

# Part I

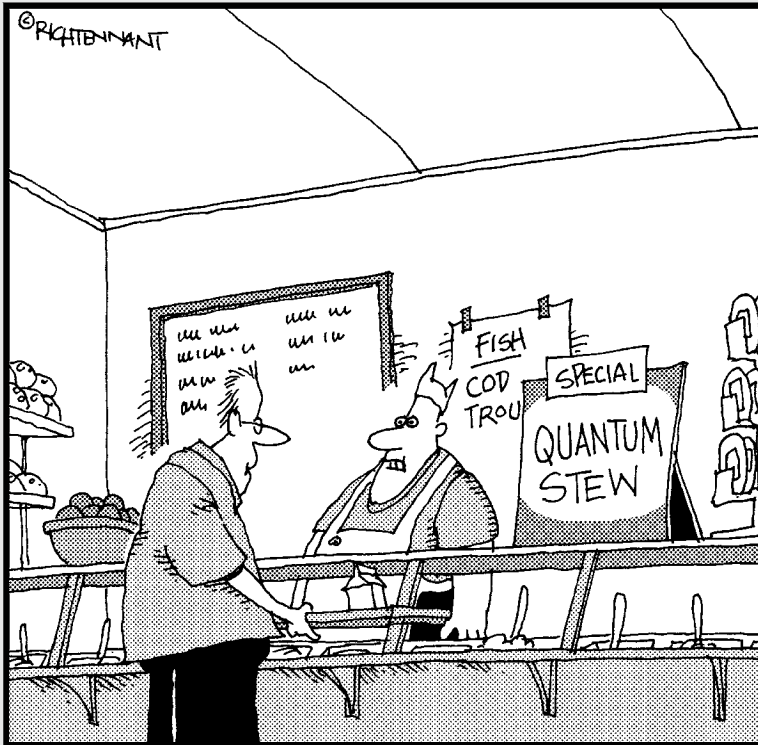
# Small World, Huh?

# Essential Quantum

# Physics

The 5<sup>th</sup> Wave

By Rich Tennant



"It's just like the regular stew only it's got some bits of matter in it we can't identify."

## ***In this part . . .***

**T**his part is designed to give you an introduction to the ways of quantum physics. You see the issues that gave rise to quantum physics and the kinds of solutions it provides. I also introduce you to the kind of math that quantum physics requires, including the notion of state vectors.

## Chapter 1

# Discoveries and Essential Quantum Physics

---

### *In This Chapter*

- ▶ Putting forth theories of quantization and discrete units
  - ▶ Experimenting with waves acting as particles
  - ▶ Experimenting with particles acting as waves
  - ▶ Embracing uncertainty and probability
- 

According to classical physics, particles are particles and waves are waves, and never the twain shall mix. That is, particles have an energy  $E$  and a momentum vector  $\mathbf{p}$ , and that's the end of it. And waves, such as light waves, have an amplitude  $A$  and a wave vector  $\mathbf{k}$  (where the magnitude of  $\mathbf{k} = \frac{2\pi}{\lambda}$ , where  $\lambda$  is the wavelength) that points in the direction the wave is traveling. And that's the end of that, too, according to classical physics.

But the reality is different — particles turn out to exhibit wave-like properties, and waves exhibit particle-like properties as well. The idea that waves (like light) can act as particles (like electrons) and vice versa was the major revelation that ushered in quantum physics as such an important part of the world of physics. This chapter takes a look at the challenges facing classical physics around the turn of the 20th century — and how quantum physics gradually came to the rescue. Up to that point, the classical way of looking at physics was thought to explain just about everything. But as those pesky experimental physicists have a way of doing, they came up with a bunch of experiments that the theoretical physicists couldn't explain.

That made the theoretical physicists mad, and they got on the job. The problem here was the microscopic world — the world that's too tiny to see. On

the larger scale, classical physics could still explain most of what was going on — but when it came to effects that depended on the micro-world, classical physics began to break down. Taking a look at how classical physics collapsed gives you an introduction to quantum physics that shows why people needed it.

## *Being Discrete: The Trouble with Black-Body Radiation*

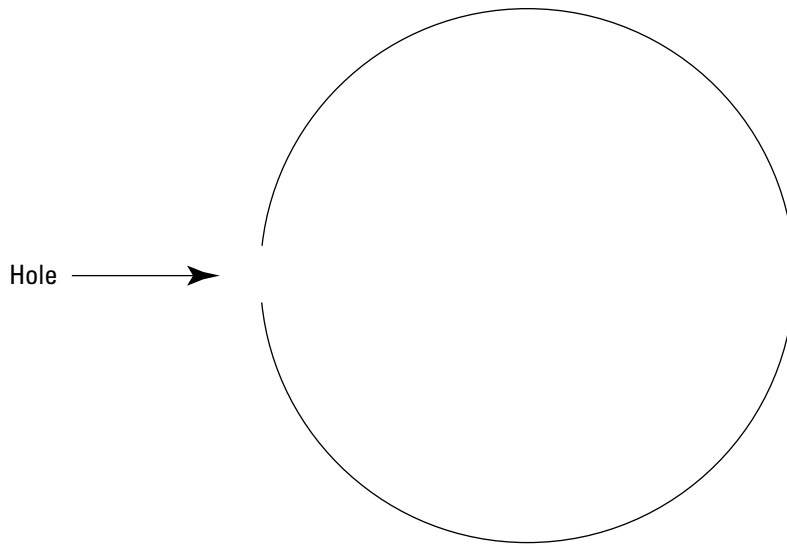
One of the major ideas of quantum physics is, well, *quantization* — measuring quantities in discrete, not continuous, units. The idea of quantized energies arose with one of the earliest challenges to classical physics: the problem of black-body radiation.

When you heat an object, it begins to glow. Even before the glow is visible, it's radiating in the infrared spectrum. The reason it glows is that as you heat it, the electrons on the surface of the material are agitated thermally, and electrons being accelerated and decelerated radiate light.

Physics in the late 19th and early 20th centuries was concerned with the spectrum of light being emitted by black bodies. A *black body* is a piece of material that radiates corresponding to its temperature — but it also absorbs and reflects light from its surroundings. To make matters easier, physics postulated a black body that reflected nothing and absorbed all the light falling on it (hence the term *black body*, because the object would appear perfectly black as it absorbed all light falling on it). When you heat a black body, it would radiate, emitting light.

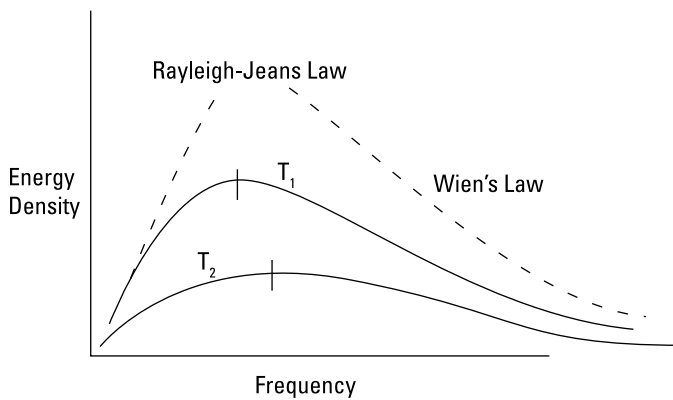
Well, it was hard to come up with a physical black body — after all, what material absorbs light 100 percent and doesn't reflect anything? But the physicists were clever about this, and they came up with the hollow cavity you see in Figure 1-1, with a hole in it.

When you shine light on the hole, all that light would go inside, where it would be reflected again and again — until it got absorbed (a negligible amount of light would escape through the hole). And when you heated the hollow cavity, the hole would begin to glow. So there you have it — a pretty good approximation of a black body.



**Figure 1-1:**  
A black  
body.

You can see the spectrum of a black body (and attempts to model that spectrum) in Figure 1-2, for two different temperatures,  $T_1$  and  $T_2$ . The problem was that nobody was able to come up with a theoretical explanation for the spectrum of light generated by the black body. Everything classical physics could come up with went wrong.



**Figure 1-2:**  
Black-body  
radiation  
spectrum.

## *First attempt: Wien's Formula*

The first one to try to explain the spectrum of a black body was Wilhelm Wien, in 1889. Using classical thermodynamics, he came up with this formula:

$$u(\nu, T) = A\nu^5 e^{-\beta\nu/T}$$

where  $u(\nu, T)$  is the intensity distribution of the light spectrum at frequency  $\nu$  of a black body at the temperature  $T$ , and  $A$  and  $\beta$  are constants which can be measured in experiments. (The spectrum is given by  $u[\nu, T]$ , which is the energy density of the emitted light as a function of frequency and temperature.)

This equation, Wien's formula, worked fine for high frequencies, as you can see in Figure 1-2; however, it failed for low frequencies.

## *Second attempt: Rayleigh-Jeans Law*

Next up in the attempt to explain the black-body spectrum was the Rayleigh-Jeans Law, introduced around 1900. This law predicted that the spectrum of a black body was

$$u(\nu, T) = \frac{2\pi\nu^4}{c^3} kT$$

where  $k$  is Boltzmann's constant (approximately  $1.3807 \times 10^{-23}$  J·K<sup>-1</sup>). However, the Rayleigh-Jeans Law had the opposite problem of Wien's law: Although it worked well at low frequencies (see Figure 1-2), it didn't match the higher-frequency data at all — in fact, it diverged at higher frequencies. This was called the *ultraviolet catastrophe* because the best predictions available diverged at high frequencies (corresponding to ultraviolet light). It was time for quantum physics to take over.

## *An intuitive (quantum) leap: Max Planck's spectrum*

The black-body problem was a tough one to solve, and with it came the first beginnings of quantum physics. Max Planck came up with a radical suggestion — what if the amount of energy that a light wave can exchange with matter wasn't continuous, as postulated by classical physics, but *discrete*? In other

words, Planck postulated that the energy of the light emitted from the walls of the black-body cavity came only in integer multiples like this, where  $h$  is a universal constant:

$$E = nh\nu, \text{ where } n = 0, 1, 2, \dots$$

With this theory, crazy as it sounded in the early 1900s, Planck converted the continuous integrals used by Rayleigh-Jeans to discrete sums over an infinite number of terms. Making that simple change gave Planck the following equation for the spectrum of black-body radiation:

$$u(\nu, T) = \frac{2\pi h \nu^5}{c^3 (e^{h\nu/kT} - 1)}$$

This equation got it right — it exactly describes the black-body spectrum, both at low and high (and medium, for that matter) frequencies.

This idea was quite new. What Planck was saying was that the energy of the radiating oscillators in the black body couldn't take on just any level of energy, as classical physics allows; it could take on only specific, *quantized* energies. In fact, Planck hypothesized that that was true for *any* oscillator — that its energy was an integral multiple of  $h\nu$ .



And so Planck's equation came to be known as *Planck's quantization rule*, and  $h$  became *Planck's constant*:  $h = 6.626 \times 10^{-34}$  Joule-seconds. Saying that the energy of all oscillators was quantized was the birth of quantum physics.

One has to wonder how Planck came up with his theory, because it's not an obvious hypothesis. Oscillators can oscillate only at discrete energies? Where did that come from? In any case, the revolution was on — and there was no stopping it.

## The First Pieces: Seeing Light as Particles

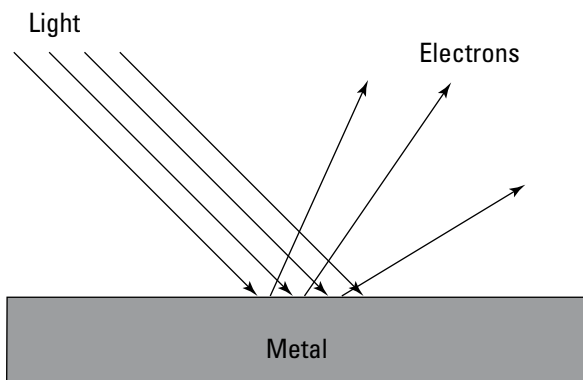
Light as particles? Isn't light made up of waves? Light, it turns out, exhibits properties of both waves and particles. This section shows you some of the evidence.

## *Solving the photoelectric effect*

The photoelectric effect was one of many experimental results that made up a crisis for classical physics around the turn of the 20th century. It was also one of Einstein's first successes, and it provides proof of the quantization of light. Here's what happened.

When you shine light onto metal, as Figure 1-3 shows, you get emitted electrons. The electrons absorb the light you shine, and if they get enough energy, they're able to break free of the metal's surface. According to classical physics, light is just a wave, and it can exchange any amount of energy with the metal. When you beam light on a piece of metal, the electrons in the metal should absorb the light and slowly get up enough energy to be emitted from the metal. The idea was that if you were to shine more light onto the metal, the electrons should be emitted with a higher kinetic energy. And very weak light shouldn't be able to emit electrons at all, except in a matter of hours.

But that's not what happened — electrons were emitted as soon as someone shone light on the metal. In fact, no matter how weak the intensity of the incident light (and researchers tried experiments with such weak light that it should have taken hours to get any electrons emitted), electrons *were* emitted. Immediately.

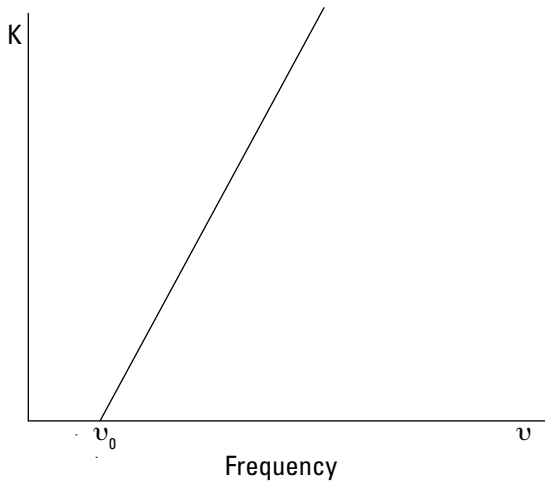


**Figure 1-3:**  
The photo-  
electric  
effect.

Experiments with the photoelectric effect showed that the kinetic energy,  $K$ , of the emitted electrons depended only on the frequency — not the intensity — of the incident light, as you can see in Figure 1-4.



**Figure 1-4:**  
Kinetic  
energy of  
emitted  
electrons  
versus  
frequency  
of the inci-  
dent light.



In Figure 1-4,  $\nu_0$  is called the *threshold frequency*, and if you shine light with a frequency below this threshold on the metal, no electrons are emitted. The emitted electrons come from the pool of free electrons in the metal (all metals have a pool of free electrons), and you need to supply these electrons with an energy equivalent to the metal's work function,  $W$ , to emit the electron from the metal's surface.

The results were hard to explain classically, so enter Einstein. This was the beginning of his heyday, around 1905. Encouraged by Planck's success (see the preceding section), Einstein postulated that not only were oscillators quantized but so was light — into discrete units called *photons*. Light, he suggested, acted like particles as well as waves.

So in this scheme, when light hits a metal surface, photons hit the free electrons, and an electron completely absorbs each photon. When the energy,  $h\nu$ , of the photon is greater than the work function of the metal, the electron is emitted. That is,

$$h\nu = W + K$$

where  $W$  is the metal's work function and  $K$  is the kinetic energy of the emitted electron. Solving for  $K$  gives you the following:

$$K = h\nu - W$$

You can also write this in terms of the threshold frequency this way:

$$K = h(\nu - \nu_0)$$

So apparently, light isn't just a wave; you can also view it as a particle, the photon. In other words, light is quantized.

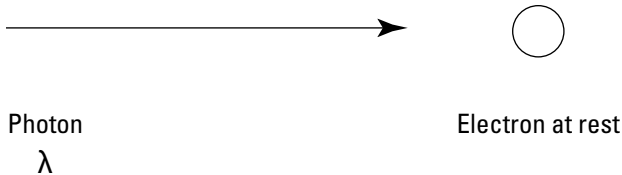
That was also quite an unexpected piece of work by Einstein, although it was based on the earlier work of Planck. Light *quantized*? Light coming in discrete energy packets? What next?

## *Scattering light off electrons: The Compton effect*

To a world that still had trouble comprehending light as particles (see the preceding section), Arthur Compton supplied the final blow with the Compton effect. His experiment involved scattering photons off electrons, as Figure 1-5 shows.

**Figure 1-5:**

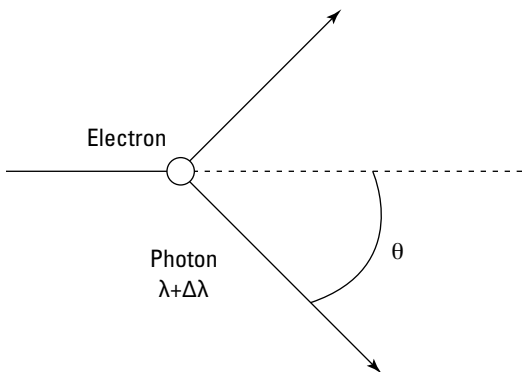
Light  
incident on  
an electron  
at rest.



Incident light comes in with a wavelength of  $\lambda$  and hits the electron at rest. After that happens, the light is scattered, as you see in Figure 1-6.

**Figure 1-6:**

Photon  
scattering  
off an  
electron.



Classically, here's what should've happened: The electron should've absorbed the incident light, oscillated, and emitted it — with the same wavelength but with an intensity depending on the intensity of the incident light. But that's not what happened — in fact, the wavelength of the light is actually changed by  $\Delta\lambda$ , called the *wavelength shift*. The scattered light has a wavelength of  $\lambda + \Delta\lambda$  — in other words, its wavelength has increased, which means the light has lost energy. And  $\Delta\lambda$  depends on the scattering angle,  $\theta$ , not on the intensity of the incident light.

Arthur Compton could explain the results of his experiment only by making the assumption that he was actually dealing with two particles — a photon and an electron. That is, he treated light as a discrete particle, not a wave. And he made the assumption that the photon and the electron collided elastically — that is, that both total energy and momentum were conserved.

Making the assumption that both the light and the electron were particles, Compton then derived this formula for the wavelength shift (it's an easy calculation if you assume that the light is represented by a photon with energy  $E = h\nu$  and that its momentum is  $p = E/c$ ):

$$\Delta\lambda = \frac{h}{m_e c} (1 - \cos\theta)$$

where  $h$  is Planck's constant,  $m_e$  is the mass of an electron,  $c$  is the speed of light, and  $\theta$  is the scattering angle of the light.

You also see this equation in the equivalent form:

$$\Delta\lambda = 2\lambda_c (\theta/2)$$

where  $\lambda_c$  is the Compton wavelength of an electron,  $\lambda_c = h/m_e c$ . And experiment confirms this relation — both equations.

Note that to derive the wavelength shift, Compton had to make the assumption that here, light was acting as a particle, not as a wave. That is, the particle nature of light was the aspect of the light that was predominant.

## ***Proof positron? Dirac and pair production***

In 1928, the physicist Paul Dirac posited the existence of a positively charged anti-electron, the *positron*. He did this by taking the newly evolving field of quantum physics to new territory by combining relativity with quantum

mechanics to create relativistic quantum mechanics — and that was the theory that predicted, through a plus/minus-sign interchange — the existence of the positron.

It was a bold prediction — an *anti-particle* of the electron? But just four years later, physicists actually saw the positron. Today's high-powered elementary particle physics has all kinds of synchrotrons and other particle accelerators to create all the elementary particles they need, but in the early 20th century, this wasn't always so.

In those days, physicists relied on cosmic rays — those particles and high-powered photons (called gamma rays) that strike the Earth from outer space — as their source of particles. They used *cloud-chambers*, which were filled with vapor from dry ice, to see the trails such particles left. They put their chambers into magnetic fields to be able to measure the momentum of the particles as they curved in those fields.

In 1932, a physicist noticed a surprising event. A pair of particles, oppositely charged (which could be determined from the way they curved in the magnetic field) appeared from apparently nowhere. No particle trail led to the origin of the two particles that appeared. That was *pair-production* — the conversion of a high-powered photon into an electron and positron, which can happen when the photon passes near a heavy atomic nucleus.

So experimentally, physicists had now seen a photon turning into a pair of particles. Wow. As if everyone needed more evidence of the particle nature of light. Later on, researchers also saw *pair annihilation*: the conversion of an electron and positron into pure light.

Pair production and annihilation turned out to be governed by Einstein's newly introduced theory of relativity — in particular, his most famous formula,  $E = mc^2$ , which gives the pure energy equivalent of mass. At this point, there was an abundance of evidence of the particle-like aspects of light.

## *A Dual Identity: Looking at Particles as Waves*

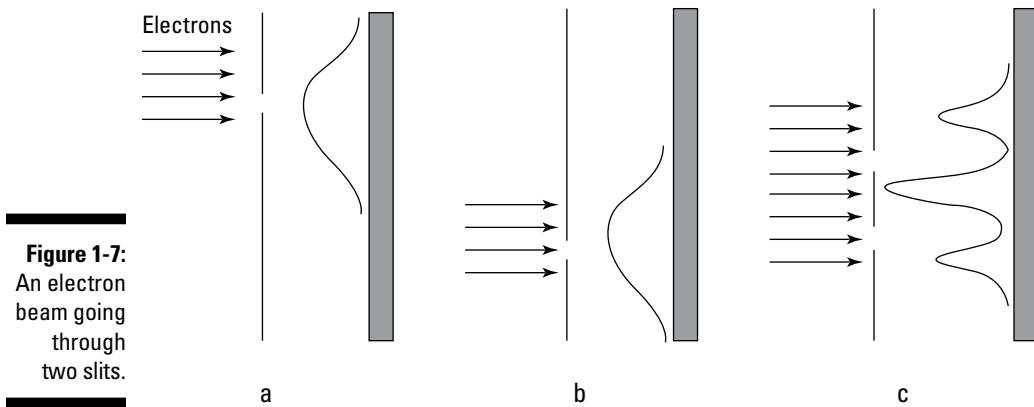
In 1923, the physicist Louis de Broglie suggested that not only did waves exhibit particle-like aspects but the reverse was also true — all material particles should display wave-like properties.

How does this work? For a photon, momentum  $p = h\nu/c = h/\lambda$ , where  $\nu$  is the photon's frequency and  $\lambda$  is its wavelength. And the wave vector,  $\mathbf{k}$ , is equal to  $\mathbf{k} = \mathbf{p}/\hbar$ , where  $\hbar = h/2\pi$ . De Broglie said that the same relation should hold for all material particles. That is,

$$\lambda = \frac{h}{p}$$

$$k = \frac{p}{\hbar}$$

De Broglie presented these apparently surprising suggestions in his Ph.D. thesis. Researchers put these suggestions to the test by sending a beam through a dual-slit apparatus to see whether the electron beam would act like it was made up of particles or waves. In Figure 1-7, you can see the setup and the results.



In Figure 1-7a, you can see a beam of electrons passing through a single slit and the resulting pattern on a screen. In Figure 1-7b, the electrons are passing through a second slit. Classically, you'd expect the intensities of Figure 1-7a and 1-7b simply to add when both slits are open:

$$I = I_1 + I_2$$

But that's not what happened. What actually appeared was an interference pattern when both slits were open (Figure 1-7c), not just a sum of the two slits' electron intensities.

The result was a validation of de Broglie's invention of matter waves. Experiment bore out the relation that  $\lambda = h/p$ , and de Broglie was a success.



The idea of matter waves is a big part of what's coming up in the rest of the book. In particular, the existence of matter waves says that you add the waves' amplitude,  $\psi_1(r, t)$  and  $\psi_2(r, t)$ , not their intensities, to sum them:

$$\psi(r, t) = \psi_1(r, t) + \psi_2(r, t)$$

You square the amplitude to get the intensity, and the phase difference between  $\psi_1(r, t)$  and  $\psi_2(r, t)$  is what actually creates the interference pattern that's observed.

## You Can't Know Everything (But You Can Figure the Odds)



So particles apparently exhibit wave-like properties, and waves exhibit particle-like properties. But if you have an electron, which is it — a wave or a particle? The truth is that physically, an electron is just an electron, and you can't actually say whether it's a wave or a particle. The act of *measurement* is what brings out the wave or particle properties. You see more about this idea throughout the book.

Quantum mechanics lives with an uncertain picture quite happily. That view offended many eminent physicists of the time — notably Albert Einstein, who said, famously, “God does not play dice.” In this section, I discuss the idea of uncertainty and how quantum physicists work in probabilities instead.

## The Heisenberg uncertainty principle

The fact that matter exhibits wave-like properties gives rise to more trouble — waves aren't localized in space. And knowing that inspired Werner Heisenberg, in 1927, to come up with his celebrated uncertainty principle.

You can completely describe objects in classical physics by their momentum and position, both of which you can measure exactly. In other words, classical physics is completely *deterministic*.

On the atomic level, however, quantum physics paints a different picture. Here, the *Heisenberg uncertainty principle* says that there's an inherent uncertainty in the relation between position and momentum. In the  $x$  direction, for example, that looks like this:

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}$$

where  $\Delta x$  is the measurement uncertainty in the particle's  $x$  position,  $\Delta p_x$  is its measurement uncertainty in its momentum in the  $x$  direction and  $\hbar = h/2\pi$ .

That is to say, the more accurately you know the position of a particle, the less accurately you know the momentum, and vice versa. This relation holds for all three dimensions:

$$\Delta y \Delta p_y \geq \frac{\hbar}{2}$$

$$\Delta z \Delta p_z \geq \frac{\hbar}{2}$$

And the Heisenberg uncertainty principle is a direct consequence of the wave-like nature of matter, because you can't completely pin down a wave.



Quantum physics, unlike classical physics, is completely undeterministic. You can never know the *precise* position and momentum of a particle at any one time. You can give only probabilities for these linked measurements.

## ***Rolling the dice: Quantum physics and probability***

In quantum physics, the state of a particle is described by a wave function,  $\psi(r, t)$ . The wave function describes the de Broglie wave of a particle, giving its amplitude as a function of position and time. (See the earlier section “A Dual Identity: Looking at Particles as Waves” for more on de Broglie.)



Note that the wave function gives a particle's amplitude, not intensity; if you want to find the intensity of the wave function, you have to square it:  $|\psi(r, t)|^2$ . The *intensity* of a wave is what's equal to the probability that the particle will be at that position at that time.

That's how quantum physics converts issues of momentum and position into probabilities: by using a wave function, whose square tells you the *probability density* that a particle will occupy a particular position or have a particular momentum. In other words,  $|\psi(r, t)|^2 d^3r$  is the probability that the particle will be found in the volume element  $d^3r$ , located at position  $r$  at time  $t$ .

Besides the position-space wave function  $\psi(r, t)$ , there's also a momentum-space version of the wave function:  $\phi(p, t)$ .

This book is largely a study of the wave function — the wave functions of free particles, the wave functions of particles trapped inside potentials, of identical particles hitting each other, of particles in harmonic oscillation, of light scattering from particles, and more. Using this kind of physics, you can predict the behavior of all kinds of physical systems.



## Chapter 2

# Entering the Matrix: Welcome to State Vectors

---

### *In This Chapter*

- ▶ Creating state vectors
  - ▶ Using Dirac notation for state vectors
  - ▶ Working with bras and kets
  - ▶ Understanding matrix mechanics
  - ▶ Getting to wave mechanics
- 

**Q**uantum physics isn't just about playing around with your particle accelerator while trying not to destroy the universe. Sometimes, you get to do things that are a little more mundane, like turn lights off and on, perform a bit of calculus, or play with dice.

If you're actually doing physics with those dice (beyond hurling them across the room), the lab director won't even get mad at you. In quantum physics, absolute measurements are replaced by probabilities, so you may use dice to calculate the probabilities that various numbers will come up. You can then assemble those values into a vector (single-column matrix) in Hilbert space (a type of infinitely dimensional vector space with some properties that are especially valuable in quantum physics).

This chapter introduces how you deal with probabilities in quantum physics, starting by viewing the various possible states a particle can occupy as a vector — a vector of probability states. From there, I help you familiarize yourself with some mathematical notations common in quantum physics, including bras, kets, matrices, and wave functions. Along the way, you also get to work with some important operators.

## *Creating Your Own Vectors in Hilbert Space*

In quantum physics, probabilities take the place of absolute measurements. Say you've been experimenting with rolling a pair of dice and are trying to figure the relative probability that the dice will show various values. You come up with a list indicating the relative probability of rolling a 2, 3, 4, and so on, all the way up to 12:

<i>Sum of the Dice</i>	<i>Relative Probability (Number of Ways of Rolling a Particular Total)</i>
2	1
3	2
4	3
5	4
6	5
7	6
8	5
9	4
10	3
11	2
12	1

In other words, you're twice as likely to roll a 3 than a 2, you're four times as likely to roll a 5 than a 2, and so on. You can assemble these relative probabilities into a vector (if you're thinking of a "vector" from physics, think in terms of a column of the vector's components, not a magnitude and direction) to keep track of them easily:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$$

Okay, now you're getting closer to the way quantum physics works. You have a vector of the probabilities that the dice will occupy various states. However, quantum physics doesn't deal directly with probabilities but rather with *probability amplitudes*, which are the square roots of the probabilities. To find the actual probability that a particle will be in a certain state, you add wave functions — which are going to be represented by these vectors — and then square them (see Chapter 1 for info on why). So take the square root of all these entries to get the probability amplitudes:

$$\begin{bmatrix} \sqrt{1} \\ \sqrt{2} \\ \sqrt{3} \\ \sqrt{4} \\ \sqrt{5} \\ \sqrt{6} \\ \sqrt{5} \\ \sqrt{4} \\ \sqrt{3} \\ \sqrt{2} \\ \sqrt{1} \end{bmatrix}$$

That's better, but adding the squares of all these should add up to a total probability of 1; as it is now, the sum of the squares of these numbers is 36, so divide each entry by  $36^{1/2}$ , or 6:

$$\begin{bmatrix} \sqrt{1/36} \\ \sqrt{2/36} \\ \sqrt{3/36} \\ \sqrt{4/36} \\ \sqrt{5/36} \\ \sqrt{6/36} \\ \sqrt{5/36} \\ \sqrt{4/36} \\ \sqrt{3/36} \\ \sqrt{2/36} \\ \sqrt{1/36} \end{bmatrix}$$

So now you can get the probability amplitude of rolling any combination from 2 to 12 by reading down the vector — the probability amplitude of rolling a 2 is  $1/6$ , of rolling a 3 is

$\frac{\sqrt{2}}{6}$ , and so on.

## *Making Life Easier with Dirac Notation*

When you have a state vector that gives the probability amplitude that a pair of dice will be in their various possible states, you basically have a vector in *dice space* — all the possible states that a pair of dice can take, which is an 11-dimensional space. (See the preceding section for more on state vectors.)

But in most quantum physics problems, the vectors can be infinitely large — for example, a moving particle can be in an infinite number of states. Handling large arrays of states isn't easy using vector notation, so instead of explicitly writing out the whole vector each time, quantum physics usually uses the notation developed by physicist Paul Dirac — the *Dirac* or *bra-ket notation*.

## Abbreviating state vectors as kets

Dirac notation abbreviates the state vector as a *ket*, like this:  $|\psi\rangle$ . So in the dice example, you can write the state vector as a ket this way:

$$|\psi\rangle = \begin{bmatrix} 1/6 \\ \sqrt{2}/6 \\ \sqrt{3}/6 \\ 2/6 \\ \sqrt{5}/6 \\ \sqrt{6}/6 \\ \sqrt{5}/6 \\ 2/6 \\ \sqrt{3}/6 \\ \sqrt{2}/6 \\ 1/6 \end{bmatrix}$$

Here, the components of the state vector are represented by numbers in 11-dimensional dice space. More commonly, however, each component represents a function of position and time, something like this:

$$|\psi\rangle = \begin{bmatrix} 1/6 \cdot e^{i(kx-\omega t)} \\ \sqrt{2}/6 \cdot e^{2i(kx-\omega t)} \\ \sqrt{3}/6 \cdot e^{3i(kx-\omega t)} \\ 2/6 \cdot e^{4i(kx-\omega t)} \\ \sqrt{5}/6 \cdot e^{5i(kx-\omega t)} \\ \sqrt{6}/6 \cdot e^{6i(kx-\omega t)} \\ \sqrt{5}/6 \cdot e^{7i(kx-\omega t)} \\ 2/6 \cdot e^{8i(kx-\omega t)} \\ \sqrt{3}/6 \cdot e^{9i(kx-\omega t)} \\ \sqrt{2}/6 \cdot e^{10i(kx-\omega t)} \\ 1/6 \cdot e^{11i(kx-\omega t)} \end{bmatrix}$$



In general, a set of vectors  $\phi_N$  in Hilbert space is linearly independent if the only solution to the following equation is that all the coefficients  $a_n = 0$ :

$$\sum_{i=1}^N a_n \phi_i = 0$$

That is, as long as you can't write any one vector as a linear combination of the others, the vectors are linearly independent and so form a valid basis in Hilbert space.

## Writing the Hermitian conjugate as a bra

For every ket, there's a corresponding *bra*. (The terms come from *bra-ket*, or *bracket*, which should be clearer in the upcoming section titled “Grooving with Operators.”) A *bra* is the Hermitian conjugate of the corresponding ket.

Suppose you start with this ket:

$$|\psi\rangle = \begin{bmatrix} 1/6 \\ \sqrt{2}/6 \\ \sqrt{3}/6 \\ 2/6 \\ \sqrt{5}/6 \\ \sqrt{6}/6 \\ \sqrt{5}/6 \\ 2/6 \\ \sqrt{3}/6 \\ \sqrt{2}/6 \\ 1/6 \end{bmatrix}$$

The  $*$  symbol means the complex conjugate. (A *complex conjugate* flips the sign connecting the real and imaginary parts of a complex number.) So the corresponding bra, which you write as  $\langle\psi|$ , equals  $|\psi\rangle^{T*}$ . The bra is this row vector:

$$\langle\psi| = \left[ \frac{1}{6} \quad \frac{\sqrt{2}}{6} \quad \frac{\sqrt{3}}{6} \quad \frac{2}{6} \quad \frac{\sqrt{5}}{6} \quad \frac{\sqrt{6}}{6} \quad \frac{\sqrt{5}}{6} \quad \frac{2}{6} \quad \frac{\sqrt{3}}{6} \quad \frac{\sqrt{2}}{6} \quad \frac{1}{6} \right]$$



Note that if any of the elements of the ket are complex numbers, you have to take their complex conjugate when creating the associated bra. For instance, if your complex number in the ket is  $a + bi$ , its complex conjugate in the bra is  $a - bi$ .

## Multiplying bras and kets: A probability of 1

You can take the product of your ket and bra, denoted as  $\langle \psi | \psi \rangle$ , like this:

$$\langle \psi | \psi \rangle = \begin{bmatrix} \frac{1}{6} & \frac{\sqrt{2}}{6} & \frac{\sqrt{3}}{6} & \frac{2}{6} & \frac{\sqrt{5}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{5}}{6} & \frac{2}{6} & \frac{\sqrt{3}}{6} & \frac{\sqrt{2}}{6} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 1/6 \\ \sqrt{2}/6 \\ \sqrt{3}/6 \\ 2/6 \\ \sqrt{5}/6 \\ \sqrt{6}/6 \\ \sqrt{5}/6 \\ 2/6 \\ \sqrt{3}/6 \\ \sqrt{2}/6 \\ 1/6 \end{bmatrix}$$

This is just matrix multiplication, and the result is the same as taking the sum of the squares of the elements:

$$\langle \psi | \psi \rangle = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = 1$$



And that's the way it should be, because the total probability should add up to 1. Therefore, in general, the product of the bra and ket equals 1:

$$\langle \psi | \psi \rangle = 1$$

If this relation holds, the ket  $|\psi\rangle$  is said to be *normalized*.

## *Covering all your bases: Bras and kets as basis-less state vectors*

The reason ket notation,  $|\psi\rangle$ , is so popular in quantum physics is that it allows you to work with state vectors in a basis-free way. In other words, you're not stuck in the position basis, the momentum basis, or the energy basis. That's helpful, because most of the work in quantum physics takes place in abstract calculations, and you don't want to have to drag all the components of state vectors through those calculations (often you can't — there may be infinitely many possible states in the problem you're dealing with).

For example, say that you're representing your states using position vectors in a three-dimensional Hilbert space — that is, you have  $x$ ,  $y$ , and  $z$  axes, forming a position *basis* for your space. That's fine, but not all your calculations have to be done using that position basis.

You may want to, for example, represent your states in a three-dimensional momentum space, with three axes in Hilbert space,  $p_x$ ,  $p_y$ , and  $p_z$ . Now you'd have to change all your position vectors to momentum vectors, adjusting each component, and keep track of what happens to every component through all your calculations.

So Dirac's bra-ket notation comes to the rescue here — you use it to perform all the math and then plug in the various components of your state vectors as needed at the end. That is, you can perform your calculations in purely symbolic terms, without being tied to a basis.

And when you need to deal with the components of a ket, such as when you want to get physical answers, you can also convert kets to a different basis by taking the ket's components along the axes of that basis. Generally, when you have a vector  $|\psi\rangle$ , you can express it as a sum over  $N$  basis vectors,  $|\phi_i\rangle$ , like so:

$$|\psi\rangle = \sum_{i=1}^N |\phi_i\rangle \langle \phi_i | \psi \rangle$$

where  $N$  is the dimension of the Hilbert space, and  $i$  is an integer that labels the basis vectors.

## *Understanding some relationships using kets*

Ket notation makes the math easier than it is in matrix form because you can take advantage of a few mathematical relationships. For example, here's the so-called Schwarz inequality for state vectors:

$$|\langle \psi | \phi \rangle|^2 \leq \langle \psi | \psi \rangle \langle \phi | \phi \rangle$$



This says that the square of the absolute value of the product of two state vectors,  $|\langle\psi|\phi\rangle|^2$ , is less than or equal to  $\langle\psi|\psi\rangle\langle\phi|\phi\rangle$ . This turns out to be the analog of the vector inequality:

$$|\mathbf{A} \cdot \mathbf{B}|^2 \leq |\mathbf{A}|^2 |\mathbf{B}|^2$$

So why is the Schwarz inequality so useful? It turns out that you can derive the Heisenberg uncertainty principle from it (see Chapter 1 for more on this principle).

Other ket relationships can also simplify your calculations. For instance, two kets,  $|\psi\rangle$  and  $|\phi\rangle$ , are said to be *orthogonal* if

$$\langle\psi|\phi\rangle = 0$$

And two kets are said to be *orthonormal* if they meet the following conditions:

$$\checkmark \langle\psi|\phi\rangle = 0$$

$$\checkmark \langle\psi|\psi\rangle = 1$$

$$\checkmark \langle\phi|\phi\rangle = 1$$

With this information in mind, you're now ready to start working with operators.

## Grooving with Operators

What about all the calculations that you're supposed to be able to perform with kets? Taking the product of a bra and a ket,  $\langle\psi|\phi\rangle$ , is fine as far as it goes, but what about extracting some physical quantities you can measure? That's where operators come in.

### Hello, operator: How operators work

Here's the general definition of an operator  $A$  in quantum physics: An *operator* is a mathematical rule that, when operating on a ket,  $|\psi\rangle$ , transforms that ket into a new ket,  $|\psi'\rangle$  in the same space (which could just be the old ket multiplied by a scalar). So when you have an operator  $A$ , it transforms kets like this:

$$A|\psi\rangle = |\psi'\rangle$$

For that matter, the same operator can also transform bras:

$$\langle \psi | A = \langle \psi' |$$



Here are several examples of the kinds of operators you'll see:

- ✓ **Hamiltonian (H):** Applying the Hamiltonian operator (which looks different for every different physical situation) gives you  $E$ , the energy of the particle represented by the ket  $|\psi\rangle$ ;  $E$  is a scalar quantity:

$$H|\psi\rangle = E|\psi\rangle$$

- ✓ **Unity or identity (I):** The unity or identity operator leaves kets unchanged:

$$I|\psi\rangle = |\psi\rangle$$

- ✓ **Gradient ( $\nabla$ ):** The gradient operator works like this:

$$\nabla|\psi\rangle = \frac{\partial}{\partial x}|\psi\rangle \mathbf{i} + \frac{\partial}{\partial y}|\psi\rangle \mathbf{j} + \frac{\partial}{\partial z}|\psi\rangle \mathbf{k}$$

- ✓ **Linear momentum (P):** The linear momentum operator looks like this in quantum mechanics:

$$P|\psi\rangle = -i\hbar\nabla|\psi\rangle$$

- ✓ **Laplacian ( $\nabla^2$ ):** You use the Laplacian operator, which is much like a second-order gradient, to create the energy-finding Hamiltonian operator:

$$\nabla^2|\psi\rangle = \nabla \cdot \nabla|\psi\rangle = \frac{\partial^2}{\partial x^2}|\psi\rangle + \frac{\partial^2}{\partial y^2}|\psi\rangle + \frac{\partial^2}{\partial z^2}|\psi\rangle$$



In general, multiplying operators together is not the same independent of order, so for the operators  $A$  and  $B$ ,  $AB \neq BA$ .

And an operator  $A$  is said to be *linear* if

$$A(c_1|\psi\rangle + c_2|\psi\rangle) = c_1A|\psi\rangle + c_2A|\psi\rangle$$



So to find the expectation value of  $R$ , you need to calculate  $\langle \psi | R | \psi \rangle$ . Spelling that out in terms of components gives you the following:

$$\langle \psi | R | \psi \rangle = \begin{bmatrix} \frac{1}{6} & \frac{\sqrt{2}}{6} & \frac{\sqrt{3}}{6} & \frac{2}{6} & \frac{\sqrt{5}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{5}}{6} & \frac{2}{6} & \frac{\sqrt{3}}{6} & \frac{\sqrt{2}}{6} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 11 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \end{bmatrix} \begin{bmatrix} 1/6 \\ \sqrt{2}/6 \\ \sqrt{3}/6 \\ 2/6 \\ \sqrt{5}/6 \\ \sqrt{6}/6 \\ \sqrt{5}/6 \\ 2/6 \\ \sqrt{3}/6 \\ \sqrt{2}/6 \\ 1/6 \end{bmatrix}$$

Doing the math, you get

$$\langle \psi | R | \psi \rangle = 7$$

So the expectation value of a roll of the dice is 7. Now you can see where the terms *bra* and *ket* come from — they “bracket” an operator to give you expectation values. In fact, the expectation value is such a common thing to find that you’ll often find  $\langle \psi | R | \psi \rangle$  abbreviated as  $\langle R \rangle$ , so

$$\langle R \rangle = 7$$

## Looking at linear operators

An operator  $A$  is said to be *linear* if it meets the following condition:

$$A(c_1|\psi\rangle + c_2|\phi\rangle) = c_1A|\psi\rangle + c_2A|\phi\rangle$$

For instance, the expression  $|\phi\rangle\langle\chi|$  is actually a linear operator. In order for us to see this we shall need to know just a little more about what happens when we take the products of bras and kets. Firstly, if we take the product of the bra,  $\langle\chi|$ , with the ket,  $c|\psi\rangle$ , where  $c$  is a complex number, then we get the answer,

$$\langle \chi | c | \psi \rangle = c \langle \chi | \psi \rangle$$

Secondly, if we take the product of the bra,  $\langle \chi |$ , with the sum of two kets,  $|\psi_1\rangle + |\psi_2\rangle$ , then we get the answer,

$$\langle \chi | (|\psi_1\rangle + |\psi_2\rangle) = \langle \chi | \psi_1 \rangle + \langle \chi | \psi_2 \rangle$$

Now that we know this we can test to see if  $|\phi\rangle\langle\chi|$  is actually a linear operator. OK then, let's apply  $|\phi\rangle\langle\chi|$  to a linear combination of bras, like so,

$$|\phi\rangle\langle\chi| (c_1|\psi_1\rangle + c_2|\psi_2\rangle)$$

where  $c_1$  and  $c_2$  are complex numbers. Now that you know how the product of a bra with a sum of two kets goes, you can say,

$$|\phi\rangle\langle\chi| (c_1|\psi_1\rangle + c_2|\psi_2\rangle) = |\phi\rangle\langle\chi| c_1|\psi_1\rangle + |\phi\rangle\langle\chi| c_2|\psi_2\rangle$$

Then, as you know,  $\langle \chi | c | \psi \rangle = c \langle \chi | \psi \rangle$ , you can finally write this as,

$$|\phi\rangle\langle\chi| (c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1|\phi\rangle\langle\chi|\psi_1\rangle + c_2|\phi\rangle\langle\chi|\psi_2\rangle$$

This is exactly what a linear operator should do — if you replace  $A$  in the above equation defining a linear operator, with  $|\phi\rangle\langle\chi|$ , then the result is the same as the one you just found. So  $|\phi\rangle\langle\chi|$  is indeed a linear operator — though I would agree, it's a pretty funny looking one!

## Going Hermitian with Hermitian Operators and Adjoint

The *Hermitian adjoint* — also called the *adjoint* or *Hermitian conjugate* — of an operator  $A$  is denoted  $A^\dagger$ . To find the Hermitian adjoint, follow these steps:

### 1. Replace complex constants with their complex conjugates.

The Hermitian adjoint of a complex number is the complex conjugate of that number:

$$a^\dagger = a^*$$



**2. Replace kets with their corresponding bras, and replace bras with their corresponding kets.**

You have to exchange the bras and kets when finding the Hermitian adjoint of an operator, so finding the Hermitian adjoint of an operator is not just the same as mathematically finding its complex conjugate.

**3. Replace operators with their Hermitian adjoints.**

In quantum mechanics, operators that are equal to their Hermitian adjoints are called *Hermitian operators*. In other words, an operator is Hermitian if

$$A^\dagger = A$$

Hermitian operators appear throughout the book, and they have special properties. For instance, the matrix that represents them may be *diagonalized* — that is, written so that the only nonzero elements appear along the matrix's diagonal. Also, the expectation value of a Hermitian operator is guaranteed to be a real number, not complex (see the earlier section “I expected that: Finding expectation values”).

**4. Write your final equation.**

$$(\langle \psi | A | \phi \rangle)^* = \langle \phi | A^\dagger | \psi \rangle$$



Here are some relationships concerning Hermitian adjoints:

$$\checkmark (aA)^\dagger = a^* A^\dagger$$

$$\checkmark (A^\dagger)^\dagger = A$$

$$\checkmark (A + B)^\dagger = A^\dagger + B^\dagger$$

$$\checkmark (AB)^\dagger = B^\dagger A^\dagger$$

$$\checkmark (AB | \psi \rangle)^\dagger = \langle \psi | B^\dagger A^\dagger$$

## Forward and Backward: Finding the Commutator



The measure of how different it is to apply operator A and then B, versus B and then A, is called the operators' *commutator*. Here's how you define the commutator of operators A and B:

$$[A, B] = AB - BA$$

## *Commuting*

Two operators *commute* with each other if their commutator is equal to zero. That is, it doesn't make any difference in what order you apply them:

$$[A, B] = 0$$

Note in particular that any operator commutes with itself:

$$[A, A] = 0$$

And it's easy to show that the commutator of A, B is the negative of the commutator of B, A:

$$[A, B] = -[B, A]$$

It's also true that commutators are linear:

$$[A, B + C + D + \dots] = [A, B] + [A, C] + [A, D] + \dots$$

And the Hermitian adjoint of a commutator works this way:

$$[A, B]^\dagger = [B^\dagger, A^\dagger]$$

You can also find the anticommutator,  $\{A, B\}$ :

$$\{A, B\} = AB + BA$$

## *Finding anti-Hermitian operators*

Here's another one: What can you say about the Hermitian adjoint of the commutator of two Hermitian operators? Here's the answer. First, write the adjoint:

$$[A, B]^\dagger$$

The definition of commutators tells you the following:

$$[A, B]^\dagger = (AB - BA)^\dagger$$

You know  $(AB)^\dagger = B^\dagger A^\dagger$  (see the earlier section “Going Hermitian with Hermitian Operators and Adjoints” for properties of adjoints). Therefore,

$$[A, B]^\dagger = (AB - BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger$$

But for Hermitian operators,  $A = A^\dagger$ , so remove the  $^\dagger$  symbols:

$$[A, B]^\dagger = (AB - BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB$$

But  $BA - AB$  is just  $-[A, B]$ , so you have the following:

$$[A, B]^\dagger = -[A, B]$$



A and B here are Hermitian operators. When you take the Hermitian adjoint of an expression and get the same thing back with a negative sign in front of it, the expression is called *anti-Hermitian*, so the commutator of two Hermitian operators is anti-Hermitian. (And by the way, the expectation value of an anti-Hermitian operator is guaranteed to be completely imaginary.)

## Starting from Scratch and Ending Up with Heisenberg

If you’ve read through the last few sections, you’re now armed with all this new technology: Hermitian operators and commutators. How can you put it to work? You can come up with the Heisenberg uncertainty relation starting virtually from scratch.

Here’s a calculation that takes you from a few basic definitions to the Heisenberg uncertainty relation. This kind of calculation shows how much easier it is to use the basis-less bra and ket notation than the full matrix version of state vectors. This isn’t the kind of calculation that you’ll need to do in class, but follow it through — knowing how to use kets, bras, commutators, and Hermitian operators is vital in the coming chapters.

The uncertainty in a measurement of the Hermitian operator named A is formally given by

$$\Delta A = \left( \langle A^2 \rangle - \langle A \rangle^2 \right)^{1/2}$$



That is,  $\Delta A$  is equal to the square root of the expectation value of  $A^2$  minus the squared expectation value of  $A$ . If you've taken any math classes that dealt with statistics, this formula may be familiar to you. Similarly, the uncertainty in a measurement using Hermitian operator  $B$  is

$$\Delta B = \left( \langle B^2 \rangle - \langle B \rangle^2 \right)^{1/2}$$

Now consider the *operators*  $\Delta A$  and  $\Delta B$  (not the uncertainties  $\Delta A$  and  $\Delta B$  anymore), and assume that applying  $\Delta A$  and  $\Delta B$  as operators gives you measurement values like this:

$$\Delta A = A - \langle A \rangle$$

$$\Delta B = B - \langle B \rangle$$

Like any operator, using  $\Delta A$  and  $\Delta B$  can result in new kets:

$$\Delta A |\psi\rangle = |\chi\rangle$$

$$\Delta B |\psi\rangle = |\phi\rangle$$

Here's the key: The Schwarz inequality (from the earlier section "Understanding some relationships using kets") gives you

$$\langle \chi | \chi \rangle \langle \phi | \phi \rangle \geq |\langle \chi | \phi \rangle|^2$$

So you can see that the inequality sign,  $\geq$ , which plays a big part in the Heisenberg uncertainty relation, has already crept into the calculation.

Because  $\Delta A$  and  $\Delta B$  are Hermitian,  $\langle \chi | \chi \rangle$  is equal to  $\langle \psi | \Delta A^2 | \psi \rangle$  and  $\langle \phi | \phi \rangle$  is equal to  $\langle \psi | \Delta B^2 | \psi \rangle$ . Because  $\Delta A^\dagger = \Delta A$  (the definition of a Hermitian operator), you can see that

$$\langle \chi | \chi \rangle = \langle \psi | \Delta A^\dagger \Delta A | \psi \rangle$$

This means that

$$\langle \chi | \chi \rangle = \langle \psi | \Delta A^\dagger \Delta A | \psi \rangle = \langle \psi | \Delta A^2 | \psi \rangle$$

That is,  $\langle \chi | \chi \rangle$  is equal to  $\langle \Delta A^2 \rangle$  and  $\langle \phi | \phi \rangle$  is equal to  $\langle \Delta B^2 \rangle$ . So you can rewrite the Schwarz inequality like this:

$$\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle \geq |\langle \Delta A \Delta B \rangle|^2$$

Okay, where has this gotten you? It's time to be clever. Note that you can write  $\Delta A \Delta B$  as

$$\Delta A \Delta B = \frac{1}{2} [\Delta A, \Delta B] + \frac{1}{2} \{ \Delta A, \Delta B \}$$

Here,  $\{ \Delta A, \Delta B \} = \Delta A \Delta B + \Delta B \Delta A$  is the anticommutator of the operators  $\Delta A$  and  $\Delta B$ . Because  $[\Delta A, \Delta B] = [A, B]$  (the constants  $\langle A \rangle$  and  $\langle B \rangle$  subtract out), you can rewrite this equation:

$$\Delta A \Delta B = \frac{1}{2} [A, B] + \frac{1}{2} \{ \Delta A, \Delta B \}$$

Here's where the math gets intense. Take a look at what you know so far:

- ✓ The commutator of two Hermitian operators,  $[A, B]$ , is anti-Hermitian.
- ✓ The expectation value of an anti-Hermitian is imaginary.
- ✓  $\{ \Delta A, \Delta B \}$  is Hermitian.
- ✓ The expectation value of a Hermitian is real.

All this means that you can view the expectation value of the equation as the sum of real ( $\{ \Delta A, \Delta B \}$ ) and imaginary ( $[A, B]$ ) parts, so

$$|\langle \Delta A \Delta B \rangle|^2 = \frac{1}{4} |\langle [A, B] \rangle|^2 + \frac{1}{4} |\{ \Delta A, \Delta B \}|^2$$

And because the second term on the right is positive or zero, you can say that the following is true:

$$|\langle \Delta A \Delta B \rangle|^2 \geq \frac{1}{4} |\langle [A, B] \rangle|^2$$

Whew! But now compare this equation to the relationship from the earlier use of the Schwarz inequality:

$$\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle \geq |\langle \Delta A \Delta B \rangle|^2$$

Combining the two equations gives you this:

$$\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle \geq \frac{1}{4} |\langle [A, B] \rangle|^2$$

This has the look of the Heisenberg uncertainty relation, except for the pesky expectation value brackets,  $\langle \rangle$ , and the fact that  $\Delta A$  and  $\Delta B$  appear squared here. You want to reproduce the Heisenberg uncertainty relation here, which looks like this:

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}$$

Okay, so how do you get the left side of the equation from  $\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle$  to  $\Delta A \Delta B$ ? Because an earlier equation tells you that  $\Delta A = A - \langle A \rangle$ , you know the following:

$$\langle \Delta A^2 \rangle = \langle A^2 + \langle A \rangle^2 - 2A \langle A \rangle \rangle$$

Taking the expectation value of the last term in this equation, you get this result:

$$\langle \Delta A^2 \rangle = \langle A^2 + \langle A \rangle^2 - 2A \langle A \rangle \rangle = \langle A^2 \rangle - \langle A \rangle^2$$

Square the earlier equation  $\Delta A = (\langle A^2 \rangle - \langle A \rangle^2)^{1/2}$  to get the following:

$$\Delta A^2 = \langle A^2 \rangle - \langle A \rangle^2$$

And comparing that equation to the before it, you conclude that

$$\langle \Delta A^2 \rangle = \Delta A^2$$

Cool. That result means that  $\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle \geq \frac{1}{4} |\langle [A, B] \rangle|^2$  becomes

$$\Delta A^2 \Delta B^2 \geq \frac{1}{4} |\langle [A, B] \rangle|^2$$

This inequality at last means that

$$\Delta A \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle|$$

Well, well, well. So the product of two uncertainties is greater than or equal to  $\frac{1}{2}$  the absolute value of the commutator of their respective operators? Wow. Is that the Heisenberg uncertainty relation? Well, take a look. In quantum mechanics, the momentum operator looks like this:

$$P = -i\hbar\nabla$$

And the operator for the momentum in the  $x$  direction is

$$P_x = -i\hbar \frac{\partial}{\partial x}$$

So what's the commutator of the  $X$  operator (which just returns the  $x$  position of a particle) and  $P_x$ ?  $[X, P_x] = i\hbar$ , so from  $\Delta A \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle|$ , you get

this next equation (remember,  $\Delta x$  and  $\Delta p_x$  here are the uncertainties in  $x$  and  $\Delta p_x$ , not the operators):

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}$$

Hot dog! That is the Heisenberg uncertainty relation. (Notice that by deriving it from scratch, however, you haven't actually constrained the physical world through the use of abstract mathematics — you've merely proved, using a few basic assumptions, that you can't *measure* the physical world with perfect accuracy.)

## ***Eigenvectors and Eigenvalues: They're Naturally Eigentastic!***

As you know if you've been following along in this chapter, applying an operator to a ket can result in a new ket:

$$A|\psi\rangle = |\chi\rangle$$



To make things easier, you can work with eigenvectors and eigenvalues (*eigen* is German for “innate” or “natural”). For example,  $|\psi\rangle$  is an *eigenvector* of the operator  $A$  if

✓ The number  $a$  is a complex constant

✓  $A|\psi\rangle = a|\psi\rangle$

Note what’s happening here: Applying  $A$  to one of its eigenvectors,  $|\psi\rangle$ , gives you  $|\psi\rangle$  back, multiplied by that eigenvector’s *eigenvalue*,  $a$ .

Although  $a$  can be a complex constant, the eigenvalues of Hermitian operators are real numbers, and their eigenvectors are orthogonal (that is,  $\langle\psi|\phi\rangle = 0$ ).

Casting a problem in terms of eigenvectors and eigenvalues can make life a lot easier because applying the operator to its eigenvectors merely gives you the same eigenvector back again, multiplied by its eigenvalue — there’s no pesky change of state, so you don’t have to deal with a different state vector.

Take a look at this idea, using the  $R$  operator from rolling the dice, which is expressed this way in matrix form (see the earlier section “I expected that: Finding expectation values” for more on this matrix):

$$R = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 11 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \end{bmatrix}$$

The  $R$  operator works in 11-dimensional space and is Hermitian, so there’ll be 11 orthogonal eigenvectors and 11 corresponding eigenvalues.

Because  $R$  is a diagonal matrix, finding the eigenvectors is easy. You can take unit vectors in the eleven different directions as the eigenvectors. Here’s what the first eigenvector,  $\xi_1$ , would look like:

$$\xi_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

And here's what the second eigenvector,  $\xi_2$ , would look like:

$$\xi_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

And so on, up to  $\xi_{11}$ :

$$\xi_{11} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Note that all the eigenvectors are orthogonal.

And the eigenvalues? They're the numbers you get when you apply the R operator to an eigenvector. Because the eigenvectors are just unit vectors in all 11 dimensions, the eigenvalues are the numbers on the diagonal of the R matrix: 2, 3, 4, and so on, up to 12.

## Understanding how they work



The eigenvectors of a Hermitian operator define a complete set of orthonormal vectors — that is, a complete basis for the state space. When viewed in this “eigenbasis,” which is built of the eigenvectors, the operator in matrix format is diagonal and the elements along the diagonal of the matrix are the eigenvalues.

This arrangement is one of the main reasons working with eigenvectors is so useful; your original operator may have looked something like this (**Note:** Bear in mind that the elements in an operator can also be functions, not just numbers):

$$R = \begin{bmatrix} 0 & 0 & 1 & 0 & 6 & 0 & 0 & 3 & 0 & 0 & 00 \\ 0 & 1 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 12 & 0 & 0 & 2 & 0 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 6 & 0 & 0 & 0 & 8 & 0 & 0 \\ 1 & 0 & 3 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 77 \\ 0 & 0 & 0 & 0 & 11 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 7 & 0 & 0 & 0 & 88 & 0 & 0 & 0 & 0 \end{bmatrix}$$

By switching to the basis of eigenvectors for the operator, you diagonalize the matrix into something more like what you’ve seen, which is much easier to work with:

$$R = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 11 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \end{bmatrix}$$

You can see why the term *eigen* is applied to eigenvectors — they form a natural basis for the operator.

If two or more of the eigenvalues are the same, that eigenvalue is said to be *degenerate*. So for example, if three eigenvalues are equal to 6, then the eigenvalue 6 is threefold degenerate.



Here’s another cool thing: If two Hermitian operators, A and B, commute, and if A doesn’t have any degenerate eigenvalues, then each eigenvector of A is also an eigenvector of B. (See the earlier section “Forward and Backward: Finding the Commutator” for more on commuting.)

## Finding eigenvectors and eigenvalues

So given an operator in matrix form, how do you find its eigenvectors and eigenvalues? This is the equation you want to solve:

$$A|\psi\rangle = a|\psi\rangle$$

And you can rewrite this equation as the following:

$$(A - aI)|\psi\rangle = 0$$



$I$  represents the identity matrix, with 1s along its diagonal and 0s otherwise:

$$I = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ & & & & & \vdots & & & & & \ddots \end{bmatrix}$$



The solution to  $(A - aI) |\psi\rangle = 0$  exists only if the determinant of the matrix  $A - aI$  is 0:

$$\det(A - aI) = 0$$

### ***Finding eigenvalues***

Any values of  $a$  that satisfy the equation  $\det(A - aI) = 0$  are eigenvalues of the original equation. Try to find the eigenvalues and eigenvectors of the following matrix:

$$A = \begin{bmatrix} -1 & -1 \\ 2 & -4 \end{bmatrix}$$

First, convert the matrix into the form  $A - aI$ :

$$A - aI = \begin{bmatrix} -1-a & -1 \\ 2 & -4-a \end{bmatrix}$$

Next, find the determinant:

$$\det(A - aI) = (-1 - a)(-4 - a) + 2$$

$$\det(A - aI) = a^2 + 5a + 6$$

And this can be factored as follows:

$$\det(A - aI) = a^2 + 5a + 6 = (a + 2)(a + 3)$$

You know that  $\det(A - aI) = 0$ , so the eigenvalues of  $A$  are the roots of this equation; namely,  $a_1 = -2$  and  $a_2 = -3$ .

***Finding eigenvectors***

How about finding the eigenvectors? To find the eigenvector corresponding to  $a_1$  (see the preceding section), substitute  $a_1$  — the first eigenvalue,  $-2$  — into the matrix in the form  $A - aI$ :

$$A - aI = \begin{bmatrix} -1 - a & -1 \\ 2 & -4 - a \end{bmatrix}$$

$$A - aI = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix}$$

$$(A - aI)|\psi\rangle = 0$$

So you have

$$\begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Because every row of this matrix equation must be true, you know that  $\psi_1 = \psi_2$ . And that means that, up to an arbitrary constant, the eigenvector corresponding to  $a_1$  is the following:

$$c \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Drop the arbitrary constant, and just write this as a matrix:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

How about the eigenvector corresponding to  $a_2$ ? Plugging  $a_2$ ,  $-3$ , into the matrix in  $A - aI$  form, you get the following:

$$A - aI = \begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix}$$

Then you have

$$\begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So  $2\psi_1 - \psi_2 = 0$ , and  $\psi_1 = \psi_2 \div 2$ . And that means that, up to an arbitrary constant, the eigenvector corresponding to  $a_2$  is

$$c \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Drop the arbitrary constant:

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

So the eigenvalues of this next matrix operator

$$A = \begin{bmatrix} -1 & -1 \\ 2 & -4 \end{bmatrix}$$

are  $a_1 = -2$  and  $a_2 = -3$ . And the eigenvector corresponding to  $a_1$  is

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The eigenvector corresponding to  $a_2$  is

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

## *Preparing for the Inversion: Simplifying with Unitary Operators*

Applying the inverse of an operator undoes the work the operator did:

$$A^{-1}A = AA^{-1} = I$$

Sometimes, finding the inverse of an operator is helpful, such as when you want to solve equations like  $Ax = y$ . Solving for  $x$  is easy if you can find the inverse of  $A$ :  $x = A^{-1}y$ .

However, finding the inverse of a large matrix often isn't easy, so quantum physics calculations are sometimes limited to working with unitary operators,  $U$ , where the operator's inverse is equal to its adjoint,  $U^{-1} = U^\dagger$ . (To find the adjoint of an operator,  $A$ , you find the transpose by interchanging the rows and columns,  $A^T$ . Then take the complex conjugate,  $A^{T*} = A^\dagger$ .) This gives you the following equation:

$$U^\dagger U = UU^\dagger = I$$

The product of two unitary operators,  $U$  and  $V$ , is also unitary because

$$(UV)(UV)^\dagger = (UV)(V^\dagger U^\dagger) = U(VV^\dagger)U^\dagger = UU^\dagger = I$$

When you use unitary operators, kets and bras transform this way:

$$\checkmark \quad |\psi'\rangle = U|\psi\rangle$$

$$\checkmark \quad \langle\psi'| = \langle\psi|U^\dagger$$

And you can transform other operators using unitary operators like this:

$$A' = UAU^\dagger$$

Note that the preceding equations also mean the following:

$$\checkmark \quad |\psi\rangle = U^\dagger|\psi'\rangle$$

$$\checkmark \quad \langle\psi| = \langle\psi'|U$$

$$\checkmark \quad A = U^\dagger A' U$$



Here are some properties of unitary transformations:

- ✓ If an operator is Hermitian, then its unitary transformed version,  $A' = UAU^\dagger$ , is also Hermitian.
- ✓ The eigenvalues of  $A$  and its unitary transformed version,  $A' = UAU^\dagger$ , are the same.
- ✓ Commutators that are equal to complex numbers are unchanged by unitary transformations:  $[A', B'] = [A, B]$ .

## Comparing Matrix and Continuous Representations

Werner Heisenberg developed the matrix-oriented view of quantum physics that you've been using so far in this chapter. It's sometimes called *matrix mechanics*. The matrix representation is fine for many problems, but sometimes you have to go past it, as you're about to see.

One of the central problems of quantum mechanics is to calculate the energy levels of a system. The energy operator is called the *Hamiltonian*,  $H$ , and finding the energy levels of a system breaks down to finding the eigenvalues of the problem:

$$H|\psi\rangle = E|\psi\rangle$$

Here,  $E$  is an eigenvalue of the  $H$  operator.

Here's the same equation in matrix terms:

$$\begin{bmatrix} H_{11}-E & H_{12} & H_{13} & H_{14} & \cdots \\ H_{21} & H_{22}-E & H_{23} & H_{24} & \cdots \\ H_{31} & H_{32} & H_{33}-E & H_{34} & \cdots \\ H_{41} & H_{42} & H_{43} & H_{44}-E & \cdots \\ & & \vdots & & \ddots \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \\ \vdots \end{bmatrix} = 0$$

The allowable energy levels of the physical system are the eigenvalues  $E$ , which satisfy this equation. These can be found by solving the characteristic polynomial, which derives from setting the determinant of the above matrix to zero, like so

$$\det \begin{bmatrix} H_{11}-E & H_{12} & H_{13} & H_{14} & \cdots \\ H_{21} & H_{22}-E & H_{23} & H_{24} & \cdots \\ H_{31} & H_{32} & H_{33}-E & H_{34} & \cdots \\ H_{41} & H_{42} & H_{43} & H_{44}-E & \cdots \\ & & \vdots & & \ddots \end{bmatrix} = 0$$

That's fine if you have a discrete basis of eigenvectors — if the number of energy states is finite. But what if the number of energy states is infinite? In that case, you can no longer use a discrete basis for your operators and bras and kets — you use a *continuous* basis.

## Going continuous with calculus

Representing quantum mechanics in a continuous basis is an invention of the physicist Erwin Schrödinger. In the continuous basis, summations become integrals. For example, take the following relation, where  $I$  is the identity matrix:

$$\sum_{n=1}^{\infty} |\phi_n\rangle \langle \phi_n| = I$$

It becomes the following:

$$\int d\phi |\phi\rangle \langle \phi| = I$$

And every ket  $|\psi\rangle$  can be expanded in a basis of other kets,  $|\phi_n\rangle$ , like this:

$$|\psi\rangle = \int d\phi |\phi\rangle \langle \phi|\psi\rangle$$

## Doing the wave

Take a look at the position operator,  $R$ , in a continuous basis. Applying this operator gives you  $\mathbf{r}$ , the position vector:

$$R|\psi\rangle = \mathbf{r}|\psi\rangle$$

In this equation, applying the position operator to a state vector returns the locations,  $r$ , that a particle may be found at. You can expand any ket in the position basis like this:

$$|\psi\rangle = \int d^3r |\mathbf{r}\rangle \langle \mathbf{r}|\psi\rangle$$

And this becomes

$$|\psi\rangle = \int d^3r \psi(\mathbf{r}) |\mathbf{r}\rangle$$



Here's a very important thing to understand:  $\psi(\mathbf{r}) = \langle \mathbf{r}|\psi\rangle$  is the *wave function* for the state vector  $|\psi\rangle$  — it's the ket's representation in the position basis. Or in common terms, it's just a function where the quantity  $|\psi(\mathbf{r})|^2 d^3r$  represents the probability that the particle will be found in the region  $d^3r$  at  $\mathbf{r}$ .

The wave function is the foundation of what's called *wave mechanics*, as opposed to matrix mechanics. What's important to realize is that when you talk about representing physical systems in wave mechanics, you don't use the basis-less bras and kets of matrix mechanics; rather, you usually use the wave function — that is, bras and kets in the position basis.

Therefore, you go from talking about  $|\psi\rangle$  to  $\langle r|\psi\rangle$ , which equals  $\psi(r)$ . This wave function appears a lot in the coming chapters, and it's just a ket in the position basis. So in wave mechanics,  $H|\psi\rangle = E|\psi\rangle$  becomes the following:

$$\langle r|H|\psi\rangle = E\langle r|\psi\rangle$$

You can write this as the following:

$$\langle r|H|\psi\rangle = E\psi(r)$$

But what is  $\langle r|H|\psi\rangle$ ? It's equal to  $H\psi(r)$ . The Hamiltonian operator,  $H$ , is the total energy of the system, kinetic ( $p^2/2m$ ) plus potential ( $V(r)$ ) so you get the following equation:

$$H = \frac{p^2}{2m} + V(r)$$

But the momentum operator is

$$p|\psi\rangle = -i\hbar\frac{\partial}{\partial x}|\psi\rangle = \mathbf{i} - i\hbar\frac{\partial}{\partial y}|\psi\rangle = \mathbf{j} - i\hbar\frac{\partial}{\partial z}|\psi\rangle = \mathbf{k}$$

Therefore, substituting the momentum operator for  $p$  gives you this:

$$H = \frac{-\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V(r)$$

Using the Laplacian operator, you get this equation:

$$\nabla^2|\psi\rangle = \frac{\partial^2}{\partial x^2}|\psi\rangle + \frac{\partial^2}{\partial y^2}|\psi\rangle + \frac{\partial^2}{\partial z^2}|\psi\rangle$$

You can rewrite this equation as the following (called the *Schrödinger equation*):

$$H\psi(r) = \frac{-\hbar^2}{2m} \nabla^2\psi(r) + V(r)\psi(r) = E\psi(r)$$

So in the wave mechanics view of quantum physics, you're now working with a differential equation instead of multiple matrices of elements. This all came from working in the position basis,  $\psi(r) = \langle r | \psi \rangle$  instead of just  $|\psi\rangle$ .

The quantum physics in the rest of the book is largely about solving this differential equation for a variety of potentials,  $V(r)$ . That is, your focus is on finding the wave function that satisfies the Schrödinger equation for various physical systems. When you solve the Schrödinger equation for  $\psi(r)$ , you can find the allowed energy states for a physical system, as well as the probability that the system will be in a certain position state.

Note that, besides wave functions in the position basis, you can also give a wave function in the momentum basis,  $\psi(p)$ , or in any number of other bases.



The Heisenberg technique of matrix mechanics is one way of working with quantum physics, and it's best used for physical systems with well-defined energy states, such as harmonic oscillators. The Schrödinger way of looking at things, wave mechanics, uses wave functions, mostly in the position basis, to reduce questions in quantum physics to a differential equation.