
Quantum Mechanics

Diez B. Borja
June 6, 2022

Preface

*This notes has been written by **B. Diez** mainly from Introduction to Quantum Mechanics by David J. Griffiths and own knowledge.*

Contents

Preface	i
1 The Wave Function	2
1.1 The Schrodinger Equation	2
1.2 The Statistical Interpretation	2
1.3 Probability	3
1.4 Continous Variables	4
1.5 Normalization	4
1.6 Momentum	5
1.7 The Uncertainty Principle	6
2 Time-Independent Schrodinger Equation	7
2.1 Stationary States	7
2.2 The Infinite Square Well	10
2.3 The Harmonic Oscillator	14
2.3.1 Algebraic Method	15
2.3.2 Analytic Method	19
3 Formalism	22
3.1 Hilbert Space	22
3.2 Obervables	24
3.2.1 Hermitian Operators	24
3.2.2 Determinate States	24
3.3 Eigenfunctions of a Hermitian Operator	25
3.3.1 Discrete Spectra	25
3.3.2 Continuous Spectra	26
3.3.3 Generalized Statistical Interpretation	26
3.4 The Uncertainty Principle	27
3.5 Dirac Notation	27
4 Quantum Mechanics In Three Dimensions	31
4.1 Schrodinger Equation In Spherical Coordinates	31
4.2 The Hydrogen Atom	33
4.3 Angular Momentum	36
4.4 Spin	41

Chapter 1

The Wave Function

1.1 The Schrodinger Equation

Imagine a particle of mass m , constrained to move along the x -axis, subject to some specified force $F(x, t)$. The program of *classical* mechanics is to determine the position of the particle at any given time: $x(t)$. Once we know that, we can figure out the velocity ($v = dx/dt$), the momentum ($p = mv$), the kinetic energy ($T = (1/2)mv^2$), or any other dynamical variable of interest. And how do we go about determining $x(t)$? We apply Newton's second law: $F = ma$. (For *conservative* systems—the only kinds we shall consider, and, fortunately, the only kinds that occur at the microscopic level—the force can be expressed as the derivative of a potential energy function¹, $F = -\partial V/\partial x$, and Newton's law reads $m d^2x/dt^2 = -\partial V/\partial x$) This, together with appropriate initial conditions (typically the position and velocity at $t = 0$), determines $x(t)$.

Quantum mechanics approaches this same problem quite differently. In this case what we're looking for is the particle's **wave function**, $\Psi(x, t)$, and we get it by solving the **Schrodinger equation**:

$$\boxed{i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V\Psi} \quad (1.1) \quad \text{Schrodinger equation}$$

This equation plays a role logically analogous to Newton's second law: Given suitable initial conditions (typically, $\Psi(x, 0)$), the Schrodinger equation determines $\Psi(x, t)$ for all future time, just as, in classical mechanics, Newton's law determines $x(t)$ for all future time.

1.2 The Statistical Interpretation

What exactly is this "wave function"? Born's **statistical interpretation** of the wave function, says that $|\Psi(x, t)|^2$ gives the **probability** of finding the particle at point x , at time t , or more precisely

$$\boxed{\int_a^b |\Psi(x, t)|^2 dx = \text{probability of finding the particle between } a \text{ and } b, \text{ at time } t} \quad (1.2) \quad \text{Statistical interpretation}$$

¹Magnetic forces are an exception, but let's not worry about them just yet. By the way, we shall assume for the moment that the motion is nonrelativistic ($v \ll c$)

Probability is the *area* under the graph of $|\Psi(x, t)|^2$

The statistical interpretation introduces a kind of **indeterminacy** into quantum mechanics, for even if you know everything the theory has to tell you about the particle /to wit: its wave function, still you cannot predict with certainty the outcome of a simple experiment to measure its position– all quantum mechanics has to offer is *statistical* information about the *possible* results.

We say that the wave function **collapses**, upon measurement, to spike some point. There are, then, two entirely distinct kinds of physical processes: "ordinary" ones, in which the wave function evolves in a leisurely fashion under the Schrodinger equation, and "measurements", in which Ψ suddenly and discontinuously collapses ².

1.3 Probability

I will give some definitions. The total number is

$$N = \sum_{j=0}^{\infty} N(j) \quad (1.3)$$

The probability of getting j is

$$P(j) = \frac{N(j)}{N} \quad (1.4)$$

In particular, the sum of all the probabilities is 1:

$$\sum_{j=0}^{\infty} P(j) = 1 \quad (1.5)$$

In general, the average value of some *function* of j es given by

$$\boxed{\langle f(j) \rangle = \sum_{j=0}^{\infty} f(j)P(j)} \quad (1.6) \quad \text{Average value of a function}$$

In quantum mechanics the average is usually the quantity of interest; in that context it has come to be called the **expectation value**.

Now we need a numerical measure of the amount of "spread" in a distribution, with respect to the average. The most obvious way to do this would be to find out how far each individual deviates from the average,

$$\Delta j = j - \langle j \rangle \quad (1.7)$$

and compute the average of Δj . Trouble is, of course, that you get zero, by the nature of the average, Δj is as often negative as positive. To avoid this irritating problem you might decide

²The role of measurement in quantum mechanics is so critical and so bizarre that you may well be wondering what precisely constitutes a measurement. Does it have to do with the interaction between a microscopic (quantum) system and a macroscopic (classical) measuring apparatus (as Bohr insisted), or is it intervention of a conscious "observer" (as Wigner proposed)? For the moment let's take the naive view: A measurement is the kind of thing that a scientist does in the laboratory, with rules, stopwatches, and so on

to average the *absolute value* of Δj . But absolute values are nasty to work with; instead, we get around the sign problem by *squaring* before averaging:

$$\sigma^2 \equiv \langle (\Delta j)^2 \rangle \quad (1.8) \quad \text{Variance}$$

This quantity is known as the **variance** of the distribution; σ itself is called the **standard deviation**.

There is a useful little theorem on variances:

$$\sigma = \sqrt{\langle j^2 \rangle - \langle j \rangle^2} \quad (1.9)$$

1.4 Continuous Variables

I have assumed that we are dealing with a *discrete* variable—that is, one that can take on only isolated values. But it is simple enough to generalize to *continuous* distributions. If the interval is sufficiently short, this probability is *proportional to the length of the interval*. This

$$\text{probability that an individual lies between } x \text{ and } (x + dx) = \rho(x)dx \quad (1.10) \quad \text{Probability density}$$

The proportionality factor $\rho(x)$, is often loosely called "the probability of getting x ", but is sloppy language; a better term is **probability density**. The probability that x lies between a and b (a finite interval) is given by the integral of $\rho(x)$:

$$P_{ab} = \int_a^b \rho(x)dx \quad (1.11)$$

and the rules we deduced for discrete distributions translate in the obvious way:

$$1 = \int_{-\infty}^{+\infty} \rho(x)dx \quad (1.12)$$

$$\langle x \rangle = \int_{-\infty}^{+\infty} x\rho(x)dx \quad (1.13)$$

$$\langle f(x) \rangle = \int_{-\infty}^{+\infty} f(x)\rho(x)dx \quad (1.14)$$

$$\sigma^2 \equiv \langle (\Delta x)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \quad (1.15)$$

1.5 Normalization

We return now to the statistical interpretation of the wave function (1.2), which says that $|\Psi(x, t)|^2$ is the probability density for finding the particle at point x , at time t . It follows (1.5) that the integral of $|\Psi|^2$ must be 1 (the particle's got to be *somewhere*):

$$\boxed{\int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx = 1} \quad (1.16)$$

If $\Psi(x, t)$ is a solution of the Schrodinger equation, so too is $A\Psi(x, t)$, where A is any (complex) constant. What we must do, then, is pick this undetermined multiplicative factor so as to ensure that (1.16) is satisfied. This process is called **normalizing** the wave function. For some solutions to the Schrodinger equation the integral is *infinite*; in that case *no* multiplicative factor is going to make it 1. The same goes for the trivial solution $\Psi = 0$. Such **non-normalizable** solutions cannot represent particles, and must be rejected. Physically realizable states correspond to the **square-integrable** solutions to Schrodinger equation.³

Fortunately, the Schrodinger equation has the remarkable property that it automatically preserves the normalization of the wave function.

1.6 Momentum

For a particle in state Ψ , the expectation value of x is

$$\langle x \rangle = \int_{-\infty}^{+\infty} x |\Psi(x, t)|^2 dx \quad (1.17) \quad \text{Expectation value of } x$$

What exactly does this mean? It emphatically does not mean that if you measure the position of one particle over and over, $\int x |\Psi|^2 dx$ is the average of the results you'll get. Rather, $\langle x \rangle$ is the average of measurements, performed on particles *all in the state* Ψ , which means that either you must find some way to returning the particle to its original state after each measurement, or else you have to prepare a whole **ensemble** of particles, each in the same state Ψ , and measure the positions of all of them: $\langle x \rangle$ is the average of *these* results.

Now, as time goes on, $\langle x \rangle$ will change (because of the time dependence of Ψ), and we might be interested in knowing how fast it moves. We see that

$$\frac{d\langle x \rangle}{dt} = -\frac{i\hbar}{m} \int \Psi^* \frac{\partial \Psi}{\partial x} dx \quad (1.18)$$

The *expectations value of the velocity is equal to the time derivative of the expectation value of position*:

$$\langle v \rangle = \frac{d\langle x \rangle}{dt} \quad (1.19)$$

(1.18) tell us, then, how to calculate $\langle v \rangle$ directly from Ψ .

Actually, it is customary to work with **momentum** ($p = mv$), rather than velocity:

$$\langle p \rangle = m \frac{d\langle x \rangle}{dt} = -i\hbar \int \left(\Psi^* \frac{\partial \Psi}{\partial x} \right) dx \quad (1.20) \quad \text{Momentum}$$

Let me write the expressions for $\langle x \rangle$ and $\langle p \rangle$ in a more suggestive way:

$$\langle x \rangle = \int \Psi^*(x) \Psi dx \quad (1.21)$$

$$\langle p \rangle = \int \Psi^* \left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right) \Psi dx \quad (1.22)$$

³Evidently $\Psi(x, t)$ must go to zero faster than $1/\sqrt{|x|}$, as $|x| \rightarrow \infty$. Incidentally, normalization only fixed the *modulus* of A : the phase remains undetermined. However, as we shall, the latter carries no physical significance anyway.

We say that the **operator** x "represents" position, and the operator $(\hbar/i)(\partial/\partial x)$ "represents" momentum, in quantum mechanics; to calculate expectation values we "sandwich" the appropriate operator between Ψ^* and Ψ , and integrate.

What about other quantities? The fact is, all classical dynamics variables can be expressed in terms of position and momentum.

To calculate the expectation value of any such quantity, $Q(x, p)$, we simply replace every p by $(\hbar/i)(\partial/\partial x)$, insert the resulting operator between Ψ^* and Ψ , and integrate:

$$\langle Q(x, p) \rangle = \int \Psi^* Q \left(x, \frac{\hbar}{i} \frac{\partial}{\partial x} \right) \Psi dx \quad (1.23)$$

1.7 The Uncertainty Principle

The wavelength of Ψ is related to the *momentum* of the particle by the **de Broglie formula**:

$$p = \frac{\hbar}{\lambda} = \frac{2\pi\hbar}{\lambda} \quad (1.24) \quad \text{deBroglie formula}$$

Thus a spread in *wavelength* corresponds to a spread in *momentum*, and our general observation now says that the more precisely determined a particle's position is, the less precisely is its momentum. Quantitatively.

$$\sigma_x \sigma_p \geq \frac{\hbar}{2} \quad (1.25) \quad \text{Uncertainty principle}$$

where σ_x is the standard deviation in x , and σ_p is the standard deviation in p .

Please understand what the uncertainty principle means: Like position measurements, momentum measurements yield precise answers—the "spread" here refers to the fact that measurements on identically prepared systems do not yield identical results. You can, if you want, construct a state such the repeated position measurements will be very close together (by making Ψ a localized "spike"), but you still pay a price: Momentum measurements on this state will be widely scattered. Or you can prepare a state with a reproducible momentum (by making Ψ a long sinusoidal wave), but in that case, position measurements will be widely scattered.

Chapter 2

Time-Independent Schrodinger Equation

2.1 Stationary States

We talked a lot about the wave function, and how you use it to calculate various quantities of interest. But, how do you get $\Psi(x, t)$ in the first place? We need to solve the Schrodinger equation,

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V\Psi \quad (2.1)$$

for a specified potential¹ $V(x, t)$. In this chapter (and most of this book) I shall assume that V is *independent* of t . In that case the Schrodinger equation can be solved by the method of **separation of variables**: We look for solutions that are simple *products*,

$$\Psi(x, t) = \psi(x)\varphi(t) \quad (2.2)$$

where ψ is a function of x alone, and φ is a function of t alone.

For separable solutions we have

$$\frac{\partial \Psi}{\partial t} = \psi \frac{d\varphi}{dt}, \quad \frac{\partial^2 \Psi}{\partial x^2} = \frac{d^2 \psi}{dx^2} \varphi$$

(ordinary derivatives, now), and Schrodinger equation reads

$$i\hbar \frac{d\varphi}{dt} = -\frac{\hbar^2}{2m} \frac{d^2 \psi}{dx^2} \varphi + V\psi\varphi \quad (2.3)$$

Or, dividing through by $\psi\varphi$:

$$i\hbar \frac{1}{\varphi} \frac{d\varphi}{dt} = -\frac{\hbar^2}{2m} \frac{1}{\psi} \frac{d^2 \psi}{dx^2} + V \quad (2.4)$$

Now, the left side is a function of t alone, and the right side is a function of x alone². The only way this can possibly be true is if both sides are in fact *constant* – otherwise, by varying

¹It is tiresome to keep saying "potential energy function", so most people just call V the "potential", even though this invites occasional confusion with electric potential, which is actually potential energy per unit charge.

²Note that this would not be true if V were a function of t as well as x .

t , I could change the left side without touching the right side, and the two would no longer equal. For reasons that will appear in a moment, we shall call the separation constant E . Then

$$i\hbar \frac{1}{\varphi} \frac{d\varphi}{dt} = E$$

or

$$\frac{d\varphi}{dt} = -\frac{iE}{\hbar} \varphi$$

and

$$-\frac{\hbar^2}{2m} \frac{1}{\psi} \frac{d^2\psi}{dx^2} + V = E \quad (2.5)$$

or

$$\boxed{-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V\psi = E\psi} \quad (2.6) \quad \text{Time-independent Schrodinger equation}$$

Separation of variables has turned a *partial* differential equation into *two ordinary* differential equations (2.1) and (2.6). The first, (2.1) is easy to solve (just multiply through by dt and integrate); the general solution is $C \exp(-iEt/\hbar)$, but we might as well absorb the constant C into ψ (since the quantity of interest is the product $\psi\varphi$). Then

$$\varphi(t) = e^{-iEt/\hbar} \quad (2.7)$$

The second equation (2.6) is called the *time-independent Schrodinger equation* we can go further with it until the potential $V(x)$ is specified.

The rest of this chapter will be devoted to solving the time-independent Schrodinger equation, for a variety of simple potentials. But before I get to that you have every right ask: What's so great about separable solutions? After all, most solutions to the (time dependent) Schrodinger equation do not take the form $\psi(x)\varphi(t)$. I offer three answers—two of them physical, and one mathematical:

1. They are **stationary states**. Although the wave function itself,

$$\Psi(x, t) = \psi(x)e^{-iEt/\hbar} \quad (2.8)$$

does (obviously) depend on t , the *probability density*,

$$|\Psi(x, t)|^2 = \Psi^* \Psi = \psi^* e^{+iEt/\hbar} \psi e^{-iEt/\hbar} = |\psi(x)|^2 \quad (2.9)$$

does not—the time-dependence cancels out³. The same thing happens in calculating the expectation value of any dynamical variable; (1.23) reduces to

$$\langle Q(x, p) \rangle = \int \psi^* Q \left(x, \frac{\hbar}{i} \frac{d}{dx} \right) \psi dx \quad (2.10)$$

Every expectation value is constant in time; we might as well drop the factor $\varphi(t)$ altogether, and simply use ψ in place of Ψ . In particular, $\langle x \rangle$ is constant, and hence $\langle p \rangle = 0$ (See (1.20)). Nothing ever happens in a stationary state.

³For normalizable solutions, E must be real

2. They are states of *definite total energy*. In classical mechanics, the total energy (kinetic plus potential) is called the **Hamiltonian**:

$$H(x, p) = \frac{p^2}{2m} + V(x) \quad (2.11)$$

The corresponding Hamiltonian *operator*, obtained by the canonical substitution $p \rightarrow (\hbar/i)(\partial/\partial x)$, is therefore⁴

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \quad (2.12)$$

Thus the time-independent Schrodinger equation (2.6) can be written

$$\hat{H}\psi = E\psi \quad (2.13)$$

and the expectation value of the total energy is

$$\langle H \rangle = \int \psi^* \hat{H} \psi dx = E \int |\psi|^2 dx = E \int |\Psi|^2 dx = E \quad (2.14)$$

(Notice that the normalization of Ψ entails the normalization of ψ .) Moreover,

$$\hat{H}^2 \psi = \hat{H}(\hat{H}\psi) = \hat{H}(E\psi) = E(\hat{H}\psi) = E^2 \psi$$

and hence

$$\langle H^2 \rangle = \int \psi^* \hat{H}^2 \psi dx = E^2 \int |\psi|^2 dx = E^2$$

So the variance of H is

$$\sigma_H^2 = \langle H^2 \rangle - \langle H \rangle^2 = E^2 - E^2 = 0 \quad (2.15)$$

But remember, if $\sigma = 0$, then every member of the sample must share the same value (the distribution has zero spread). *Conclusion*: A separable solution has the property that *every measurement of the total energy is certain to return the value E* .

3. The general solution is a **linear combination** of separable solutions. As we're about to discover, the time-independent Schrodinger equation (2.6) yields an infinite collection of solutions $(\psi_1(x), \psi_2(x), \dots)$, each with its associated value of the separation constant (E_1, E_2, \dots) ; thus there is a different wave function for each **allowed energy**:

$$\Psi_1(x, t) = \psi_1(x)e^{-iE_1t/\hbar}, \quad \Psi_2(x, t) = \psi_2(x)e^{-iE_2t/\hbar}, \dots$$

Now the time-dependent Schrodinger equation (2.1) has the property that any linear combination⁵ of solutions is itself a solution. Once we have found the separable solutions, then, we can immediately construct a much more general solution, on the form

⁴Whenever confusion might arise. I'll put a "hat" on the operator, to distinguish it from the dynamical variable it represents.

⁵A **linear combination** of the functions $f_1(z), f_2(z), \dots$ is an expression of the form

$$f(z) = c_1 f_1(z) + c_2 f_2(z) + \dots$$

where c_1, c_2, \dots are any (complex) constants.

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \psi_n(x) e^{-iE_n t/\hbar} \quad (2.16)$$

It so happens that *every* solution to the (time-dependent) Schrodinger equation can be written in this form –it is simply a matter of finding the right constants (c_1, c_2, \dots) so as to fit the initial conditions for the problem at hand.

Let me recapitulate, from a somewhat different perspective. Here's the generic problem: You're given a (time-independent) potential $V(x)$ and the starting wave function $\Psi(x, 0)$; your job is to find the wave function, $\Psi(x, t)$, for any subsequent time t . To do this you must solve the (time-dependent) Schrodinger equation (2.1); this yields, in general, an infinite set of solutions ($\psi_1(x), \psi_2(x), \dots$), each with its own associated energy (E_1, E_2, \dots). To fit $\Psi(x, 0)$ you write down the general linear combination of these solutions:

$$\Psi(x, 0) = \sum_{n=1}^{\infty} c_n \psi_n(x) \quad (2.17)$$

the miracle is that you can *always* match the specified initial state by appropriate choice of the constants c_1, c_2, \dots . To construct $\Psi(x, t)$ you simply tack onto each term its characteristic time dependence, $\exp(-iE_n t/\hbar)$:

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \psi_n(x) e^{-iE_n t/\hbar} = \sum_{n=1}^{\infty} c_n \Psi_n(x, t) \quad (2.18)$$

The separable solutions themselves,

$$\Psi_n(x, t) = \psi_n(x) e^{-iE_n t/\hbar} \quad (2.19)$$

are *stationary* states, in the sense that all probabilities and expectation values are independent of time, but this property is emphatically *not* shared by the general solution (2.18); the energies are different, for different stationary states, and the exponentials do not cancel, when you calculate $|\Psi|^2$.

2.2 The Infinite Square Well

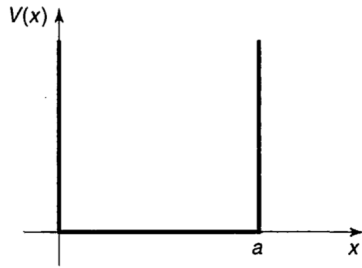


Figure 2.1: The infinite square well potential (2.20)

Suppose

$$V(x) = \begin{cases} 0 & , 0 \leq x \leq a \\ \infty & , \text{otherwise} \end{cases} \quad (2.20)$$

Fig. (2.1). A particle in this potential is completely free, except at two ends ($x = 0$ and $x = a$), where an infinite force prevents it from escaping.

Outside the well, $\psi(x) = 0$ (the probability of finding the particle there is zero) Inside the well, where $V = 0$, the time-independent Schrodinger equation (2.6) reads

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi \quad (2.21)$$

or

$$\frac{d^2\psi}{dx^2} = -k^2\psi, \quad \text{where } k \equiv \frac{\sqrt{2mE}}{\hbar} \quad (2.22)$$

(By writing it in this way, I have tacitly assumed that $E \geq 0$.) Equation (2.22) is the classical **simple harmonic oscillator** equation; the general solution is

$$\psi(x) = A \sin kx + B \cos kx \quad (2.23)$$

where A and B are arbitrary constants. Typically, these constants are fixed by the **boundary conditions** of the problem. What are the appropriate boundary conditions for $\psi(x)$? Ordinarily, both ψ and $d\psi/dx$ are continuous, but where the potential goes to infinity only the first of these applies.

Continuity of $\psi(x)$ requires that

$$\psi(0) = \psi(a) = 0 \quad (2.24)$$

so as to join onto the solution outside the well. What does this tell us about A and B ? Well,

$$\psi(0) = A \sin 0 + B \cos 0 = B$$

so $B = 0$, and hence

$$\psi(x) = A \sin kx \quad (2.25)$$

Then $\psi(a) = A \sin ka$, so either $A = 0$ (in which case we're left with the trivial non-normalizable solution $\psi(x) = 0$), or else $\sin ka = 0$, which means that

$$ka = 0, \pm\pi, \pm2\pi, \pm3\pi, \dots \quad (2.26)$$

But $k = 0$ is no good (again, that would imply $\psi(x) = 0$), and the negative solutions give nothing new, since $\sin(-\theta) = -\sin(\theta)$ and we can absorb the minus sign into A . So the *distinct* solutions are

$$k_n = \frac{n\pi}{a}, \quad \text{with } n = 1, 2, 3, \dots \quad (2.27)$$

Curiously, the boundary condition at $x = a$ does not determine the constant A , but rather the constant k , and hence the possible values of E :

$$\boxed{E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{n^2 \pi^2 \hbar^2}{2ma^2}} \quad (2.28) \quad \begin{array}{l} \text{Possible values} \\ \text{of } E \end{array}$$

In radical contrast to the classical case, a quantum particle in the infinite square well cannot have just any old energy—it has to be one of these special **allowed** values⁶. To find A we normalize ψ :

$$\int_a^a |A|^2 \sin^2(kx) dx = |A|^2 \frac{a}{2} = 1, \quad \text{so } |A|^2 = \frac{2}{a}$$

This only determines the magnitude of A , but it is simpler to pick the positive real root: $A = \sqrt{2/a}$ (the phase of A carries no physical significance anyway). Inside the well, the solutions are

$$\boxed{\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right)} \quad (2.29)$$

As a collection, the functions $\psi_n(x)$ have some interesting and important properties:

1. They are alternately **even** and **odd**, with respect to the center of the well: ψ_1 is even, ψ_2 is odd, ψ_3 is even, and so on.
2. As you go up in energy, each successive state has one more **node** (zero-crossing): ψ_1 has one (the end points don't count), ψ_4 has one, ψ_3 has two, and so on.
3. They are mutually **orthogonal**, in the sense that

$$\int \psi_m(x)^* \psi_n(x) dx = 0 \quad (2.30)$$

whenever $m \neq n$. We can combine orthogonality and normalization into a single statement⁷:

$$\boxed{\int \psi_m(x)^* \psi_n(x) dx = \delta_{mn}} \quad (2.31)$$

We say that the ψ 's are **orthonormal**.

4. They are **complete**, in the sense that any other function, $f(x)$, can be expressed as a linear combination of them:

$$f(x) = \sum_{n=1}^{\infty} c_n \psi_n(x) = \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} c_n \sin\left(\frac{n\pi}{a}x\right) \quad (2.32)$$

(2.32) is nothing but the **Fourier series** for $f(x)$, and the fact that any function can be expanded in this way is sometimes called **Dirichlet's theorem**. The coefficients c_n can be evaluated—for a given $f(x)$ —by a method I call **Fourier's trick**, which beautifully exploits the orthonormality of $\{\psi_n\}$: Multiply both sides of (2.32) by $\psi_m(x)^*$, and integrate

$$\psi_m(x)^* f(x) dx = \sum_{n=1}^{\infty} c_n \int \psi_m(x)^* \psi_n(x) dx = \sum_{n=1}^{\infty} c_n \delta_{nm} = c_n \quad (2.33)$$

⁶Notice that the quantization of energy emerged as a rather technical consequence of the boundary conditions on solutions to the time-independent Schrödinger equation

⁷In this case the ψ 's are real, so the $*$ on ψ_m is unnecessary, but for future purposes it's a good idea to get the habit of putting it there.

Thus the n th coefficient in the expansion of $f(x)$ is

$$c_n = \int \psi_n(x)^* f(x) dx \quad (2.34)$$

Completeness holds for all the potentials you are likely to encounter, but the proofs tend to be nasty and laborious.

The stationary states (2.19) of the infinite square well are evidently

$$\Psi_n(x, t) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) e^{-i(n^2\pi^2\hbar/2ma^2)t} \quad (2.35)$$

I claimed (2.18) that the most general solution to the (time-dependent) Schrodinger equation is a linear combinations of stationary states:

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) e^{-i(n^2\pi^2\hbar/2ma^2)t} \quad (2.36)$$

It remains only for me to demonstrate that I can fit any prescribed initial wave function, $\Psi(x, 0)$, by appropriate choice of the coefficients c_n :

$$\Psi(x, 0) = \sum_{n=1}^{\infty} c_n \psi_n(x)$$

The completeness of the ψ 's guarantees that I can always express $\Psi(x, 0)$ in this way, and their orthonormality licenses the use of Fourier's trick to determine the actual coefficients:

$$c_n = \sqrt{\frac{2}{a}} \int_a^a \sin\left(\frac{n\pi}{a}x\right) \Psi(x, 0) dx \quad (2.37)$$

That does it: Given the initial wave function, $\Psi(x, 0)$, we first compute the expansion coefficients c_n using (2.37), and then plug these into (2.36) to obtain $\Psi(x, t)$.

Loosely speaking, c_n tells you the "amount of ψ_n that is contained in Ψ ". Some people like to say that $|c_n|^2$ is the "probability of finding the particle in the n th stationary state", but this is a bad language; the particle is in the state Ψ , not Ψ_n , and, anyhow, in the laboratory you don't "find a particle to be in a particular state" —you measure some observable, and what you get is a number. As we'll see next, what $|c_n|^2$ tells you is the probability that a measurement of the energy would yield the value E_n (a competent measurement will always return one of the "allowed" values—hence the name—and $|c_n|^2$ is the probability of getting the particular value E_n).

Of course, the sum of these probabilities should be 1,

$$\sum_{n=1}^{\infty} |c_n|^2 = 1 \quad (2.38)$$

Indeed, this follows from the normalization of Ψ .

Moreover, the expectation value of the energy must be

$$\langle H \rangle = \sum_{n=1}^{\infty} |c_n|^2 E_n \quad (2.39) \quad \text{Expectation value of the energy}$$

Notice that the probability of getting a particular energy is independent of time, and so, a fortiori, is the expectation value of H . This is a manifestation of **conservation of energy** in quantum mechanics.

2.3 The Harmonic Oscillator

The paradigm for a classical harmonic oscillator is a mass m attached to a spring of force constant k . The motion is governed by **Hooke's law**,

$$F = -kx = m \frac{d^2x}{dt^2}$$

(ignoring friction), and the solution is

$$x(t) = A \sin(\omega t) + B \cos(\omega t)$$

where

$$\omega \equiv \sqrt{\frac{k}{m}} \quad (2.40)$$

is the (angular) frequency of the oscillation. The potential energy is

$$V(x) = \frac{1}{2} k x^2 \quad (2.41)$$

its graph is a parabola.

Virtually any oscillatory motion is approximately simple harmonic, as long as the amplitude is small.

The *quantum* problem is to solve the Schrodinger equation for the potential

$$V(x) = \frac{1}{2} m \omega^2 x^2 \quad (2.42)$$

(it is customary to eliminate the spring constant in favor of the classical frequency, using (2.40)). As we have seen, it suffices to solve the time-independent Schrodinger equation:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + \frac{1}{2} m \omega^2 x^2 \psi = E \psi \quad (2.43)$$

In the literature you will find two entirely different approaches to this problem. The first is a straightforward "brute force" solution to the differential equation, using the **power series method**; it has the virtue that the same strategy can be applied to many other potentials. The second is a diabolically clever algebraic technique, using so-called **ladder operator**. We'll see the algebraic method first, because it is quicker and simpler (and a lot more fun)⁸.

⁸We'll encounter some of the same strategies in the theory of angular momentum, and the technique generalizes to a broad class of potentials in **super-symmetric quantum mechanics**

2.3.1 Algebraic Method

To begin with, let's rewrite (2.43) in a more suggestive form:

$$\frac{1}{2m}[p^2 + (m\omega x)^2]\psi = E\psi \quad (2.44)$$

where $p \equiv (\hbar/i)d/dx$ is, of course, the momentum operator. The basic idea is to *factor* the Hamiltonian,

$$H = \frac{1}{2m}[p^2 + (m\omega x)^2] \quad (2.45)$$

If these were *numbers*, it would be easy:

$$u^2 + v^2 = (iu + v)(-iu + v)$$

Here, however, it's not quite so simple, because p and x are *operators*, and operators do not, in general, **commute** (xp is not the same as px). Still, this does motivate us to examine the quantities

$$a_{\pm} \equiv \frac{1}{\sqrt{2\hbar m\omega}}(\mp ip + m\omega x) \quad (2.46)$$

(the factor in front is just there to make the final result look nicer).

Well, what is the product a_-a_+ ?

$$\begin{aligned} a_-a_+ &= \frac{1}{2\hbar m\omega}(ip + m\omega x)(-ip + m\omega x) \\ &= \frac{1}{2\hbar m\omega}[p^2 + (m\omega x)^2 - im\omega(xp - px)] \end{aligned}$$

As anticipated, there's an extra term, involving $(xp - px)$. We call this the **commutator** of x and p ; it is a measure of how badly the *fail* to commute. In general the commutator of operators A and B (written with brackets) is

$$[A, B] \equiv AB - BA \quad (2.47)$$

In this notation,

$$a_-a_+ = \frac{1}{2\hbar m\omega}[p^2 + (m\omega x)^2] - \frac{i}{2\hbar}[x, p] \quad (2.48)$$

We need to figure out the commutator of x and p . *Warning:* Operators are notoriously slippery to work with in the abstract, and you are bound to make mistakes unless you give them a "test function", $f(x)$, to act on. At the end you can throw away the test function, and you'll be left with an equation involving the operators alone. In the present case we have:

$$[x, p]f(x) = \left[x \frac{\hbar}{i} \frac{df}{dx} - \frac{\hbar}{i} \frac{dx f}{dx} \right] = \frac{\hbar}{i} \left(x \frac{df}{dx} - x \frac{df}{dx} - x \right) = i\hbar f(x) \quad (2.49)$$

Dropping the test function, which has served its purpose,

$$\boxed{[x, p] = i\hbar} \quad (2.50) \quad \text{Canonical commutation relation}$$

This lovely and ubiquitous result is known as the **cannonical commutation relation**⁹.

With this, (2.48) becomes

$$a_- a_+ = \frac{1}{\hbar\omega} H + \frac{1}{2} \quad (2.51)$$

or

$$H = \hbar\omega \left(a_- a_+ - \frac{1}{2} \right) \quad (2.52)$$

Evidently the Hamiltonian does not factor perfectly—there's that extra $-1/2$ on the right. Notice that the ordering of a_+ and a_- is important here; the same argument, with a_+ on the left, yields

$$a_+ a_- = \frac{1}{\hbar\omega} H - \frac{1}{2} \quad (2.53)$$

In particular,

$$[a_-, a_+] = 1 \quad (2.54)$$

So the Hamiltonian can equally well be written

$$H = \hbar\omega \left(a_+ a_- + \frac{1}{2} \right) \quad (2.55)$$

In terms of a_{\pm} , then, the Schrodinger equation¹⁰ for the harmonic oscillator takes the form

$$\hbar\omega \left(a_{\pm} a_{\mp} \pm \frac{1}{2} \right) \psi = E\psi \quad (2.56)$$

(in equations like this you read the upper signs all the way across, or else the lower signs).

Now, here comes the crucial step: I claim that if ψ satisfies the Schrodinger equations with energy E , (that is: $H\psi = E\psi$), then $a_+\psi$ satisfies the Schrodinger equations with energy $(E + \hbar\omega)$: $H(a_+\psi) = (E + \hbar\omega)(a_+\psi)$

By the same token, $a_-\psi$ is a solution with energy $(E - \hbar\omega)$

$$H(a_+\psi) = (E + \hbar\omega)(a_+\psi) \quad H(a_-\psi) = (E - \hbar\omega)(a_-\psi)$$

Here, then is a wonderful machine for generating new solution, with higher and lower energies—if we could just find *one* solution, to get started! We call a_{\pm} **ladder operators**, because they allow us to climb up and down in energy; A_+ is the **raising operator**, and a_- the **lowering operator**.

But wait! What if I apply the lowering operator repeatedly? Eventually I'm going to reach a state with energy less than zero, which does not exist! At some point the machine must fail. How can that happen? We know that $a_-\psi$ is a new solution to the Schrodinger equation, but *there is no guarantee that it will be normalizable*—it might be zero, or its square-integral might be infinite. In practice it is the former: There occurs a "lowest rung" (call it ψ_0) such that

$$a_-\psi_0 = 0 \quad (2.57)$$

⁹In deep sense all of the mysteries of quantum mechanics can be traced to the fact that position and momentum do not commute. Indeed, some authors take the cannonical commutation relation as an *axiom* of the theory, and use it to *derive* $p = (\hbar/i)d/dx$

¹⁰I'm getting tired of writing "time-independent Schrodinger equation", so when it's clear from the context which one I mean, I'll just call it the "S equation".

We can use this to determine $\psi_0(x)$:

$$\frac{1}{\sqrt{2\hbar m\omega}} \left(\hbar \frac{d}{dx} + m\omega x \right) \psi_0 = 0$$

or

$$\frac{d\psi_0}{dx} = -\frac{m\omega}{\hbar} x \psi_0$$

The differential equation is easy to solve:

$$\int \frac{d\psi_0}{\psi_0} = -\frac{m\omega}{\hbar} \int x dx \Rightarrow \ln \psi_0 = -\frac{m\omega}{2\hbar} x^2 + \text{constant}$$

so

$$\psi_0(x) = A e^{-\frac{m\omega}{2\hbar} x^2}$$

We might as well normalize it right away:

$$= |A|^2 \int_{-\infty}^{\infty} e^{-m\omega x^2/\hbar} dx = |A|^2 \sqrt{\frac{\pi\hbar}{m\omega}}$$

so $A^2 = \sqrt{m\omega/\pi\hbar}$, and hence

$$\boxed{\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega}{2\hbar} x^2}} \quad (2.58)$$

To determine the energy of this state we plug it into the Schrodinger equation (in the form of (2.56)), $\hbar\omega(a_+a_- + 1/2)\psi_0 = E_0\psi_0$. and exploit the fact that $a_-\psi_0 = 0$:

$$E_0 = \frac{1}{2} \hbar\omega \quad (2.59)$$

With our foot now securely planted on the bottom rung (the ground state of the quantum oscillator), we simply apply the raising operator (repeatedly) to generate the excited states, increasing the energy by $\hbar\omega$ with each step:

$$\boxed{\psi_n(x) = A_n (a_+)^n \psi_0(x), \quad \text{with } E_n = \left(n + \frac{1}{2}\right) \hbar\omega} \quad (2.60)$$

where A_n is the normalization constant. By applying the raising operator (repeatedly) to ψ_0 , then, we can (in principle) construct all the stationary states of the harmonic oscillator. Meanwhile, without ever doing that explicitly, we have determined the allowed energies.

You can even get the normalization algebraically, but it takes some fancy footwork, so watch closely. we know that $a_{\pm}\psi_n$ is *proportional* to $\psi_{n\pm 1}$,

$$a_+\psi_n = c_n\psi_{n+1}, \quad a_-\psi_n = d_n\psi_{n-1} \quad (2.61)$$

but what are the proportionality factors, c_n and d_n ? First note that for "any" functions $f(x)$ and $g(x)$,

$$\int_{-\infty}^{\infty} f^*(a_{\pm}g)dx = \int_{-\infty}^{\infty} (a_{\mp}f)^*gdx \quad (2.62)$$

(In the language of linear algebra, a_{\mp} is the **hermitian conjugate** of A_{\pm}). In particular,

$$\int_{-\infty}^{\infty} (a_{\pm}\psi_n)^*(a_{\pm}\psi_n)dx = \int_{-\infty}^{\infty} (a_{\mp}a_{\pm}\psi_n)^*\psi_n dx$$

But (invoking (2.56) and (2.60))

$$a_+a_-\psi_n = n\psi_n, \quad a_-a_+\psi_n = (n+1)\psi_n \quad (2.63)$$

so

$$\begin{aligned} \int_{-\infty}^{\infty} (a_+\psi_n)^*(a_+\psi_n)dx &= |c_n|^2 \int_{-\infty}^{\infty} |\psi_{n+1}|^2 dx = (n+1) \int_{-\infty}^{\infty} |\psi_n|^2 dx \\ \int_{-\infty}^{\infty} (a_-\psi_n)^*(a_-\psi_n)dx &= |d_n|^2 \int_{-\infty}^{\infty} |\psi_{n-1}|^2 dx = n \int_{-\infty}^{\infty} |\psi_n|^2 dx \end{aligned}$$

But since ψ_n and $\psi_{n\pm 1}$ are normalized, it follows that $|c_n|^2 = n+1$ and $|d_n|^2 = n$ and hence.

$$\boxed{a_+\psi_n = \sqrt{n+1}\psi_{n+1}, \quad a_-\psi_n = \sqrt{n}\psi_{n-1}} \quad (2.64)$$

Thus

$$\begin{aligned} \psi_1 &= a_+\psi_0, \quad \psi_2 = \frac{1}{\sqrt{2}}a_+\psi_1 = \frac{1}{\sqrt{2}}(a_+)^2\psi_0 \\ \psi_3 &= \frac{1}{\sqrt{3}}a_+\psi_2 = \frac{1}{\sqrt{3 \cdot 2}}(a_+)^3\psi_0, \quad \psi_4 = \frac{1}{\sqrt{4}}a_+\psi_3 = \frac{1}{\sqrt{4 \cdot 3 \cdot 2}}(a_+)^4\psi_0 \end{aligned}$$

and so on .Clearly

$$\boxed{\psi_n = \frac{1}{\sqrt{n!}}(a_+)^n\psi_0} \quad (2.65)$$

which is to say that the normalization factor in (2.60) is $A_n = 1/\sqrt{n!}$.

As in the case of the infinite square well, the stationary states of the harmonic oscillator are orthogonal:

$$\int_{-\infty}^{\infty} \psi_m^* \psi_n dx = \delta_{mn} \quad (2.66)$$

This can be proved using (2.63) and (2.62) twice—first moving a_+ and then moving a_- :

$$\begin{aligned} \int_{-\infty}^{\infty} \psi_m^* (a_+a_-)\psi_n dx &= n \int_{-\infty}^{\infty} \psi_m^* \psi_n dx \\ &= \int_{-\infty}^{\infty} (a_-\psi_m)^*(a_-\psi_n) dx = \int_{-\infty}^{\infty} (a_-a_-\psi_m)^* \psi_n dx \\ &= m \int_{-\infty}^{\infty} \psi_m^* \psi_n dx \end{aligned}$$

Unless $m = n$, then, $\int \psi_m^* \psi_n dx$ must be zero. Orthonormality means that we can again use Fourier's trick (2.34) to evaluate the coefficients, when we expand $\Psi(x, 0)$ as a linear combination of stationary states (2.17), and $|c_n|^2$ is again the probability that a measurement of energy would yield the value E_n .

2.3.2 Analytic Method

We return now to the Schrodinger equation for the harmonic oscillator,

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2 x^2 \psi = E\psi \quad (2.67)$$

and solve it directly, by the series method. Thing loock a little cleanes if we introduce the dimensionless variable

$$\xi \equiv \sqrt{\frac{m\omega}{\hbar}} x \quad (2.68)$$

in terms of ξ the Schrodinger equation reads

$$\frac{d^2\psi}{d\xi^2} = (\xi^2 - K)\psi \quad (2.69)$$

where K is the energy, in units of $(1/2)\hbar\omega$:

$$k \equiv \frac{2E}{\hbar\omega} \quad (2.70)$$

Our problem is to solve (2.69), and in the process obtain the "allowed" values of K (and hence of E).

To begin with, note that at very large ξ (which is to say, at very large x), ξ^2 completely dominates over the constant K , so in this regime

$$\frac{d^2\psi}{d\xi^2} \approx \xi^2 \psi \quad (2.71)$$

which has the approximate solution

$$\psi(\xi) \approx Ae^{-\xi^2/2} + Be^{+\xi^2/2} \quad (2.72)$$

The B term is clearly not normalizable (it blows up as $|x| \rightarrow \infty$); the physucally acceptable solutions, then, have the asymptotic form

$$\psi(\xi) \rightarrow ()e^{-\xi^2/2}, \quad \text{at large } \xi \quad (2.73)$$

This suggests that we "peel off" the exponential part,

$$\psi(\xi) = h(\xi)e^{-\xi^2/2} \quad (2.74)$$

in hopes that what remains, $h(\xi)$, has a simpler functional form than $\psi(\xi)$ itself. Differentiating (2.74)

$$\frac{d\psi}{d\xi} = \left(\frac{dh}{d\xi} - \xi h \right) e^{-\xi^2/2}$$

and

$$\frac{d^2\psi}{d\xi^2} = \left(\frac{d^2h}{d\xi^2} - 2\xi \frac{dh}{d\xi} + (\xi^2 - 1)h \right) e^{-\xi^2/2}$$

so the Schrodinger equation (2.69) becomes

$$\frac{d^2 h}{d\xi^2} - 2\xi \frac{dh}{d\xi} + (K - 1)h = 0 \quad (2.75)$$

I propose to look for solutions to (2.75) in the form of *power series* in ξ ¹¹:

$$h(\xi) = a_0 + a_1\xi + a_2\xi^2 + \cdots = \sum_{j=0}^{\infty} a_j \xi^j \quad (2.76)$$

Differentiating the series term by term,

$$\frac{dh}{d\xi} = a_1 + 2a_2\xi + 3a_3\xi^2 + \cdots = \sum_{j=0}^{\infty} j a_j \xi^{j-1}$$

and

$$\frac{d^2 h}{d\xi^2} = 2a_2 + 2 \cdot 3a_3\xi + 3 \cdot 4a_4\xi^2 + \cdots = \sum_{j=0}^{\infty} (j+1)(j+2)a_{j+2}\xi^j$$

Putting these into (2.75), we find

$$\sum_{j=0}^{\infty} [(j+1)(j+2)a_{j+2} - 2ja_j + (K-1)a_j]\xi^j = 0 \quad (2.77)$$

It follows (from the uniqueness of power series expansions) that the coefficient of *each power* of ξ must vanish,

$$(j+1)(j+2)a_{j+2} - 2ja_j + (K-1)a_j = 0$$

and hence that

$$a_{j+2} = \frac{(2j+1-K)}{(j+1)(j+2)} a_j \quad (2.78)$$

This **recursion formula** is entirely equivalent to the Schrodinger equation. Starting with a_0 , it generates all the even-numbered coefficients

$$a_2 = \frac{(a-K)}{2} a_0 \quad a_4 = \frac{(5-K)}{12} a_2 = \frac{(5-K)(1-K)}{24} a_0, \quad \dots$$

and starting with a_1 , it generates the odd coefficients:

$$a_3 = \frac{(3-K)}{6} a_1, \quad a_5 = \frac{(7-K)}{20} a_3 = \frac{(7-K)(3-K)}{120} a_1, \quad \dots$$

We write the complete solution as

$$h(\xi) = h_{\text{even}}(\xi) + h_{\text{odd}}(\xi) \quad (2.79)$$

¹¹This is known as the **Frobenius method** for solving a differential equation. According to Taylor's theorem, any reasonably well-behaved function can be expressed as a power series, so (2.76) ordinarily involves no loss of generality.

where

$$h_{\text{even}}(\xi) \equiv a_0 + a_2\xi^2 + a_4\xi^4 + \dots$$

is an even function of ξ , built on a_0 , and

$$h_{\text{odd}}(\xi) \equiv a_1\xi + a_3\xi^3 + a_5\xi^5 + \dots$$

is an odd function, built on a_1 (2.78) determines $h(\xi)$ in terms of two arbitrary constants (a_0 and a_1)—which is just what we would expect, for a second-order differential equation.

However, not all the solutions so obtained are *normalizable*. For at very large j , the recursion formula becomes (approximately)

$$a_{j+2} \approx \frac{2}{j} a_j$$

with the (approximate) solution

$$a_j \approx \frac{C}{(j/2)!}$$

for some constant C , and this yields (at large ξ , where the higher powers dominate)

$$h(\xi) \approx C \sum \frac{1}{(j/2)!} \xi^j \approx C \sum \frac{1}{j!} \xi^{2j} \approx C e^{\xi^2}$$

Now, if h goes like $\exp(\xi^2/2)$, then ψ (remember ψ ?—that's what we're trying to calculate) goes like $\exp(\xi^2/2)$ (2.74), which is precisely the asymptotic behavior we didn't want¹². There is only one way to wiggle out of this: For normalizable solutions the *power series must terminate*. There must occur some "highest" j (call it n), such that the recursion formula spits out $a_{n+2} = 0$ (this will truncate either series h_{even} or the series h_{odd} ; the other one must be zero from that start: $a_1 = 0$ if n is even, and $a_0 = 0$ if n is odd). For physically acceptable solutions, then, (2.78) requires that

$$K = 2n + 1$$

for some non-negative integer n , which is to say (referring to (2.74)) that the *energy* must be

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega, \quad \text{for } n = 0, 1, 2, \dots \quad (2.80)$$

Thus we recover, by a completely different method, the fundamental quantization condition we found algebraically in (2.60).

¹²It's no surprise that the ill-behaved solutions are still in (2.78): this recursion relation is equivalent to the Schrodinger equation, so it's got to include both the asymptotic forms we found in (2.72)

Chapter 3

Formalism

3.1 Hilbert Space

The purpose of this chapter is to recast the theory in a more powerful form.

Quantum theory is based on two constructs: *wave functions* and *operators*. The state of a system is represented by its wave function, observables are represented by operators. Mathematically, wave functions satisfy the defining conditions for abstract **vectors**, and operators act on them as **linear transformations**. So the natural language of quantum mechanics is **linear algebra**.

But it is not, I suspect, a form of linear algebra with which you are immediately familiar. In an N -dimensional space it is simplest to represent a vector $|\alpha\rangle$, by the N -tuple of its components, $\{a_n\}$, with respect to a specified orthonormal basis:

$$|\alpha\rangle \rightarrow \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} \quad (3.1)$$

The **inner product**, $\langle\alpha|\beta\rangle$, of two vectors (generalizing the dot product in three dimensions) is the complex number,

$$\langle\alpha|\beta\rangle = a_1^* b_1 + a_2^* b_2 + \cdots + a_N^* b_N \quad (3.2)$$

Linear transformations, T , are represented by **matrices** (with respect to the specified basis), which act on vectors (to produce new vectors) by the ordinary rules of matrix multiplication:

$$|\beta\rangle = T|\alpha\rangle \rightarrow \mathbf{b} = \mathbf{T}\mathbf{a} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1N} \\ t_{21} & t_{22} & \cdots & t_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ t_{N1} & t_{N2} & \cdots & t_{NN} \end{pmatrix} \quad (3.3)$$

But the "vectors" we encounter in quantum mechanics are (for the most specified) *functions*, and they live in infinite-dimensional spaces. For them the N -tuple/matrix notation is awkward, at best, and manipulations that are well-behaved in the finite dimensional case can be problematic. (The underlying reason is that whereas the finite sum in (3.2) always exists, an

infinite sum—or a integral— may not converge, in which case the inner product does not exists, and any argument involving inner products is immediately suspect.).

The collection of all functions of x constitutes a vector space, but for our purposes it is much too large. To represent a possible physical state, the function Ψ must be normalized:

$$\int |\Psi|^2 dx = 1$$

The set of all **square-integrable functions**, on a specified interval¹,

$$f(x) \quad \text{such that} \quad \int_a^b |f(x)|^2 dx < \infty \quad (3.4)$$

constitutes a (much smaller) vector space. Mathematicians call it $L_2(a, b)$; physicist call it **Hilbert space**. In quantum mechanics then

$$\boxed{\text{Wave functions live in Hilbert space.}} \quad (3.5)$$

We define the **inner product of two functions**, $f(x)$ and $g(x)$, as follows:

$$\langle f|g \rangle \equiv \int_a^b f(x)^* g(x) dx \quad (3.6)$$

If f and g are both square-integrable (that is, if they are both in Hilbert spaces), their inner product is guaranteed to exist. This follows from the integral **Schwartz inequality**:

$$\left| \int_a^b f(x)^* g(x) dx \right| \leq \sqrt{\int_a^b |f(x)|^2 dx \int_a^b |g(x)|^2 dx} \quad (3.7)$$

Notice in particular that

$$\langle g|f \rangle = \langle f|g \rangle^* \quad (3.8)$$

Moreover, the inner product of $f(x)$ with itself,

$$\langle f|f \rangle = \int_a^b |f(x)|^2 dx \quad (3.9)$$

is real and non-negative; it's zero only when $f(x) = 0$.

A function is said to be **normalized** if its inner product with itself is 1; two functions are **orthogonal** if their inner product is 0; and a set of functions, $\{f_n\}$, is **orthonormal** if they are normalized and mutually orthogonal

$$\langle f_m|f_n \rangle = \delta_{mn} \quad (3.10)$$

Finally, a set of functions is **complete** if any other function (in Hilbert space) can be expressed as a linear combination of them:

$$f(x) = \sum_{n=1}^{\infty} c_n f_n(x) \quad (3.11)$$

If the functions $\{f_n(x)\}$ are orthonormal, the coefficients are given by Fourier's trick

$$c_n = \langle f_n|f \rangle \quad (3.12)$$

¹For is, the limits (a and b) will almost always be $-\infty$ and $+\infty$, but we might as well keep things more general for the moment

3.2 Obervables

3.2.1 Hermitian Operators

The expectation value of an observable $Q(x, p)$ can be expressed very neatly in inner-product notation:

$$\langle Q \rangle = \int \Psi^* \hat{Q} \Psi dx = \langle \Psi | \hat{Q} \Psi \rangle \quad (3.13)$$

Now, the outcome of a measurement has got to be *real*, and so, a fortiori, is the average of many measurements:

$$\langle Q \rangle = \langle Q \rangle^* \quad (3.14)$$

But the complex conjugate of an inner product reverses the order (3.8), so

$$\langle \Psi | \hat{Q} \Psi \rangle = \langle \hat{Q} \Psi | \Psi \rangle \quad (3.15)$$

and this must hold true for any wave function Ψ . Thus operators representing *obervables* have the very special property that

$$\langle f | \hat{Q} f \rangle = \langle \hat{Q} f | f \rangle \quad \text{for all } f(x) \quad (3.16)$$

We call such operators **hermitian**.

The essentiak point is that a hermitian operator can be applied either to the first member of an inner product or to the second, with the same result, and hermitian operators naturally arise in quantum mechanics because their expectations values are real:

Observables are represented by hermitian operators.

(3.17)

Is the momentum operator, for example, hermitian?

$$\langle f | \hat{p} g \rangle = \int_{-\infty}^{\infty} \frac{\hbar}{i} \frac{dg}{dx} dx = \frac{\hbar}{i} f^* g \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \left(\frac{\hbar}{i} \frac{df}{dx} \right)^* g dx = \langle \hat{p} f | g \rangle \quad (3.18)$$

Notice if $f(x)$ and $g(x)$ are squared integrable, they must go to zero at $\pm\infty$.

3.2.2 Determinate States

Ordinarily, when you measure an obervable Q on an ensamble of identically prepared system, all in the same state Ψ , you do not ger the same reuslt each time—this is the indeterminacy of quantum mechanics. *Question:* Would it be possible to prepare a state such that *every* measurement of Q is certain to return the *same* value (call it a)? This would be, if youlike, a **determinate state**, for the obervable Q (Actually, we already know one example: Stationary states are dterminate states of Hamiltonian; a measurement of the total enery, on a particle in the stationary state Ψ_n , is certain to yield the corresponding "allowed" energy E_n).

Well, the standard deviation of Q , in a determiante state, would to be zero, shich is to say,

$$\sigma^2 = \langle (\hat{Q} - \langle Q \rangle)^2 \rangle = \langle \Psi | (\hat{Q} - q)^2 \Psi \rangle = \langle (\hat{Q} - q) \Psi | (\hat{Q} - q) \Psi \rangle = 0 \quad (3.19)$$

But the only function whose inner product with itself vanishes is 0, so

$$\hat{Q}\Psi = q\Psi \quad (3.20)$$

This is the **eigenvalue equation** for the operator \hat{Q} ; Ψ is an **eigenfunction** of \hat{Q} , and q is the corresponding **eigenvalue**. Thus

Determinant states are eigenfunctions of \hat{Q}

(3.21)

Measurements of Q on such states is certain to yield the eigenvalue, q .

Note that the eigenvalue is a number (not a operator or a function). You can multiply any eigenfunction by a constant, and it is still an eigenfunction, with the same eigenvalue. Zero does not count as an eigenfunction. But there's nothing wrong with zero as an eigenvalue. The collection of all the eigenvalues of an operator is called its **spectrum**. Sometimes two (or more) linearly independent eigenfunctions share the same eigenvalue; in that case the spectrum is said to be **degenerate**.

3.3 Eigenfunctions of a Hermitian Operator

Our attention is thus directed to the *eigenfunction of hermitian operators* (physically: determine states of observables). These fall into two categories: If the spectrum is **discrete** (i.e., the eigenvalues are separated from one another) then the eigenfunctions lie in Hilbert space and they constitute physically realizable states. If the spectrum is **continuous** (i.e., the eigenvalues fill out an entire range) then the eigenfunctions are non normalizable, and they do not represent possible wave functions (*though linear combinations* of them—involving necessarily a spread in eigenvalues—may be normalizable). Some operators have a discrete spectrum only (for example, the Hamiltonian for the harmonic oscillator). Some have only a continuous spectrum (for example, the free particle Hamiltonian), and some have both discrete part and a continuous part (for example, the Hamiltonian for a finite square well). The discrete case is easier to handle, because the relevant inner products are guaranteed to exist—in fact, it is very similar to the finite-dimensional theory (the eigenvectors of a hermitian matrix). I'll treat the discrete case first, and then the continuous one:

3.3.1 Discrete Spectra

Mathematically, the normalizable eigenfunctions of a hermitian have two important properties:

Theorem 1 *Their eigenvalues are real.*

This is comforting: If you measure an observable on a particle in a determinate state, you will at least get a real number.

Theorem 2 *Eigenfunctions belonging to distinct eigenvalues are orthogonal.*

In a *finite-dimensional* vector space the eigenvectors of a hermitian matrix have a third fundamental property: They span the space (every vector can be expressed as a linear combination of them). Unfortunately, the proof does not generalize to infinite-dimensional spaces. But the property itself is essential to the internal consistency of quantum mechanics, so we will take it as an *axiom* (or, more precisely, as a restriction on the class of hermitian operators that can represent observables):

Axiom 1 *The eigenvalues of an observable operator are complete: Any function (in Hilbert space) can be expressed as a linear combination of them.*

3.3.2 Continuous Spectra

If the spectrum of a hermitian operator is *continuous*, the eigenfunctions are not normalizable, and the proofs of Theorems 1 and 2 fail, because the inner products may not exist. Nevertheless, there is a sense in which the three essential properties (reality, orthogonality, and completeness) still hold. I think it's best to approach this subtle case through specific examples.

Example 1 *Find the eigenfunctions and eigenvalues of the momentum operator*

Solution 1

Let $f_p(x)$ be the eigenfunction and p the eigenvalue:

$$\frac{\hbar}{i} \frac{d}{dx} f_p(x) = p f_p(x) \quad (3.22)$$

The general solution is

$$f_p(x) = A e^{ipx/\hbar}$$

This is not square-integrable, for any (complex) value of p —the momentum operator has no eigenfunctions in Hilbert space. And yet, if we restrict ourselves to *real* eigenvalues, we do recover a kind of *ersatz* "orthonormality". ...

3.3.3 Generalized Statistical Interpretation

Generalized statistical interpretation: If you measure an observable $Q(x, p)$ on a particle in the state $\Psi(x, t)$, you are certain to get *one of the eigenvalues* of the hermitian operator $\hat{Q}(x, -i\hbar d/dx)$. If the spectrum of \hat{Q} is discrete, the probability of getting the particular eigenvalue q_n associated with the orthonormalized eigenfunction $f_n(x)$ is

$$|c_n|^2, \quad \text{where} \quad c_n = \langle f_n | \Psi \rangle \quad (3.23)$$

If the spectrum is continuous, with real eigenvalues $q(z)$ and associated Dirac-orthonormalized eigenfunctions $f_z(x)$, the probability of getting a result in the range dz is

$$|c(z)|^2 dz \quad \text{where} \quad c(z) = \langle f_z | \Psi \rangle \quad (3.24)$$

Upon measurement, the wave function "collapses" to the corresponding eigenstate.

The statistical interpretation is radically different from anything we encounter in classical physics. A somewhat different perspective helps to make it plausible: The eigenfunctions of an observable operator are *complete*, so the wave function can be written as a linear combination of them:

$$\Psi(x, t) = \sum_n c_n f_n(x) \quad (3.25)$$

(For simplicity, I'll assume that the spectrum is discrete; it's easy to generalise this argument to the continuous case.) Because the eigenfunctions are *orthonormal*, the coefficients are given by Fourier's trick:

$$c_n = \langle f_n | \Psi \rangle = \int f_n(x)^* \Psi(x, t) dx \quad (3.26)$$

Qualitatively, c_n tells you "how much f_n is contained in Ψ ", and given that a measurement has to return one of the eigenvalues of \hat{Q} , it seems reasonable that the probability of getting the particular eigenvalue q_n would be determined by the "amount of f_n " in Ψ . But because probabilities are determined by the absolute square of the wave function, the precise measure is actually $|c_n|^2$. That's the essential burden of the generalized statistical interpretation.

Of course, the total probability (summed over all outcomes) has got to be one:

$$\sum_n |c_n|^2 = 1 \quad (3.27)$$

Similarly, the expectation value of Q should be the sum over all possible outcomes of the eigenvalue times the probability of getting the eigenvalue:

$$\langle Q \rangle = \sum_n q_n |c_n|^2 \quad (3.28)$$

The **momentum space wave function**, $\Phi(p, t)$. It is essentially the *Fourier transform* of the (**position space**) wave function $\Psi(x, t)$ —which, by Plancherel's theorem, is its *inverse* Fourier transform:

$$\Phi(p, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \Psi(x, t) dx \quad (3.29)$$

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{ipx/\hbar} \Phi(p, t) dp \quad (3.30)$$

According to the generalized statistical interpretation, the probability that a measurement of momentum would yield a result in the range dp is

$$|\Phi(p, t)|^2 dp \quad (3.31)$$

3.4 The Uncertainty Principle

3.5 Dirac Notation

Imagine an ordinary vector \mathbf{A} in two dimensions. How would you describe this vector to someone? The most convenient way is to set up cartesian axes, x and y , and specify the components of \mathbf{A} : $a_x = \hat{i} \cdot \mathbf{A}$, $A_y = \hat{j} \cdot \mathbf{A}$. Of course, your sister might have drawn a different set of axes, x' and y' , and she would report different components: $A'_x = \hat{i}' \cdot \mathbf{A}$, $A'_y = \hat{j}' \cdot \mathbf{A}$. But it's all the same vector—we're simply expressing it with respect to two different bases ($\{\hat{i}, \hat{j}\}$ and $\{\hat{i}', \hat{j}'\}$). The vector itself lives "out there in space", independent of anybody's (arbitrary) choice of coordinates.

The same is true for the state of a system in quantum mechanics. It is represented by a vector, $|s(t)\rangle$, that lives "out there in Hilbert space", but we can express it with respect to

to any number of different bases. The wave function $\Psi(x, t)$ is actually the coefficient in the expansion of $|s(t)\rangle$ in the basis of position eigenfunctions:

$$\Psi(x, t) = \langle x|s(t)\rangle \quad (3.32)$$

(with $|x\rangle$ standing for the eigenfunctions of \hat{x} with eigenvalue x), whereas the momentum space wavefunction $\Phi(p, t)$ is the expansion of $|s\rangle$ in the basis of momentum eigenfunctions:

$$\Phi(p, t) = \langle p|s(t)\rangle \quad (3.33)$$

(with $|p\rangle$ standing for the eigenfunction of \hat{p} with eigenvalue p). Or we could expand $|s\rangle$ in the basis of energy eigenfunctions (supposing for simplicity that the spectrum is discrete):

$$c_n(t) = \langle n|s(t)\rangle \quad (3.34)$$

(with $|n\rangle$ standing for the n th eigenfunction of \hat{H})—(3.26). But it's all the same state; the functions Ψ and Φ , and the collection of coefficients $\{c_n\}$, contain exactly the same information—they are simply three different ways of describing the same vector:

$$\begin{aligned} \Psi(x, t) &= \int \Psi(y, t) \delta(x - y) dy = \int \Phi(p, t) \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} dp \\ &= \sum c_n e^{-iE_n t/\hbar} \psi_n(x) \end{aligned} \quad (3.35)$$

Operators (representing observables) are linear transformations—they "transform" one vector into another:

$$|\beta\rangle = \hat{Q} |\alpha\rangle \quad (3.36)$$

Just as vectors are represented, with respect to a particular basis $\{|e_n\rangle\}$ ², by their components,

$$|\alpha\rangle = \sum_n a_n |e_n\rangle, \text{ with } a_n = \langle e_n|\alpha\rangle : \quad |\beta\rangle = \sum_n b_n |e_n\rangle, \text{ with } b_n = \langle e_n|\beta\rangle \quad (3.37)$$

operators are represented (with respect to a particular basis) by their **matrix elements**³

$$\langle e_m|\hat{Q}|e_n\rangle \equiv Q_{mn} \quad (3.38)$$

In this notation (3.36) takes the form

$$\sum_n b_n |e_n\rangle = \sum_n a_n \hat{Q} |e_n\rangle \quad (3.39)$$

or, taking the inner product with $|e_m\rangle$,

$$\sum_n b_n \langle e_m|e_n\rangle = \sum_n a_n \langle e_m|\hat{Q}|e_n\rangle \quad (3.40)$$

²I'll assume the basis is discrete: otherwise n becomes a continuous index and the sums are replaced by integrals

³This terminology is inspired, obviously, by the finite-dimensional case, but the "matrix" will now typically have an infinite (maybe even uncountable) number of elements.

and hence

$$b_m = \sum_n Q_{mn} a_n \quad (3.41)$$

Thus the matrix elements tell you how the components transform.

Later on we will encounter systems that admit only a finite number (N) of linearly independent states. In that case $|s(t)\rangle$ lives in an N -dimensional vector space; it can be represented as a column of (N) components (with respect to a given basis), and operators take the form of ordinary ($N \times N$) matrices. These are the simplest quantum systems—none of the subtleties associated with infinite-dimensional vector spaces arise.

Dirac proposed to chop the bracket notation for the inner product, $\langle\alpha|\beta\rangle$, into two pieces, which he called **bra**, $\langle\alpha|$, and **ket** $|\beta\rangle$. The latter is a vector, but what exactly is the former? It's a *linear function* of vectors, in the sense that when it hits a vector (to its right) it yields a (complex) number—the inner product. (When an *operator* hits a vector, it delivers another vector; when a *bra* hits a vector, it delivers a number.) In a function space, the bra can be thought of as an instruction to integrate:

$$\langle f| = \int f^*[\dots] dx$$

with the ellipsis $[\dots]$ waiting to be filled by whatever function the bra encounters in the jet to its right. In a finite-dimensional vector space, with the vectors expressed as columns,

$$|\alpha\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad (3.42)$$

the corresponding bra is a row vector:

$$\langle\alpha| = (a_1^* a_2^* \dots a_n^*) \quad (3.43)$$

The collection of all bras constitutes another vector space—the so-called **dual space**.

The license to treat bras as separate entities in their own right allows for some powerful and pretty notation. For example, if $|\alpha\rangle$ is a normalized vector, the operator

$$\hat{P} \equiv |\alpha\rangle\langle\alpha| \quad (3.44)$$

picks out the portion of any vector that "lies along" $|\alpha\rangle$:

$$\hat{P}|\beta\rangle = \langle\alpha|\beta\rangle |\alpha\rangle$$

we call it the **projection operator** onto the one-dimensional subspace spanned by $|\alpha\rangle$. If $\{|e_n\rangle\}$ is a discrete orthonormal basis,

$$\langle e_m|e_n\rangle = \delta_{mn} \quad (3.45)$$

then

$$\sum_n |e_n\rangle\langle e_n| = 1 \quad (3.46)$$

(the identity operator). For if we let this operator act on any vector $|\alpha\rangle$, we recover the expansion of $|\alpha\rangle$ in the $\{|e_n\rangle\}$ basis:

$$\sum_n |e_n\rangle \langle e_n|\alpha\rangle = |\alpha\rangle \quad (3.47)$$

Similarly, if $\{|e_z\rangle\}$ is a *Dirac* orthonormalized continuous basis,

$$\langle e_a|e_{z'}\rangle = \delta(z - z') \quad (3.48)$$

then

$$\int |e_z\rangle \langle e_z| dz = 1 \quad (3.49)$$

(3.46) and (3.49) are the tidest ways to express completeness.

Chapter 4

Quantum Mechanics In Three Dimensions

4.1 Schrodinger Equation In Spherical Coordinates

The generalization to three dimensions is straightforward. Schrodinger's equation says

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi \quad (4.1)$$

the Hamiltonian operator¹ H is obtained from the classical energy

$$\frac{1}{2}mv^2 + V = \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) + V$$

by the standard prescription (applied now to y and z , as well as x):

$$p_x \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial x}, \quad p_y \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial y}, \quad p_z \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial z} \quad (4.2)$$

or

$$\mathbf{p} \rightarrow \frac{\hbar}{i} \nabla \quad (4.3)$$

for short. Thus

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + V\Psi \quad (4.4)$$

where

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (4.5)$$

is the **Laplacian**, in cartesian coordinates.

¹Where confusion might otherwise occur I have been putting "hats" on operators, to distinguish the from the corresponding classical observables. I don't think there will be much occasion for ambiguity in this chapter, and the hats get to be cumbersome, so I going to leave the off from now on.

The potential energy V and the wave function Ψ are now functions of $\mathbf{r} = (x, y, z)$ and t . The probability of finding the particle in the infinitesimal volume $d^3\mathbf{r} = dx dy dz$ is $|\Psi(\mathbf{r}, t)|^2 d^3\mathbf{r}$, and the normalization condition reads

$$\int |\Psi|^2 d^3\mathbf{r} = 1 \quad (4.6)$$

with the integral taken over all space. If the potential is independent of time, there will be a complete set of stationary states

$$\Psi_n(\mathbf{r}, t) = \psi_n(\mathbf{r}) e^{-iE_n t/\hbar} \quad (4.7)$$

where the spatial wave function ψ_n satisfies the time-independent Schrodinger equation:

$$\boxed{-\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi = E\psi} \quad (4.8)$$

The general solution to the (time-dependent) Schrodinger equation is

$$\Psi(\mathbf{r}, t) = \sum c_n \psi_n(\mathbf{r}) e^{-iE_n t/\hbar} \quad (4.9)$$

where the constants c_n determined by the initial wave function, $\Psi(\mathbf{r}, 0)$, in the usual way. (If the potential admits continuum states, the the sum in (4.9) becomes an integral.)

Separation of Variables

Typically, the potential is a function only of the distance from the origin. In that case it is natural to adopt **spherical coordinates**, (r, θ, ϕ) . In spherical coordinates the Laplacian takes the form²

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \left(\frac{\partial^2}{\partial \phi^2} \right) \quad (4.10)$$

In spherical coordinates, then, the time-independent Schrodinger equation reads

$$-\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \left(\frac{\partial^2 \psi}{\partial \phi^2} \right) \right] + V\psi = E\psi \quad (4.11)$$

We begin by looking for solutions that are separable into products:

$$\psi(r, \theta, \phi) = R(r)Y(\theta, \phi) \quad (4.12)$$

Putting this into (4.11), we have

$$-\frac{\hbar^2}{2m} \left[Y \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{R}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{R}{r^2 \sin^2 \theta} \frac{\partial^2 Y}{\partial \phi^2} \right] + VRY = ERY$$

Dividing by RY and multiplying by $-2mr^2/\hbar^2$:

$$\left\{ \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) - \frac{2mr^2}{\hbar^2} [V(r) - E] \right\} + \frac{1}{Y} \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \phi^2} \right\} = 0$$

²In principle, this can be obtained by change of variables from the cartesian expression (4.5). However, there are much more efficient ways of getting it.

The term in the first curly bracket depends only on r , whereas the remainder depends only on θ and ϕ ; accordingly, each must be a constant. For reasons that will appear on due course³, I will write this "separation constant" in the form $l(l+1)$

4.2 The Hydrogen Atom

The hydrogen atom consists of a heavy, essentially motionless proton (we may as well put it at the origin), of charge e , together with a much lighter electron (charge $-e$) that orbits around it, bound by the mutual attraction of opposite charges. From Coulomb's law, the potential energy (in SI units) is

$$V(r) = -\frac{e^2}{4\pi\epsilon_0} \frac{1}{r} \quad (4.13)$$

and the radial equation (??) says

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dr^2} + \left[-\frac{e^2}{4\pi\epsilon_0} \frac{1}{r} + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2} \right] u = Eu \quad (4.14)$$

Our problem is to solve this equation for $u(r)$, and determine the allowed energies, E . The hydrogen atom is such an important case that I'm not going hand you the solutions this time—we'll work them out in detail, by the method we used in the analytical solution to the harmonic oscillator.

Incidentally, the Coulomb potential (4.14) admits *continuum* states (with $E > 0$), describing electron-proton scattering, as well as discrete *bound* states, representing the hydrogen atom, but we shall confine our attention to the latter.

The Radial Wave Function

Our first task is to tidy up the notation. Let

$$\kappa \equiv \frac{\sqrt{-2mE}}{\hbar} \quad (4.15)$$

(For bound states, E is negative, so κ is *real*.) Dividing (4.14) by E , we have

$$\frac{1}{\kappa^2} \frac{d^2u}{dr^2} = \left[1 - \frac{me^2}{2\pi\epsilon_0\hbar^2\kappa} \frac{1}{(\kappa r)} + \frac{l(l+1)}{(\kappa r)^2} \right] u$$

This suggests that we introduce

$$\rho \equiv \kappa r, \quad \text{and} \quad \rho_0 \equiv \frac{me^2}{2\pi\epsilon_0\hbar^2\kappa} \quad (4.16)$$

so that

$$\frac{d^2u}{d\rho^2} = \left[1 - \frac{\rho_0}{\rho} + \frac{l(l+1)}{\rho^2} \right] u \quad (4.17)$$

³Note that is no loss of generality here—at this stage l could be any complex number. Later on we'll discover that l must in fact be an *integer*, and it is in anticipation of that result that I express the separation constant in a way that looks peculiar now.

Next we examine the asymptotic form of the solutions. As $\rho \rightarrow \infty$, the constant term in the brackets dominates, so (approximately)

$$\frac{d^2 u}{d\rho^2} = u$$

The general solution is

$$u(\rho) = Ae^{-\rho} + Be^{\rho} \quad (4.18)$$

but e^{ρ} blows up (as $\rho \rightarrow \infty$), so $B = 0$. Evidently

$$u(\rho) \sim Ae^{-\rho} \quad (4.19)$$

for large ρ . On the other hand, as $\rho \rightarrow 0$ the centrifugal term dominates⁴; approximately, then:

$$\frac{d^2 u}{d\rho^2} = \frac{l(l+1)}{\rho^2} u$$

The general solution (check it!) is

$$u(\rho) = C\rho^{l+1} + D\rho^{-l}$$

but ρ^{-l} blows up (as $\rho \rightarrow 0$), so $D = 0$. Thus

$$u(\rho) \sim \rho^{l+1} \quad (4.20)$$

for small ρ .

The next step is to peel off the asymptotic behavior, introducing the new function $v(\rho)$:

$$u(\rho) = \rho^{l+1} e^{-\rho} v(\rho) \quad (4.21)$$

in the hope that $v(\rho)$ will turn out to be simpler than $u(\rho)$. The first indications are not auspicious:

$$\frac{du}{d\rho} = \rho^l e^{-\rho} \left[(l+1-\rho)v + \rho \frac{dv}{d\rho} \right]$$

and

$$\frac{d^2 u}{d\rho^2} = \rho^l e^{-\rho} \left\{ \left[-2l - 2 + \rho + \frac{l(l+1)}{\rho} \right] v + 2(l+1-\rho) \frac{dv}{d\rho} + \rho \frac{d^2 v}{d\rho^2} \right\}$$

In terms of $v(\rho)$, then, the radial equation (4.17) reads

$$\rho \frac{\partial^2 v}{\partial \rho^2} + 2(l+1-\rho) \frac{dv}{d\rho} + [\rho_0 - 2(l+1)]v = 0 \quad (4.22)$$

Finally, we assume the solution, $v(\rho)$, can be expressed as a power series in ρ :

$$v(\rho) = \sum_{j=1}^{\infty} c_j \rho^j \quad (4.23)$$

⁴This argument does not apply when $l = 0$ (although the conclusion, (4.20), is in fact valid for that case too). But never mind: All I am trying to do is provide some *motivation* for a change of variables (4.21))

Our problem is to determine the coefficients (c_0, c_1, c_2, \dots) . Differentiating term by term

$$\frac{dv}{d\rho} = \sum_{j=0}^{\infty} j c_j \rho^{j-1} = \sum_{j=0}^{\infty} (j+1) c_{j+1} \rho^j$$

[In the second summation I have renamed the "dummy index": $j \rightarrow j+1$. If this troubles you, write out the first few terms explicitly, and *check* it. You may object that the sum should now begin at $j = -1$, but the factor $(j+1)$ kills that term anyway, so we might as well start at zero.] Differentiating again

$$\frac{d^2v}{d\rho^2} = \sum_{j=1}^{\infty} j(j+1) c_{j+1} \rho^{j-1}$$

Inserting these into (4.22), we have

$$\begin{aligned} & \sum_{j=1}^{\infty} j(j+1) c_{j+1} \rho^j + 2(l+1) \sum_{j=0}^{\infty} (j+1) c_{j+1} \rho^j \\ & - 2 \sum_{j=0}^{\infty} j c_j \rho^j + [\rho_0 - 2(l+1)] \sum_{j=0}^{\infty} c_j \rho^j = 0 \end{aligned}$$

Equating the coefficients of like powers yields

$$j(j+1) c_{j+1} + 2(l+1)(j+1) c_{j+1} - 2j c_j + [\rho_0 - 2(l+1)] c_j = 0$$

or

$$c_{j+1} = \left\{ \frac{2(j+l+1) - \rho_0}{(j+1)(j+2l+2)} \right\} c_j \quad (4.24)$$

This recursion formula determines the coefficients, and hence the function $v(\rho)$: We start with c_0 (this becomes an overall constant, to be fixed eventually by normalization), and (4.24) gives us c_1 ; putting this back in, we obtain c_2 , and so on⁵.

Now let's see what the coefficients look like for large j (this corresponds to large ρ , where the higher powers dominate). In this regime the recursion formula says

$$c_{j+1} \approx \frac{2j}{j(j+1)} c_j = \frac{2}{j+1} c_j$$

Suppose for a moment that this were *exact*. Then

$$c_j = \frac{2^j}{j!} c_0 \quad (4.25)$$

so

$$v(\rho) = C_0 \sum_{j=0}^{\infty} \frac{2^j}{j!} \rho^j = c_0 e^{2\rho}$$

⁵You might wonder why I didn't use the series method directly on $u(\rho)$ —why factor out the asymptotic behavior before applying this procedure? Well, the reason for peeling off ρ^{l+1} is largely aesthetic: Without this, the sequence would begin with a long string of zeros (the first nonzero coefficient begin c_{l+1}): by factoring out ρ^{l+1} we obtain a series that starts out with ρ^0 . The $e^{-\rho}$ factor is more critical—if you don't pull that out, you get a three-term recursion formula, involving c_{j+1} , c_{j+1} , and c_j and that is enormously more difficult to work with.

and hence

$$u(\rho) = c_0 \rho^{l+1} e^\rho \quad (4.26)$$

which blows up at large ρ . The positive exponential is precisely the asymptotic behavior we didn't want, in (4.18). (It's no accident that it reappears here; after all, it does represent the asymptotic form of *some* solutions to the radial equation—they just don't happen to be the ones we're interested in, because they aren't normalizable.) There is only one way out of this dilemma: The *series must terminate*. There must occur some maximal integer, j_{max} , such that

$$c(j_{max} + 1) = 0 \quad (4.27)$$

(and beyond which all coefficients vanish automatically). Evidently (4.24)

$$2(j_{max} + l + 1) - \rho_0 = 0$$

Defining

$$n \equiv j_{max} + l + 1 \quad (4.28)$$

4.3 Angular Momentum

Classically, the angular momentum of a particle (with respect to the origin) is given by the formula

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad (4.29)$$

which is to say,

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x \quad (4.30)$$

The corresponding quantum operators are obtained by the standard prescription $p_x \rightarrow -\hbar\partial/\partial x$, $p_y \rightarrow -\hbar\partial/\partial y$, $p_z \rightarrow -\hbar\partial/\partial z$.

Eigenvalues

The operators L_x and L_y do not commute; in fact⁶

$$\begin{aligned} [L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\ &= [yp_z, zp_x] - [yp_z, xp_z] - [zp_y, zp_x] + [zp_y, xp_z] \end{aligned} \quad (4.31)$$

From the canonical commutation relations⁷ we know that the only operators here that *fail* to commute are x with p_x , y with p_y , and z with p_z . So the two middle terms drop out, leaving

$$[L_x, L_y] = yp_x[p_z, z] + xp_y[z, p_z] = i\hbar(xp_y - yp_x) = i\hbar L_z \quad (4.32)$$

Of course, we could have started out with $[L_y, L_z]$ or $[L_z, L_x]$, but there is no need to calculate these separately—we can get them immediately by cyclic permutation of the indices ($x \rightarrow y, y \rightarrow z, z \rightarrow x$)

$$\boxed{[L_x, L_y] = i\hbar L_z; \quad [L_y, L_z] = i\hbar L_x; \quad [L_z, L_x] = i\hbar L_y} \quad (4.33)$$

Fundamental
commutation
relations for
angular
momentum

⁶Note that all the operators we encounter in quantum mechanics are distributive with respect to addition: $A(B + C) = AB + AC$. In particular, $[A, B + C] = [A, B] + [A, C]$.

⁷The canonical commutation relations are $[r_i, p_j] = -[p_i, r_j] = i\hbar\delta_{ij}$ $[r_i, r_j] = [p_i, p_j] = 0$

These are the fundamental commutation relations for angular momentum; everything else follows from them.

Notice that L_x, L_y and L_z are *incompatible* observables. According to the generalized uncertainty principle (??)

$$\sigma_{L_x}^2 \sigma_{L_y}^2 \geq \left(\frac{1}{2i} \langle i\hbar L_z \rangle \right)^2 = \frac{\hbar^2}{4} \langle L_z \rangle^2$$

or

$$\sigma_{L_x} \sigma_{L_y} \geq \frac{\hbar}{2} |\langle L_z \rangle| \quad (4.34)$$

It would therefore be futile to look for states that are simultaneously eigenfunctions of L_x and L_y . On the other hand, the square of the total angular momentum,

$$L^2 \equiv L_x^2 + L_y^2 + L_z^2 \quad (4.35)$$

does commute with L_x :

$$\begin{aligned} [L^2, L_x] &= [L_x^2, L_x] + [L_y^2, L_x] + [L_z^2, L_x] \\ &= L_y[L_y, L_x] + [L_y, L_x]L_y + L_z[L_z, L_x] + [L_z, L_x]L_z \\ &= L_y(-i\hbar L_z) + (-i\hbar L_z)L_y + L_z(i\hbar L_y) + (i\hbar L_y)L_z \\ &= \end{aligned}$$

(I used (??)) to simplify the commutator; note also that any operator commutes with itself. It follows, of course, that L^2 also commutes with L_y and L_z :

$$[L^2, L_x] = 0, \quad [L^2, L_y] = 0, \quad [L^2, L_z] = 0 \quad (4.36)$$

or, more compactly,

$$[L^2, \mathbf{L}] = 0 \quad (4.37)$$

So L^2 is compatible with each component of \mathbf{L} , and we can hope to find simultaneous eigenstates of L^2 and (say) L_z :

$$L^2 f = \lambda f \quad \text{and} \quad L_z f = \mu f \quad (4.38)$$

We'll use a "ladder operator" technique, very similar to the one we applied to the harmonic oscillator back in Section 2.3.1. Let

$$\boxed{L_{\pm} \equiv L_x \pm iL_y} \quad (4.39)$$

The commutator with L_z is

$$[L_z, L_{\pm}] = [L_z, L_x] \pm i[L_z, L_y] = i\hbar L_y \pm i(-i\hbar L_x) = \pm\hbar(L_x \pm iL_y)$$

so

$$[L_z, L_{\pm}] = \pm\hbar L_{\pm} \quad (4.40)$$

And, of course,

$$[L^2, L_{\pm}] = 0 \quad (4.41)$$

I claim that if f is an eigenfunction of L^2 and L_z , so also is $L_{\pm}f$: (4.41) says

$$L^2(L_{\pm}f) = L_{\pm}(L^2f) = L_{\pm}(\lambda f) = \lambda(L_{\pm}f) \quad (4.42)$$

so $L_{\pm}f$ is an eigenfunction of L^2 , with the same eigenvalue λ , and (4.40) says

$$\begin{aligned} L_z(L_{\pm}f) &= (L_zL_{\pm} - L_{\pm}L_z)f + L_{\pm}L_zf = \pm\hbar L_{\pm}f + L_{\pm}(\mu f) \\ &= (\mu \pm \hbar)(L_{\pm}f) \end{aligned} \quad (4.43)$$

so $L_{\pm}f$ is an eigenfunction of L_z with the *new* eigenvalue $\mu \pm \hbar$. We call L_+ the "raising" operator, because it *increases* the eigenvalue of L_z by \hbar , and L_- the "lowering" operator, because it *lowers* the eigenvalue by \hbar .

For a given value of λ , then we obtain a "ladder" of states, with each "rung" separated from its neighbors by one unit of \hbar in the eigenvalue of L_z . To ascend the ladder we apply the raising operator, and to descend, the lowering operator. But this process cannot go on forever. Eventually we're going to reach a state for which the z -component exceeds the *total*, and that cannot be. There must exist a "top rung", such that

$$L_+f_t = 0 \quad (4.44)$$

Let $\hbar l$ be the eigenvalue of L_z at this top rung (the appropriateness of the letter " l " will appear in a moment):

$$L_zf_t = \hbar lf_t; \quad L^2f_t = \lambda f_t \quad (4.45)$$

Now,

$$\begin{aligned} L_{\pm}L_{\mp} &= (L_x \pm iL_y)(L_x \mp iL_y) = L_x^2 + L_y^2 \mp i(L_xL_y - L_yL_x) \\ &= L^2 - L_z^2 \mp i(\hbar L_z) \end{aligned}$$

or, putting in the other way around,

$$L^2 = L_{\pm}L_{\mp} + L_z^2 \mp \hbar L_z \quad (4.46)$$

It follows that

$$L^2f_t = (L_-L_+ + L_z^2 + \hbar L_z)f_t = (0 + \hbar^2l^2 + \hbar^2l)f_t = \hbar^2l(l+1)f_t$$

and hence

$$\lambda = \hbar^2l(l+1) \quad (4.47)$$

This tells us the eigenvalue of L^2 in terms of the *maximum* eigenvalue of L_z .

Meanwhile, there is also (for the same reason) a *bottom* rung, f_b , such that

$$L_-f_b = 0 \quad (4.48)$$

Let $\hbar \bar{l}$ be the eigenvalue of L_z at this bottom rung:

$$L_zf_b = \hbar \bar{l}f_b; \quad L^2f_b = \lambda f_b \quad (4.49)$$

Using (4.46), we have

$$L^2f_b = (L_+L_- + L_z^2 - \hbar L_z)f_b = (0 + \hbar^2\bar{l}^2 - \hbar^2\bar{l})f_b = \hbar^2\bar{l}(\bar{l}-1)f_b$$

and therefore

$$\lambda = \hbar^2 \bar{l}(\bar{l} - 1) \quad (4.50)$$

Comparing (4.47) and (4.116), we see that $l(l+1) = \bar{l}(\bar{l}-1)$, so either $\bar{l} = l+1$ (which is absurd—the bottom rung would be higher than the top rung!) or else

$$\bar{l} = -l \quad (4.51)$$

Evidently the eigenvalues of L_z are $m\hbar$ (the appropaites of this letter will also be clear in a moment) goes from $-l$ to $+l$ in a N integer steps. In a particular, it follows that $l = -l+N$, and hence $l = N/2$, so l must be *an integer or a half-integer*. The eigenfunctions are characterized by the numbers l and m :

$$\boxed{L^2 f_l^m = \hbar^2 l(l+1) f_l^m; \quad L_z f_l^m = \hbar m f_l^m} \quad (4.52)$$

where

$$l = 0, 1/2, 1, 3/2, \dots; \quad m = -l, -l+1, \dots, l-1, l \quad (4.53)$$

For a given value of l , there are $2l+1$ different values of m (i.e., $2l+1$ "rungs" on the "ladder").

I hope you're impressed: By *purely algebraic means*, startin with the fundamental commutation relations for angular momentum (4.33), we have determined the eigenvalues of L^2 and L_z —without ever seeing the eigenfunctions themselves! We turn now to the problem of constructing the eigenfunctions. but we're headed. I'll begin with the punch line: $f_l^m = Y_l^m$ the eigenfunctions of L^2 and L_z are nothing but the old spherical harmonics, whics we came upon by a quite different route (that's ehy I chose the letters l and m , of course). And I can now tell you whi the spherical harmonics are orthogonal: They are eigenfucntions of hermitian operatos (L^2 and L_z) belonging to distinct eigenvalues.

Eigenfunctions

First of all we need no rewrite L_x, L_y and L_z in spherical coordinates. Now, $\mathbf{L} = (\hbar/i)(\mathbf{r} \times \nabla)$, and the gradient, in spherical coordinates, is:

$$\nabla = \hat{r} \frac{\partial}{\partial r} + \hat{\theta} \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\phi} \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \quad (4.54)$$

meanwhile, $\mathbf{r} = r\hat{r}$, so

$$\mathbf{L} = \frac{\hbar}{i} \left[r(\hat{r} \times \hat{r}) \frac{\partial}{\partial r} + (\hat{r} \times \hat{\theta}) \frac{\partial}{\partial \theta} + (\hat{r} \times \hat{\phi}) \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right]$$

. But $(\hat{r} \times \hat{r}) = 0$, $(\hat{r} \times \hat{\theta}) = \hat{\phi}$, and $(\hat{r} \times \hat{\phi}) = -\hat{\theta}$, and hence

$$\mathbf{L} = \frac{\hbar}{i} \left(\hat{\phi} \frac{\partial}{\partial \theta} - \hat{\theta} \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right) \quad (4.55)$$

The unit vectors $\hat{\theta}$ and $\hat{\phi}$ can be resolved into their cartesian components:

$$\hat{\theta} = (\cos \theta \cos \phi) \hat{i} + (\cos \theta \sin \phi) \hat{j} + (\sin \theta) \hat{k} \quad (4.56)$$

$$\hat{\phi} = -(\sin \theta)\hat{i} + (\cos \phi)\hat{j} \quad (4.57)$$

Thus

$$\mathbf{L} = \frac{\hbar}{i} \left[(-\sin \phi \hat{i} + \cos \phi \hat{j}) \frac{\partial}{\partial \theta} - (\cos \theta \cos \phi \hat{i} + \cos \theta \sin \phi \hat{j} - \sin \theta \hat{k}) \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right]$$

Evidently

$$L_x = \frac{\hbar}{i} \left(-\sin \phi \frac{\partial}{\partial \theta} - \cos \phi \cot \theta \frac{\partial}{\partial \phi} \right) \quad (4.58)$$

$$L_y = \frac{\hbar}{i} \left(+\cos \phi \frac{\partial}{\partial \theta} - \sin \phi \cot \theta \frac{\partial}{\partial \phi} \right) \quad (4.59)$$

and

$$L_z = \frac{\hbar}{i} \frac{\partial}{\partial \phi} \quad (4.60)$$

We shall also need the raising and lowering operators:

$$L_{\pm} = L_x \pm iL_y = \frac{\hbar}{i} \left[(-\sin \phi \pm i \cos \phi) \frac{\partial}{\partial \theta} - (\cos \phi \pm i \sin \phi) \cot \theta \frac{\partial}{\partial \phi} \right]$$

But $\cos \phi \pm i \sin \phi = e^{\pm i\phi}$, so

$$L_{\pm} = \pm \hbar e^{\pm i\phi} \left(\frac{\partial}{\partial \theta} \pm i \cot \theta \frac{\partial}{\partial \phi} \right) \quad (4.61)$$

In particular

$$L_+ L_- = -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \cot^2 \theta \frac{\partial^2}{\partial \phi^2} + i \frac{\partial}{\partial \phi} \right) \quad (4.62)$$

and hence

$$L^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] \quad (4.63)$$

We are now in position to determine $f_l^m(\theta, \phi)$. It's an eigenfunction of L^2 , with eigenvalue $\hbar^2 l(l+1)$:

$$L^2 f_l^m = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] f_l^m = \hbar^2 l(l+1) f_l^m$$

But this is precisely the "angular equation" (??). And it's also an eigenfunction of L_z , with the eigenvalue $m\hbar$:

$$L_z f_l^m = \frac{\hbar}{i} \frac{\partial}{\partial \phi} f_l^m = \hbar m f_l^m$$

, but this is equivalent to the azimuthal equation (??). We have already solved this system of equations: The result (appropriately normalized) is the spherical harmonic, $Y_l^m(\theta, \phi)$. *Conclusion*: Spherical harmonics are eigenfunctions of L^2 and L_z . When we solved the Schrodinger

equation by separation of variables, in Section ??, we were inadvertently constructing simultaneous eigenfunctions of the three commuting operators H, L^2 and L_z :

$$H\psi = E\psi, \quad L^2\psi = \hbar^2 l(l+1)\psi, \quad L_z\psi = \hbar m\psi \quad (4.64)$$

Incidentally, we can use (4.63) to rewrite the Schrodinger equation (4.11) more compactly:

$$\frac{1}{2mr^2} \left[-\hbar^2 \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + L^2 \right] \psi + V\psi = E\psi$$

There is a curious final twist to this story, for the *algebraic* theory of angular momentum permits l (and hence also m) to take on *half-integer* values (4.53), whereas separation of variables yielded eigenfunctions only for *integer* values (??). You might suppose that the half-integer solutions are spurious, but it turns out that they are of profound importance, as we shall see in the following sections.

4.4 Spin

In classical mechanics, a rigid object admits two kinds of angular momentum: **orbital** ($\mathbf{L} = \mathbf{r} \times \mathbf{p}$), associated with the motion of the center of mass, and **spin** ($\mathbf{S} = I\boldsymbol{\omega}$), associated with motion about the center of mass. For example, the earth has orbital angular momentum attributable to its revolution around the sun, and spin angular momentum coming from its daily rotation about the north-south axis. In the classical context this distinction is largely a matter of convenience, for when you come right down to it, \mathbf{S} is nothing but the sum total of the "orbital" angular momenta of all the rocks and dirt clods that happens in quantum mechanics, and here the distinction is absolutely fundamental. In addition to orbital angular momentum, associated (in the case of hydrogen) with the motion of the electron around the nucleus (and described by the spherical harmonics), the electron also carries *another* form of angular momentum, which has nothing to do with motion in space (and which is not, therefore, described by any function of the position variables r, θ, ϕ) but which is somewhat analogous to classical spin (and for which, therefore, we use the same word.) It doesn't pay to press this analogy too far: The electron (as far as we know) is a structureless point particle, and its spin angular momentum cannot be decomposed into orbital angular momenta of constituent parts. Suffice it to say that elementary particles carry **intrinsic** angular momentum (\mathbf{S}) in addition to their "extrinsic" angular momentum (\mathbf{L}).

The *algebraic* theory of spin is a carbon copy of the theory of orbital angular momentum, beginning with the fundamental commutation relations.

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y \quad (4.65)$$

It follows (as before) that the eigenvectors of S^2 and S_z satisfy⁸

$$S^2 |sm\rangle = \hbar^2 s(s+1) |sm\rangle; \quad S_z |sm\rangle = \hbar m |sm\rangle \quad (4.66)$$

⁸Because the eigensates of spin are not *functions*, I will use the "ket" notation for them. (I could have done the same in Section ??, writing $|lm\rangle$ in place of Y_l^m , but in that context the function notation seems more natural.) By the way, I'm running out of letters, so I'll use m for the eigenvalue of S_z just as I did for L_z (some authors write m_l and m_s at this stage, just to be absolutely clear).

and

$$S_{\pm} |sm\rangle = \hbar \sqrt{s(s+1) - m(m \pm 1)} |s(m \pm 1)\rangle \quad (4.67)$$

where $S_{\pm} \equiv S_x \pm i S_y$. But this time the eigenfunctions are not spherical harmonics (they're not functions of θ and ϕ at all), and there is no *a priori* reason to exclude the half-integer values of s and m :

$$s = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots; \quad m = -s, -s+1, \dots, s-1, s \quad (4.68)$$

It so happens that every elementary particle has a *specific and immutable* value of s , which we call **the spin** of that particular species: pi meson have spin 0; electrons have spin 1/2; photons have spin 1; deltas have spin 3/2; gravitons have spin 2; and so on. By contrast; the *orbital* angular momentum quantum number l (for an electron in a hydrogen atom, say) can take on any (integer) values you please, and will change from one to another when the system is perturbed. But s is *fixed*, for any given particle, and this makes the theory of spin comparatively simple.

Spin 1/2

By far the most important case is $s = 1/2$, for this is the spin of the particles that make up ordinary matter (protons, neutrons, and electrons), as well as all quarks and all leptons. Moreover, once you understand spin 1/2, it is a simple matter to work out the formalism for any higher spin. There are just *two* eigenstates: $|\frac{1}{2} \frac{1}{2}\rangle$, which we call **spin up** (informally, \uparrow), and $|\frac{1}{2} (-\frac{1}{2})\rangle$, which we call **spin down** (\downarrow). Using these as basis vectors, the general state of a spin $-1/2$ particle can be expressed as a two-element column matrix (or **spinor**):

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} = a\chi_+ + b\chi_- \quad (4.69)$$

with

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (4.70)$$

representing spin up, and

$$\chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (4.71)$$

for spin down.

Meanwhile, the spin operators become 2×2 matrices, which we can work out by noting their effect on χ_+ and χ_- . (4.66) says

$$\mathbf{S}^2 \chi_+ = \frac{3}{4} \hbar^2 \chi_+ \quad \text{and} \quad \mathbf{S}^2 \chi_- = \frac{3}{4} \hbar^2 \chi_- \quad (4.72)$$

If we write \mathbf{S}^2 as a matrix with (as yet) undetermined elements,

$$\mathbf{S}^2 = \begin{pmatrix} c & d \\ e & f \end{pmatrix}$$

then the first equation says

$$\begin{pmatrix} c & d \\ e & f \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{3}{4} \hbar^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} c \\ e \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \hbar^2 \\ 0 \end{pmatrix}$$

so $c = (3/4)\hbar^2$ and $e = 0$. The second equation says

$$\begin{pmatrix} c & d \\ e & f \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{3}{4}\hbar^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} d \\ f \end{pmatrix} \begin{pmatrix} 0 \\ \frac{3}{4}\hbar^2 \end{pmatrix}$$

so $d = 0$ and $f = (3/4)\hbar^2$. *Conclusion:*

$$\mathbf{S}^2 = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (4.73)$$

Similarly,

$$\mathbf{S}\chi_+ = \frac{\hbar}{2}\chi_+, \quad \mathbf{S}_z\chi_- = -\frac{\hbar}{2}\chi_- \quad (4.74)$$

from which it follows that

$$\mathbf{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (4.75)$$

Meanwhile, (4.67) says

$$\mathbf{S}_+\chi_- = \hbar\chi_+ \quad \mathbf{S}_z\chi_+\hbar\chi_-, \quad \mathbf{S}_+\chi_+ = \mathbf{S}_-\chi_- = 0$$

so

$$\mathbf{S}_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{S}_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (4.76)$$

Now $S_{\pm} = S_x \pm iS_y$, so $S_x = (1/2)(S_+ + S_-)$ and $S_y = (1/2i)(S_+ - S_-)$, and hence

$$\mathbf{S}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{S}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (4.77)$$

Since $\mathbf{S}_x, \mathbf{S}_y$ and \mathbf{S}_z all carry a factor of $\hbar/2$, it is tidier to write $\mathbf{S} = (\hbar/2)\sigma$, where

$$\boxed{\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \quad (4.78) \quad \text{Pauli spin matrices}$$

These are the famous **Pauli spin matrices**. Notice that $\mathbf{S}_x, \mathbf{S}_y, \mathbf{S}_z$ and \mathbf{S}^2 are all *hermitian* (as they should be, since they represent observables). On the other hand, \mathbf{S}_+ and \mathbf{S}_- are *not* hermitian—evidently they are not observables.

The eigenspinors of \mathbf{S}_z are (of course):

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, (\text{eigenvalue } +\frac{\hbar}{2}); \quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, (\text{eigenvalue } -\frac{\hbar}{2}) \quad (4.79)$$

If you measure S_z on a particle in the general state χ (4.69), you could get $+\hbar/2$, with probability $|a|^2$, or $-\hbar/2$, with probability $|b|^2$. Since these are the *only* possibilities,

$$|a|^2 + |b|^2 = 1 \quad (4.80)$$

(i.e., the spinor must be *normalized*).⁹

⁹People often say that $|a|^2$ is the "probability that the particle is in the spin-up state", but this is sloppy language; what they *mean* is that if you *measured* S_z , $|a|^2$ is the probability you'd get $\hbar/2$

But what if, instead, you chose to measure S_x ? What the possible results, and what are thei respective probabilities? According to the generalized statistical interpretation, we need to know the eigenvalues and eigenspinors of \mathbf{S}_x . The characteristic equation is

$$\begin{vmatrix} -\lambda & \hbar/2 \\ \hbar/2 & -\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 = \left(\frac{\hbar}{2}\right)^2 \Rightarrow \lambda = \pm \frac{\hbar}{2}$$

Not surprisingly, the possible values for S_x are the same as those for S_z . The eigenspinor are obtained in the usual way:

$$\frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \pm \frac{\hbar}{2} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \Rightarrow \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \pm \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

, so $\beta = \pm\alpha$. Evidently the (normalized) eigenspinors of \mathbf{S}_x are

$$\chi_+^{(x)} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, (\text{eigenvalue } +\frac{\hbar}{2}); \quad \chi_-^{(x)} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, (\text{eigenvalue } -\frac{\hbar}{2}) \quad (4.81)$$

As the eigenvectors of hermitian matrix, they span the space; the generic spinor χ (4.69) can be expressed as a linear combination of them:

$$\chi = \left(\frac{a+b}{\sqrt{2}}\right) \chi_+^{(x)} + \left(\frac{a-b}{\sqrt{2}}\right) \chi_-^{(x)} \quad (4.82)$$

If you measure S_x , the probability of getting $+\hbar/2$ is $(1/2)|a+b|^2$, and the probability of getting $-\hbar/2$ is $(1/2)|a-b|^2$. (You should check for yourself that these probabilities add up to 1).

I'd like now to walk you through an imaginary measurement scenario involving spin 1/2, because it serves to illustrate in very concrete terms some of the abstract ideas we discussed back in Chapter 1. Let's say we start out with a particle in the state χ_+ . If someone asks, "What is the z -component of that particle's spin angular momentum?", we could answer unambiguously: $+\hbar/2$. For a measurement of S_z is certain to return that value. But if our interrogator asks instead, "What is the x -component of that particle's spin angular momentum?" we are obliged to equivocate: If you measure S_x , the chances are fifty-fifty (in the sense of Section ??), he will regard this as an inadequate—not to say impertinent—response: "Are you telling me that you *don't know* the true state of that particle?" On the contrary; I know *precisely* what the state of the particle is: χ_+ . "Well, then, how come you can't tell me what the x -component of its spin is?" Because it simply *does not have* a particular x -component of spin. Indeed, it *cannot*, for if both S_x and S_z were well-defined, the uncertainty principle would be violated.

At this point our challenger grabs the test-tube and *measures* the x -component of its spin; let's say he gets the value $+\hbar/2$. "Aha!" (he shouts in triumph), "You *lied*! This particle has a perfectly well-defined value of $S_x = \hbar/2$!" Well, sure—it does *now*, but that doesn't prove it *had* the value, prior to your measurement. "You have obviously been reduced to splitting hairs. And anyway, what happened to your uncertainty principle? I now know both S_x and S_z !" I'm sorry, but you do *not*: In the course of your measurement, you altered the particle's state; it is now in the state $\chi_+^{(x)}$, and whereas you know the value of S_x , you no longer know the value of S_z . "But I was extremely careful not to disturb the particle when I measured S_x !" Very well, if you don't believe me, *check it out*: Measure S_z , and see what you get. (Of course, he *may*

get $+\hbar/2$, which will be embarrassing to my case—but if we repeat this whole scenario over and over, half the time he will get $-\hbar/2$.)

To the layman, the philosopher, or the classical physicist, a statement of the form "this particle doesn't have a well-defined position" (or momentum, or x -component of spin angular momentum, or whatever) sound vague, incompetent, or (worst of all) profound. It is none of these. But its precise meaning is, I think, almost impossible to convey to anyone who has not studied quantum mechanics in some depth. If you find your own comprehension slipping, from time to time (if you *don't*, you probably haven't understood the problem), come back to the spin-1/2 system: It is the simplest and cleanest context for thinking through the conceptual paradoxes of quantum mechanics.

Electron in a Magnetic Field

A spinning charged particle constitutes a magnetic dipole. Its **magnetic dipole moment**, μ is proportional to its spin angular momentum, S

$$\mu = \gamma S \quad (4.83)$$

the proportionality constant, γ , is called the **gyromagnetic ratio**. When a magnetic dipole is placed in a magnetic field \mathbf{B} , it experiences a torque, $\mu \times \mathbf{B}$, which tends to line it up parallel to the field (just like a compass needle). The energy associated with this torque is

$$H = -\mu \cdot \mathbf{B} \quad (4.84)$$

so the Hamiltonian of a spinning charged particle, at rest in a magnetic field \mathbf{B} is

$$H = -\gamma \mathbf{B} \cdot \mathbf{S} \quad (4.85)$$

Example 2 Larmor precession:

Example 3 The Stern-Gerlach experiment:

Addition of Angular Momenta

Suppose now that we have *two* spin -1/2 particles—for example, the electron and the proton in the ground state¹⁰ of hydrogen. Each can have spin up or spin down, so there are four possibilities in all¹¹

$$\uparrow\uparrow, \quad \uparrow\downarrow, \quad \downarrow\uparrow, \quad \downarrow\downarrow \quad (4.86)$$

¹⁰I put them in the ground state so there won't be any *orbital* angular momentum to worry about.

¹¹More precisely, each particle is in a *linear combination* of spin up and spin down, and the composite system is in a *linear combination* of the four states listed.