repulsion even when the first-order diagonal shifts vanish. The truncated series follows the exact answer only while the coupling remains small relative to the unperturbed gap.

Example 10.2: an anharmonic correction that grows with the state

Take the harmonic oscillator of Chapter 7, $\hat{H}_0 = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})$, perturbed by a quartic term $\hat{V} = \hat{x}^4$ with strength λ, so $\hat{H} = \hat{H}_0 + \lambda\hat{x}^4$. When we write $\hat{x} = \sqrt{\hbar/(2m\omega)}\,(\hat{a} + \hat{a}^\dagger)$ and expand $(\hat{a} + \hat{a}^\dagger)^4$, only terms that return to $|n\rangle$ survive the diagonal matrix element, and their combinatorics give

$$E_n^{(1)} = \lambda\langle n|\hat{x}^4|n\rangle = \frac{3\lambda\hbar^2}{4m^2\omega^2}\left(2n^2 + 2n + 1\right). \tag{10.8}$$

Two features are worth pausing on. First, unlike Example 10.1, no diagonal matrix element vanishes here—the quartic term is even, like the unperturbed potential, so nothing forbids a first-order shift. Second, and more important, the correction is not a fixed small number: it grows quadratically with n. A denominator or matrix element being "small" is a statement about one particular calculation, not a permanent property of λ.

A small coupling is not a small correction at every level

Fix λ as small as you like. Equation (10.8) shows $E_n^{(1)}$ growing like n^2, while the unperturbed level spacing $\hbar\omega$ stays fixed. For large enough n, $E_n^{(1)}$ eventually becomes comparable to $\hbar\omega$ itself, and first-order perturbation theory can no longer be trusted there even though it worked well for the low-lying levels with the very same λ. "The perturbation is weak" must be checked against the specific state being corrected, not asserted once for the whole spectrum.

Concept check 10.2

For the quartic oscillator of Example 10.2, does making λ smaller guarantee that first-order perturbation theory is accurate for every level n?

Answer

No. Equation (10.8) grows like n^2 for fixed λ, so for any nonzero λ there is always some large enough n at which the correction rivals the level spacing $\hbar\omega$. Making λ smaller only pushes that breakdown to higher n; it does not remove it.

10.3 Degenerate perturbation theory

Equation (10.4) fails when a denominator vanishes: dividing by $E_n^{(0)} - E_m^{(0)} = 0$ is not a small correction but a breakdown of the expansion. Whenever two or more unperturbed states share an energy, the zeroth-order states are not yet uniquely determined—any combination within the degenerate subspace is equally an eigenstate of $\hat{H}_0$—and the perturbation itself must be used to select the right combination before a first-order energy shift means anything. Within a degenerate subspace, first diagonalize the matrix

$$V_{ab} = \langle a^{(0)}|\hat{V}|b^{(0)}\rangle. \tag{10.9}$$

Its eigenvectors are the correct zeroth-order combinations, and its eigenvalues are their first-order energy shifts.

Example 10.3: splitting a degenerate pair

Suppose $|1\rangle$ and $|2\rangle$ both have energy E_0, and within their subspace

$$\hat{V} \doteq \begin{pmatrix} \epsilon & v \\ v^* & \epsilon \end{pmatrix}. \tag{10.10}$$

The shifts are $\epsilon \pm |v|$. If $v = |v|e^{i\chi}$, corresponding normalized states can be chosen as

$$\frac{|1\rangle \pm e^{-i\chi}|2\rangle}{\sqrt{2}}. \tag{10.11}$$

The perturbation both selects a preferred basis and lifts the degeneracy.

10.4 The variational principle

Sections 10.2 and 10.3 both correct a Hamiltonian that is already close to a solved one. The variational principle needs no such reference Hamiltonian: it converts a guess about the general shape of a ground state directly into a number, at the price of no longer knowing how close that number is to the truth without further work.

For a Hamiltonian bounded below, any normalized trial state obeys

$$E_{\text{trial}} = \langle\psi_{\text{trial}}|\hat{H}|\psi_{\text{trial}}\rangle \geq E_0. \tag{10.12}$$

Expanding the trial state in exact energy eigenstates proves the result: $E_{\text{trial}} = \sum_n |c_n|^2 E_n \geq E_0$. Varying parameters in a sensible trial family gives the best upper bound available within that family.

Example 10.4: estimate a confined ground state

For a one-dimensional Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + a|x|, \tag{10.13}$$

choose the normalized Gaussian

$$\psi_b(x) = \frac{1}{(\pi b^2)^{1/4}} e^{-x^2/(2b^2)}. \tag{10.14}$$

Its energy is

$$E(b) = \frac{\hbar^2}{4mb^2} + \frac{ab}{\sqrt{\pi}}. \tag{10.15}$$

Minimizing $E(b)$ balances localization kinetic energy against potential energy and determines both an energy upper bound and an approximate spatial scale. Dimensional analysis already predicts $b \sim (\hbar^2/ma)^{1/3}$.

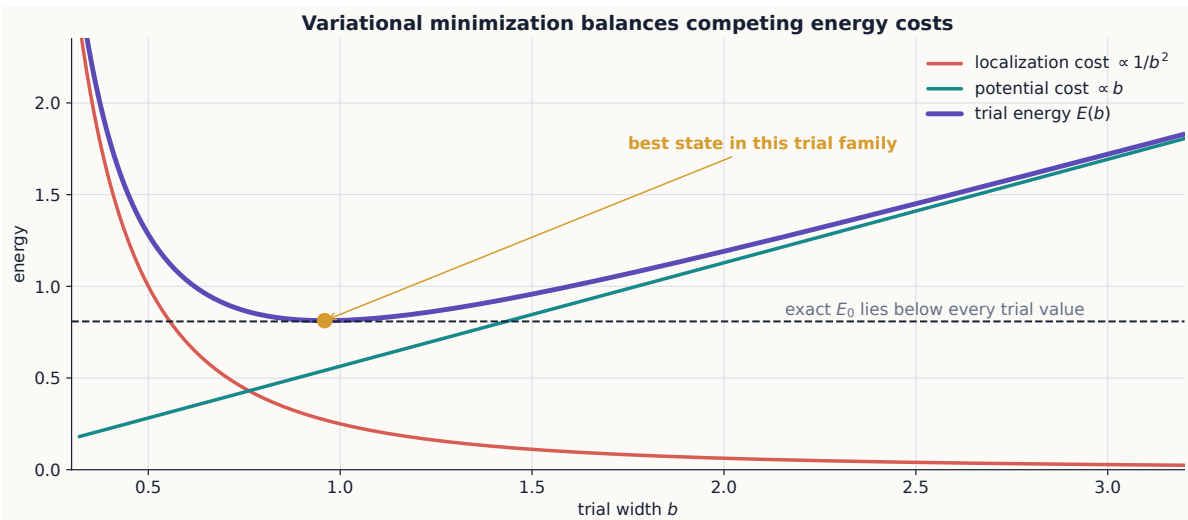

Figure 10.3: Variational optimization balances competing costs: a narrow state pays high kinetic energy, while a broad state samples more potential energy. The minimum is the best upper bound supplied by the chosen family, not necessarily the exact ground energy.

The variational method is safest for the ground state. The reason is worth making explicit rather than taking on faith. Expanding any normalized trial state in the exact eigenbasis, $|\psi_{\text{trial}}\rangle = \sum_n c_n|n\rangle$, gives $E_{\text{trial}} = \sum_n |c_n|^2 E_n$, a weighted average of *all* the exact energies with nonnegative weights summing to one. Such an average can never fall below the smallest term E_0, which is exactly equation (10.12). There is no equally automatic bound relative to the first excited energy E_1: nothing stops a trial state from having some overlap $|c_0|^2 > 0$ with the true ground state, which pulls E_{trial} below E_1 without violating anything. Excited-state estimates therefore require trial functions built to be exactly orthogonal to the true ground state—typically enforced by symmetry, such as parity or a node structure, rather than by numerically minimizing overlap. If the exact ground state is not fully known, an excited-state trial function orthogonal only to an *approximate* ground state offers no guaranteed bound at all.

Concept check 10.3

Why is a variational energy below a known exact ground energy a warning?

Answer

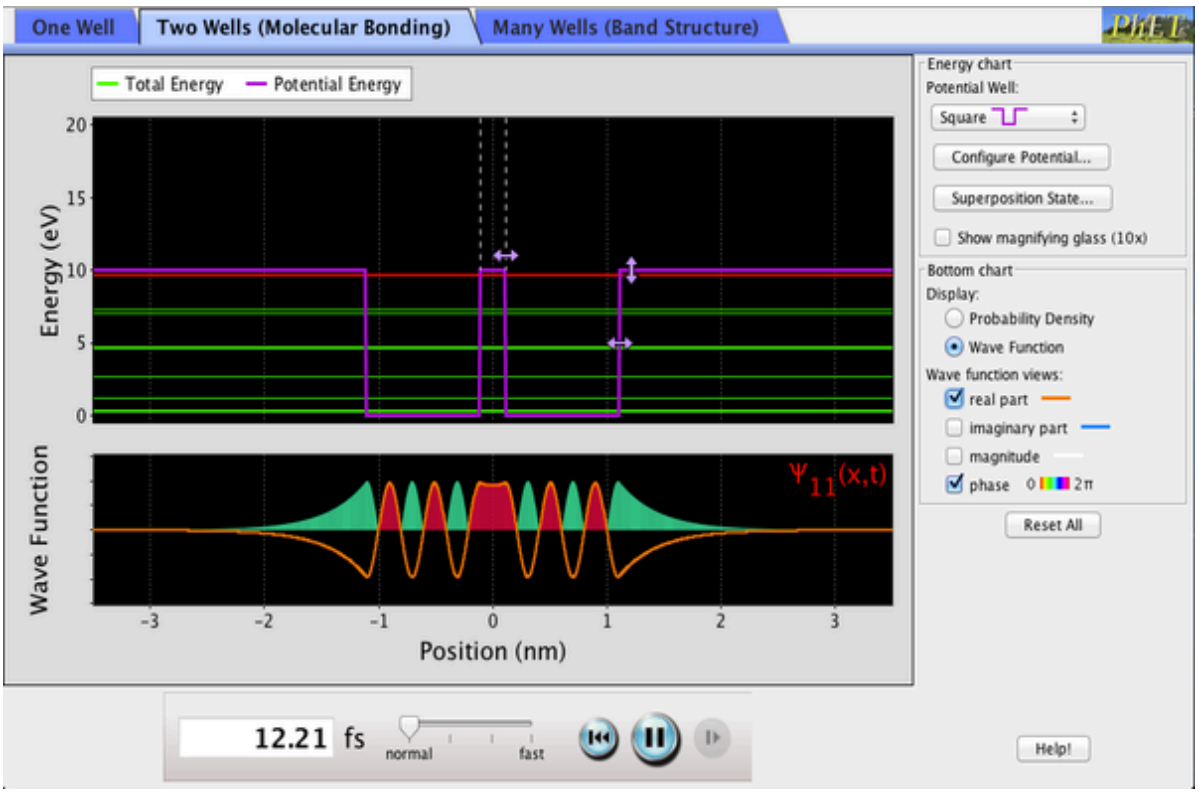

Figure 10.4: Build a well and read off its ground-state energy directly, then compare that exact number against the trial-Gaussian estimate from equation (10.12)—the same simulation introduced for the finite well in Chapter 7, now used to check a variational guess rather than to explore an exactly solvable potential. Interactive simulation: Bound States

For a normalized trial state and the correct Hamiltonian it is impossible. Such a result signals an algebra, normalization, boundary-condition, or numerical error.

10.5 The WKB approximation

The variational principle traded a small parameter for a guessed function shape. WKB reintroduces a small parameter, but a geometric one: not a weak coupling, but a potential that varies slowly compared with the local de Broglie wavelength. Write a stationary wavefunction locally as

$$\psi(x) = A(x)e^{iS(x)/\hbar}. \tag{10.16}$$

When the potential varies slowly over a local wavelength, the leading semiclassical solution in an allowed region is

$$\psi(x) \approx \frac{1}{\sqrt{p(x)}} \left[C_+ e^{(i/\hbar)\int^x p(x')dx'} + C_- e^{-(i/\hbar)\int^x p(x')dx'} \right], \tag{10.17}$$

where $p(x) = \sqrt{2m[E - V(x)]}$. In a forbidden region, set $\kappa(x) = \sqrt{2m[V(x) - E]}/\hbar$ to obtain growing and decaying exponentials.

Across a broad forbidden region $x_1 < x < x_2$, the leading tunneling factor is

$$T \sim \exp\left[-2 \int_{x_1}^{x_2} \kappa(x)\, dx \right]. \tag{10.18}$$

For bound motion between two smooth turning points, connection formulas give

$$\int_{x_1}^{x_2} p(x)\, dx = \left(n + \frac{1}{2} \right) \pi\hbar. \tag{10.19}$$

WKB fails near a turning point if used without connection formulas and fails when the potential changes substantially within one wavelength.

Concept check 10.4

Does WKB require high total energy in every problem?

Answer

No. It requires the wavelength to vary slowly on the wavelength scale. High energy often helps, but slowly varying potentials can also create a valid semiclassical regime.

Example 10.5: the quantum bouncer

A particle of mass m rests above an impenetrable floor under a uniform downward force F: $V(x) = Fx$ for $x > 0$ and $V(x) = \infty$ for $x < 0$. This potential has only one smooth classical turning point, at $x_t = E/F$; the other boundary is not a turning point at all but a hard wall, where the exact wavefunction must vanish outright rather than decay smoothly.

Equation (10.19) was derived for two smooth turning points, each contributing a connection-formula phase of $\pi/4$. Standard connection-formula analysis shows that a hard wall contributes $\pi/2$ instead, since the wave is forced to a node exactly at the wall rather than merely turning around over an extended region. Replacing one $\pi/4$ by $\pi/2$ changes the quantization condition to

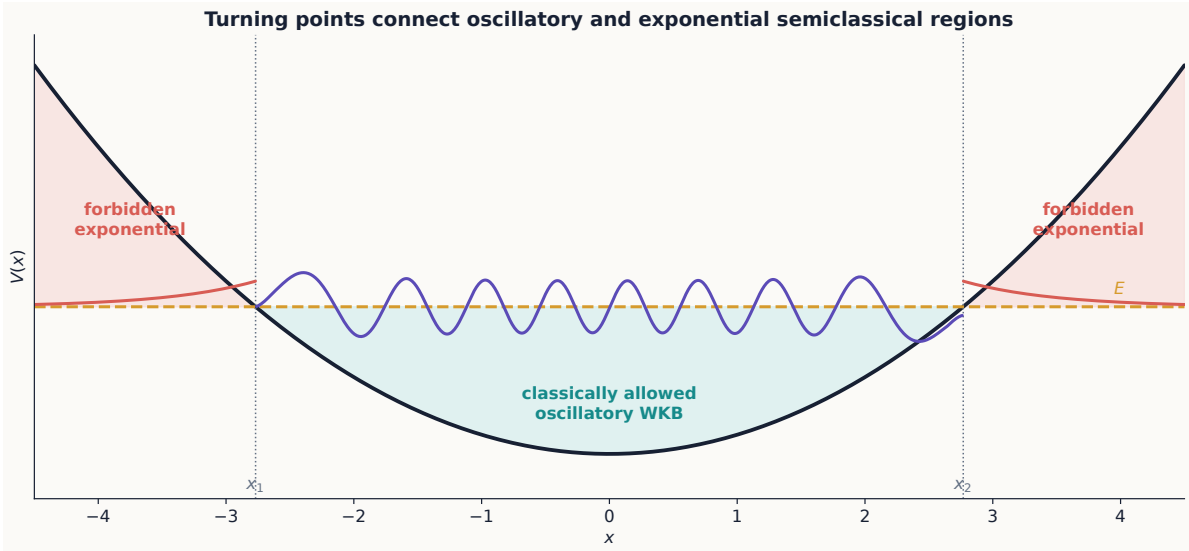

Figure 10.5: Turning points divide oscillatory, classically allowed motion from exponential, forbidden behavior. Connection formulas—not the separate local expressions—are what carry one approximation smoothly into the other.

$$\int_0^{x_t} p(x)\,dx = \left(n + \frac{3}{4}\right)\pi\hbar, \qquad n = 0, 1, 2, \ldots \tag{10.20}$$

With $p(x) = \sqrt{2m(E - Fx)}$, substituting $u = E - Fx$ gives

$$\int_0^{E/F} \sqrt{2m(E - Fx)}\,dx = \frac{2\sqrt{2m}}{3F} E^{3/2}. \tag{10.21}$$

Solving equation (10.20) for E gives

$$E_n = \left[\frac{9\pi^2\hbar^2F^2}{8m}\left(n + \frac{3}{4}\right)^2\right]^{1/3}, \qquad n = 0, 1, 2, \ldots \tag{10.22}$$

so $E_n \propto (F^2\hbar^2n^2)^{1/3}$ for large n. Two features distinguish this spectrum from the harmonic oscillator of Chapter 7: the energies grow as $n^{2/3}$ rather than linearly, so the fractional level spacing $(E_{n+1} - E_n)/E_n$ *shrinks* as n grows instead of staying fixed. WKB reaches this result without ever constructing the bouncer's exact Airy-function wavefunction—a genuine payoff of the method beyond the textbook-symmetric wells where an exact solution is available for comparison.

Concept check 10.5

Why isn't the quantum bouncer's quantization condition the same $(n + \frac{1}{2})\pi\hbar$ rule used for the harmonic oscillator?

Answer

The oscillator has two smooth classical turning points, each contributing a connection-formula phase of $\pi/4$, for a total of $\pi/2$. The bouncer has one smooth turning point and one hard wall; the wall enforces an exact node and contributes $\pi/2$ by itself, for the same total phase but a different split, giving the $(n + \frac{3}{4})\pi\hbar$ condition of equation (10.20).

10.6 Time-dependent perturbations

Stationary perturbation theory and the variational method both estimate properties of energy eigenstates. Time-dependent perturbation theory asks a different kind of question: given a system prepared in one eigenstate of $\hat{H}_0$, with what probability does a time-dependent coupling later find it in another? Let

$$\hat{H}(t) = \hat{H}_0 + \lambda \hat{V}(t). \tag{10.23}$$

Starting in $|i\rangle$, the first-order amplitude to find $|f\rangle$ is

$$c_f^{(1)}(t) = -\frac{i}{\hbar}\int_0^t e^{i\omega_{fi}t'}\langle f|\hat{V}(t')|i\rangle\, dt', \qquad \omega_{fi} = \frac{E_f - E_i}{\hbar}. \tag{10.24}$$

Example 10.6: a coupling switched on and held constant

The simplest case is a time-independent matrix element $V_{fi} = \langle f|\hat{V}|i\rangle$ switched on suddenly at $t = 0$ and left constant afterward. Equation (10.24) then integrates directly:

$$c_f^{(1)}(T) = -\frac{i}{\hbar}V_{fi}\int_0^T e^{i\omega_{fi}t'}\, dt' = -\frac{V_{fi}}{\hbar}\,\frac{e^{i\omega_{fi}T} - 1}{\omega_{fi}}. \tag{10.25}$$

Taking $|c_f^{(1)}(T)|^2$ and using $|e^{i\theta} - 1|^2 = 2 - 2\cos\theta = 4\sin^2(\theta/2)$ gives

$$P_{i\to f}(T) = \frac{|V_{fi}|^2}{\hbar^2}\,\frac{\sin^2(\omega_{fi}T/2)}{(\omega_{fi}/2)^2}. \tag{10.26}$$

At exact resonance, $\omega_{fi} = 0$, this reduces to $P_{i\to f}(T) = |V_{fi}|^2T^2/\hbar^2$: the probability grows without bound as T^2, a reminder that first-order perturbation theory is itself only a short-time approximation and must break down—by unitarity, P cannot exceed one—before T grows too large. Away from resonance, equation (10.26) oscillates in T without growing, confined between 0 and $4|V_{fi}|^2/(\hbar\omega_{fi})^2$.

Viewed as a function of ω_{fi} at fixed T, equation (10.26) is a single central peak of height $|V_{fi}|^2T^2/\hbar^2$ and width (first zero) at $|\omega_{fi}| = 2\pi/T$, surrounded by decaying side lobes—the same sinc^2 line shape that appears in single-slit diffraction. The peak narrows as $1/T$ while its height grows as T^2; a longer observation window makes the transition simultaneously more probable at exact resonance and more sharply selective in energy. This time–frequency trade-off is what allows the discrete result of this example to reproduce the smooth, irreversible-looking rate of equation (10.27) once many closely spaced final states are summed over.

Concept check 10.6

Does making the interaction time T longer in Example 10.6 make a transition more or less selective in energy?

Answer

More selective. The peak of equation (10.26) narrows in ω_{fi} as $1/T$ even as its height grows as T^2. In the limit of very long T this sharply peaked, narrow function is what allows a sum over many closely spaced final states to collapse into the energy-conserving rate of Fermi's golden rule, equation (10.27).

An oscillating perturbation contains phases $e^{\pm i\omega t}$. The integral grows most strongly when the driving frequency is near $|\omega_{fi}|$: resonance is constructive accumulation of transition amplitude over time.

For a long interaction and a dense set of final states, the transition rate often takes the form

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar}|V_{fi}|^2\rho(E_f), \tag{10.27}$$

known as Fermi's golden rule. The density of available final states $\rho(E_f)$ matters alongside the coupling strength.

10.7 Choosing and checking a method

Use stationary perturbation theory for weak changes to known discrete eigenstates, degenerate perturbation theory when small denominators signal a degenerate subspace, the variational method for ground-state bounds, WKB for slowly varying potentials or semiclassical tunneling, and time-dependent perturbation theory for weakly driven transitions.

An approximation should report:

1. the parameter assumed small or large;
2. the order retained;
3. the symmetry restrictions used;
4. the range where the expression is expected to work; and
5. at least one independent check.

Summary

- Perturbation theory organizes corrections in powers of a small coupling, but "small" must be checked state by state: a fixed coupling can produce an arbitrarily large correction at a high enough quantum number.
- Degenerate spaces must be diagonalized before nondegenerate formulas apply.
- The variational principle gives an upper bound to the ground energy for any normalized trial state; the same guarantee does not extend to excited states without enforced orthogonality to the true ground state.
- WKB describes slowly varying waves, semiclassical quantization, and exponential tunneling; a hard wall and a smooth turning point contribute different phases to the same connection formula.
- Time-dependent perturbations create transitions by coherent amplitude accumulation, especially near resonance, and a longer interaction time trades a broader, weaker response for a narrower, more resonance-selective one.
- Every approximation needs an explicit control parameter and validity check.

Exercises

1. Derive equations (10.3) and (10.4) by matching first powers of λ.
2. Expand the exact eigenvalues in Example 10.1 through fourth order in g/Δ.

3. For an infinite well perturbed by $V_1(x) = \epsilon x/L$, calculate the first-order energy shift.

4. Use parity to decide which matrix elements connect even and odd states for even and odd perturbations.

5. Diagonalize the matrix in Example 10.3 and verify its phase-dependent eigenvectors.

6. Minimize the trial energy in Example 10.4 and give the numerical coefficient multiplying its natural energy scale.

7. Prove the variational bound by expanding a normalized trial state in exact eigenstates.

8. Apply equation (10.19) to the harmonic oscillator and recover its exact energies.

9. Show that equation (10.18) reduces to the exponential part of the rectangular-barrier result.

10. Redo the calculation of Example 10.6 for a matrix element that is switched on at $t = 0$ and switched off again at $t = T/2$, remaining zero afterward. Find $c_f^{(1)}(t)$ for $t > T/2$ and compare $|c_f^{(1)}|^2$ with equation (10.26) evaluated at the same total "on" time.

11. State which approximation method is best suited to four scenarios: a weak Stark shift, a ground energy of an anharmonic well, alpha decay, and weak resonant light absorption.

12. Identify a dimensionless validity parameter for each method used in this chapter.

13. Verify equation (10.8) for $n = 0$ and $n = 1$ directly, by expanding $(\hat{a} + \hat{a}^\dagger)^4$ in normal-ordered form and evaluating each surviving term between $\langle n|$ and $|n\rangle$.

14. For fixed small λ in Example 10.2, find the quantum number n^* at which $E_n^{(1)}$ first becomes comparable to the unperturbed spacing $\hbar\omega$. Explain why n^* depends on λ even though λ is "small" by assumption.

15. A perturbation $\hat{V} = \mu\hat{x}$ (linear, not quadratic) is added to the harmonic oscillator. Show that first-order perturbation theory predicts a zero energy shift at every level, then find the exact shifted spectrum by completing the square in the full Hamiltonian, and explain why the "vanishing" first-order result is not a contradiction.

16. Using the eigenstate-expansion argument of Section 10.4, explain why no two normalized trial states can both give an expectation value strictly below the true ground energy E_0 of the same Hamiltonian.

17. Give a concrete example of a trial state that is exactly orthogonal to an *approximate* (numerically obtained, not exact) ground state, and explain why minimizing its energy need not bound the true first excited energy from above.

18. Derive the classical turning point $x_t = E/F$ and the WKB integral for the quantum bouncer of Example 10.5, and confirm both the quantization condition (10.20) and the resulting scaling $E_n \propto (F^2\hbar^2 n^2)^{1/3}$.

19. Compare the quantum-bouncer energies at $n = 0$ and $n = 10$ using equation (10.22): by what factor does the energy increase? Does the fractional spacing $(E_{n+1} - E_n)/E_n$ grow or shrink with n?

20. Explain, in terms of the boundary condition on the wavefunction, why a hard wall and a smooth classical turning point contribute different phases to a WKB connection formula.

21. Starting from equation (10.24), rederive equation (10.26), and show that it reduces to $P_{i\to f}(T) \approx |V_{fi}|^2 T^2/\hbar^2$ in the limit $\omega_{fi}T \ll 1$.

22. Show that the peak value of equation (10.26) grows as T^2 while its width in ω_{fi} shrinks as $1/T$, so that the area under the peak approaches a constant. Explain in one sentence how this supports treating the transition rate as time-independent once many closely spaced final states are summed, as in equation (10.27).

23. A perturbation is switched on with a smooth ramp instead of the abrupt step used in Example 10.6. Without repeating the full integral, explain qualitatively how a smoother switch-on changes the side lobes of the sinc^2 line shape far from resonance.

24. For each of the five methods developed in this chapter, write one sentence describing the physical situation where it is the natural first method to try, and a second sentence describing a warning sign—in the setup or in the answer—that the method has been pushed past its domain of validity.

Chapter 11

Identical Particles

11.1 Identical particles

Learning objectives

After this chapter, you should be able to:

- explain why exchanging identical particles cannot create a new observable configuration;
- construct symmetric boson and antisymmetric fermion states;
- apply the Pauli exclusion principle;
- show that the permutation operator's eigenvalues force every species into a bosonic or a fermionic sector;
- build two-particle Slater determinants;
- distinguish exchange effects from ordinary forces, including the direct and exchange energies behind Hund's rule;
- use occupation-number notation; and
- connect quantum statistics to atomic structure and to degeneracy pressure in many-body and astrophysical systems.

11.1 Indistinguishability is physical

Every measurement discussed so far has involved distinguishable degrees of freedom: two separately labeled spins, or a single particle at a definite location. Real matter is built from many particles of the same species, and every electron in the universe has exactly the same mass, charge, and spin. This chapter asks what changes when the objects being described are fundamentally identical—not merely similar, but interchangeable in principle—and shows that the answer is not a new force or interaction. It is a restriction on which states of several particles are allowed to occur at all.

Quick review: two very different meanings of "can't tell them apart"
Skip this box if it is already clear why quantum indistinguishability is a new postulate and not just an admission of experimental sloppiness.

Take two identical classical billiard balls, prepared at positions x_1 and x_2 at some initial time. Even if the balls are manufactured to agree in every measurable respect—same mass, same color, same everything—classical mechanics still assigns each ball its own continuous trajectory. An observer who tracks both trajectories without blinking always

knows which ball started at x_1: "ball 1" and "ball 2" remain meaningful labels for the rest of the experiment, even though no measurement performed on either ball *alone*, after the fact, could ever tell them apart. Classical indistinguishability is a practical limit on bookkeeping, not a limit built into the state itself—nothing in the mathematics prevents assigning permanent labels; only our patience for tracking them runs out.

Quantum mechanics removes even that in-principle handle. A wavefunction does not follow a single sharp trajectory: two identical particles' wave packets can spread out and genuinely overlap, so that there is no fact of the matter, even in principle, about which particle is "the one that started on the left." There is no hidden trajectory that an all-seeing observer could have followed instead. This is why the rest of this chapter treats exchange symmetry as a postulate about which *states* are physically allowed to occur—equation (11.3) below—rather than as a statement about the limits of measurement. Working out the consequences of that one difference is the entire content of this chapter.

For two distinguishable particles, a product wavefunction can assign ψ_a to particle 1 and ψ_b to particle 2:

$$\Psi(1,2) = \psi_a(1)\psi_b(2). \tag{11.1}$$

For identical particles, the labels 1 and 2 are bookkeeping devices, not measurable properties. Exchanging them cannot change any probability:

$$|\Psi(2,1)|^2 = |\Psi(1,2)|^2. \tag{11.2}$$

In three spatial dimensions, consistency permits two fundamental exchange rules:

$$\Psi(2,1) = +\Psi(1,2) \quad \text{(bosons)}, \qquad \Psi(2,1) = -\Psi(1,2) \quad \text{(fermions)}. \tag{11.3}$$

Particles with integer spin are bosons; particles with half-integer spin are fermions. This spin–statistics connection is a result of relativistic quantum field theory and is taken as an empirical rule in nonrelativistic mechanics.

Concept check 11.1

Is an antisymmetric wavefunction itself an observable negative probability?

Answer

No. Exchange changes an amplitude's sign. Probabilities remain nonnegative, but the sign changes interference with the exchanged amplitude.

The permutation operator

The fact that only the squared magnitude of Ψ is constrained suggests introducing a definite operator for exchange itself, rather than reasoning about probabilities alone. Define the **permutation operator** $\hat{P}_{12}$ by its action on any two-particle wavefunction,

$$\hat{P}_{12}\Psi(1,2) = \Psi(2,1). \tag{11.4}$$

Relabeling which coordinate is read first and which second is applied twice by $\hat{P}_{12}^2$, and doing so must return the original function unchanged:

$$\hat{P}_{12}^2 = \hat{I}. \tag{11.5}$$

Exchanging labels also cannot change the inner product of any two wavefunctions built the same way, so $\hat{P}_{12}$ is unitary; combined with $\hat{P}_{12}^2 = \hat{I}$, this makes it Hermitian as well ($\hat{P}_{12}^\dagger = \hat{P}_{12}^{-1} = \hat{P}_{12}$). An eigenvalue λ of $\hat{P}_{12}$ must then satisfy $\lambda^2 = 1$, so $\lambda = \pm 1$: exchange has exactly two possible eigenvalues, and the mathematics permits no others.

For identical particles, the Hamiltonian cannot contain any term that treats "particle 1" differently from "particle 2," since the two share every intrinsic property and enter every interaction symmetrically. Consequently

$$[\hat{H}, \hat{P}_{12}] = 0, \tag{11.6}$$

so energy eigenstates can always be chosen to be simultaneous eigenstates of $\hat{P}_{12}$. Nature realizes only one of the two eigenvalues for a given species—+1 for every observed boson, -1 for every observed fermion, never a mixture and never an intermediate value. This is the substance behind equation (11.3): it is less an independent postulate about wavefunctions than a statement of which permutation eigenvalue a given kind of particle occupies, a choice forced to be one of only two possibilities by the operator argument just given.

Concept check 11.2

Why does $[\hat{H}, \hat{P}_{12}] = 0$ for identical particles, and what does it let us conclude about energy eigenstates?

Answer

The Hamiltonian of identical particles cannot contain any term that treats "particle 1" differently from "particle 2," since the two share every intrinsic property. Because $\hat{H}$ and $\hat{P}_{12}$ commute, energy eigenstates can be chosen to also be eigenstates of $\hat{P}_{12}$, with eigenvalue +1 (bosons) or -1 (fermions). A given species is always found with one eigenvalue, never a mixture.

11.2 Symmetrized two-particle states

If one particle occupies a and the other b, with orthonormal one-particle states, the allowed states are

$$|\Psi_+\rangle = \frac{|a\rangle_1|b\rangle_2 + |b\rangle_1|a\rangle_2}{\sqrt{2}}, \tag{11.7}$$

$$|\Psi_-\rangle = \frac{|a\rangle_1|b\rangle_2 - |b\rangle_1|a\rangle_2}{\sqrt{2}}. \tag{11.8}$$

The first is bosonic and the second fermionic. Neither factors into a state for particle 1 and a state for particle 2. Exchange symmetry creates correlations even without an interaction potential. Both combinations still live inside the same two-particle tensor-product space built for distinguishable spins in Chapter 5; what is new is that, once the particles are identical, an unsymmetrized product such as $|a\rangle_1|b\rangle_2$ alone is no longer a physically allowed state, because it implicitly—and wrongly—asserts which particle is in which state.

If $a = b$, the antisymmetric expression vanishes. Two identical fermions cannot occupy the same complete one-particle state:

Pauli exclusion principle
No two identical fermions may occupy the same one-particle quantum state.

Any number of bosons may occupy one state. This difference produces radically different collective behavior: it is ultimately why bosonic and fermionic matter organize into such different phases at low temperature, a theme this chapter returns to in Section 11.6.

Example 11.1: position detection

For orbitals $\psi_a(x)$ and $\psi_b(x)$, the joint detection density is

$$|\Psi_\pm(x_1, x_2)|^2 = \frac{1}{2}\left[|\psi_a(x_1)\psi_b(x_2)|^2 + |\psi_b(x_1)\psi_a(x_2)|^2 \pm 2\,\mathrm{Re}\{\psi_a^*(x_1)\psi_b^*(x_2)\psi_b(x_1)\psi_a(x_2)\}\right]. \tag{11.9}$$

The last term is an exchange-interference term. For fermions, the density vanishes at $x_1 = x_2$ when the internal spin states are also identical. This "exchange hole" does not require a repulsive force. The same cross term reappears, with the opposite role, once *energies* rather than mere detection probabilities are computed in Section 11.3: there it becomes the exchange integral responsible for Hund's rule.

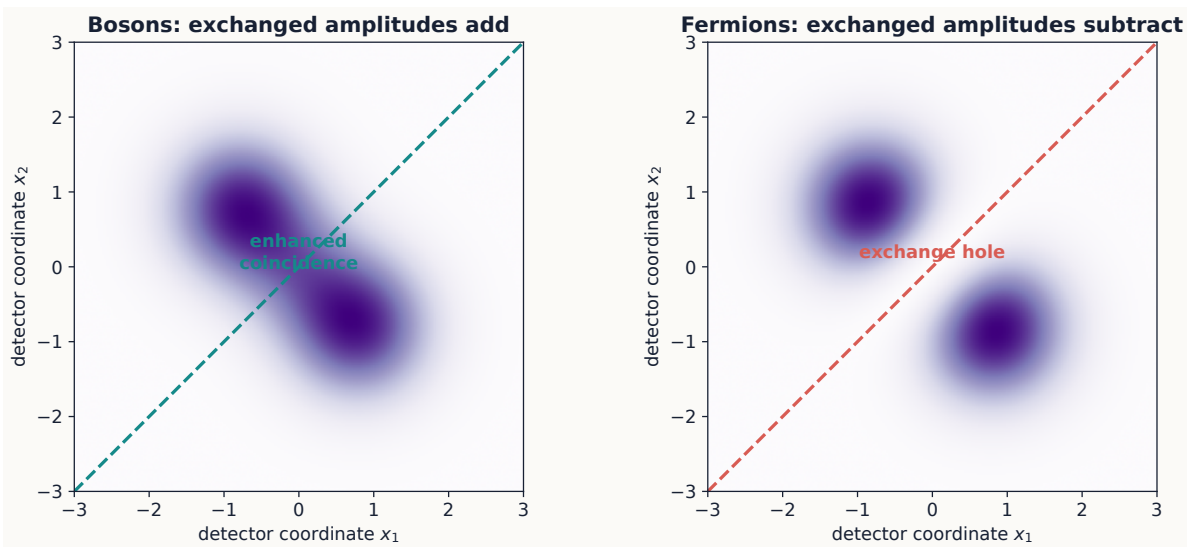

Figure 11.1: Direct and exchanged alternatives are physically indistinguishable, so their amplitudes interfere. The plus sign enhances coincidence for bosons; the minus sign forces the fermionic density to zero when the complete one-particle states coincide.

Concept check 11.3

Does an exchange hole prove that two fermions repel through a new force?

Answer

No. It can occur even for a Hamiltonian with no interaction between the particles. It is a correlation required by antisymmetry.

11.3 Spin and spatial symmetry

The **complete** state—including space and spin—must have the required exchange symmetry. For two electrons it must be antisymmetric. This requirement multiplies rather than adds the constraints on the two factors: it is the product of the spatial part and the spin part, together, that must change sign under exchange, so a symmetric spatial factor demands an antisymmetric spin factor, and an antisymmetric spatial factor demands a symmetric spin factor.

The singlet spin state is antisymmetric:

$$\chi_{00} = \frac{|+z,-z\rangle - |-z,+z\rangle}{\sqrt{2}}. \tag{11.10}$$

It must be multiplied by a symmetric spatial wavefunction. The three triplet spin states are symmetric and must be paired with antisymmetric spatial wavefunctions.

Example 11.2: two electrons in a well

In the lowest-energy configuration, both electrons can occupy the same spatial orbital $u_1(x)$ only if their spin state is the singlet. The spatial product $u_1(x_1)u_1(x_2)$ is symmetric, and symmetric space times antisymmetric spin is antisymmetric overall.

If the spin state is a triplet, the spatial part must be antisymmetric, so one electron must occupy a higher orbital:

$$\Psi_{\text{space}} = \frac{u_1(x_1)u_2(x_2) - u_2(x_1)u_1(x_2)}{\sqrt{2}}. \tag{11.11}$$

Exclusion therefore affects energy even when the Hamiltonian contains no spin-dependent term.

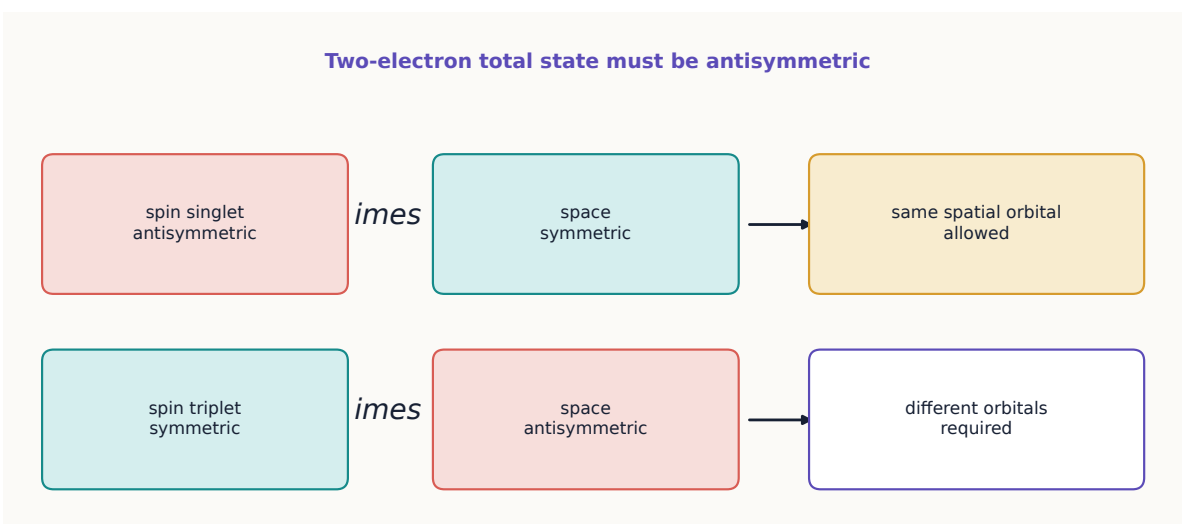

Figure 11.2: Only the symmetry of the complete two-electron state is fixed. Spin and space therefore compensate: the singlet permits shared spatial occupation, whereas the triplet forces an antisymmetric spatial factor with an exchange hole.

Concept check 11.4

Can two electrons occupy the same spatial orbital?

Answer

Yes, if their combined spin state is antisymmetric—the singlet. They do not occupy the same complete spin-orbital.

Direct and exchange energies

Example 11.2 already shows that the singlet and triplet configurations differ in spatial symmetry before any interaction between the electrons is even switched on. Once the electrons repel one another through a Coulomb interaction V, this symmetry difference feeds directly into the energy. Evaluating $\langle V \rangle$ in the symmetric or antisymmetric spatial pair state of equation (11.7) splits the result into a piece that would be present even for distinguishable particles and a piece that survives only because of exchange:

$$E_{\pm} = \langle u_1 u_2 | \hat{V} | u_1 u_2 \rangle \pm \langle u_1 u_2 | \hat{V} | u_2 u_1 \rangle \equiv J \pm K. \tag{11.12}$$

Here J is the **direct** (Coulomb) integral—the ordinary electrostatic repulsion between two classical-looking charge distributions $|u_1|^2$ and $|u_2|^2$—and K is the **exchange** integral, built from the same cross term already encountered in Example 11.1, now weighted by the interaction rather than left as a bare interference term in a detection probability. The $+$ sign belongs to the symmetric spatial state, paired with the antisymmetric spin singlet; the $-$ sign belongs to the antisymmetric spatial state, paired with a symmetric spin triplet.

For a repulsive Coulomb interaction, K is typically positive. The triplet configuration then lies **lower** in energy than the singlet configuration built from the same pair of orbitals, even though the Hamiltonian contains no spin-dependent term at all: the triplet's antisymmetric spatial wavefunction already keeps the two electrons farther apart on average—the exchange hole of Example 11.1—which lowers their mutual repulsion. This is the microscopic origin of **Hund's first rule**: for a given electron configuration, the term with the largest total spin usually has the lowest energy, because parallel spins force the spatial wavefunction to keep electrons apart and reduce their Coulomb repulsion. No new force is required, and none was invented—only the antisymmetry postulate together with an ordinary repulsive interaction.

Not a universal rule

Hund's first rule is a reliable guide for filling atomic subshells, but J and K depend on the specific orbitals involved, and their ranking can differ in other systems—some excited-state configurations and some molecular orbitals reverse it. Treat $E_{\pm} = J \pm K$ as the correct general structure, and check the sign of K case by case rather than assuming it.

Concept check 11.5

Using $E_{\pm} = J \pm K$, explain why the triplet configuration of two electrons in different orbitals is usually lower in energy than the singlet configuration built from the same orbitals.

Answer

The triplet's symmetric spin state forces an antisymmetric spatial wavefunction, which already vanishes when the two electrons coincide—the exchange hole. This keeps the electrons farther apart on average and lowers their mutual Coulomb repulsion, giving the triplet energy

$J - K$, with $K > 0$ typically, below the singlet's $J + K$. No spin-dependent force is needed, only antisymmetry combined with ordinary repulsion.

11.4 Slater determinants

For N identical fermions in orthonormal one-particle spin-orbitals $\chi_1, \ldots, \chi_N$, an antisymmetric state is

$$\Psi(1, \ldots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(1) & \chi_2(1) & \cdots & \chi_N(1) \\ \chi_1(2) & \chi_2(2) & \cdots & \chi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(N) & \chi_2(N) & \cdots & \chi_N(N) \end{vmatrix}. \tag{11.13}$$

Exchanging two particle labels swaps two rows and changes the sign. If two spin-orbitals are identical, two columns coincide and the determinant is zero. Both antisymmetry and exclusion are built into this one object.

A single Slater determinant is often a useful starting approximation for many-electron atoms and molecules. General correlated fermion states require superpositions of determinants.

A determinant is a model, not a postulate

Writing a many-electron state as one Slater determinant—the Hartree–Fock approximation—is a computational choice, not an additional physical law. It exactly enforces antisymmetry, but by construction it cannot capture correlation beyond exchange: the extent to which electrons of unlike spin still avoid each other because of their mutual Coulomb repulsion. Better approximations combine several Slater determinants.

11.5 Occupation-number notation

When particle labels carry no physical meaning, it is often clearer to state how many particles occupy each one-particle mode:

$$|n_1, n_2, n_3, \ldots\rangle. \tag{11.14}$$

For fermions, $n_j = 0$ or 1. For bosons, $n_j = 0, 1, 2, \ldots$. Creation and annihilation operators change occupations. For a bosonic mode,

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \qquad \hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \tag{11.15}$$

with $[\hat{a}, \hat{a}^\dagger] = 1$. Fermionic operators instead anticommute:

$$\{\hat{c}_i, \hat{c}_j^\dagger\} = \delta_{ij}, \qquad \{\hat{c}_i, \hat{c}_j\} = 0. \tag{11.16}$$

The relation $(\hat{c}_i^\dagger)^2 = 0$ is exclusion in operator form.

For each mode, the **number operator**

$$\hat{n}_j = \hat{a}_j^\dagger \hat{a}_j \quad \text{(bosons)}, \qquad \hat{n}_j = \hat{c}_j^\dagger \hat{c}_j \quad \text{(fermions)} \tag{11.17}$$

counts how many particles occupy mode j, since $\hat{n}_j|n_j\rangle = n_j|n_j\rangle$ in either case. Summing over all modes gives the total-number operator $\hat{N} = \sum_j \hat{n}_j$. Section 11.1 argued that "particle

1" and "particle 2" are not physically meaningful labels; occupation-number notation makes that argument precise by never introducing such labels at all. It is $\hat{N}$, not any single-particle identity, that is the well-defined observable.

11.6 Consequences for matter

Electrons fill atomic orbitals subject to exclusion. Including spin, an orbital specified by (n, ℓ, m) holds at most two electrons, with opposite spin labels. Shell structure and the periodic table emerge from this filling together with Coulomb interactions and energy splittings. Within a given subshell, the exchange energy of Section 11.3 further determines how electrons distribute among degenerate orbitals: Hund's rule predicts that electrons first occupy separate orbitals with parallel spins, maximizing the (typically positive) exchange energy K, before any orbital is doubly occupied.

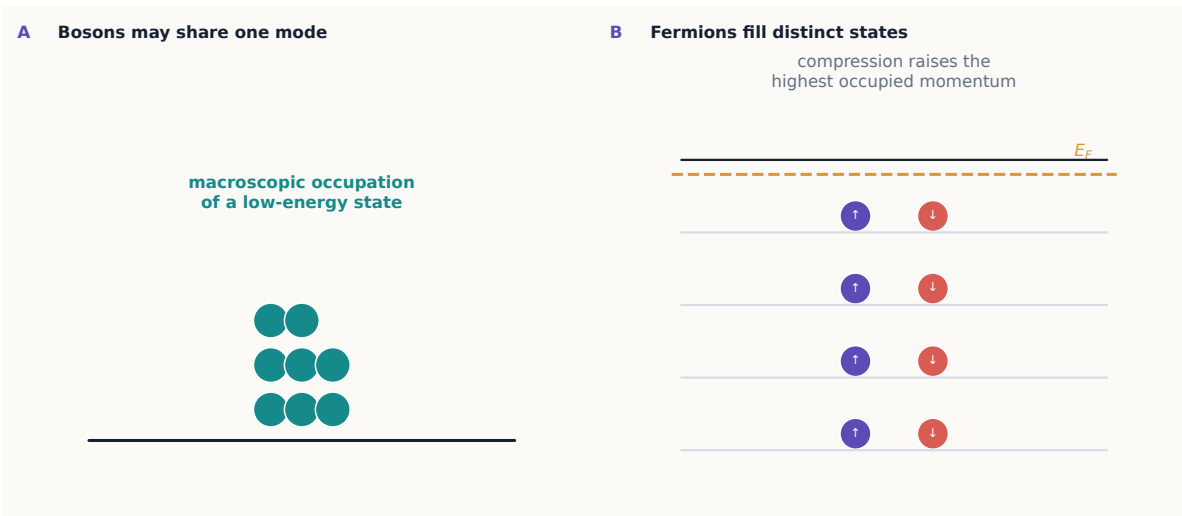

Figure 11.3: Bosonic occupations can become arbitrarily large in one mode. Exclusion makes fermions occupy a ladder of distinct states, creating a nonzero Fermi energy—and therefore pressure—even at zero temperature.

In a macroscopic fermion system, low-energy states fill up to a Fermi energy. Compression forces particles into higher momentum states even at zero temperature, producing degeneracy pressure. For bosons, many particles can accumulate in one mode, enabling coherent phenomena such as Bose–Einstein condensation.

These are statistical effects rooted in state-space symmetry. They are not optional corrections applied only when particles collide; they constrain the allowed state from the start.

Example 11.3: degeneracy pressure holds up dying stars

Consider N spin-$\frac{1}{2}$ fermions of mass m confined to volume V at essentially zero temperature. Exclusion forces them to fill distinct momentum states, two spin orientations per momentum value, up to a Fermi momentum p_F fixed by counting the occupied states:

$$n \equiv \frac{N}{V} = \frac{p_F^3}{3\pi^2\hbar^3} \quad \Longrightarrow \quad p_F = \hbar\,(3\pi^2 n)^{1/3}. \tag{11.18}$$

The corresponding Fermi energy is

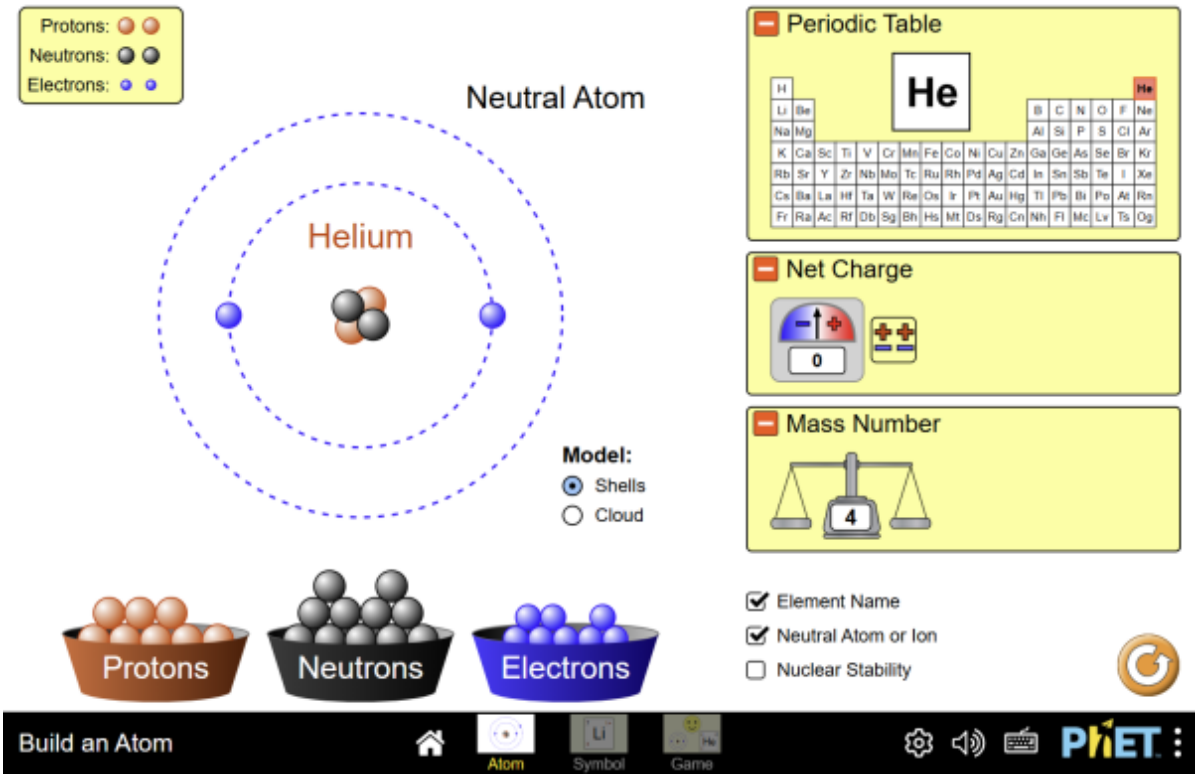

Figure 11.4: Add electrons one at a time and watch them fill shells subject to exclusion, building up the periodic table's structure orbital by orbital. Try filling a subshell with parallel spins before doubling any orbital, and check the result against Hund's rule above. Interactive simulation: Build an Atom

$$E_F = \frac{p_F^2}{2m} = \frac{\hbar^2}{2m}(3\pi^2 n)^{2/3}. \tag{11.19}$$

Crucially, E_F depends only on the density n, not on the temperature: exclusion alone forces most particles into states of substantial kinetic energy, and compressing the gas into a smaller volume raises that energy further. The resulting outward degeneracy pressure exists even at $T = 0$, unlike the pressure of an ordinary classical gas.

For the electron gas inside a white dwarf, a representative density is $n \sim 10^{36}$ m^{-3}, giving $p_F \sim 3 \times 10^{-22}$ kg m/s and $E_F \sim 0.4$ MeV from equation (11.19)—already a sizable fraction of the electron rest energy $m_e c^2 \approx 0.511$ MeV. Electron degeneracy pressure, not thermal pressure, is what balances gravity in a white dwarf. Because the electrons are only mildly nonrelativistic at best, a fully relativistic treatment becomes necessary near the Chandrasekhar mass limit ($\approx$ 1.4 solar masses), beyond which electron degeneracy pressure can no longer support the star; further collapse can produce a neutron star, supported in turn by neutron degeneracy pressure at far higher density. None of this requires any interaction beyond exclusion applied to a very dense, very cold Fermi gas.

Concept check 11.6

Does the degeneracy pressure of Example 11.3 require a nonzero temperature?

Answer

No. Equation (11.19) depends only on the particle density, not on temperature. Exclusion alone forces fermions into high-momentum states even at $T = 0$; an ordinary classical gas, by contrast, would exert zero pressure at zero temperature.

Distinguishable in practice

Two particles of the same species may behave approximately distinguishably when their wave packets remain well separated and no experiment recombines them. Then exchange terms are negligibly small. As wavefunctions overlap, particle-by-particle labels cease to support correct predictions and the symmetrized description becomes essential.

Summary

- Identical-particle labels are not observable, so exchange constrains the joint state.
- Boson states are symmetric and fermion states antisymmetric.
- Antisymmetry implies Pauli exclusion and exchange correlations.
- The permutation operator's eigenvalues are restricted to ± 1; since $[\hat{H}, \hat{P}_{12}] = 0$ for identical particles, a given species occupies one eigenvalue exclusively, fixing it as a boson or a fermion.
- For electrons, spatial and spin symmetries combine to give an antisymmetric complete state.
- Direct and exchange energies split singlet and triplet configurations oppositely, and the sign of the exchange integral underlies Hund's rule.
- Slater determinants enforce fermionic antisymmetry for many particles.
- Occupation numbers replace unphysical particle labels and lead naturally to creation, annihilation, and number operators.
- Exclusion alone produces a temperature-independent degeneracy pressure that stabilizes white dwarfs and neutron stars against gravitational collapse.

Exercises

1. Normalize the symmetric and antisymmetric states when a and b are orthogonal.
2. Apply the exchange operator twice and explain why its eigenvalues must be ± 1.
3. Show that the antisymmetric two-particle spatial wavefunction vanishes when $x_1 = x_2$.
4. Construct all allowed two-electron states when only one spatial orbital is available.
5. For two spatial orbitals a, b, list the singlet and triplet possibilities and their required spatial symmetries.
6. Verify that a two-particle Slater determinant changes sign under exchange.
7. Explain why a determinant vanishes if two occupied spin-orbitals coincide.
8. Write occupation-number states corresponding to two bosons in three modes and to two fermions in three modes.
9. Derive $(\hat{c}_i^\dagger)^2 = 0$ from the fermionic anticommutation rule.
10. Two noninteracting fermions occupy a one-dimensional infinite well. Compare the lowest total energies for a spin singlet and a spin triplet.
11. Explain when two identical particles can be treated approximately as distinguishable and how that approximation could fail.
12. Distinguish exchange correlation, Coulomb repulsion, and entanglement in a two-electron state.
13. Show directly from $\hat{P}_{12}\Psi(1,2) = \Psi(2,1)$ that $\hat{P}_{12}^2 = \hat{I}$, and use this to show that its eigenvalues must be ± 1.
14. For $\hat{H} = h(1) + h(2) + V(|\mathbf{r}_1 - \mathbf{r}_2|)$, where h is the same one-particle Hamiltonian for both particles, show explicitly that $[\hat{H}, \hat{P}_{12}] = 0$.
15. Using the cross term identified in Example 11.1, explain why it enters the two-particle detection density with a plus sign for the symmetric (bosonic) spatial state and a minus sign for the antisymmetric (fermionic) spatial state, and connect this sign difference to the exchange integral K in equation (11.12).
16. Rank the singlet and triplet configurations of two electrons in orbitals $a \neq b$ from lowest to highest energy, assuming $K > 0$, and state which configuration Hund's rule predicts for the corresponding real atomic ground state.
17. Show that $\hat{n}_j|n_j\rangle = n_j|n_j\rangle$ follows from the fermionic or bosonic operator algebra, and explain why $\hat{N} = \sum_j \hat{n}_j$ is a meaningful observable even though no individual particle carries a meaningful label.
18. Estimate the electron Fermi energy of a white dwarf with electron number density 10^{36} m^{-3} using equation (11.19), and compare your result with the electron rest energy $m_e c^2$.

19. Explain why increasing the density of a fermion gas increases its degeneracy pressure even at $T = 0$, while increasing the pressure of an ordinary classical gas at fixed density requires raising its temperature.

20. For three identical fermions distributed among four available spin-orbitals, how many distinct Slater determinant states can be formed?

21. Using the exchange-hole picture of Example 11.1, explain why two identical fermions with the same spin state are never found at exactly the same location, while two identical fermions with opposite spin states can be.

22. A student claims that antisymmetrization matters only for computing energies, and that the positions of noninteracting identical fermions are otherwise statistically independent. Use Example 11.1 to show that this claim is false even with no interaction present.

23. In helium, the ground configuration $1s^2$ admits only a spin singlet, but the excited configuration $1s2s$ admits both a singlet and a triplet term, with the triplet lying lower in energy (orthohelium below parahelium). Explain both halves of this statement using the results of Section 11.3.

24. In the Slater determinant of equation (11.13), exclusion is enforced by two *columns* coinciding, not two rows. Explain what each row and each column of the determinant represents, and why it is specifically the column structure, and not the row structure, that encodes the Pauli principle.

Chapter 12

Modern Quantum Applications

12.1 Modern quantum applications

Learning objectives

After this chapter, you should be able to:

- represent quantum logic operations as unitary gates;
- explain no-cloning, teleportation, and superdense coding;
- distinguish decoherence from measurement uncertainty;
- state the purpose of quantum error correction;
- connect two-state dynamics to clocks, magnetic resonance, and sensing;
- illustrate how phase kickback lets a quantum algorithm answer a global question about a function with fewer queries than any classical algorithm (Deutsch's algorithm); and
- evaluate quantum-technology claims in terms of states, transformations, measurements, and resources.

12.1 From spin states to qubits

A **qubit** is any controlled two-state quantum system:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \qquad |\alpha|^2 + |\beta|^2 = 1. \tag{12.1}$$

The physical realization might use spin, atomic energy levels, photon polarization, superconducting circuits, or another pair of states. Different hardware has different control and noise, but the same two-dimensional formalism applies.

Common one-qubit gates include

$$\hat{X} = \sigma_x, \quad \hat{Z} = \sigma_z, \quad \hat{H}_{\mathrm{q}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \hat{S} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \tag{12.2}$$

They are unitary transformations, not measurements. The X gate exchanges $|0\rangle$ and $|1\rangle$; Z changes a relative sign; Hadamard converts between z and x bases; and S adds a quarter-cycle phase.

12.2 Circuits and quantum information

A circuit diagram is read in time order. Gates compose as matrix products, and measurement converts final amplitudes into classical records. A useful calculation always specifies the input,

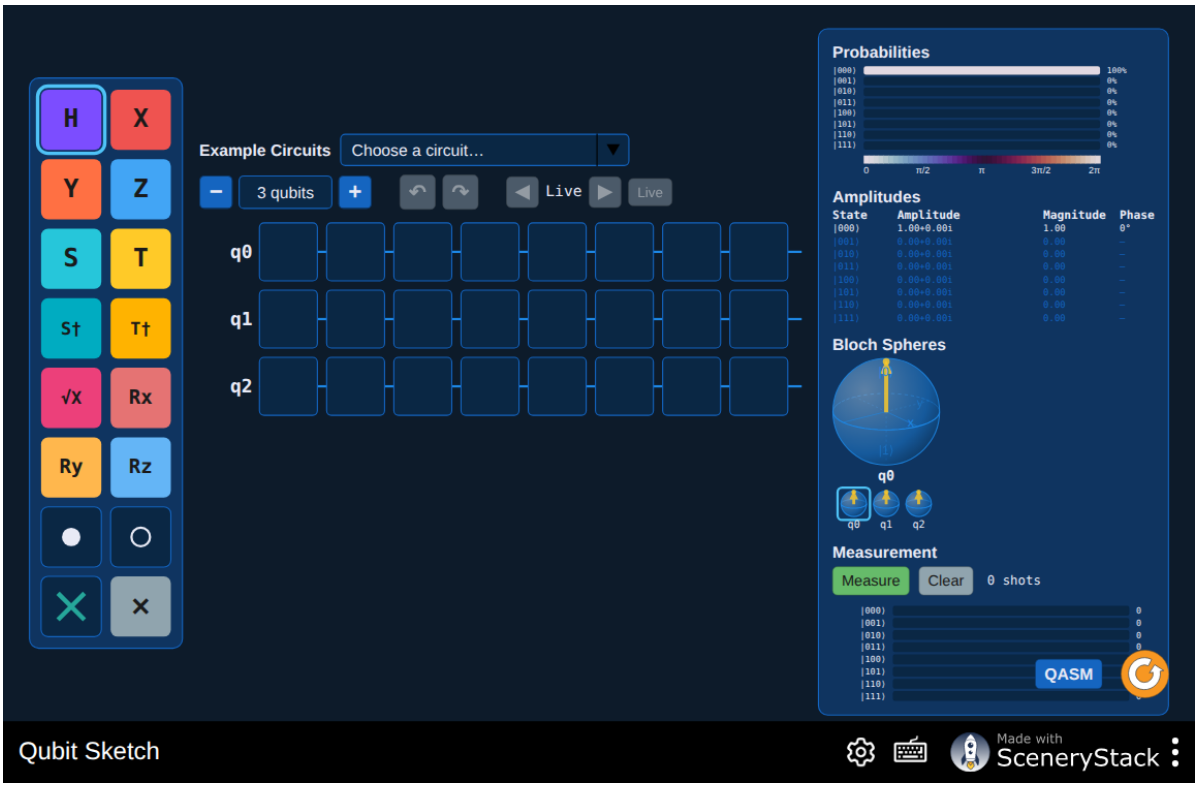

Figure 12.1: Drag X, Z, H, and S gates onto a single qubit and watch its Bloch sphere and amplitudes α, β update after each one—the same circuit builder used in Chapter 5 to entangle two qubits, now applied one gate at a time to a single qubit's state. Interactive simulation: Qubit Sketch

applies gates in order, and projects in the stated measurement basis.

The no-cloning theorem

Suppose one unitary $\hat{U}$ could copy every unknown state using a single fixed, fresh ancilla $|0\rangle$:

$$\hat{U}|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle, \qquad \hat{U}|\phi\rangle|0\rangle = |\phi\rangle|\phi\rangle. \tag{12.3}$$

Unitary evolution preserves inner products, so the overlap $\langle\phi|\langle 0|\hat{U}^\dagger\hat{U}|\psi\rangle|0\rangle$ can be evaluated two ways. Using $\hat{U}^\dagger\hat{U} = \hat{I}$ gives one value; substituting the cloning hypothesis gives another:

$$\langle\phi|\langle 0|\hat{U}^\dagger\hat{U}|\psi\rangle|0\rangle = \langle\phi|\psi\rangle, \qquad \langle\phi|\langle 0|\hat{U}^\dagger\hat{U}|\psi\rangle|0\rangle = \big(\langle\phi|\langle\phi|\big)\big(|\psi\rangle|\psi\rangle\big) = (\langle\phi|\psi\rangle)^2. \tag{12.4}$$

Both expressions describe the same quantity, so writing $x = \langle\phi|\psi\rangle$ turns equation (12.4) into $x = x^2$, or $x(x-1) = 0$. Only $x = 0$ or $x = 1$ solves this: a single fixed $\hat{U}$ can reproduce the cloning relation for a pair of states that are already orthogonal or already identical up to normalization, never for a generic unknown pair with $0 < |x| < 1$. A device can copy a known orthogonal basis, such as $|0\rangle, |1\rangle$, but no physical operation perfectly clones an arbitrary unknown qubit.

No-cloning is not a technological inconvenience. It follows from linearity alone and protects the distinction between quantum information and an unknown classical description: if an unknown qubit could always be silently duplicated, no protocol whose security or correctness depends on that impossibility could work. It is also why the teleportation protocol below must destroy the sender's local access to $|\psi\rangle$ as the price of reconstructing it elsewhere—transmitting quantum information can never mean copying it.

Concept check 12.1

Why can a controlled-NOT gate copy $|0\rangle$ and $|1\rangle$ without violating no-cloning?

Answer

No-cloning forbids copying every unknown state. CNOT copies the selected orthogonal basis, but an input superposition produces entanglement rather than two independent copies.

12.3 Teleportation

Quantum teleportation transfers an unknown qubit state using shared entanglement and two classical bits. Let qubit 1 contain $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, while qubits 2 and 3 share $|\Phi^+\rangle$. Reexpressing the three-qubit state in the Bell basis of qubits 1 and 2 gives

$$|\psi\rangle_1|\Phi^+\rangle_{23} = \frac{1}{2}\big[|\Phi^+\rangle_{12}|\psi\rangle_3 + |\Phi^-\rangle_{12}\hat{Z}|\psi\rangle_3 + |\Psi^+\rangle_{12}\hat{X}|\psi\rangle_3 + |\Psi^-\rangle_{12}\hat{X}\hat{Z}|\psi\rangle_3\big]. \tag{12.5}$$

A Bell-basis measurement on 1 and 2 gives one of four random outcomes. Sending its two-bit label allows the holder of qubit 3 to apply the corresponding correction and recover $|\psi\rangle$.

The original qubit is not copied: the measurement destroys its prior local state, consistent with no-cloning. The receiver cannot recover $|\psi\rangle$ before the classical message arrives, so teleportation does not send information faster than light.

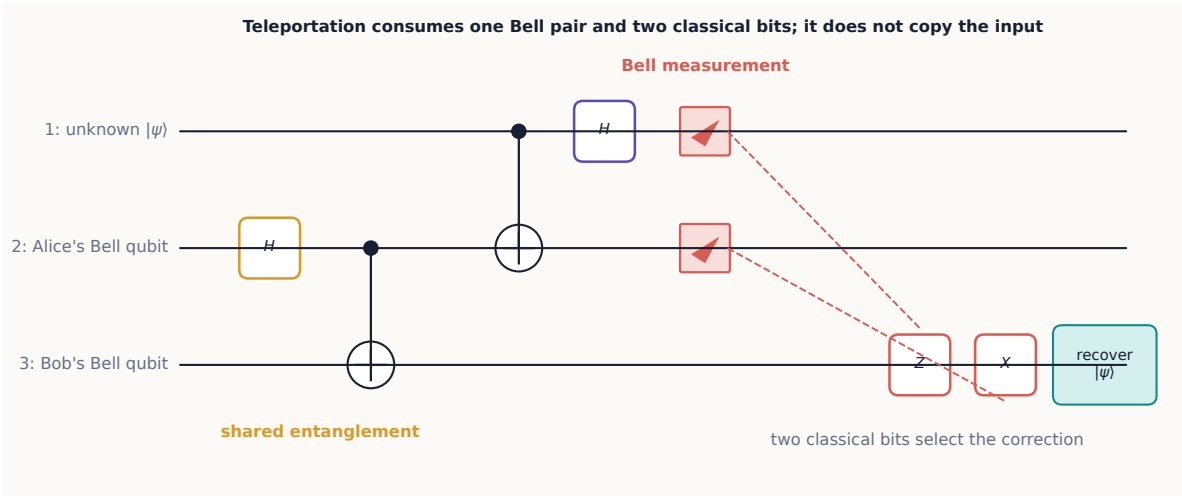

Figure 12.2: Teleportation relocates a preparation by combining a shared Bell pair with two classical bits. The Bell measurement removes the sender's local copy, and the receiver obtains the state only after the outcome-dependent correction.

Concept check 12.2

Does teleportation eliminate the need to transmit information?

Answer

No. It consumes shared entanglement and requires two ordinary classical bits. Without them, the receiver's local state contains no accessible copy of the unknown input.

Example 12.1: a concrete correction

Let $|\psi\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$, so $\alpha = 1/\sqrt{2}$ and $\beta = i/\sqrt{2}$ in equation (12.5). Suppose the Bell measurement on qubits 1 and 2 returns the outcome $|\Psi^-\rangle_{12}$. Immediately afterward—before any correction is applied—qubit 3 is left in

$$\hat{X}\hat{Z}|\psi\rangle = \hat{X}\left(\frac{|0\rangle - i|1\rangle}{\sqrt{2}}\right) = \frac{-i|0\rangle + |1\rangle}{\sqrt{2}}. \tag{12.6}$$

This is not $|\psi\rangle$, and by itself it carries no information about which of the four outcomes occurred at the sender. The two classical bits naming $|\Psi^-\rangle_{12}$ tell the receiver to undo $\hat{X}\hat{Z}$. Because $\hat{X}$ and $\hat{Z}$ do not commute, the order of the correction matters: applying $\hat{X}$ and then $\hat{Z}$,

$$\hat{Z}\hat{X}\left(\frac{-i|0\rangle + |1\rangle}{\sqrt{2}}\right) = \hat{Z}\left(\frac{|0\rangle - i|1\rangle}{\sqrt{2}}\right) = \frac{|0\rangle + i|1\rangle}{\sqrt{2}} = |\psi\rangle, \tag{12.7}$$

exactly recovers the original state. Applying the same two corrections in the opposite order returns $-|\psi\rangle$ instead—physically indistinguishable from $|\psi\rangle$, since an overall phase carries no observable consequence (Chapter 2), but not identical as a vector, so a careful statement of the protocol fixes one convention. In the vocabulary that will organize the rest of this chapter (Section 12.9), the *preparation* reconstructed at qubit 3 depends on both the shared entangled resource and the classical record of the *measurement* at the sender; neither the entanglement alone nor the two bits alone would be evidence of a transferred qubit.

Concept check 12.3

In Example 12.1, does the sender end up holding a usable copy of $|\psi\rangle$ after the Bell measurement?

Answer

No. The Bell measurement projects qubits 1 and 2 into an entangled two-qubit outcome state, leaving qubit 1 with no local state that recovers $|\psi\rangle$. If the sender also retained a usable copy while qubit 3 ended up with $|\psi\rangle$, two independent copies of an unknown state would exist, contradicting no-cloning.

Superdense coding

The resource accounting can be reversed. If two parties share a Bell pair, one party can apply one of $\hat{I}, \hat{X}, \hat{Z}, \hat{X}\hat{Z}$ to encode two classical bits in the joint Bell state, then transmit one qubit. A Bell measurement recovers the two-bit message. The protocol consumes prior entanglement plus one transmitted qubit; it does not encode arbitrarily many bits in one unassisted two-state system.

12.4 Decoherence and open systems

Useful quantum behavior depends on controlled relative phases. Unmonitored degrees of freedom can become correlated with a system:

$$(\alpha|0\rangle + \beta|1\rangle)|E_0\rangle \longrightarrow \alpha|0\rangle|E_0'\rangle + \beta|1\rangle|E_1'\rangle. \tag{12.8}$$

If the environment states become distinguishable and are not observed, tracing them out suppresses the qubit's off-diagonal density-matrix elements. This is **decoherence**. It explains the loss of interference in an open subsystem without requiring a person to inspect the environment.

Two common time scales are:

- T_1, energy relaxation toward thermal populations; and
- T_2, loss of phase coherence.

Typically $T_2 \leq 2T_1$, and additional low-frequency noise can make it much shorter. Gate times must be small compared with relevant coherence times, but that ratio alone does not fully characterize correlated errors, leakage, or measurement faults.

Example 12.2: comparing a gate time with coherence times

A two-qubit gate takes $t_{\text{gate}} = 200$ ns on a qubit with $T_1 = 50$ μs and $T_2 = 30$ μs. Modeling the decay as simple exponentials e^{-t/T_1} and e^{-t/T_2},

$$e^{-t_{\text{gate}}/T_1} = e^{-0.2/50} \approx 0.9960, \qquad e^{-t_{\text{gate}}/T_2} = e^{-0.2/30} \approx 0.9934. \tag{12.9}$$

A single gate loses under half a percent of population and about two-thirds of a percent of coherence—consistent with the qualitative requirement above. But small per-gate losses compound over a circuit: after 1,000 such gates run back to back, the coherence factor is roughly $e^{-1000\times 0.00667} \approx e^{-6.7} \approx 0.001$, which would destroy essentially all interference long before T_2 is reached in absolute time. A short ratio t_{gate}/T_2 is necessary for one gate to be useful, but a circuit of any depth needs many gates, which is exactly why quantum error correction—not coherence time by itself—is needed to run long computations. In the language

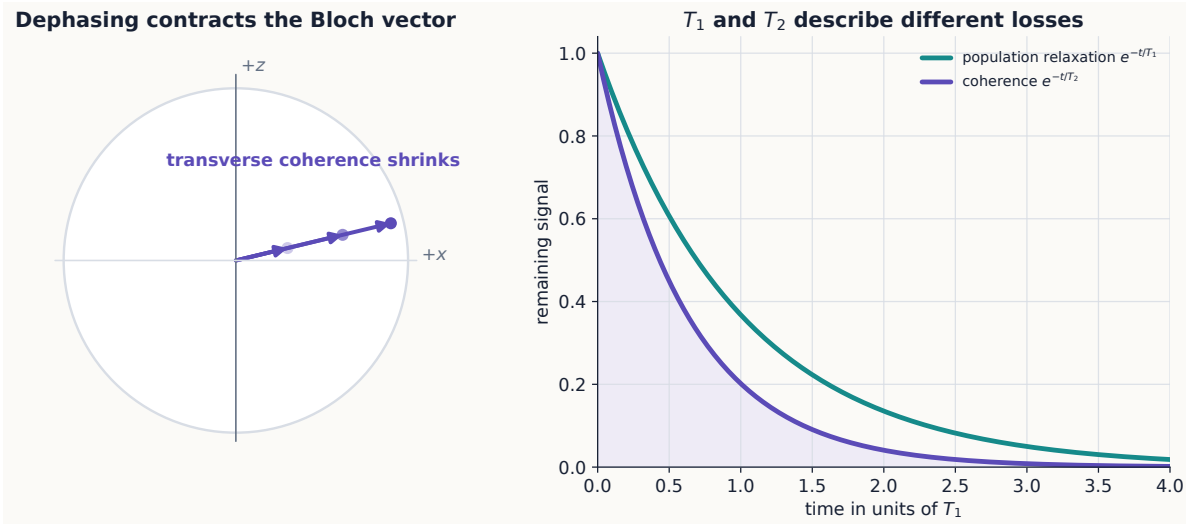

Figure 12.3: Decoherence turns a sharp Bloch vector into a shorter mixed-state vector. T_1 tracks population relaxation, while T_2 tracks transverse phase coherence; a device must characterize both rather than quote one lifetime.

of Section 12.9, decoherence is a *transformation* imposed by an uncontrolled environment rather than a chosen one, and its accumulated effect over a circuit is part of the *evidence* that must be checked against any performance claim.

12.5 Quantum error correction

Classical repetition cannot be applied by first cloning an unknown qubit. Quantum error correction instead encodes one logical state into an entangled subspace of several physical qubits. Carefully chosen **syndrome measurements** reveal which error occurred without revealing the logical amplitudes α, β.

For example, the three-qubit repetition code maps

$$\alpha|0\rangle + \beta|1\rangle \longmapsto \alpha|000\rangle + \beta|111\rangle. \tag{12.10}$$

Parity checks can locate one bit-flip error, after which a correction restores the encoded state. This code does not correct phase flips; more complete codes combine checks for both error types.

Example 12.3: locating an error without reading the logical state

Take the two stabilizer checks $\hat{Z}_1\hat{Z}_2$ and $\hat{Z}_2\hat{Z}_3$, each with eigenvalues ± 1. On an uncorrupted codeword $\alpha|000\rangle+\beta|111\rangle$, both checks return $+1$ regardless of α, β: $\hat{Z}_1\hat{Z}_2|000\rangle = (+1)(+1)|000\rangle$ and $\hat{Z}_1\hat{Z}_2|111\rangle = (-1)(-1)|111\rangle$, both equal to $+1$ times the state, and $\hat{Z}_2\hat{Z}_3$ behaves the same way.

A bit-flip $\hat{X}_1$ sends the codeword to $\alpha|100\rangle + \beta|011\rangle$. Direct evaluation gives $\hat{Z}_1\hat{Z}_2|100\rangle = (-1)(+1)|100\rangle$ and $\hat{Z}_1\hat{Z}_2|011\rangle = (+1)(-1)|011\rangle$: both terms agree on the eigenvalue -1, so the syndrome is well defined even though α, β are unknown and different for the two terms. Repeating this check for every single bit flip gives

error	$\hat{Z}_1\hat{Z}_2$	$\hat{Z}_2\hat{Z}_3$
none	+1	+1
$\hat{X}_1$	-1	+1
$\hat{X}_2$	-1	-1
$\hat{X}_3$	+1	-1

Each possibility gives a distinct pair of eigenvalues, so measuring only these two joint operators—never $\hat{Z}_1$, $\hat{Z}_2$, or $\hat{Z}_3$ individually—identifies which qubit, if any, flipped. The syndrome takes the same value on the $\alpha|000\rangle$ term and the $\beta|111\rangle$ term of a corrupted codeword, so this measurement collapses only the error information and leaves the encoded superposition of α and β untouched. Measuring $\hat{Z}_1$ alone, by contrast, would project each qubit individually onto $|0\rangle$ or $|1\rangle$ and destroy the logical superposition outright. This is a measurement engineered to reveal only the *transformation* that occurred (which error, if any) while leaving the encoded *preparation* untouched—the same operational split between reading a result and merely narrowing a preparation that organized Chapter 1.

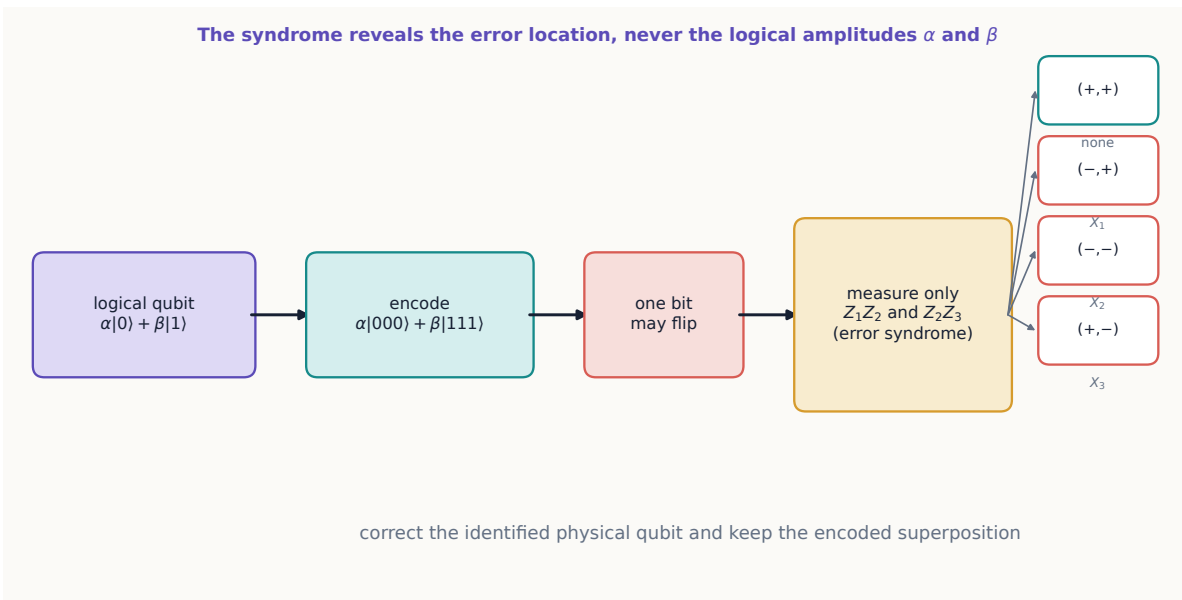

Figure 12.4: The repetition code stores logical information nonlocally. Two joint parity checks distinguish the four error locations but take the same values on both terms of the encoded superposition, so they reveal no information about α or β.

Concept check 12.4

In Example 12.3, a measured syndrome is $(-1, -1)$. Which physical qubit has a bit flip, and has this measurement revealed anything about α or β?

Answer

Qubit 2. The table shows $(-1, -1)$ occurs only for the $\hat{X}_2$ row. The amplitudes are not revealed: the same syndrome occurs for the $\alpha|000\rangle$ term and the $\beta|111\rangle$ term of the corrupted codeword, so the measurement distinguishes only which error occurred, not the encoded superposition.

Fault tolerance is the architecture that prevents imperfect correction operations from spreading errors uncontrollably. A scalable device needs error rates, connectivity, control, decoding, and overhead that together support logical operations—not merely a large physical qubit count.

Concept check 12.5

Is a larger number of physical qubits alone evidence of a more capable quantum computer?

Answer

No. Fidelity, connectivity, coherence, leakage, control speed, measurement, error correction, and algorithmic overhead all affect useful logical capability.

12.6 Quantum computation

A quantum algorithm uses interference so that amplitudes of unwanted answers cancel while amplitudes of useful outcomes reinforce. Superposition alone does not provide all answers to a problem: a final measurement returns one classical record, and the circuit must arrange that useful records have high probability.

Entangling gates plus arbitrary one-qubit rotations form a universal control set. Computational advantage is problem-dependent. Known examples exploit structure in integer factoring, quantum simulation, and some search or linear-algebra tasks; they do not imply a speedup for every computation.

When assessing an algorithmic claim, ask:

1. What input is loaded, and at what cost?
2. Which operations are assumed available?
3. How do time, qubit count, precision, and error-correction overhead scale?
4. What classical baseline is used?
5. What measurement produces the promised answer?

Example 12.4: Deutsch's algorithm

Let $f : \{0,1\} \to \{0,1\}$ be **constant** ($f(0) = f(1)$) or **balanced** ($f(0) \neq f(1)$), with access to f only through the oracle $\hat{U}_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$. Classically, deciding which case holds requires evaluating f at both inputs. One query suffices quantum mechanically.

Prepare $|0\rangle|1\rangle$ and apply $\hat{H}_\mathrm{q} \otimes \hat{H}_\mathrm{q}$:

$$|0\rangle|1\rangle \longrightarrow \frac{|0\rangle + |1\rangle}{\sqrt{2}} \cdot \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{12.11}$$

Because the second qubit is $(|0\rangle - |1\rangle)/\sqrt{2}$, applying $\hat{U}_f$ to one term $|x\rangle(|0\rangle - |1\rangle)/\sqrt{2}$ reproduces the same two-term state up to an overall sign:

$$\hat{U}_f\,|x\rangle\frac{|0\rangle - |1\rangle}{\sqrt{2}} = |x\rangle\frac{|f(x)\rangle - |1 \oplus f(x)\rangle}{\sqrt{2}} = (-1)^{f(x)}|x\rangle\frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{12.12}$$

This is **phase kickback**: the oracle leaves the target qubit unchanged and writes $f(x)$ into a sign attached to $|x\rangle$ instead of into the target qubit's state. Applying it to both terms gives

$$\frac{(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle}{\sqrt{2}} \cdot \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{12.13}$$

If f is constant, the first factor in equation (12.13) is $(-1)^{f(0)}(|0\rangle + |1\rangle)/\sqrt{2}$, an overall phase times $\hat{H}_{\mathrm{q}}|0\rangle$; a final $\hat{H}_{\mathrm{q}}$ on the first qubit returns $|0\rangle$ with certainty. If f is balanced, the first factor is $\pm(|0\rangle - |1\rangle)/\sqrt{2}$, an overall phase times $\hat{H}_{\mathrm{q}}|1\rangle$; the final $\hat{H}_{\mathrm{q}}$ returns $|1\rangle$ with certainty. Measuring the first qubit therefore reveals whether f is constant or balanced—a single global property—after exactly one query to $\hat{U}_f$, even though neither $f(0)$ nor $f(1)$ separately becomes known. That one measured bit is direct *evidence* distinguishing the two hypotheses, not merely a result consistent with both. Nothing here followed one branch more than the other with higher probability; the two branches instead interfered constructively for one final measurement outcome and destructively for the other, exactly the mechanism named at the start of this section.

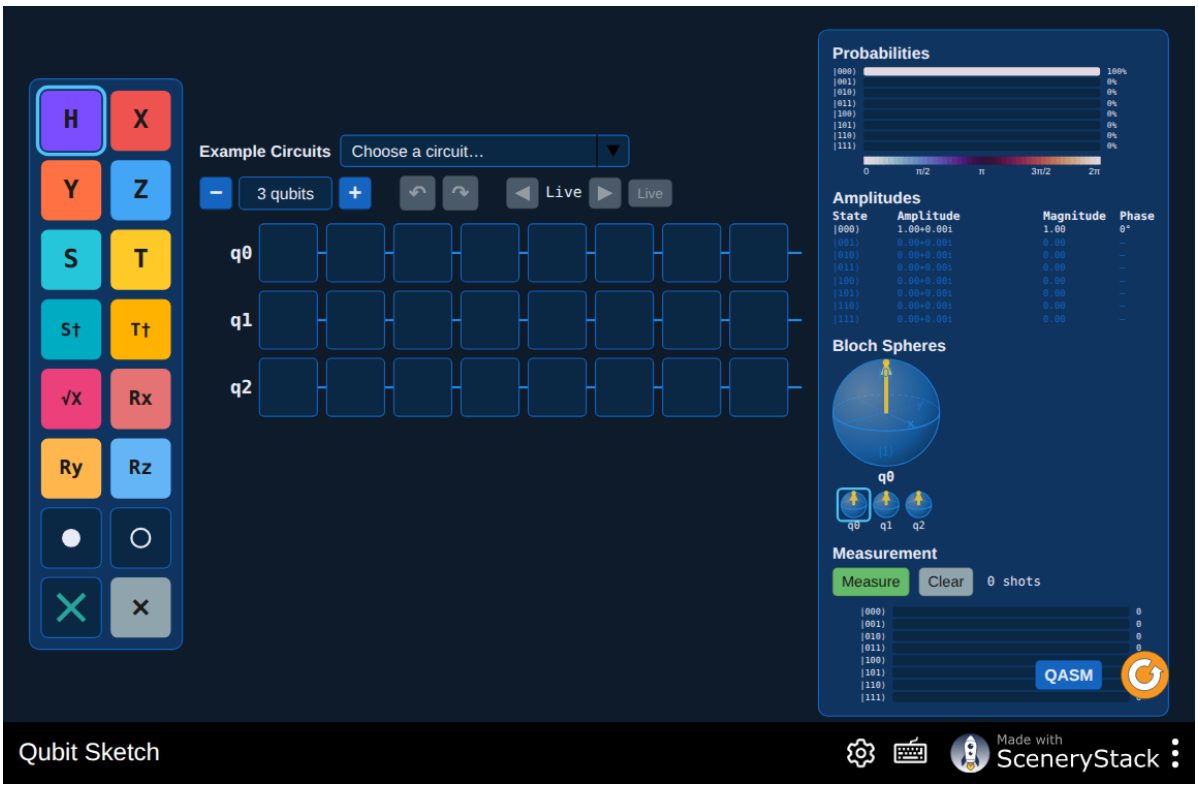

Figure 12.5: Build this circuit on two wires: Hadamard both qubits, insert an oracle, then Hadamard the first qubit again before measuring it. All four possible oracles are built from gates already in the palette—identity or X on the target qubit for a constant f, and CNOT (with an extra X for the other balanced case) for a balanced f. Confirm the first qubit always reads 0 for the constant oracles and 1 for the balanced ones. Interactive simulation: Qubit Sketch

Concept check 12.6

After running Deutsch's algorithm in Example 12.4 and measuring the first qubit, do we know the value of $f(0)$?

Answer

No. Only the global constant-or-balanced property is determined. The individual values $f(0)$ and $f(1)$ remain unknown from that one run—the algorithm answers a question about the function as a whole, not about any particular input.

12.7 Clocks, resonance, and sensing

The Ramsey sequence of Chapter 4 compares a controlled oscillator with a quantum transition frequency. Feedback steers the oscillator toward the center of the fringe, forming the basis of atomic clocks.

Magnetic resonance uses precession and pulses to infer magnetic fields and local environments. In nuclear magnetic resonance and magnetic-resonance imaging, spatially varying fields connect resonance frequency and phase to position. The measurable signal is an ensemble response, but its building blocks are the rotations, phases, and projections developed for one spin.

A two-state sensor accumulates phase

$$\phi = \int_0^T \gamma B(t)\, dt. \tag{12.14}$$

A final analyzer converts ϕ into a probability. Sensitivity improves with phase-accumulation time until decoherence reduces contrast. Pulse sequences can reject slowly varying unwanted fields or select a narrow frequency band.

Example 12.5: projection noise

If a binary readout has probability p, N independent trials estimate it with standard deviation

$$\sigma_p = \sqrt{\frac{p(1-p)}{N}}. \tag{12.15}$$

Near the steepest point of an interference fringe, probability uncertainty becomes phase uncertainty. Repeating four times as many independent trials improves this standard quantum scaling by a factor of two. Entangled sensor states can alter the scaling under ideal conditions, but state preparation, loss, and decoherence determine whether the advantage survives.

12.8 Quantum simulation

Quantum systems are often difficult to simulate classically because the state space dimension grows exponentially with subsystem count. A controllable quantum device can encode another Hamiltonian and measure its dynamics.

Two broad strategies are:

- **digital simulation**, which decomposes evolution into gates; and
- **analog simulation**, which engineers a physical Hamiltonian with the same relevant structure as the target.

Validation remains essential. Conservation laws, small instances accessible to classical calculation, symmetry checks, and measurements in multiple bases help distinguish target physics from control error.

12.9 Reading claims with the textbook's framework

Every quantum technology can be analyzed with four questions:

1. **Preparation:** Which density operator is actually produced?
2. **Transformation:** Which Hamiltonian, channel, or gate acts, and for how long?
3. **Measurement:** Which observable or generalized detector record is read?
4. **Evidence:** Which statistics distinguish the quantum model from noise or a classical alternative?

Terms such as *superposition*, *entanglement*, and *quantum advantage* name specific mathematical and experimental resources. They should lead to testable probability distributions, not serve as substitutes for a mechanism.

Summary

- Qubits are physical two-state systems controlled with unitary gates and projective measurements.
- Linearity forbids universal cloning of an unknown state.
- Entanglement plus classical communication enables teleportation and dense coding without faster-than-light signaling.
- Decoherence arises when uncontrolled correlations hide phase information from a subsystem.
- Error correction protects logical amplitudes by measuring error syndromes that reveal which error occurred without collapsing the encoded state.
- Phase kickback lets an oracle mark a global property of a function in the phase of a query register, letting interference answer a global question with fewer queries than any classical algorithm, as in Deutsch's algorithm.
- Computation, sensing, clocks, resonance, and simulation all use controlled preparation, phase evolution, interference, and readout.

Exercises

1. Find the action of $\hat{H}_{\mathrm{q}}$, $\hat{S}$, and $\hat{X}$ on each computational basis state and on $|+x\rangle$.
2. Show that CNOT applied to $(\alpha|0\rangle + \beta|1\rangle)|0\rangle$ does not produce two copies unless the first state is a computational-basis state.
3. Derive the teleportation identity by expanding both sides in the computational basis.
4. Make a table of the four Bell-measurement outcomes and the receiver's correcting gate.
5. Work through superdense coding for all four two-bit messages.
6. Trace out the environment in the decoherence model and show that the off-diagonal terms are multiplied by $\langle E_1'|E_0'\rangle$.
7. Show how a single bit flip affects each codeword of the three-qubit repetition code and propose parity checks that locate it.

8. Explain why measuring every physical qubit directly would destroy the logical superposition even though syndrome measurements need not.

9. For $P = (1 + \cos\phi)/2$, find $|dP/d\phi|$ and identify its steepest operating points.

10. A sensor has contrast C, so $P = [1 + C\cos(\gamma BT)]/2$. Explain how C and T affect small-field sensitivity.

11. Compare the resources consumed by teleportation, superdense coding, and direct qubit transmission.

12. Choose one quantum-technology claim and analyze it using the four questions in Section 12.9.

13. Suppose a unitary $\hat{U}$ satisfies the cloning hypothesis for two states with $\langle\phi|\psi\rangle = 1/\sqrt{2}$. Show that equation (12.4) requires $1/\sqrt{2} = 1/2$, and state which assumption fails.

14. Redo Example 12.1 for the Bell-measurement outcome $|\Phi^-\rangle_{12}$ instead of $|\Psi^-\rangle_{12}$: find qubit 3's state before correction and the operator order that recovers $|\psi\rangle$ exactly.

15. For $T_1 = 80$ μs and $T_2 = 45$ μs, find the number of 150 ns gates, run back to back, after which the coherence factor e^{-t/T_2} first falls below 0.5.

16. Verify directly, as in Example 12.3, that an $\hat{X}_3$ error gives the syndrome $(+1, -1)$.

17. A phase-flip error $\hat{Z}_2$ acts on $\alpha|000\rangle + \beta|111\rangle$. Show that both stabilizers $\hat{Z}_1\hat{Z}_2$ and $\hat{Z}_2\hat{Z}_3$ still return $+1$, and explain what this implies about the limits of the bit-flip code stated in the text.

18. Work through the phase-kickback calculation of Example 12.4 explicitly for the balanced assignment $f(0) = 1, f(1) = 0$, tracking the overall phase, and confirm that the first qubit is still measured to be 1.

19. Suppose the target qubit in Deutsch's algorithm were prepared in $|0\rangle$ instead of $|1\rangle$, so no $\hat{H}_{\mathrm{q}}|1\rangle$ state is created. Show that $\hat{U}_f$ then leaves the target entangled with the first qubit's basis states, and explain why the final measurement on the first qubit no longer reliably distinguishes constant from balanced f.

20. The Deutsch–Jozsa algorithm generalizes Example 12.4 to $f : \{0,1\}^n \to \{0,1\}$, promised to be constant or balanced, distinguished with one query using n query qubits and the same target-qubit trick. Explain qualitatively, without a full derivation, why the phase-kickback mechanism still applies with n query qubits and why a classical deterministic algorithm can require up to $2^{n-1} + 1$ queries in the worst case.

21. Using the orthonormality of the Bell basis and of $\{|\psi\rangle, \hat{Z}|\psi\rangle, \hat{X}|\psi\rangle, \hat{X}\hat{Z}|\psi\rangle\}$, verify that the four terms on the right of equation (12.5) are mutually orthonormal, and explain why this is required for the identity to represent a valid change of basis.

22. A vendor claims a 1,000-physical-qubit device demonstrates "exponential quantum advantage" on an optimization problem. Apply the four questions of Section 12.9 to list the specific information needed before evaluating the claim.

23. Using the per-gate coherence factor from Example 12.2, estimate how many gates a device could run before losing half its coherence, and compare this with the 1,000-gate estimate already given in the text.

24. Explain, using the vocabulary introduced across Sections 12.3–12.5 (preparation, transformation, measurement, evidence), what distinguishes a genuine demonstration of quantum error correction from one that merely shows a syndrome-measurement circuit runs without crashing.