Quantum Mechanics

Diez B. Borja
December 2, 2022

Preface

This notes has been written by *B. Diez* *mainly from Introduction to Quantum Mechanics by David J. Griffiths and own knowledge.*
Since these notes were written for self-study, this is an unfinished version.

Contents

Chapter 1

The Wave Function

1.1　The Schrodinger Equation

Imagine a particle of mass m, constrained to move along the x-axis, subjet to some to specifies force $F(x,t)$. The program of *classical* mechanics is to determine the position of the particle at any give time: $x(t)$. Once we know that, we can figure out the velocity ($v = \mathrm{d}x/\mathrm{d}t$), 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 kins we shall consider, and, fortunately, the only kins that *occur* at the microscopic level–the force can be expresed as the derivative of a potential energy function [1], $F = -\partial V/\partial x$, and Newton's law reads $m\mathrm{d}^2x/\mathrm{d}t^2 = -\partial V/\partial x$) This, together with appropiate 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 loocking for ir the particle's **wave function**, $\Psi(x,t)$, and we get it by solving the **Schrodinger equaation**:

$$\boxed{i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V\Psi}$$

(1.1) Schrodinger equation

This equation plays a role logically analogous to Newton's second law: Given suitable initial conditions (typically, $\Psi(x,0)$), the Schrodignes 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

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

(1.2) Statistical interpretation

---

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

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

Yhe statistical interpretation introduces a kind of **indeterminacy** into quantum mechanics, for even if you know evertything 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" oenes, in which the wave function evolves in a leisurely fashion inder the Schrodinger equation, and "measurements", in which Ψ suddenly and discontinously collapses [2].

1.3 Probability

I will give some definitions. The total number is

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

The probability of getting j is

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

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

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

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

$$\boxed{< f(j) >= \sum_{j=0}^{\infty} f(j)P(j)} \tag{1.6}$$

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

$$\Delta j = j - < j > \tag{1.7}$$

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

---

[2]The 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 miscroscopic (quantum) system and a mecorscopic (classical) measuring apparatus (as Bohr insisted), or is it intervention of a conscious "oberver" (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 de *absolute value* of Δj. But absolute values are nasty to work with; instead, we get around the sign problem by *squaring* before averaging:

$$\sigma^2 \equiv\ <(\Delta j)^2> \tag{1.8}$$

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

There is a useful little theorem pn variances:

$$\sigma = \sqrt{<j^2> - <j>^2} \tag{1.9}$$

1.4   Continous Variables

I hace 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 generaliza to *continous* distributions. If the interval is sufficiently short, this probabilty is *proportional to the length of the interval*. This

$$\text{probabilty that an individual lies between } x \text{ and } (x+\mathrm{d}x) = \rho(x)\mathrm{d}x \tag{1.10}$$

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

$$P_{ab} = \int_a^b \rho(x)\mathrm{d}x \tag{1.11}$$

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

$$1 = \int_{-\infty}^{+\infty} \rho(x)\mathrm{d}x \tag{1.12}$$

$$<x> = \int_{-\infty}^{+\infty} x\rho(x)\mathrm{d}x \tag{1.13}$$

$$<f(x)> = \int_{-\infty}^{+\infty} f(x)\rho(x)\mathrm{d}x \tag{1.14}$$

$$\sigma^2 \equiv\ <(\Delta x)^2> = <x^2> - <x>^2 \tag{1.15}$$

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*):

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

4

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 $\Psi = 0$. Shuch **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 rematkable 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

$$\boxed{< x >= \int_{-\infty}^{+\infty} x|\Psi(x,t)|^2 \mathrm{d}x}$$

(1.17)   Expectation value of x

What exactly does this mean? It emphatically does not meand that if you measure the position of one particle over and over, $\int x|\Psi 1|^2 \mathrm{d}x$ is the average of the results you'll get. Rather, $< x >$ is the average of measurements, performes on particles *all in the satate* 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 **ensamble** of particles, each in the same state Ψ, and measure the positions of all of them: $< x >$ is the average of *these* results.

Now, as time goes on, $< x >$ will cahnge (because of the time dependence of Ψ), ando we might be interested in knowing how fast ir moves. We see that

$$\frac{\mathrm{d} < x >}{\mathrm{d}t} = -\frac{i\hbar}{m} \int \Psi^* \frac{\partial \Psi}{\partial x} \mathrm{d}x$$

(1.18)

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

$$< v >= \frac{\mathrm{d} < x >}{\mathrm{d}t}$$

(1.19)

(1.18) tell us, then, how to calculate $< v >$ directly from Ψ.

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

$$< p >= m\frac{\mathrm{d} < x >}{\mathrm{d}t} = -i\hbar \int \left(\Psi^* \frac{\partial \Psi}{\partial x} \right) \mathrm{d}x$$

(1.20)   Momentum

Let me write the expressions for $< x >$ and $< p >$ in a more siggestive way:

$$< x >= \int \Psi^*(x)\Psi \mathrm{d}x$$

(1.21)

$$< p >= \int \Psi^* \left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right) \Psi \mathrm{d}x$$

(1.22)

---

[3]Evidently $\Psi(x,t)$ must go to zero faster than $1/\sqrt{|x|}$, as $|x| \to \infty$. Incidentaly, normalziation only fixed the *modulus* of A: the phase remains undetermiend. 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 appropiate 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 quantuty, $Q(x, p)$, we simply replace every p by $(\hbar/i)(\partial/\partial x)$, insert the resulting operator between Ψ^* and Ψ, and integrate:

$$\boxed{< Q(x, p) >= \int \Psi^* Q\left(x, \frac{\hbar}{i}\frac{\partial}{\partial x}\right)\Psi \mathrm{d}x} \tag{1.23}$$

1.7   The Uncertaintly Principle

Teh wavelenght of Ψ is related to the *momentum* of the particle by the **de Broglie formula**:

$$p = \frac{\hbar}{\lambda} = \frac{2\pi\hbar}{\lambda} \tag{1.24}$$

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 ñedd precisely is its momentum. Quantitatively.

$$\boxed{\sigma_x \sigma_p \geq \frac{\hbar}{2}} \tag{1.25}$$

Uncertanty principle

where σ_x es 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 yiled precise answers–the "spread" here refers to the fact that measurements on identically prepared systems do not yiled identical results. You can, if you want, construct a state such the repeated position measurements will be vey close together (bu making Ψ a localized "spike"), but you sill pay a price: Momentum measurements on this state will be widely scattered. Oer you can prepare a state with a reproducible momentum (by making Ψ a long sinusoidal wave), but in that case, position measurements will de widely scattered.

Chapter 2

Time-Independent Schrodinger Equation

2.1 Stationary States

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

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

for a specified potential[1] $V(x,t)$. 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*,

$$\Psi(x,t) = \psi(x)\varphi(t) \tag{2.2}$$

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

For separable solutions we have

$$\frac{\partial \Psi}{\partial t} = \psi \frac{\mathrm{d}\varphi}{\mathrm{d}t}, \qquad \frac{\partial^2 \Psi}{\partial x^2} = \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}\varphi$$

(*ordinaty* derivatives, now), and Schrodinger equation reads

$$i\hbar \frac{\mathrm{d}\varphi}{\mathrm{d}t} = -\frac{\hbar^2}{2m} \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}\varphi + V\psi\varphi \tag{2.3}$$

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

$$i\hbar \frac{1}{\varphi} \frac{\mathrm{d}\varphi}{\mathrm{d}t} = -\frac{\hbar^2}{2m} \frac{1}{\psi} \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V \tag{2.4}$$

Now, the left side is a function of t alone, and the right side is a function of x alone[2]. The only way this can possibly be true is if both sides are in fact *constant* − otherwise, by varying

---

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

[2]Note 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 lonegr equal. For reasons that will appear in a moment, we shall call the separation constan E. Then

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

or

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = -\frac{iE}{\hbar}\varphi$$

and

$$-\frac{\hbar^2}{2m}\frac{1}{\psi}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V = E \tag{2.5}$$

or

$$\boxed{-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V\psi = E\psi} \tag{2.6}$$

Time-independent Schrodinger equation

Separation of variables hs 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 $\mathrm{d}t$ 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

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

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

The rest of thus chapter will be devoted to solving the time-independent Schrodinger equation, fot 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 so 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 **stationaty states**. Although the wave function itself,

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

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

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

does not–the time-dependence cancels out[3]. The same thing happens in calculate the expectation value of any dynamil variable; (1.23) reduces to

$$<Q(x,p)> = \int \psi^* Q\left(x, \frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}x}\right)\psi\,\mathrm{d}x \tag{2.10}$$

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, $<x>$ is constant, and hence $<p> = 0$ (See (1.20)). Nothing ever happens in a stationary state.

---

[3]For normalizable solutions, E must be real

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

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

The corresponding Hamiltonian *operator*, obtained by the canonical substitution $p \rightarrow (\hbar/i)(\partial/\partial x)$, is therefore[4]

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

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

$$\hat{H}\psi = E\psi \tag{2.13}$$

and the expectation value of the total energy is

$$< H >= \int \psi^* \hat{H}\psi \mathrm{d}x = E \int |\psi|^2 \mathrm{d}x = E \int |\Psi|^2 \mathrm{d}x = E \tag{2.14}$$

(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

$$< H^2 >= \int \psi^* \hat{H}^2\psi \mathrm{d}x = E^2 \int |\psi|^2 \mathrm{d}x = E^2$$

So the variance of H is

$$\sigma_H^2 =< H^2 > - < H >^2 = E^2 - E^2 = 0 \tag{2.15}$$

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

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

Now the time-dependent Schrodinger equation (2.1) has the property that any linear combination[5] 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

---

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

[5]A **linear combination** of the functions $f_1(z), f_z(z), ...$ is an expression of the form

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

where $c_1, c_2, ...$ are any (complex) constants.

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

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

Let me recapitulate, from a somewhat different persperctive. Here's the generic problem: You're given a (time-independent) potencial $V(x)$ 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 $(\psi_1(x), \psi_2(x), ...)$, each with its own assocaited energy $(E_1, E_2, ...)$. To fit $\Psi(x, 0)$ you write down the general linear combination of these solutions:

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

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

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

The separable solutions themselves,

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

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 enrgies are different, for different stationary states, and the exponentials do not cancel, when you calculate $|\Psi|^2$.
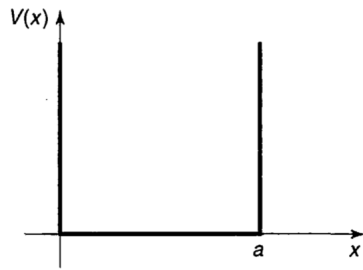
2.2   The Infinite Square Well

Figure 2.1: The infinte square well potential (2.20)

10

Suppose

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

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

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

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = E\psi \tag{2.21}$$

or

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = -k^2\psi, \qquad \text{where } k \equiv \frac{\sqrt{2mE}}{\hbar} \tag{2.22}$$

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

$$\psi(x) = A\sin kx + B\cos kx \tag{2.23}$$

where A and B are arbitraty constants. Tipically, these constants are fixed by the **boundary conditions** of the problem. What are the appropiate boundary conditions for $\psi(x)$? Ordinatily, both ψ and $\mathrm{d}\psi/\mathrm{d}x$ are continuous, but where the potential goes to infinity only the first of these applies.

Continuity of $\psi(x)$ requieres that

$$\psi(0) = \psi(a) = 0 \tag{2.24}$$

so as to join onto the solution outisde 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

$$\psi(x) = A\sin kx \tag{2.25}$$

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

$$ka = 0, \pm\pi, \pm 2\pi, \pm 3\pi, ... \tag{2.26}$$

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

$$k_n = \frac{n\pi}{a}, \qquad \text{with } n = 1, 2, 3, ... \tag{2.27}$$

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:

$$\boxed{E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{n^2\pi^2\hbar^2}{2ma^2}} \tag{2.28}$$

Possible values of E

11

In radical contrast to the classical case, a quantum particle in the infinite square well cannot have just any old energy–it has be one of these special **allowed** values[6]. To find A we normalize ψ:

$$\int_a^a |A|^2 \sin^2(kx)\mathrm{d}x = |A|^2 \frac{a}{2} = 1, \qquad \text{so } |A|^2 = \frac{2}{a}$$

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

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

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

1. They are alternately **even** and **odd**, with respecto 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

$$\int \psi_m(x)^* \psi_n(x)\mathrm{d}x = 0 \tag{2.30}$$

   whenever $m \neq n$ We can combine orthogonality and normalization into a single satatement[7]:

$$\boxed{\int \psi_m(x)^* \psi_n(x)\mathrm{d}x = \delta_{mn}} \tag{2.31}$$

   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:

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

(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

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

---

[6]Notice that the quantization of energy emerged as a rather technical consequence of the boudnary conditions on solutions to the time-independent Schrodinger equation

[7]In 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 habir of putting it there.

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

$$\boxed{c_n = \int \psi_n(x)^* f(x)\mathrm{d}x}$$
(2.34)

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

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

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

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

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

It remaisn only for me to demosntrate that I can fit any prescribed initial wave fucntion, $\Psi(x,0)$, by appropiate 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:

$$c_n = \sqrt{\frac{2}{a}} \int_a^a \sin\left(\frac{n\pi}{a}x\right) \Psi(x,0)\mathrm{d}x$$
(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 the 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 nth stationary state", but this is a bad language; the particle is in the state Ψ, not Ψ_n, and, anuhow, 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 og the "allowed" values–hebce the name–and $|c_n|^2$ is the probability of getting the particular value E_n).

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

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

Indeed, this follows follows from the normalization of Ψ.

Moreover, the expectation value of the energy must be

$$< H >= \sum_{n=1}^{\infty} |c_n|^2 E_n$$ (2.39)

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 **conservatio 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**,

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

(ignoring friction), and the solution is

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

where

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

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

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

its graph is a parabola.

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

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

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

(it it 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:

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \frac{1}{2}m\omega^2 x^2 \psi = E\psi$$ (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, usin 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 se the algebraic methos first, because it is quicker and simpler (and a lot more fun)[8].

---

[8]We'll encounter some of the same strategies in the theory og 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:

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

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

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

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 de same as px). Still, this does motivate us to examine the quantities

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

(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 commuatator of operators A and B (written with brackets) is

$$[A, B] \equiv AB - BA \tag{2.47}$$

In this notation,

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

We need to figure out the commuatator 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:

$$[x, p]f(x) = \left[x\frac{\hbar}{i}\frac{\mathrm{d}f}{\mathrm{d}x} - \frac{\hbar}{i}\frac{\mathrm{d}xf}{\mathrm{d}x}\right] = \frac{\hbar}{i}\left(x\frac{\mathrm{d}f}{\mathrm{d}x} - x\frac{\mathrm{d}f}{\mathrm{d}x} - x\right) = i\hbar f(x) \tag{2.49}$$

Dropping the test function, which has served its purpose,

$$\boxed{[x, p] = i\hbar} \tag{2.50}$$

Cannonical commutation relation

15

This lovely and obiquitous result is known as the **cannonical commutation relation**[9].

With this, (2.48) becomes

$$a_- a_+ = \frac{1}{\hbar\omega} H + \frac{1}{2} \tag{2.51}$$

or

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

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

$$a_+ a_- = \frac{1}{\hbar\omega} H - \frac{1}{2} \tag{2.53}$$

In particular,

$$[a_-, a_+] = 1 \tag{2.54}$$

So the Hamiltonian can equally well be written

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

In terms of $a_\pm$, then, the Schrodinger equation[10] for the harmonic oscilaltor takes the form

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

(in equations like this you rrad 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 enrgy 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 $(E - \hbar\omega)$

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

Here, then is a wonderful machine for generating new solution, with higer and lower energies–if we could just find *one* solution, to ger started! We call $a_\pm$ **ladder operators**, because they allow us to climb up and doen 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 forme: There occurs a "lowest rung" (call it ψ_0) such that

$$a_-\psi_0 = 0 \tag{2.57}$$

---

[9]In 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 cannoncial commuation relation as an *axiom* of the theory, and use ir to *derive* $p = (\hbar/i)\mathrm{d}/\mathrm{d}x$

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

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

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

or

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

The differential equation is easy to solve:

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

so

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

We might as well normalize it right away:

$$= |A|^2 \int_{-\infty}^{\infty} e^{-m\omega x^2/\hbar}\mathrm{d}x = |A1|^2\sqrt{\frac{\pi\hbar}{m\omega}}$$

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

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

To determine the energy of this satte 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 $a_-\psi_0 = 0$:

$$E_0 = \frac{1}{2}\hbar\omega \tag{2.59}$$

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:

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

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 foorwork, wo watch closely. we know that $a_\pm\psi_n$ is *proportional* to $\psi_{m\pm1}$,

$$a_+\psi_n = c_n\psi_{n+1}, \qquad a_-\psi_n = d_n\psi_{n-1} \tag{2.61}$$

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

$$\int_{-\infty}^{\infty} f^*(a_\pm g)\mathrm{d}x = \int_{-\infty}^{\infty}(a_\mp f)^* g\mathrm{d}x \tag{2.62}$$

17

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

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

But (invoking (2.56) and (2.60))

$$a_+a_-\psi_n = n\psi_n, \qquad a_-a_+\psi_n = (n+1)\psi_n \tag{2.63}$$

so

$$\int_{\infty}^{\infty} (a_+\psi_n)^*(a_+\psi_n)\mathrm{d}x = |c_n|^2 \int_{-\infty}^{\infty} |\psi_{n+1}|^2\mathrm{d}x = (n+1)\int_{-\infty}^{\infty} |\psi_n|^2\mathrm{d}x$$

$$\int_{-\infty}^{\infty} (a_-\psi_n)^*(a_-\psi_n)\mathrm{d}x = |d_n|^2 \int_{-\infty}^{\infty} |\psi_{n-1}|^2\mathrm{d}x = n\int_{-\infty}^{\infty} |\psi_n|^2\mathrm{d}x$$

But since ψ_n and $\psi_{n\pm1}$ are normalized, it follows that $|c_n|^2 = n+1$ and $|d_n|^2 = n$ and hence.

$$\boxed{a_+\psi_n\sqrt{n+1}\psi_{n+1}, \qquad a_-\psi_n = \sqrt{n}\psi_{n-1}} \tag{2.64}$$

Thus

$$\psi_1 = a_+\psi_0, \qquad \psi_2 = \frac{1}{\sqrt{2}}a_+\psi_1 = \frac{1}{\sqrt{2}}(a_+)^2\psi_2$$

$$\psi_3 = \frac{1}{\sqrt{3}}a_+\psi_2 = \frac{1}{\sqrt{3\cdot2}}(a_+)^3\psi_0, \qquad \psi_4 = \frac{1}{\sqrt{4}}a_+\psi_3 = \frac{1}{\sqrt{4\cdot3\cdot2}}(a_+)^4\psi_0$$

and so on .Clearly

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

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 if the harmonic oscillator are orthogonal:

$$\int_{-\infty}^{\infty} \psi_m^*\psi_n\mathrm{d}x = \delta_{mn} \tag{2.66}$$

This can be proved usign (2.63) and (2.62) twice–fist moving a_+ and then moving a_-:

$$\int_{-\infty}^{\infty} \psi_m^*(a_+a_-)\psi_n\mathrm{d}x = n\int_{-\infty}^{\infty} \psi_m^*\psi_n\mathrm{d}x$$

$$= \int_{-\infty}^{\infty} (a_-\psi_m)^*(a_-\psi_n)\mathrm{d}x = \int_{-\infty}^{\infty} (a_-a_-\psi_m)^*\psi_n\mathrm{d}x$$

$$= m\int_{-\infty}^{\infty} \psi_m^*\psi_n\mathrm{d}x$$

Unless $m = n$, then, $\int \psi_m^*\psi_n\mathrm{d}x$ must be zero. Orthonomality means that we can again use Forier's trcik (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 vakue E_n.

2.3.2 Analytic Method

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

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2 x^2\psi = E\psi \tag{2.67}$$

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

$$\xi \equiv \sqrt{\frac{m\omega}{\hbar}}x \tag{2.68}$$

in terms of ξ the Schrodinger equation reads

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

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

$$k \equiv \frac{2E}{\hbar\omega} \tag{2.70}$$

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

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}\xi^2} \approx \xi^2\psi \tag{2.71}$$

which has the approximate solution

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

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

$$\psi(\xi) \to (\,)e^{-\xi^2/2}, \qquad \text{at large } \xi \tag{2.73}$$

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

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

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

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

and

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

19

so the Schrodinger equation (2.69) becomes

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

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

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

Differentiating the seris term by term,

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

and

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

Putting these into (2.75), we find

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

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

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

and hence that

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

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

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

and starting with a_2, it generates the odd coefficients:

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

We write the complete solution as

$$h(\xi) = h_{\mathsf{even}}(\xi) + h_{\mathsf{odd}}(\xi) \tag{2.79}$$

---

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

where

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

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

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

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 differentual equation.

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

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

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)

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

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's want[12]. There is only one way to wiggle out of this: For normalziable solutions the *power series must terminate.* There must occur some "highest" j (call it n), such that the recursion formula spits out $a_{a+2} = 0$ (this woll 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

$$K = 2_n + 1$$

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

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

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

---

[12]It's no surprise that the ill-behaven solutions are still in (2.78):this recursion relation is equivalent to the Schrodinger equation, so it's got to include both the asymtotic 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, obervables are represented by operators. Mathematically, weave functions satify 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:

$$|\alpha\rangle \rightarrow \boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} \tag{3.1}$$

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

$$\langle\alpha|\beta\rangle = a_1^* b_1 + a_2^* b_2 + \cdots + a_N^* b_N \tag{3.2}$$

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:

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

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 fot our purposes it is much too latge. To represent a possible physical state, the function Ψ must be normalized:

$$\int |\Psi|^2 \mathrm{d}x = 1$$

The set of all **square-integrable functions**, on a specified ineterval[1],

$$f(x) \qquad \text{such that} \qquad \int_a^b |f(x)|^2 \mathrm{d}x < \infty \qquad (3.4)$$

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

$$\boxed{\textbf{Wave functions live in Hilbert space.}} \qquad (3.5)$$

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

$$\langle f | g \rangle \equiv \int_a^b f(x)^* g(x) \mathrm{d}x \qquad (3.6)$$

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

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

Notice in particular that

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

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

$$\langle f | f \rangle = \int_a^b |f(x)|^3 \mathrm{d}x \qquad (3.9)$$

is real and 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** of their inner product is 0; and a set of functions, $\{f_n\}$, is **orthonormal** of they are normalized and mutually orthogonal

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

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

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

If rge functuons $\{f_n(x)\}$ are orthonrmal, the coefficients are given by Fourier's trick

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

---

[1] Fot is, the limits (a and b) will almost always be $pm\infty$, but we might as well keep thing 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:

$$< Q >= \int \Psi^* \hat{Q} \Psi \mathrm{d}x = \left\langle \Psi \middle| \hat{Q} \Psi \right\rangle \tag{3.13}$$

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

$$< Q >=< Q >^* \tag{3.14}$$

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

$$\left\langle \Psi \middle| \hat{Q} \Psi \right\rangle = \left\langle \hat{Q} \Psi \middle| \middle| \Psi \right\rangle \tag{3.15}$$

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

$$\left\langle f \middle| \hat{Q} f \right\rangle = \left\langle \hat{Q} f \middle| f \right\rangle \qquad \text{for all } f(x) \tag{3.16}$$

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:

$$\boxed{\textbf{Obervables are represented by hermitian operators.}} \tag{3.17}$$

Is the momentum operator, for example, hermitian?

$$\langle f | \hat{p} g \rangle = \int_{-\infty}^{\infty} \frac{\hbar}{i} \frac{\mathrm{d}g}{\mathrm{d}x} \mathrm{d}x = \frac{\hbar}{i} \left. f^* g \right|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \left(\frac{\hbar}{i} \frac{\mathrm{d}f}{\mathrm{d}x} \right)^* g \mathrm{d}x = \langle \hat{p} f | g \rangle \tag{3.18}$$

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 indeterminancy 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 younlike, 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,

$$\sigma^2 =< (\hat{Q} - < Q >)^2 >= \left\langle \Psi \middle| (\hat{Q} - q)^2 \Psi \right\rangle = \left\langle (\hat{Q} - q) \Psi \middle| (\hat{Q} - q) \Psi \right\rangle = 0 \tag{3.19}$$

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

$$\hat{Q}\Psi = q\Psi \tag{3.20}$$

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

$$\boxed{\text{Determiante states are eigenfunctions of } \hat{Q}} \tag{3.21}$$

Measurements of Q a¡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 multuply 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 **specturm** . 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 obervables). 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 **contonuous** (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 cominations* of them–involving neccessarily a spread in eigenvalues–may be normalizable). Some operator have a discrete spectrum only (for example, the Hamiltonian for the harmonic oscillator). some have only a continuous spectrum (for example, the free partivle 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 product 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 eqigenvalues are real.*

This is comforting: If you measure an obervable on a particle ina dterminate state, you will at leat least get a real number.

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

In a *finite*-diemsnional vector space the eigenvectos of a hermitian matrix have a third fundamental property: The span the space (every vector can be expressed as a linar combination of them). Unfortunately, the proof does not generalize to infinite-dimensional spaces. But the property itslef 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 obervables):

Axiom 1 *The eigenvalues of an obervable operator are compelte: Any function (in Hilber 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 exists. Nevertheless, there is a sense in which the three essential properties (reality, orthogonality, and complereness) 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:

$$\frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}x}f_p(x) = pf_p(x) \tag{3.22}$$

The general solution is

$$f_p(x) = Ae^{ipc/\hbar}$$

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 obervable $Q(x,p)$ on a particle in the state $\Psi(x,t)$, you are certain to get *one of the eigenvalues* of the hermitian operator $\hat{q}(x, -i\hbar\mathrm{d}/\mathrm{d}x)$. If the spectrum of $\hat{Q}$ is discrete, the probability of getting the particular eigenvalue q_n associated with the orthonormalized eifengunction $f_n(x)$ is

$$|c_n|^2, \qquad \text{where} \qquad c_n = \langle f_n|\Psi\rangle \tag{3.23}$$

If the sprectum is continuous, with real eigenvalues $q(z)$ and associated Dirac-orthonormalized eigenfunctions $f_z(x)$, the probability of gettin a result in the range $\mathrm{d}z$ is

$$|c(z)|^2\mathrm{d}z \text{where} \qquad c(z) = \langle f_z|\Psi\rangle \tag{3.24}$$

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

The statistical interpretation is raadically different from anything we encounter in classical physics. A somewhat different persperctive 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:

$$\Psi(x,t) = \sum_n c_n f_x(x) \tag{3.25}$$

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

$$c_n = \langle f_n | \Psi \rangle = \int f_n(x)^* \Psi(x,t) \mathrm{d}x \tag{3.26}$$

Qualitatively, c_n tells you "how much f_n is contained in Ψ", and given that a mesuremnet 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:

$$\sum_n |c_n|^2 = 1 \tag{3.27}$$

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

$$<Q> = \sum_n q_n |c_n 1^2 \tag{3.28}$$

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'stheorem, is its *inverse* Fourier transform:

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

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

According to the generalized statistical interpretationm, the probability that a measurement of momentum would yiled a result in the range $\mathrm{d}p$ is

$$|\Phi(p,t)|^2 \mathrm{d}p \tag{3.31}$$

3.4   The Uncertaintly Principle

3.5   Dirac Notation

Imagine an ordinary vactor $\mathbf{A}$ in two dimensions. How would you describe this vector to someone? The most convenient way is to set up catesian 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 axed, x' and y', and she would report differen 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 respecto to two different bases ($\{\hat{i}, \hat{j}\}$ and $\{\hat{i}', \hat{j}'\}$). The vector itself lives "out there is space", independent of anybody's (arbitraty) choice of coordinates.

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

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:

$$\Psi(x,t) = \langle x|s(t)\rangle \tag{3.32}$$

(with $|x\rangle$ standing fot the eigenfucntions of $\hat{x}$ with eigenvalue x), whereas the momentum sapce wavefunction $\Phi(x, t)$ is the expansion of $|s\rangle$ in the absis of momentum eigenfucntions:

$$\Phi(p,t) = \langle p|s(t)\rangle \tag{3.33}$$

(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):

$$c_n(t) = \langle n|s(t)\rangle \tag{3.34}$$

(with $|n\rangle$ standing for thr nth 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:

$$\Psi(x,t) = \int \Psi(y,t)\delta(x-y)\mathrm{d}y = \int \Phi(p,t)\frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}\mathrm{d}p$$
$$= \sum c_n e^{-iE_n t/\hbar}\psi_n(x) \tag{3.35}$$

Operators (representing obervables) are linear transformations–they "transform" one vector into another:

$$|\beta\rangle = \hat{Q}|\alpha\rangle \tag{3.36}$$

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

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

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

$$\langle e_m|\hat{Q}|e_n\rangle \equiv Q_{mn} \tag{3.38}$$

In this notations (3.36) takes the form

$$\sum_n b_n|e_n\rangle = \sum_n a_n\hat{Q}|e_n\rangle \tag{3.39}$$

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

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

---

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

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

28

and hence

$$b_m = \sum_n Q_{mn} a_n \tag{3.41}$$

Thus the matrix elements tell you how the components transform.

Later on we will encounter sysrem that admit only a finite number (N) of linearly independet states. In that case $|s(t)\rangle$ lives un an N-dimensional vector space; it can be represented as a column of (N) components (with respecto to a given bases), and operators take the form of ordinary $(N \times N)$ matrices. These are the simplest quantum systems–none of the subtleties associocted 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:

$$\langle f | = \int f^*[\cdots] \mathrm{d}x$$

with the ellipsis $[\cdot]$ 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,

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

the corresponding bra is a row vector:

$$\langle \alpha | = \begin{pmatrix} a_1^* a_2^* \cdots a_n^* \end{pmatrix} \tag{3.43}$$

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

$$\hat{P} \equiv |a\rangle\langle a| \tag{3.44}$$

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,

$$\langle e_m | e_n \rangle = \delta_{mn} \tag{3.45}$$

then

$$\sum_n |e_n\rangle\langle e_n| = 1 \tag{3.46}$$

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

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

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

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

then

$$\int |e_z\rangle\langle e_z|\,\mathrm{d}z = 1 \qquad (3.49)$$

(3.46) and (3.49) are the tidiest ways ti 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

$$i\hbar\frac{\partial \Psi}{\partial t} = H\Psi \tag{4.1}$$

the Hamiltonian operator[1] 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):

$$p_x \to \frac{\hbar}{i}\frac{\partial}{\partial x}, \quad p_y \to \frac{\hbar}{i}\frac{\partial}{\partial y}, \quad p_z \to \frac{\hbar}{i}\frac{\partial}{\partial z} \tag{4.2}$$

or

$$\boxed{\mathbf{p} \to \frac{\hbar}{i}\nabla} \tag{4.3}$$

for short. Thus

$$\boxed{i\hbar\frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi} \tag{4.4}$$

where

$$\nabla^2 \equiv \frac{\partial^2}{\psi x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{4.5}$$

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

---

[1]Where confusion might otherwise occur I have been putting "hats" on operators, to distinguish the from the corresponding classical observables. I don't think therw 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 finsing the particle in the infinitesimal volume $\mathrm{d}^3\mathbf{r} = \mathrm{d}x\mathrm{d}y\mathrm{d}z$ is $|\Psi(\mathbf{r}, t)|^2\mathrm{d}^3\mathbf{r}$, and the normalization condition reads

$$\int |\Psi|^2 \mathrm{d}^3\mathbf{r} = 1 \tag{4.6}$$

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

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

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

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

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

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

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

Separation of Variables

Typically, the potential s a function only of the distance from the origin. In that case it is natural to adopt **spherical coordinates**, (r, θ, ϕ). In spherical cooridinates the Laplacian takes the form[2]

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

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

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

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

$$\psi(r, \theta, \phi) = R(r)Y(\theta, \phi) \tag{4.12}$$

Putting this into (4.11), we have

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

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

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

---

[2]In principle, this can be obtained by change of variables from the cartesian expression (4.5). Howevere, 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 course[3], I will write this "separation constant" in the form $l(l+1)$

The Angular Equation

(??) determines the dependence of ψ on θ and ϕ; multiplying by $Y \sin^2 \theta$, it becomes

$$\sin \theta \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{\partial^2 Y}{\partial \phi^2} = -l(l+1) \sin^2 \theta Y \tag{4.13}$$

You might recognize this equation–it occurs in the solution to Laplace's equation in classical electrodynamics. As always, we try separations of variables:

$$Y(\theta, \phi) = \Theta(\theta) \Phi(\phi) \tag{4.14}$$

Plugging this in, and dividing vy $\Theta\Phi$, we find:

$$\left\{ \frac{1}{\Theta} \left[\sin \theta \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) \right] + l(l+1) \sin^2 \theta \right\} + \frac{1}{\Theta} \frac{d^2\Theta}{d\theta^2} = 0$$

The first term is a function only of θ, and de second is a function only of ϕ, so each must be a constant. This time I'll call the separation constant m^2:

$$\frac{1}{\Theta} \left[\sin \theta \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) \right] + l(l+1) \sin^2 \theta = m^2 \tag{4.15}$$

$$\frac{1}{\Theta} \frac{d^2\Theta}{d\theta^2} = -m^2 \tag{4.16}$$

The ϕ equation is easy

$$\frac{d^2\Phi}{d\phi^2} = -m^2\Phi \Rightarrow \Phi(\phi)e^{im\phi} \tag{4.17}$$

[Actually, there are two solutions: $\exp(im\phi)$ and $\exp(-im\phi)$, but we'll cover the latter by allowing m to run negative.]

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

$$V(r) = -\frac{e^2}{4\pi\epsilon_0} \frac{1}{r} \tag{4.18}$$

---

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

and the radial equation (**??**) says

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u}{\mathrm{d}r^2} + \left[-\frac{e^2}{4\pi\epsilon_0}\frac{1}{r} + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right]u = Eu \tag{4.19}$$

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 analytica solution to the harmonic oscillator.

Insidentally, the Coulomb potential (4.19) 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

$$\kappa \equiv \frac{\sqrt{-2mE}}{\hbar} \tag{4.20}$$

(For bound states, E is negative, so κ is *real*.) Dividing (4.19) by E, we have

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

This suggests that we introduce

$$\rho \equiv \kappa r, \quad \text{and} \quad \rho_0 \equiv \frac{me^2}{2\pi\epsilon_0\hbar^2\kappa} \tag{4.21}$$

so that

$$\frac{\mathrm{d}^2u}{\mathrm{d}\rho^2} = \left[1 - \frac{\rho_0}{\rho} + \frac{l(l+1)}{\rho^2}\right]u \tag{4.22}$$

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

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

The general solution is

$$u(\rho) = Ae^{-\rho} + Be^{\rho} \tag{4.23}$$

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

$$u(\rho) \sim Ae^{-\rho} \tag{4.24}$$

for large ρ. On the other hand, as $\rho \to 0$ the centrifugal term dominates[4]; approximately, then:

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

---

[4]This argument does not apply when $l = 0$ (although the conclusion, (4.25, 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 (4.26))

The general solution (check it!) is

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

but ρ^{-1} blows op (as $\rho \to 0$), so $D = 0$. Thus

$$u(\rho) \sim \rho^{l+1} \tag{4.25}$$

for small ρ.

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

$$u(\rho) = \rho^{l+1} e^{-\rho} v(\rho) \tag{4.26}$$

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

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

and

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

In terms of $v(\rho)$, then, the radial equation (4.22) reads

$$\rho \frac{\partial^2 v}{\partial \rho^2} + 2(l + 1 - \rho)\frac{\mathrm{d}v}{\mathrm{d}\rho} + [\rho_0 - 2(l + 1)]v = 0 \tag{4.27}$$

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

$$v(\rho) = \sum_{j=1}^{\infty} c_j \rho^j \tag{4.28}$$

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

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

[In the second summation I have renamed the "dummy index": $j \to j+1$. 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

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

Inserting these into (4.27), we have

$$\sum_{j=1}^{\infty} j(j + 1)c_{j+1} \rho^j + 2(l + 1) \sum_{j=0}^{\infty} (j + 1)c_{j+1} \rho^j$$

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

Equating the coefficients of like powers yields

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

or

$$c_{j+1} = \left\{ \frac{2(j+l+1) - \rho_0}{(j+1)(j+2l+2)} \right\} c_j \tag{4.29}$$

This recursion formular determines the coefficients, and hence de function $v(\rho)$: We start with c_0 (this becomes an overall constant, to be fixde eventually by normalization), and (4.29) gives is c_1; putting this back in, we obtain c_2, and so on[5].

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

$$c_j = \frac{2^j}{j!} c_0 \tag{4.30}$$

so

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

and hence

$$u(\rho) = c_0 \rho^{l+1} e^{\rho} \tag{4.31}$$

which blows uo at large ρ. The positive exponential is precisely the asymptotic behavior we didn't want, in (4.23). (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 thay aren't normalizable.) There is only one way out of this dilema: The *series must terminate*. There must occur some maximal integer, j_{max}, such that

$$c(j_{max} + 1) = 0 \tag{4.32}$$

(and beyond which all coefficients vanish automatically). Evidently (4.29)

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

Defining

$$n \equiv j_{max} + l + 1 \tag{4.33}$$

---

[5]You 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 sequnce would begin with a long string of zeris (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.

4.3   Angular Momentum

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

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \tag{4.34}$$

which os to say,

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x \tag{4.35}$$

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

Eigenvalues

The operators L_x and L_y do not commute; in fact[6]

$$\begin{aligned} [L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\ &= [yp_z, zp_x] - [yp_z, zp_z] - [zp_y, zp_x] + [zp_y, xp_z] \end{aligned} \tag{4.36}$$

From the canonical commutation relations[7] we know that the inly operators here that *fail* to commute are x with p_x, y with p_y, and z with p_z. So the two middles terms drop out, leaving

$$[L_x, L_y] = yp_x[p_z, z] + xp_y[z, p_z] = i\hbar(xp_y - yp_x) = i\hbar L_z \tag{4.37}$$

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 ($x \to y, y \to z, z \to x$)

$$\boxed{[L_x, L_y] = i\hbar L_z; \quad [L_y, L_z] = i\hbar L_x; \quad [L_z, L_x] = i\hbar L_y} \tag{4.38}$$

Fundamental commutation relations for angular momentum

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* obervables. 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

$$\sigma_{L_x}\sigma_{L_y} \geq \frac{\hbar}{2}|\langle L_z\rangle| \tag{4.39}$$

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

$$L^2 \equiv L_x^2 + L_y^2 + L_z^2 \tag{4.40}$$

---

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

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

does commute with L_x:

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

(I used (??)) to simplify the commutator; note also that any operator commutes with itself. It follows, of course, that L^4 also commute with L_y and L_z:

$$[L^2, L_x] = 0, \quad [L^2, L_y] = 0, \quad L^2, L_z = 0 \tag{4.41}$$

or, more compactly,

$$[L^2, \mathbf{L}] = 0 \tag{4.42}$$

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

$$L^2 f = \lambda f \quad \text{and} \quad L_z f = \mu f \tag{4.43}$$

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

$$\boxed{L_\pm \equiv L_x \pm iL_y} \tag{4.44}$$

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

$$[L_z, L\pm] = \pm\hbar L_\pm \tag{4.45}$$

And. of course,

$$[L^2, L_\pm] = 0 \tag{4.46}$$

I claim that if f es an eigenfunction of L^2 and L_z, so also is $L_\pm f$: (4.46) says

$$L^2(L_\pm f) = L_\pm(L^2 f) = L_\pm(\lambda f) = \lambda(L_\pm f) \tag{4.47}$$

so $L_\pm f$ is an eigenfunction of L^2, with the same eigenvalue λ, and (4.45) says

$$
\begin{aligned}
L_z(L_\pm f) &= (L_z L_\pm - L_\pm L_z)f + L_\pm L_z f = \pm\hbar L_\pm f + L_\pm(\mu f) \\
&= (\mu \pm \hbar)(L_\pm f)
\end{aligned}
\tag{4.48}
$$

so $L_\pm f$ is an eigenfunction of L_z with the *new* eigenvalue $\mu \pm \hbar$. We call L_+ the "raising" operator, because is *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 procces cannot go on forever.

Eventually we're going to reach a state gor which the z-component exceeds the *total*, and that cannot be. There must exist a "top rung", such that

$$L_+ f_t = 0 \qquad (4.49)$$

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

$$L_z f_t = \hbar l f_t; \quad L^2 f_t = \lambda f_t \qquad (4.50)$$

Now,

$$L_\pm L_\mp = (L_x \pm i L_y)(L_x \mp i L_y) = L_x^2 + L_y^2 \mp i(L_x L_y - L_y L_x)$$
$$= L^2 - L_z^2 \mp i(i\hbar L_z)$$

or, putting in the other way around,

$$L^2 = L_\pm L_\mp + L_z^2 \mp \hbar L_z \qquad (4.51)$$

It follows that

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

and hence

$$\lambda = \hbar^2 l(l+1) \qquad (4.52)$$

This tells is 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

$$L_- f_b = 0 \qquad (4.53)$$

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

$$L_z f_b = \hbar \bar{l} f_b; \quad L^2 f_b = \lambda f_b \qquad (4.54)$$

Using (4.51), we have

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

and therefore

$$\lambda = \hbar^2 \bar{l}(\bar{l} - 1) \qquad (4.55)$$

Comparing (4.52) 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

$$\bar{l} = -l \qquad (4.56)$$

Evidently the eigenvalues of L_z are $m\hbar$ (the appropaites of this letter will aso 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:

$$\boxed{L^2 f_l^m = \hbar^2 l(l+1) f_l^m; \quad L_z f_l^m = \hbar m f_l^m} \qquad (4.57)$$

where

$$l = 0, 1/2, 1, 3/2, ...; \quad m = -l, -l+1, ..., l-1, l \tag{4.58}$$

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 (4.38), we have determined the eigenvalues of L^2 and L_z–without ever seeing the eigenfunctions themselves! We turn now to the problem of constucting 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 cooridinates. Now, $\mathbf{L} = (\hbar/i)(\mathbf{r} \times \nabla)$, and the gradient, in spherical cooridinates, is:

$$\nabla = \hat{r}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial \theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial \phi} \tag{4.59}$$

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

$$\mathbf{L} = \frac{\hbar}{i}\left(\hat{\phi}\frac{\partial}{\partial \theta} - \hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial \phi} \right) \tag{4.60}$$

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

$$\hat{\theta} = (\cos\theta\cos\phi)\hat{i} + (\cos\theta\sin\phi)\hat{j} + (\sin\theta)\hat{k} \tag{4.61}$$

$$\hat{\phi} = -(\sin\theta)\hat{i} + (\cos\phi)\hat{j} \tag{4.62}$$

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

$$L_x = \frac{\hbar}{i}\left(-\sin\phi\frac{\partial}{\partial \theta} - \cos\phi\cot\theta\frac{\partial}{\partial \phi} \right) \tag{4.63}$$

$$L_y = \frac{\hbar}{i}\left(+\cos\phi\frac{\partial}{\partial \theta} - \sin\phi\cot\theta\frac{\partial}{\partial \phi} \right) \tag{4.64}$$

and

$$\boxed{L_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi}}$$

(4.65)

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

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

(4.66)

In particular

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

(4.67)

and hence

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

(4.68)

We are now in position to determine $f_l^m(\theta,\phi)$. It's an eigenfucntion 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" (4.13). And it's also an eigenfucntion 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 (4.16). 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:

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

(4.69)

Incidentally, we can use (4.68) 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* theoryof angular momentum permits l (and hence also m) to take on *half*-integer values (4.58), 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 profuound 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** ($\mathbf{S} = I\omega$), associated with motion about the center of mass. For example, the earth has orbital angular momentum attribuitable to its revolution around the sun, and spin angular momentum coming from its daily rotation about the north-south axis. In the classical econtext this distinction is largely a matter of convenience, for when you come right down to is, $\mathbf{S}$ is nothing but the sum total of the "orbital" angulat 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, assocaited (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 known) 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 realtions.

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y \tag{4.70}$$

It follows (as before) that the eigenvectors of S^2 and S_z satisfy[8]

$$S^2 \left|sm\right\rangle = \hbar^2 s(s+q) \left|sm\right\rangle; \quad S_z \left|sm\right\rangle = \hbar m \left|sm\right\rangle \tag{4.71}$$

and

$$S_{\pm} \left|sm\right\rangle = \hbar\sqrt{s(s+1) - m(m\pm 1)} \left|s(m\pm 1)\right\rangle \tag{4.72}$$

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

$$s = 0, \frac{1}{2}, 1, \frac{3}{2}...; \quad m = -s, -s+1, ..., s-1, s \tag{4.73}$$

It so happens that every elementary particle has a *specific and immutable* value of s, which we call **the spin** of that particular apecies: 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 changa from one to another when the system is pertubed. But s is *fixed*, for any given particle, and this makes the theory of spin comparatively simple.

---

[8]Because the eigensates of spin are not *functions*, I will the "ket" notation for them. (I could have done the same in Section **??**, writing $\left|lm\right\rangle$ 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).

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: $\left|\frac{1}{2}\frac{1}{2}\right\rangle$, which we call **spin up** (informally, ↑), and $\left|\frac{1}{2}(-\frac{1}{2})\right\rangle$, which we call **spin down** (↓). 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**):

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} = a\chi_+ + b\chi_- \tag{4.74}$$

with

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{4.75}$$

representing spin up, and

$$\chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.76}$$

for spin down.

Meanshile, the spin operators become 2×2 matrices, which we can work out by noting their effect on χ_+ and χ_-. (4.71) says

$$\mathbf{S}^2 \chi_+ = \frac{3}{4}\hbar^2 \chi_+ \quad \text{and} \quad \mathbf{S}^2 \chi_- = \frac{3}{4}\hbar^2 \chi_- \tag{4.77}$$

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*:

$$\mathbf{S}^2 = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.78}$$

Similarly,

$$\mathbf{S}\chi_+ = \frac{\hbar}{2}\chi_+, \quad \mathbf{S}_z\chi_- = -\frac{\hbar}{2}\chi_- \tag{4.79}$$

from which it follows that

$$\mathbf{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.80}$$

Meanwhile, (4.72) says

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

so

$$\mathbf{S}_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{S}_- = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{4.81}$$

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

$$\mathbf{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{4.82}$$

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

$$\boxed{\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \tag{4.83}$$

Pauli spin matrices

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

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

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{(eigenvalue } +\tfrac{\hbar}{2}); \quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{(eigenvalue } -\tfrac{\hbar}{2}) \tag{4.84}$$

If you measure S_z on a particle in the general state χ (4.74), you could get $+\hbar/2$, with probability $|a|^2$, or $-\hbar/2$, with probability $|b|^2$. Since these are the *only* possibilities,

$$|a|^2 + |b|^2 = 1 \tag{4.85}$$

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

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 surprisongly, 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

$$\chi_+^{(x)}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \text{(eigenvalue } +\tfrac{\hbar}{2}); \quad \chi_-^{(x)} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix}, \text{(eigenvalue } -\tfrac{\hbar}{2}) \tag{4.86}$$

---

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

As the eigenvectors of hermitian matrix, they span the space; the generic spinor χ (4.74) can be expressed as a linear combination of them:

$$\chi = \left(\frac{a+b}{\sqrt{2}}\right) \chi_+^{(x)} + \left(\frac{a-b}{\sqrt{2}}\right) \chi_-^{(x)} \tag{4.87}$$

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 shouls check for yourself that these probabilities add up to 1).

I'd loke now to walk you through an imaginary mesurement scenario involving spin $1/2$, becuse it serves to ollustrate 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 satte χ_+. If someone asks, "What is the z-component of that particle's spin angular momentum?", we could answer unambiguosly: $+\hbar/2$. 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 ypu telling me that you *don't know* the true state of that particle?" On the countrary; 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 ell-defines, the uncertainty principle would be violated.

At this point our challegner grabs the test-tube and *measures* the x-component of its spin; let's say he gets the value $+\hbar/2$. "Aha!" (he shouts it 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 measuerement. "You have obviously been reduced to splittin hairs. And anyway, what happened to your ucnertainty 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 noy in the state $\chi_-^{(x)}$, and whereas you know the value of S_x, you no longer know the calue of S_z. "But I was extremely careful not to disturb the particle when I measured S_x". Veru well, if you don't belive me, *check it out*: Measure S_z, and see what you get. (Of course, he *may* get $+\hbar/2$, which will be embarrasing to my case–but if we repeat this whole scenario over and over, half the time he weill 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 od 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 conecptual paradoxes of quantum mechanics.

Electron in a Magnetic Field

A spinning charged particle constitutes a magnetic dipole. Its **magnetic dipolemoment**, $\boldsymbol{\mu}$ is proportional to its spin angular momentum, $\mathbf{S}$

$$\boldsymbol{\mu} = \gamma \mathbf{S} \tag{4.88}$$

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, $\boldsymbol{\mu} \times \mathbf{B}$, which tends to line it up parallel to the fields (just like a compass needle). The energy assocaited with this torque is

$$H = -\boldsymbol{\mu} \cdot \mathbf{B} \tag{4.89}$$

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

$$H = -\gamma \mathbf{B} \cdot \mathbf{S} \tag{4.90}$$

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 state[10] of hydrogen. Each can have spin up or spin down, so there are four possibilities in all[11]

$$\uparrow\uparrow, \quad \uparrow\downarrow, \quad \downarrow\uparrow, \quad \downarrow\downarrow \tag{4.91}$$

where the first arrow refers to the electron and the second to the proton. *Question:* What is the *total* angular momentum of the atom? Let

$$\mathbf{S} \equiv \mathbf{S}^1 + \mathbf{S}^{(2)} \tag{4.92}$$

Each of the four composite states is an eigensate of S_z–the z-components simly *add*

$$S_z \chi_1 \chi_2 = (S_z^{(1)} + S_z^{(2)}) \chi_z \chi_2 = (S_z^{(1)} \chi_1) \chi_2 + \chi_1 (S_z^{(2)} \chi_2)$$
$$= (\hbar m \chi_1) \chi_2 + \chi_1 (\hbar m_2 \chi_2) = \hbar (m_1 + m_2) \chi_1 \chi_2$$

[note that $\mathbf{S}^{(1)}$ acts only on χ_1, and $\mathbf{S}^{(2)}$ acst only on χ_2]. So m (the quantum number for the composite system) os just $m_1 + m_2$:

$$\uparrow\uparrow: m = 1$$
$$\uparrow\downarrow: m = 0$$
$$\downarrow\uparrow: m = 0$$
$$\downarrow\downarrow: m = -1$$

At first glance, this doesn't look right: m is supposed to advance in integer steps, from $-s$ to $+s$, so it appears that $s = 1$ but there is an extra state with $m = 0$. One way to untangle this problem is to apply the lowering operator $S_- = S_-^{(1)} + S_-^{(2)}$ to the state $\uparrow\uparrow$, using (4.78):

$$S_-(\uparrow\uparrow) = (S_-^{(1)} \uparrow) \uparrow + \uparrow (S_-^{(2)} \uparrow)$$
$$= (\hbar \downarrow) \uparrow + \uparrow (\hbar \downarrow) = \hbar(\downarrow\uparrow + \uparrow\downarrow)$$

---

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

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

Evidently the three states with $s = 1$ are (in the notation $|sm\rangle$):

$$\left\{ \begin{array}{rcl} |11\rangle & = & \uparrow\uparrow \\ |10\rangle & = & \frac{1}{\sqrt{2}}(\uparrow\downarrow + \downarrow\uparrow) \\ |1-1\rangle & = & \downarrow\downarrow \end{array} \right\} \qquad s = 1 \quad \text{(triplet)} \tag{4.93}$$

This is called the **triplet** combination, for the obvious reason. Meanwuile, the orthogonal state with $m = 0$ carries $s = 0$:

$$\left\{ \quad |00\rangle \quad = \quad \frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow) \quad \right\} \qquad s = 0 \quad \text{(singlet)} \tag{4.94}$$

I claim, then, that te combination of two spin-$1/2$ particles can carry a total spin of 1 or 0, depending on wheter they occupy the triplet or the singlet configuration. To confirm this, I need to prove that the triplet states are eigenvectors of S^2 with eigenvalue $2\hbar^2$ and the singlet is an eigenvector of S^2 with eigenvalue 0. Now

$$S^2 = (\mathbf{S}^{(1)} + \mathbf{S}^{(2)}) \cdot (\mathbf{S}^{(1)} + \mathbf{S}^{(2)}) = (S^{(1)})^2 + (s^{(2)})^2 + 2\mathbf{S}^{(1)} \cdot \mathbf{S}^{(2)} \tag{4.95}$$

Using (4.77) and (??), we have

$$\mathbf{S}^{81} \cdot \mathbf{S}^{(2)}(\uparrow\downarrow) = (S_x^{(1)} \uparrow)(S_x^{(2)} \downarrow) + (S_y^{(1)} \uparrow)(S_y^{(2)} \downarrow) + (S_z^{(1)} \uparrow)(S_z^{(2)} \downarrow) \tag{4.96}$$

$$= \left(\frac{\hbar}{2} \downarrow \right)\left(\frac{\hbar}{2} \uparrow \right) + \left(\frac{i\hbar}{2} \downarrow \right)\left(\frac{-\hbar}{2} \uparrow \right) + \left(\frac{\hbar}{2} \uparrow \right)\left(\frac{-\hbar}{2} \downarrow \right) \tag{4.97}$$

$$= \frac{\hbar^4}{4}(2 \downarrow\uparrow - \uparrow\downarrow) \tag{4.98}$$

Similarly,

$$\mathbf{S}^{(1)} \cdot \mathbf{S}^{(2)}(\downarrow\uparrow) = \frac{\hbar^2}{4}(2 \uparrow\downarrow - \downarrow\uparrow).$$

It follows that

$$\mathbf{S}^{(1)} \cdot \mathbf{S}^{(2)} |10\rangle = \frac{\hbar^2}{4}\frac{1}{\sqrt{2}}(2 \downarrow\uparrow - \uparrow\downarrow + 2 \uparrow\downarrow - \downarrow\uparrow) = \frac{\hbar^2}{4}|10\rangle \tag{4.99}$$

and

$$\mathbf{S}^{(1)} \cdot \mathbf{S}^{(2)} |00\rangle = \frac{\hbar^2}{4}\frac{1}{\sqrt{2}}(2 \downarrow\uparrow - \uparrow\downarrow - 2 \uparrow\downarrow - \downarrow\uparrow) = \frac{-3\hbar^2}{4}|00\rangle \tag{4.100}$$

Returning to (4.95) (and again using (4.77)), we conclude that

$$S^2 |10\rangle \left(\frac{3\hbar^2}{4} + \frac{3\hbar^2}{4} + 2\frac{\hbar^2}{4} \right)|10\rangle = 2\hbar^2 |10\rangle \tag{4.101}$$

Chapter 5

Identical Particles

5.1 Two-Particle Systems

For a single particle, $\Psi(\mathbf{r}, t)$ is a function of the spatial coordinates, $\mathbf{r}$, and thr time t (we'll ignore spin, for the moment). The state of two-particle system is a function of the coordinates of particle one ($\mathbf{r}_1$), the coordinates of pa¡rticle two ($\mathbf{r}_2$) and the time:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t) \tag{5.1}$$

Its time evolution is determined (as always) by the Schrodinger equation:

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi \tag{5.2}$$

where H is the Hamiltonian for the whole system:

$$H = -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(\mathbf{r}_1, \mathbf{r}_2, t) \tag{5.3}$$

(the subscript on ∇ indicates differentiation with respect to the coordinates of particle 1 or particle 2, as the case may be). The statistical interpretation carries over in the obvious way:

$$|\Psi(\mathbf{r}_1, \mathbf{r}_2, t)|^2 \mathrm{d}^3\mathbf{r}_1 \mathrm{d}^3\mathbf{r}_2 \tag{5.4}$$

is the probability of finding particle q in the volume $\mathrm{d}^3\mathbf{r}_1$ and the particle 2 in the volume $\mathrm{d}^3\mathbf{r}_2$; evidently Ψ must be normalized in such a way that

$$\int |\Psi(\mathbf{r}_1, \mathbf{r}_2, t)|^2 \mathrm{d}^3\mathbf{r}_1\mathbf{r}_2 = 1 \tag{5.5}$$

For time-independet potentials, we obtain a complete set of solutions by separation of variables:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t) = \psi(\mathbf{r}_1, \mathbf{r}_2)e^{-iEt/\hbar} \tag{5.6}$$

where the spatial function (ψ) satisfies the time-independent Schrodinger equation:

$$-\frac{\hbar^2}{2m_1}\nabla_1^2\psi - \frac{\hbar^2}{2m_2}\nabla_2^2\psi + V\psi = E\psi \tag{5.7}$$

and E is the total energy of the system.

5.1.1 Bosons and Fermions

Suppose particle 1 is in the (one-particle) state $\psi_a(\mathbf{r})$, and particle 2 is in the state $\psi_b(\mathbf{r})$. (Remember: I'm ignoring spin, for the moment). In that case $\psi(\mathbf{r}_1, \mathbf{r}_2)$ is a simple *product*:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \tag{5.8}$$

Of course, this assumes that we can tell the particles apart–otherwise it wouldn't make any sense to claim that number 1 is in state ψ_a and number 2 is in state ψ_b; all we could say is that one of them is in state ψ_a and the other is in state ψ_b, but we wouldn't know which is which. If we were talking classical mechanics this would be a silly objetion: Yous can always tell the particles apart, in principle–just paint one of them red and the other one blue, or stamp identification numbers on them, or hire private detectives to follow them around. But in quantum mechanics the situation is fundamentally different: You can't paint an electron red, or pin a label on it, and a detentive's observations will inevitably and unpredictably alter its state, raising doubts as to whether the two had perhaps switched places. The fact is, all electrons are *utterly identical* , in a way that no two classical objets can ever be. It's not just that we don't happen to know which electron is which; God doesn't know which is which, because there is no such thing as 'this' electron, or 'that' electron; all we can legitimately speak about is 'an' electron.

Quantum mechanics neatly accommodates the existence of particles that are indistinguishable in principle: We sumpy construct a wave function that is non-committal as to which particle is in which state. There are actually two ways to do it:

$$\psi_\pm(\mathbf{r}_1, \mathbf{r}_2) = A[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \pm \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)] \tag{5.9}$$

Thus the thoery admits two kinds of identical particles: **bosons**, for which we use the plus sign, and **fermions**, for which we use the minus sign. Photons and mesons are bosons; protons and electrons are fermions. It so happens that

> all particles with *integer* spin are bosons, and
>
> all particles with *half integer* spin are fermions $\qquad$ (5.10)

This conection between **spin and statistics** (as we shall see, bosons and fermions have quite different statistical properties) can be proved in relativistic quantum mechanics; in the nonrelativistic theory it is taken as an axiom.

It follows, in particular, that two identical fermions (for example, two electrons) cannot occupy the same state. For if $\psi_a = \psi_b$, then

$$\psi_-(\mathbf{r}_1, \mathbf{r}_2) = A[\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2) - \psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2)] = 0$$

and we are left with no wave function at all. This is the famous **Pauli exclusion priciple**. It is not (as you may have been led to belive) a weird ad hoc assumption applying only to electrons, but rather a consequance of the rules for constructing two-particle wave functions, applying to *all* identical fermions.

I assumed, for the sake of argument, that one particle was in the state ψ_a and the other in state ψ_b, but there is a more general (and more sophisticated) way to formulate the problem. Let us define the **exchange operator**, P, which interchanges the two particles:

$$Pf(\mathbf{r}_1, \mathbf{r}_2) = f(\mathbf{r}_2, \mathbf{r}_1) \tag{5.11}$$

Clearly, $P^2 = 1$, and it follows (prove it for yourself) that the eigenvalues of P are ± 1. Now, if the two particles are identical, the Hamiltonian must treat them the same: $m_1 = m_2$ and $V(\mathbf{r}_1, \mathbf{r}_2) = V(\mathbf{r}_2, \mathbf{r}_1)$. It follows that P and H are compatible observables,

$$[P, H] = 0 \tag{5.12}$$

and hence we can find a complete set of functions that are simultaneous eigenstates of both. That is to say, we can find solutions to the Schrodinger equation that are either symmetric (eigenvalue $+1$) or antisymmetric (eigenvalue -1) under exchange:

$$\boxed{\psi(\mathbf{r}_1, \mathbf{r}_2) = \pm\psi(\mathbf{r}_2, \mathbf{r}_1)} \tag{5.13}$$

Moreover, if a system starts out in such a state, it will remain in such a state. The *new* law (I'll call it the **symmetrization requeriment**) is that for identical particles the eave function is not merely *allowed*, but *requiered* to satisfy (5.13), with the plus sign for bosons, and the minus sign for fermions. This is the *general* statement, of which (5.9) is a special case.

Chapter 6

Particle in a Central Potential. The Hydrogen Atom

[1] In this chapter, we shall consider the quantum mechanical properties of a particle placed in a central potential [that is, a potential $V(r)$ which depends only on the distance r from the origin]. This problem is closely related to the study of angular momentum. As we shall know, the fact that $V(r)$ es invariant under any rotation about the origin means that the Hamiltonian H of the particle conmutes with the three components of the orbital angular momentum operator $\mathbf{L}$. This cosiderably simplifies the determination of the eigenfunctions and eigenvalues of $\mathbf{L}^2$ and L_z as well.

6.1 Statioanary states of a particle in a central potential

In this section, we consider a (spinless) particle of mass μ, subjected to a central force derived from the potential $V(r)$ (the center of force is chosen as the origin).

6.1.1 Outline of the problem

Review of some classical results

The force acring on the classical particle situated at the point M (with $\mathbf{OM} = \mathbf{r}$) is equal to:

$$\mathbf{F} = -\nabla V(r) = -\frac{\mathrm{d}V}{\mathrm{d}r}\frac{\mathbf{r}}{r} \tag{6.1}$$

$\mathbf{F}$ is always directed towards O, and its mimentum with respect to this point is therefore always zero. If:

$$\mathcal{L} = \mathbf{r} \times \mathbf{p} \tag{6.2}$$

is the angular momentum of the particle with respecto to O, the anfular momentum theorem implies that:

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} = \mathbf{0} \tag{6.3}$$

$\mathcal{L}$ is therefore a *constant of the motion*, so that the particle's trajectory is necessarily situated in the plane passing through O and perpendicular to $\mathcal{L}$.

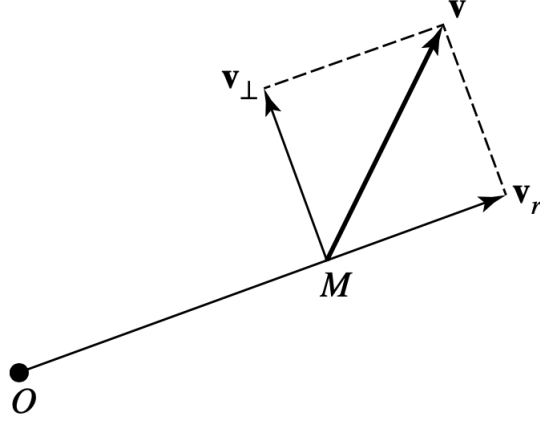[1]Quantum Mechanics - Cohen-Tannoudji, Diau, Laloe.

Figure 6.1: Radial component $\mathbf{v}_r$ and tangential component $\mathbf{v}_\perp$ of a particle's velocity.

Now let us consider Fig. 6.1 the position (denoted by $\mathbf{OM} = \mathbf{r}$) and velocity $\mathbf{v}$ of the particle at the instant t. Thw two vectors $\mathbf{r}$ and $\mathbf{v}$ lie in the plane of the trajectory and the velocity $\mathbf{v}$ can be decomposed into the radial component $\mathbf{v}_r$ (along the axis defined by $\mathbf{r}$) and the tangential component $\mathbf{v}_\perp$ (along the axis perpendicular to $\mathbf{r}$). The radual velocity, the algebraic value of $\mathbf{v}_r$, is the time derivative of the distance of the particle from the point O:

$$v_r = \frac{\mathrm{d}r}{\mathrm{d}t} \tag{6.4}$$

The tangential velociry can be expressed in term of r and the angular momentum $\mathcal{L}$, since:

$$|\mathbf{r} \times \mathbf{v}| = r|\mathbf{v}_\perp| \tag{6.5}$$

so that the modulus of the angular momentum $\mathcal{L}$ is equal to:

$$|\mathcal{L}| = |\mathbf{r} \times \mu\mathbf{v}| = \mu r|\mathbf{v}_\perp| \tag{6.6}$$

The total energy of the particle:

$$E = \frac{1}{2}\mu\mathbf{v}^2 + V(r) = \frac{1}{2}\mu\mathbf{v}_r^2 + \frac{1}{2}\mu\mathbf{v}_\perp^2 + V(r) \tag{6.7}$$

can be written:

$$E = \frac{1}{2}\mu\mathbf{v}_r^2 + \frac{\mathcal{L}^2}{2\mu r^2} + V(r) \tag{6.8}$$

The classical Hamiltonian if the system is then:

$$\mathcal{H} = \frac{p_r^2}{2\mu} + \frac{\mathcal{L}^2}{2\mu r^2} + V(r) \tag{6.9}$$

where:

$$p_r = \mu\frac{\mathrm{d}r}{\mathrm{d}t} \tag{6.10}$$

is the conjugate momentum of r, and $\mathcal{L}^2$ must be expressed in terms of the variables r, θ, φ and their conjugate momenta p_r, p_θ, p_φ. One finds:

$$\mathcal{L}^2 = p_\theta^2 + \frac{1}{\sin^2 \theta} p_\varphi^2 \tag{6.11}$$

In expression (6.9), the kinetic energy is broken into two terms: the radial kinetic energy and the kinetic energy of rotation about O . The reason is that, since $V(r)$ is independent of θ and φ in this case, the angular variables and their conjugate momenta appear only in the $\mathcal{L}^2$ term. In fact, if we are interested in the evolution of r, we can use the fact that $\mathcal{L}$ is a constant of the motion, and replace $\mathcal{L}^2$ by a constant in expression (6.9). The Hamiltonian $\mathcal{H}$ then appears as a function only of the radial variables r and p_r ($\mathcal{L}^2$ plays the role of a parameter), and the result is a differential equation involving only one variable, r:

$$\frac{\mathrm{d}p_r}{\mathrm{d}t} = \mu \frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = -\frac{\partial \mathcal{H}}{\partial r} \tag{6.12}$$

that is:

$$\mu \frac{\partial^2 r}{\partial t^2} = \frac{\mathcal{L}^2}{\mu r^2} - \frac{\mathrm{d}V}{\mathrm{d}r} \tag{6.13}$$

It is just as if we had a one-dimensional problem (with r varyng only between 0 and $+\infty$), with a particle of mass μ subjected to the "efective potential":

$$V_{eff}(r) = V(r) + \frac{\mathcal{L}^2}{2\mu r^2} \tag{6.14}$$

We shall see that the situation is analogous in quantum mechanics.

The quantum mechanical Hamiltonian

In quantum mechanics, we want to solve the eigenvalue equation of the Hamiltonian $\mathcal{H}$, the observable associated with the total energy. This equation is written, in the $\{|\mathbf{r}\rangle\}$ representation:

$$\left[-\frac{\hbar^2}{2\mu}\Delta + V(r) \right] \varphi(\mathbf{r}) = E\varphi(\mathbf{r}) \tag{6.15}$$

Since the potential V depends only on the distance r of the particle from the origin, spherical coordinates are best adapted to the problem. We therefore express the Laplacian Δ in spherical coordinates [2]

$$\Delta = \frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{r^2}\left(\frac{\partial^2}{\partial \theta^2} + \frac{1}{\tan \theta}\frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta}\frac{\partial^2}{\partial \varphi^2} \right) \tag{6.16}$$

and look for eigenfucntions $\varphi(\mathbf{r})$ that are functions of the variables r, θ, φ.

If we compare expression (6.16) with the one for the operator $\mathcal{L}^2$, wee see that the quantum mechanical Hamiltonian $\mathcal{H}$ can be put in a form completely analogous to (6.9):

$$\boxed{ H = -\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\partial^2}{\partial r^2} + \frac{1}{2\mu r^2}\mathbf{L}^2 + V(r) } \tag{6.17}$$

---

[2]Expression (6.16) gives the Laplacian only for non-zero r. This is because of the privileged position of the origin in spherical coordinates; it can be seen, moreover, that expression (6.16) is not defined for $r = 0$.

The angular dependence of the Hamiltonian is contained entirely in the $\mathbf{L}^2$ term, which is an operator here. We could, in fact, perfect the analogy by defining an operator P_r, which would allow us to write the first term of (6.17) like the one in (6.9).

We shall now show how one can solve the eigenvalue equation:

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\partial^2}{\partial r^2} + \frac{1}{2\mu r^2}\mathbf{L}^2 + V(r) \right] \varphi(r,\theta,\varphi) = E\varphi(r,\theta,\varphi) \qquad (6.18)$$

6.1.2 Separation of variables

Angular dependence of the eigenfunctions

We know that the three components of the angular momentum operator $\mathbf{L}$ act only on the angular variables θ and φ ; consequently, they commute with all operators acting only on the r-dependence. In addition, they commute with $\mathbf{L}^2$. Therefore, according to expression (6.17) for the Hamiltonian, *the three components of L are constants of the motion*[3] in the quantum mechanical sense:

$$[H, \mathbf{L}] = \mathbf{0} \qquad (6.19)$$

Obviously, H also commutes with $\mathbf{L}^2$.

Although we have at our disposition four constants of the motion (L_x, L_y, L_z and $\mathbf{L}^2$), we cannot use all four of them to solve equation (6.18) because they do not commute with each other; we shall use only $\mathbf{L}^2$ and L_z. Since the three observables H, $\mathbf{L}^2$ and L_z commute, we can find a basis of the state space $\mathcal{E}_r$ of the particle composed of eigenfunctions common to these three observables. We can, therefore, require the functions $\varphi(r,\theta,\varphi)$, solutions of equation (6.18), to be eigenfunctions of $\mathbf{L}^2$ and L_z as well. We must then solve the system of differential equations:

$$H\varphi(\mathbf{r}) = E\varphi(\mathbf{r}) \qquad (6.20)$$

$$\mathbf{L}^2\varphi(\mathbf{r}) = l(l+1)\hbar^2\varphi(\mathbf{r}) \qquad (6.21)$$

$$L_z\varphi(\mathbf{r}) = m\hbar\varphi(\mathbf{r}) \qquad (6.22)$$

But we already know the general form of the common eigenfunctions of $\mathbf{L}^2$ and L_z: the solutions $\varphi(\mathbf{r})$ of equations (6.20), (6.21), (6.22) corresponding to fixed values l of and m , are necessarily products of a function of r alone and the spherical harmonic $Y_l^m(\theta,\varphi)$:

$$\varphi(\mathbf{r}) = R(r)Y_l^m(\theta,\varphi) \qquad (6.23)$$

Whatever the radial function $R(r)$, $\varphi(\mathbf{r})$ is a solution of equations (6.21) and (6.22). The only problem which remains to be solved is therefore how to determine $R(r)$ such that $\varphi(\mathbf{r})$ is also an eigenfunction of H [equation (6.20)].

The radial equation

We shall now substitute expressions (6.17) and (6.23) into (6.20). Since $\varphi(\mathbf{r})$ is an eigenfunction of $\mathbf{L}^2$ with the eigenvalue $l(l+1)\hbar^2$, we see that $Y_l^m(\theta,\varphi)$ is a common factor on both sides.

---

[3](6.19) express the fact that H is a scalar operator with respect to rotations about the point O. This is true because the potential energy is invariant under rotations about O.

After simplifyng, we obtain the radial equation:

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{d^2}{dr^2}r + \frac{l(l+1)\hbar^2}{2\mu r^2} + V(r) \right] R(r) = ER(r) \tag{6.24}$$

Actually, a solution of (6.24), substituted into (6.23), does not necessarily yield a solution of the eigenvalue equation (6.15) of the Hamiltonian. As we have already pointed out, expression (6.16) for the Laplacian is not necessarily valid at $r = 0$. We must therefore make sure that the behavior of the solutions $R(r)$ of (6.24) at the origin is sufficiently regular for (6.23) to be in fact a solution of (6.15).

Instead of solving the partial differential equation (6.18) involving the three variables $r, \theta\varphi$ we must now solve a differential equation involving only the variable r, but dependent on a parameter l: we are looking for eigenvalues and eigenfunctions of an operator H_l which is different for each value of l.

In other words, we consider separately, in the state space $\mathcal{E}_r$, the subspaces $\mathcal{E}(l, m)$ corresponding to fixed values of l and m, studying the eigenvalue equation of H in each of these subspaces (which is possible because H commutes with $\mathbf{L}^2$ and L_z). The equation to be solved depends on l, but not on m; it is therefore the same in the $(2l+1)$ subspaces $\mathcal{E}(l, m)$ associated with a given value of l. We shall denote by $_{k,l}$ the eigenvalues of H_l, that is, the eigenvalues of the Hamiltonian H inside a given subspace $\mathcal{E}(l, m)$. The index k, which can be discrete or continuous, represents the various eigenvalues associated with the same value of l. As for the eigenfunctions of H_l, we shall label them with the same two indices as the eigenvalues: $R_{k,l}(r)$. It is not obvious that this is sufficient: several radial functions might exist and be eigenfunctions of the same operator H_l with the same eigenvalue $E_{k,l}$; we shall see that this is not the case and that, consequently, the two indices k and l are sufficient to characterize the different radial functions. We shall therefore rewrite equation (6.24) in the form:

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{d^2}{dr^2} + \frac{l(l+1)\hbar^2}{2\mu r^2} + V(r) \right] R_{k,l}(r) = E_{k,l}R_{k,l}(r) \tag{6.25}$$

We can simplify the differential operator to be studied by a change in functions.

We set:

$$R_{k,l}(r) = \frac{1}{r}u_{k,l}(r) \tag{6.26}$$

Multiplying both sides of (6.25) by r, we obtain for $u_{k,l}(r)$ the following differential equation:

$$\boxed{\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + \frac{l(l+1)\hbar^2}{2\mu r^2} + V(r) \right] u_{k,l}(r) = E_{k,l}(r)u_{k,l}(r)} \tag{6.27}$$

This equation is analogous to the one we would have to solve if, in a *one-dimensional* problem, a particle of mass μ were moving in an *effective potential* $V_{eff}(r)$:

$$V_{eff}(r) = V(r) + \frac{l(l+1)\hbar^2}{2\mu r^2} \tag{6.28}$$

Nevertheless, we must not lose sight of the fact that the variable r can take on only *non-negative* real values. The term $l(l+1)\hbar^2/2\mu r^2$ which is added to the potentual $V(r)$ is always positive or zero; the corresponding force (equal to minus the gradient of this term) always tend to repel the particle from the force center O; this is whay this term is called the *centrifugal potential* (or centrifugal barrier).

Behavior of the solutions of the radial equation at the origin

We have already pointed out that it is necessary to examine the behavior of the solutions $R(r)$ of the radial equation (6.24) at the origin in order to know if they are really solutions of (6.15). We shall assume that when r approaches zero, the potential $V(r)$ remains finite, or at least approaches infinity less rapidly than $1/r$ (this hypothesis is true in most cases encountered in physics and, in particular, in the case of the Coulomb potential. We shall consider a solution of (6.25) and assume that it behaves at the origin like r^s:

$$R_{k,l}(r) \sim Cr^s \quad \text{,as } r \to 0 \tag{6.29}$$

Substituting (6.29) into (6.25), and setting the coefficient of the dominant term equal to zero, we obtain the equation:

$$-s(s+1) + l(l+1) = 0 \tag{6.30}$$

and, consequently:

$$\begin{cases} \text{either} & s = l \\ \text{or} & s = -(l+1) \end{cases} \tag{6.31}$$

For a given value of $E_{k,l}$, there are therefore two linearly independent solutions of the second-order equation (6.25), behaving at the origin like r^l and $1/R^{l+1}$, respectively. But those which behave like $1/r^{l+1}$ must be rejected, since it can be shown[4] that $\frac{1}{r^{l+1}} Y_l^m(\theta, \varphi)$ is not a solution of the eigenvalue equation (6.15) for $r = 0$. From this, we see that acceptable solution of (6.27) go to zero at the origin for all l, since:

$$u_{k,l}(r) \sim Cr^{l+1} \quad \text{,as } r \to 0 \tag{6.32}$$

Consequantly, to (6.27) must be added the condition:

$$\boxed{u_{k,l}(0) = 0} \tag{6.33}$$

6.1.3   Statioanary states of a particle in a central potential

Quantum numbers

We can summarize the results as follows: the fact that the potential $V(r)$ is independent of θ and φ makes it possible:

1. to require the eigenfunctions of H to be *simultaneous eigenfunctions* of $\mathbf{L}^2$ and L_z, *which determines their angular dependence*:

$$\varphi_{k,l,m}(\mathbf{r}) = R_{k,l}(r) Y_l^m(\theta, \varphi) = \frac{1}{r} u_{k,l}(r) Y_l^m(\theta, \varphi) \tag{6.34}$$

2. *to replace the eigenvalue equation of H*, an equation involving partial derivatives with respect to r, θ, φ by a *differential equation involving only the variable r* and depending on a parameter l [(6.27)], with condition (6.33) imposed.

---

[4]This is because the Laplacian of $\frac{1}{r^{l+1}} Y_l^m(\theta, \varphi)$ involves the lth derivatives of $\delta(\mathbf{r})$.

In principle, the functions $\varphi_{k,l,m}(r,\theta,\varphi)$ must be squre-integrable, that is, normalizable:

$$\int |\varphi_{k,l,m}(r,\theta,\varphi)|^2 r^2 \mathrm{d}r \mathrm{d}\Omega = 1 \qquad (6.35)$$

Their form (6.34) allows us to separate radial and angular integrations:

$$\int |\varphi_{k,l,m}(r,\theta,\varphi)|^2 r^2 \mathrm{d}r \mathrm{d}\Omega = \int_0^\infty r^2 \mathrm{d}r |R_{k,l}(r)|^2 \int \mathrm{d}\Omega |Y_l^m(\theta,\varphi)|^2 \qquad (6.36)$$

But the spherical harmonics $Y_l^m(\theta,\varphi)$ are normalized with respect to θ and φ; condition (6.35) therefore reduces to:

$$\int_0^\infty r^2 \mathrm{d}r |R_{k,l}(r)|^2 = \int_0^\infty \mathrm{d}r |u_{k,l}(r)|^2 = 1 \qquad (6.37)$$

Actually, we know that it is often convenient to accept eigenfunctions of the Hamil- tonian that are not square-integrable. If the spectrum of H has a continuous part, we shall require only that the corresponding eigenfunctions be orthonormalized in the ex- tended sense, that is, that they satisfy a condition of the form:

$$\int_0^\infty r^2 \mathrm{d}r R_{k',l}^*(r) R_{k,l}(r) = \int_0^\infty \mathrm{d}r u_{k',l}^*(r) u_{k,l}(r) = \delta(k'-k) \qquad (6.38)$$

where k is a continuous index. In (6.37) and (6.38), the integrals converge at their lower limit, $r = 0$ [condition (6.33)]. This is physically satisfying since the probability of finding the particle in any volume of finite dimensions is then always finite. It is therefore only because of the behavior of the wave functions for $r \to \infty$ that, in the case of a continuous spectrum, the normalization integrals (6.38) diverge if $k = k'$.

Finally, the eigenfunctions of the Hamiltonian H of a particle placed in a central potential $V(r)$ depend on at least three indices [formula (6.34)]: $\varphi_{k,l,m}(r,\theta,\varphi) = R_{k,l}(r) Y_l^m(\theta,\varphi)$ is a somultaneous eigenfunction of $H, \mathbf{L}^2$ and L_z with the respective eigenvalues $E_{k,l}, l(l+1)\hbar^2$ and $m\hbar$. k is called the *radial* quantum number; l, the *azimuthal* quantum number; and m the *magnetic* quantum number. The radial part $R_{k,l}(r) = \frac{1}{r} u_{k,l}(r)$ of the eigenfunction and the eigenvalue $E_{k,l}$ of H are independent of the magnetic quantum number and are given by the radial equation (6.27). The angular part of the eigenfunction depends only on l and m and not on k; it does not depend on the form of the potentual $V(r)$.

Degeneracy of the energy levels

6.2 Potion of the center of mass and relative motion for a system fo two interacting particles

6.2.1 Motion od the center of masss and relative motion in classical mechanics

6.2.2 Separation of variables in quantum mechanics

Eigenvalues and eigenfunctions of the Hamiltonian

6.3 The hydrogen atom

6.3.1 Introduction

The hydrogen atom consists of a proton, of mass:

$$m_p \approx 1.7 \times 10^{-27} \text{ kg} \tag{6.39}$$

and charge:

$$q \approx 1.6 \times 10^{-19} \text{ Coulomb} \tag{6.40}$$

and of and electron, of mass:

$$m_e \approx 0.91 \times 10^{-30} \text{k}g \tag{6.41}$$

and a charge $-q$. The interaction between these two particles is essentially electrostatic. The corresponding potential energy is:

$$V(r) \approx -\frac{q^2}{4\pi\epsilon_0}\frac{1}{r} = -\frac{e^2}{r} \tag{6.42}$$

where r denotes the distance between the two particles, and:

$$\frac{q^2}{4\pi\epsilon_0} = e^2 \tag{6.43}$$

We confine ourselves to the study of this system in the center of mass frame. The classical Hamiltonian that describes the relative motion of the two particles is then:

$$\mathcal{H}(\mathbf{r}, \mathbf{p}) = \frac{\mathbf{p}^2}{2\mu} - \frac{e^2}{r} \tag{6.44}$$

Since $m_p \gg m_e$ [formulas (6.39) and (6.41)], the reduced mass μ of the system is very close to m_e:

$$\mu = \frac{m_e m_p}{m_e + m_p} \approx m_e \left(1 - \frac{m_e}{m_p}\right) \tag{6.45}$$

(the correction term m_e/m_p is on the order of $1/1800$). This means that the center of mass of the system is practically at the position of the proton, and that the relative particle can be identified, to a very good approximation, with the electron. This is why we shall adopt the slightly inaccurate convention of calling the relative particle the electron and the center of mass the proton.

6.3.2 The Bohr model

We shall briefly review the results of the Bohr model, which relate to the hydrogen atom. This model, which is based on the concept of a trajectory, is incompatible with the ideas of quantum mechanics. However, it allows us to introduce, in a very simple way, fundamental quantities such as the ionization energy E_I of the hydrogen atom and a parameter which characterizes atomic dimensions (the Bohr radius a_0). In addition, it so happens that the energies E_n given by the Bohr theory are the same as the eigenvalues of the Hamiltonian we shall calculate in 6.3.3. Finally, quantum mechanical theory is in agreement with some of the intuitive images of the Bohr model.

This semi-classical model is based on the hypotesis that the electron describes a *circular orbit* of radius r about the proton, obeying the following equations:

$$E = \frac{1}{2}\mu v^2 - \frac{e^2}{r} \tag{6.46}$$

$$\frac{\mu v^2}{r} = \frac{e^2}{r^2} \tag{6.47}$$

$$\mu v r = n\hbar \quad ; \quad n \text{ a positive number} \tag{6.48}$$

The first two equations are classical ones. (6.46) expresses the fact that the total energy E of the electron is the sum of its kinetic energy $\mu v^2/2$ and its potential energy $-e^2/r$. (6.47) is none other than the fundamental equation of Newtonian dynamics (e^2/r^2 is the Coulomb force exerted on the electron, and v^2/r is the acceleration of its uniform circular motion). The third equation expresses the quantization condition, introduced empirically by Bohr in order to explain the existence of discrete energy levels: he postulated that only circular orbits satisfying this condition are possible trajectories for the electron. The different orbits, as well as the corresponding values of the various physical quantities, are labeled by the integer n associated with them.

A very simple algebraic calculation then yields the expressions for E_n, r_n and v_n:

$$E_n = -\frac{1}{n^2}E_I \tag{6.49}$$

$$r_n = n^2 a_0 \tag{6.50}$$

$$v_n = \frac{1}{n}v_0 \tag{6.51}$$

with

$$E_I = \frac{\mu e^4}{2\hbar^2} \tag{6.52}$$

$$a_0 = \frac{\hbar^2}{\mu e^2} \tag{6.53}$$

$$v_0 = \frac{e^2}{\hbar} \tag{6.54}$$

When this model was proposed by Bohr, it marked an important step towards the understanding of atomic phenomena, since it yielded the correct values for the energy level of the hydrogen atom. These values indeed follow the $1/n^2$ (the Balmer formula) law indicated by expression

([6.49](#)). Moreover, the experimentally measured *ionization energy* (the energy which must be supplied to the hydrogen atom in its ground state in order to remove the electron) is equal to the numerical value of E_I:

$$E_I \approx 13.6 \text{ eV} \tag{6.55}$$

Finally, the *Bohr radius* a_0 indeed characterized atomix¡c dimensions:

$$a_0 \approx 0.52 \text{ Å} \tag{6.56}$$

6.3.3   Quantum mechanical theory of the hydrogen atom

We shall now take up the question of the determination of the eigenvalues and eigenfunctions of the Hamiltonian H describing the relative motion of the proton and the electron in the center of mass frame [formula ([6.44](#))]. In the $\{|\mathbf{r}\rangle\}$ representation, the eigenvalue equation of the Hamiltonian H is written:

$$\left[-\frac{\hbar^2}{2\mu}\Delta - \frac{e^2}{r} \right] \varphi(\mathbf{r}) = E\varphi(\mathbf{r}) \tag{6.57}$$

Since the potential $-e^2/r$ is central, we can apply the results of [6.1](#): the eigenfunctions $\varphi(\mathbf{r})$ are of the form:

$$\varphi_{k,l,m}(\mathbf{r}) = \frac{1}{r}u_{k,l}(r)Y_l^m(\theta, \varphi) \tag{6.58}$$

$u_{k,l}(r)$ is given by the radial equation, ([6.27](#)), that is:

$$\left[-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{e^2}{r} \right] u_{k,l}(r) = E_{k,l}(r)u_{k,l}(r) \tag{6.59}$$

We add to this equation condition ([6.33](#)):

$$u_{k,l}(0) = 0 \tag{6.60}$$

Change of variables

To simplify the reasoning, we shall choose a_0 and [formulas ([6.52](#)), ([6.53](#)) and ([6.54](#))] as the units of length and energy. That is, we shall introduce the dimensionless quantities:

$$\rho = r/a_0 \tag{6.61}$$

$$\lambda_{k,l} = \sqrt{-E_{k,l}/E_I} \tag{6.62}$$

(the quantity under the radical sign is positive, since we are looking for the bound states).
  With expressions ([6.52](#)) and ([6.53](#)) for E_I and a_0 taken into account, the radial equation ([6.59](#)) becomes simply:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} - \frac{l(l+1)}{\rho^2} + \frac{2}{\rho} \right] u_{k,l}(\rho) = 0 \tag{6.63}$$

Solving the radial equation

In order to solve (6.63), we shall expanding $u_{k,l}(\rho)$ in a power series.

- Asymptotic behavior

  Let us determine the asymptotic behavior of $u_{k,l}(\rho)$ qualitatively. When ρ approaches infinity, the terms in $1/\rho$ and $1/\rho^2$ become negligible compared to the constant term $\lambda_{k,l}^2$, so that (6.63) practically reduces to:

  $$\left[\frac{d^2}{d\rho^2} - \lambda_{k,l}^2\right] u_{k,l}(\rho) = 0 \tag{6.64}$$

  whose solutions are $e^{\pm\rho\lambda_{k,l}}$. This argument is not rigorous, since we have completely neglected the terms in $1/\rho$ and $1/\rho^2$; actually, it can be shown that $u_{k,l}(\rho)$ is equal to $E^{\pm\rho\lambda_{k,l}}$ multiplied by a power of ρ.

  We shall later be led by physical considerations to require the function $u_{k,l}(\rho)$ to be bounded at infinity, and hence to reject the solutions of (6.63) whose asymptotic behavior is governed by $e^{+\rho\lambda_{kl}}$. This is why we perform the change of function:

  $$u_{k,l}(\rho) = e^{-\rho\lambda_{k,l}} yk, l(\rho) \tag{6.65}$$

  Although this change of function singles out $e^{-\rho\lambda_{k,l}}$, it clearly does not eliminate solutions in $e^{+\rho\lambda_{k,l}}$, which must be identified and then rejected at the end of the calculation. The differential equation that $y_{k,l}(\rho)$ must satisfy can easily be derived from (6.63):

  $$\left\{\frac{d^2}{d\rho^2} - 2\lambda_{k,l}\frac{d}{d\rho} + \left[\frac{2}{\rho} - \frac{l(l+1)}{\rho^2}\right]\right\} y_{k,l}(\rho) = 0 \tag{6.66}$$

  Condition (6.60) must be associated with this equation, that is:

  $$y_{k,l}(\rho) = 0 \tag{6.67}$$

- Solutions in the form of power series

  Consider de expansion of $y_{k,l}(\rho)$ in a powers of ρ:

  $$y_{k,l}(\rho) = \rho^s \sum_{q=0}^{\infty} c_q \rho^q \tag{6.68}$$

  By definition, c_0 is the first non-zero coefficient of this expansion:

  $$c_0 \neq 0 \tag{6.69}$$

  Condition (6.67) implies that s is stricly positive.

  We calculate $\frac{d}{d\rho} y_{k,l}(\rho)$ and $\frac{d^2}{d\rho^2} y_{k,l}(\rho)$ from (6.68):

  $$\frac{d}{d\rho} = \sum_{q0}^{\infty}(q+s)c_q\rho^{q+s-1} \tag{6.70}$$

  $$\frac{d^2}{d\rho^2} y_{k,l}(\rho) = \sum_{q=0}^{\infty}(q+s)(q+s-1)c_1\rho^{q+s-2} \tag{6.71}$$

61

To obtain the left-hand side of (6.66), we multiply expressions (6.68), (6.70) and (6.71) respectively by the factors $\left[\frac{2}{\rho} - \frac{-(l+1)}{\rho^2}\right]$, $-2\lambda_{k,l}$ and 1. According to (6.66), the series to determined must ve identically zero, that is, all ist coefficients must be zero.

The lowest term is in ρ^{s-2}. Taking its coefficient as zero, we obtain:

$$[-(l+1) + s(s-1)]c_0 = 0 \tag{6.72}$$

If we take (6.69) into account, we see that s can take on one of two values:

$$\begin{cases} s & = l+1 \\ s & = -l \end{cases} \tag{6.73}$$

(in agreement with the general result of 6.1.2). We have seen that only (??) gives a behavior at the origin that can lead to an acceptable solution [condition (6.67)]. Setting the coefficient of the general term in ρ^{q+s-2} equal to zero, we obtain (with $s = l+1$) the following recurrence relation:

$$q(q+2l+1)c_q = 2[(q+l)\lambda_{k,l} - 1]c_{q-1} \tag{6.74}$$

If we fix c_0, this relation enables us to calculate c_1, then c_2m and thus by recurrence all the coefficients c_q. Since c_1/c_{q-1} approaches zero when $q \to 0$, the corresponding series is convergent for all ρ. Thus we have determined, for any value of $\lambda_{k,l}$, the solution of (6.66) that satisfies condition (6.67).

Energy quantization. Radial functions

We are now going to require the preceding solution to have a physically acceptable asymptotic behavior. This will involve quantization of the possible values of $\lambda_{k,l}$.

If the term in brackets on the right-hand side of (6.74) does not go to zero for any integer q, expansion (6.68) is a true infinite series, for which:

$$\frac{c_1}{c_{q-1}} \sim \frac{2\lambda_{k,l}}{q} \quad \text{,as } q \to \infty \tag{6.75}$$

Now, the power series expansion of the function $e^{2\lambda_{k,l}}$ is written:

$$\begin{cases} e^{2\rho\lambda_{k,l}} & = \sum_{q=0}^{\infty} d_q \rho^q \\ d_q & = \frac{(2\lambda_{k,l})^q}{q!} \end{cases} \tag{6.76}$$

which implies:

$$\frac{d_1}{d_{q-1}} = \frac{2\lambda_{k,l}}{q} \tag{6.77}$$

If we compare (6.75) and (6.77), we see that, for large values of ρ, the series being considered behaves like $e^{2\rho\lambda_{k,l}}$. The corresponding function $u_{k,l}$ [formula (6.65)] is then proportional to $e^{+\rho\lambda_{k,l}}$, which is not physically acceptable.

Consequently, we must reject all cases in which expansion (6.68) is an infinite series. The only possible values of $\lambda_{k,l}$ are those for which (6.68) has only a finite number of terms, that is, those for which $y_{k,l}$ reduces to a polynomial. The corresponding function $u_{k,l}$ is then physically

acceptable, since its asymptotic behavior is dominated by $e^{-\rho\lambda_{k,l}}$. Therefore, all we need is an integer k such that the term in brackets of the right-hand side of (6.74) goes to zero for $q = k$: the corresponding coefficient c_k is then zero, as are all those of higher order, since the fact that c_k is zero means that c_{k+1} is as well, and so on. For fixed l, we label the corresponding values of $\lambda_{k,l}$ by this integer k (note that k is greater than or equal to 1, since c_0 never goes to zero). We then have, according to (6.74):

$$\lambda_{k,l} = \frac{1}{k+l} \tag{6.78}$$

For a given l, the only negative energies possible are therefore [formula (6.62)]:

$$E_{k,l} = \frac{-E_I}{(k+l)^2}; \qquad k = 1, 2, 3, ... \tag{6.79}$$

We shall discuss this result in 6.3.4

$y_{k,l}$ is therefore a polynomial, whose term of lowest order is in ρ^{l+1} and whose term of highest order is in ρ^{k+l} . Its various coefficients can be calculated in terms of c_0 by solving recurrence relation (6.74), which can be written, using (6.78):

$$c_q = -\frac{2(k-q)}{q(q+2l+1)(k+l)}c_{q-1} \tag{6.80}$$

It is easy to show that:

$$c_q = (-1)^q \left(\frac{2}{k+l}\right)^q \frac{(k-1)!}{(k-q-1)!}\frac{(2l+1)!}{q!(q+2l+1)!}c_0 \tag{6.81}$$

$u_{k,l}$ is then given by formula (6.65), and c_0 is determined (to within a phase factor) by normalization condition (6.37) [we must first, of course, return to the variable r by using (6.61)]. Finally, we obtain the true function $R_{k,l}(r)$ by dividing $u_{k,l}(r)$ by r. The following three examples give an idea of the form of these radial functions:

$$R_{k=1,l=0}(r) = 2(a_0)^{-3/2}e^{-r/a_0} \tag{6.82}$$

$$R_{k=2,l=0}(r) = 2(a_0)^{-3/2}\left(1 - \frac{r}{2a_0}\right)e^{-r/2a_0} \tag{6.83}$$

$$R_{k=1,l=1}(r) = (2a_0)^{-3/2}\frac{1}{\sqrt{3}}\frac{r}{a_0}e^{-r/2a_0} \tag{6.84}$$

6.3.4   Discussion of the results

Order if magnitude of atomic parameters

Formulas (6.79) and (??) show that, for the hydrogen atom, the ionization energy , defined by (6.52), and the Bohr radius, given by (6.53), play an important role. These quantities give an order of magnitude of the energies and spatial extensions of the wave functions associated with the bound states of the hydrogen atom.

Relations (6.52) and (6.53) can be written in the form:

$$E_I = \frac{1}{2}\alpha^2\mu c^2 \tag{6.85}$$

$$a_0 = \frac{1}{\alpha}\lambda_c \tag{6.86}$$

where α is the *fine structure constant*, a dimensionless constant which plays a very important role in physics:

$$\alpha = \frac{e^2}{\hbar c} = \frac{q^2}{4\pi\epsilon_0 \hbar c} \approx \frac{1}{137} \tag{6.87}$$

and where $\lambdabar_c$ is defined by:

$$\lambdabar_c = \frac{\hbar}{\mu c} \tag{6.88}$$

Sinse μ is almost the same as m_e, the rest mass of the electron, $\lambdabar_c$ is practically equal to the *Compton wavelength* of the electron, which is given by:

$$\frac{\hbar}{m_e c} \approx 3.8 \times 10^{-3}\,\text{Å} \tag{6.89}$$

Relation (6.86) therefore indicates that a_0 is on the order of one hundred times the Compton wavelength of the electron. Relation (6.85) shows that the order of magnitude of the binding energy of the electron is between $10^{-4}\mu c^2$ and $10^{-5}\mu c^2$, where μc^2 is practically equal to the rest energy of the electron:

$$m_e c^2 \approx 0.51 \times 10^6\,\text{eV} \tag{6.90}$$

It follows that:

$$E_I \ll m_e c^2 \tag{6.91}$$

Energy levels

- *Possible values of the quantum numbers; degeneracies*

  For fixed l, there exists an infinite number of possible energy values [formula (6.79)], corresponding to $k = 1, 2, 3, \ldots$ Each of them is at least $(2+1)$-fold degenerate: this is an *essential degeneracy* related to the fact that the radial equation depends only on the quantum number l and not on m (2.1)). But, in addition, there exist *accidental degeneracies*: equation (6.79) indicates that two eigenvalues $E_{k,l}$ and $E_{k',l'}$ corresponding to different radial equations $(l' \neq l)$ are equal if $k + l = k' + l'$.

  In the special case of the hydrigen atom, $E_{k,l}$ does not depend on k and l separately, but only on their sum. We set:

  $$n = k + l \tag{6.92}$$

  The various energy states are labeled by the integer n (greater than or equal to 1), and (6.79) becomes:

  $$E_n = -\frac{1}{n^2} E_I \tag{6.93}$$

  According to (6.92), it is equivalent to specify k and l or n and l to determine the eigenfunctions. Following convention, from now on we shall use the quantum numbers n and l. The energy is fixed by n, which is called the *principal quantum number*; a given value of n characterizes what is called an *electron shell*.

  Since k is necessarily an integer which is greater than or equal to 1, there is only a finite number of values of l associated with the same valejue of n. According to (6.92), if n is fixed, one can have:

  $$l = 0, 1, 2, ..., n - 1 \tag{6.94}$$

The shell characterized by n is said to contain n *sub-shells*, each one corresponding to one of the values l of given in (6.94). Finally, each sub-shell contains $(2l + 1)$ distinct states, associated with the $(2l + 1)$ possible values of m for fixed l.

The total degeneracy of the energy level E_n is therefore:

$$g_n = \sum_{l=0}^{n-1}(2l + 1) = 2\frac{(n-1)n}{2} + n = n^2 \qquad (6.95)$$

We sall see that the existence of electron spin multiplies this number by 2 (if we also take into account the proton spin, which is equal to that of the electron, we obtain another factor of 2).

- *Spectroscopic notation*

  For historical reasons (dating from the period, before the development of quantum mechanics, in which the study of spectra resulted in an empirical classification of the numerous lines observed), letters of the alphabet are associated with the various values of l. The correspondence is as follows:

  $$l = 0 \leftrightarrow s$$
  $$l = 1 \leftrightarrow p$$
  $$l = 2 \leftrightarrow d$$
  $$l = 3 \leftrightarrow f$$
  $$l = 4 \leftrightarrow g$$
  $$\vdots$$

  Therefore, *spectroscopic notation* labels a sub-shell by the corresponding number n followed by the letter that characterizes the value of l. Thus, the ground level [which is non-degenerate, according to (6.95)], sometimes called the "K shell", includes only the $1s$ sub-shell; the first excited level, or "L shell", includes the $2s$ and $2p$ sub-shells; the second excited level ("M shell") includes the $3s$, $3p$ and $3d$ sub-shells, etc. (The capital letters sometimes associated with the successive shells follow an alphabetical order, starting with the letter K.)

Wave functions

The wave functions associated with the eigenstates common to $\mathbf{L}^2$, L_z and the Hamiltonian H of the hydrogen atom are generally labeled, not by the three quantum numbers k, l, m as we have done until now, but by n, l and m [passage from one set to the other simply involves use of relation (6.92)]. Since the operators $H, \mathbf{L}^2$ and L_z constitute a C.S.C.O. (cf. 6.3), specification of the three integers n, l and m which is equivalent to that of the eigenvalues of H, $\mathbf{L}^2$ and L_z, unambiguously determines the corresponding eigenfunction $\varphi_{n,l,m}(\mathbf{r})$.

- *Angular dependence*

  As is the case for any central potential, the functions $\varphi_{n,l,m}(\mathbf{r})$ are products of a radial function and a spherical harmonic $Y_l^m(\theta, \varphi)$. To visualize their angular dependence on the

axis characterized by the polar angles θ and φ, we can measure off a distance that is pro-porcional to $|\varphi_{n,l,m}(r, \theta, \varphi)|^2$ for any *fixed* r, that is, proportional to $|Y_l^m(\theta, \varphi)|^2$. Thys, we obtain a surface of revolution about Oz axis, since we know squently, $|Y_l^m(\theta, \varphi)|^2$ is independent of φ. We can therefore represent its cross-section by a plane containing Oz.

- *Radial dependence*

The radial functions $R_{n,l}(r)$, each of which characterizes a sub-sehll, can be calculated from the results of **??**, paying attention, however, to the change in notation introduced by formula (6.92).

Finally, we can derive formula (6.50) for the successive Bohr radio. To do so, consider the various states for which $l = n$. We calculate the variation with of the probability density for each of the preceding levels in an infinitesimal solid angle $d\Omega$ about a fixed direction of polar angles θ and φ. In general, the position probability for the electron in the volume element $d^3r = r^2 dr d\Omega$ situated at the point (r, θ, φ) is given by:

$$d^3\mathcal{P}_{n,l,m}(r, \theta, \varphi) = |\varphi_{n,l,.}(r, \theta, \varphi)|^2 r^2 dr d\Omega$$
$$= |R_{n,l}(r)|^2 r^2 dr \times |Y_l^m(\theta, \varphi)|^2 d\Omega \tag{6.96}$$

Here, we have fixed θ, φ and $d\Omega$. The probability of finding the electron between r and $r + dr$, inside the solid angle under consideration, is then proportional to $r^2 |R_{n,l}(r)|^2 dr$. The corresponding density is therefore, to within a constant factor, $r^2 |R_{n,l}(r)|^2$ (the factor r^2 arises from the expression for the volume element in spherical coordinates). We are interested in the case where $l = n - 1$, that is, $k = n - l = 1$; **??** then indicates that the polynomial which enters into $R_{n,l}(r)$ contains only one term in $(r/a_0)^{n-1}$ desired probability density is therefore proportional to:

$$f_n(r) = \frac{r^2}{a_0^2} \left[\left(\frac{r}{a_0} \right)^{n-1} e^{r/na_0} \right]^2$$
$$= \left(\frac{r}{a_0} \right)^{2n} e^{-2r/na_0} \tag{6.97}$$

This function has a maximum for:

$$r = r_n = n^2 a_0 \tag{6.98}$$

which is the radius of the Bohr orbit corresponding to the energy E_n.

66

Chapter 7

Electron Spin

7.1   Introduction of electron spin

7.1.1   Experimental evidence

Fine structure of spectral lines

"Anomalous" Zeeman effect

Existence of half-integral angular momenta

7.1.2   Quantum description: postulates of the Pauli theory

7.2   Special properties of an angular momentum $1/2$

We shall restrict ourselves from now on to the case of the electron, which is a spin $1/2$ particle. From the preceding chapters, we know how to handle its orbital variables. We are now going to study in more detail its spin degrees of freedom.

The spin state space is two-dimensional. We shall take as a basis the orthonormal system $\{|+\rangle, |-\rangle\}$ of eigenkets common to $\mathbf{S}^2$ and S_z which satisfy the equations:

$$\begin{cases} \mathbf{S}^2 |\pm\rangle & = \frac{3}{4}\hbar^2 |\pm\rangle \\ S_z |\pm\rangle & = \pm\frac{1}{2}\hbar |\pm\rangle \end{cases} \tag{7.1}$$

$$\begin{cases} \langle+|-\rangle & = 0 \\ \langle+|+\rangle & = \langle-|-\rangle = 1 \end{cases} \tag{7.2}$$

$$|+\rangle \langle+| + |-\rangle \langle-| = \mathbb{1} \tag{7.3}$$

where $\mathbb{1}$ is the unit operator. The most general spin state is described by an arbitrary vector of $\mathcal{E}_c$:

$$|\xi\rangle = c_+ |+\rangle + c_- |-\rangle \tag{7.4}$$

---

[1]Quantum Mechanics - Cohen-Tannoudji, Diau, Laloe.

where c_+ and c_- are complex numbers. According to (7.1), all the kets of $\mathcal{E}_s$ are eigenvectors of $\mathbf{S}^2$ with the same eigenvalue $3\hbar^2/4$, which causes $\mathbf{S}^2$ to be proportional to the identity operator of $\mathcal{E}_s$:

$$\mathbf{S}^2 = \frac{3}{4}\hbar^2 \tag{7.5}$$

(in the right hand side of this equation, as is usually done, we have not written the unit operator $\mathbb{1}$ explicitly). Since $\mathbf{S}$ is, by definition, an angular momentum, it possesses all the general properties derived before. The action of the operators:

$$S_\pm = S_x \pm iS_i \tag{7.6}$$

on the basis vectors $|+\rangle$ and $|-\rangle$ is given by the general formulas (??) when one sets $j = s = 1/2$:

$$S_- |+\rangle = 0 \qquad S_+ |-\rangle = \hbar |+\rangle \tag{7.7}$$
$$S_- |+\rangle = \hbar |-\rangle \qquad S_- |-\rangle = 0 \tag{7.8}$$

Any operator acting in $\mathcal{E}_s$ can be representes, in the $\{|+\rangle, |-\rangle\}$ basis, by a 2×2 matrix. In particular, using (??) and (??), we find the matrices corresponding to S_x, S_y and S_z in the form:

$$(\mathbf{S}) = \frac{\hbar}{2}\boldsymbol{\sigma} \tag{7.9}$$

where σ designates the set of three *Pauli matrices*:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.10}$$

The Pauli matrices posses the following properties, which can easily be verifed from their explicit form (7.10)

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathbb{1} \tag{7.11}$$
$$\sigma_x\sigma_y + \sigma_y\sigma_x = 0 \tag{7.12}$$
$$[\sigma_x, \sigma] = 2i\sigma_z \tag{7.13}$$
$$\sigma_x\sigma_y = i\sigma_z \tag{7.14}$$

(to the last three formulas must bee added those obtained through cyclic permutation of the x, y, z indices). It also follow from (7.10) that: