Contents

1 Townsend Chapter 9 – Translational and rotational symmetry in the two-body problem

In chapters 6, 7 and 8, you started learn about how to solve Schroedinger's equation in 1-D. The important new concept introduced there is that

- Uncertainty principle exists.
- World is described by quantum states.
- **Change** in a quantum state occurs when an **operator** operates on the quantum state.

One cannot put enough emphasis on the uncertainty principle. The usual

$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{1}$$

and its variations

$$\Delta L \Delta \phi \geq \frac{\hbar}{2} \tag{2}$$

and

$$\Delta E \Delta t \geq \frac{\hbar}{2} \tag{3}$$

are essential in **understanding** any quantum phenomena.

In this course, you will learn a lot of math techniques to solve particular problems. These are important and must be learned. However, the **mechanical** part of, say solving a differential equation, does not necessarily have to carried out by you. If you have a good enough computer program, it should know about all known solvable differential equations and it will just tell you what the solution looks like. Are we done? No! We are no way near 'done'.

First of all, if you don't have a physical understanding of what's going on, then the given solutions mean nothing to you. In many cases, there are unphysical solutions. If you don't know your physics, you will get a nonsensical answer like something blowing up somewhere. Also, even though there are a large class of systems for which the solutions of the Schrodinger

equation is exactly known, there are many, many, many real-world problems for which we have no hope of ever getting an exact solution even if you have worlds' finest computer. But does that stop a physicist? No! No way. But to make a progress in face of such situation, you have **have an understanding of physics**. Namely, what exactly are going on, what's the most likely form of wavefunction? Which terms are more important than others? What simplification can one make? Are there similar systems in other branches of physics where things have already been studied in depth? .....

The key to understand the underlying physics is, yes, you guessed it, the uncertainty principle and a good understanding of the classical system. So many important things can be estimated, guessed at, discarded in this way. And we'll see many examples of it in this course.

Second lesson is that the world is described by a quantum state. What is a quantum state? Well, to answer that, ask what we mean by **the state of a particle** in classical mechanics. In CM, the answer is easy. A classical state of a **particle** with no other structures is given by a point in the phase space $(\mathbf{p}, \mathbf{x})$. For a given **particle**, these 6 numbers as a function of time are all you need to specify the state of a particle.

In quantum mechanics, we can't do that because of the uncertainty principle. It is impossible to specify x and p_x at the same time. In 1-D, the only possibility is to specify x and lose all information on p_x or vice versa, or have some other state where both x and p_x are uncertain. So best we can do in terms of classical variables is either

$$|x\rangle \tag{4}$$

or

$$|p_x\rangle \tag{5}$$

Here the quantum state $|x\rangle$ means that we know where the particle is at this moment, but we have no idea what its momentum is. So in effect, the number of variables that specify the state of particle has shrunken by a factor of 2 compared to the classical mechanics.

Suppose at $t = 0$, we have $|x_0\rangle$. But this means that we have no idea what the momentum is. Hence, we have no idea where the particle would be at a later time. If the elapsed time is not so long, then the particle will be most likely be found near the starting point. But the point is that now we have deal with a *probability* of finding particles at some point $P(x|x_0)$,

rather than dealing with a trajectory $x(t), p(t)$ which just gives you the 6 numbers that describe the motion of a particle at the given time. Therefore uncertainty make it necessary that we think about the motion of a particle in terms of a function of both x and t rather than the trajectory $x(t), p(t)$.

So what do we mean by a general quantum state $|\psi\rangle$? If this is classical mechanics, we know what it is. It must be a pair of x and p or some equivalent of it. In quantum mechanics, this is not true. We can only specify one of them.

So instead, we postulate: given the quantum state $|\psi\rangle$, the probability to find the particle in the vicinity of x is given by

$$P(x|\psi) = |\langle x|\psi\rangle|^2 \tag{6}$$

so that

$$\int dx\, P(x|\psi) = \int dx\, \langle\psi|x\rangle\langle x|\psi\rangle = \langle\psi|\psi\rangle = \hat{1} \tag{7}$$

which requires that

$$\hat{1} = \int dx\, |x\rangle\langle x| \tag{8}$$

is an identity operator. Furthermore, it requires that

$$\langle x|\psi\rangle^* = \langle\psi|x\rangle \tag{9}$$

since $|\langle x|\psi\rangle|^2$ must be positive definite.

The average value is given by

$$\begin{aligned} \langle x\rangle_\psi &= \int dx\, P(x|\psi) \\ &= \int dx\, x\langle\psi|x\rangle\langle x|\psi\rangle \\ &= \langle\psi|\hat{x}|\psi\rangle \end{aligned} \tag{10}$$

where we defined

$$\hat{x} = \int dx\, x|x\rangle\langle x| \tag{11}$$

which has the property of

$$\begin{aligned} \hat{x}|y\rangle &= \int dx\, x|x\rangle\langle x|y\rangle \\ &= \int dx\, x|x\rangle\delta(x-y) \\ &= y|y\rangle \end{aligned} \tag{12}$$

That is, the states $|x\rangle$ are eigenstates of the **operator** $\hat{x}$ defined in this way.

In this way, we introduced the position space. But what about the momentum? Well, we can go thru the entirely same excersize with p instead of x and arrive at

$$P(p|\psi) = |\langle p|\psi\rangle|^2 \tag{13}$$

and

$$1 = \int dp\, |p\rangle\langle p| \tag{14}$$

and

$$\hat{p} = \int dp\, p|p\rangle\langle p| \tag{15}$$

and

$$\hat{p}|p'\rangle = p'|p'\rangle \tag{16}$$

Now seat back and take a breath. How is it that we have **two** identity operators one with x and the other with p? Are they the same operator? The answer is yes. This is because both $\hat{x}$ and $\hat{p}$ are **Hermitian** operators.

Importance of Hermitian operators

- Eigenvalues are real – Observables
- The eigenfunctions are orthogonal.
- The eigenfunctions of any Hermitian operator is complete, meaning that any function can be represented by a linear combination of the eigenstates.

But so what? How do we get to the position space representation

$$\hat{p} = \frac{\hbar}{i}\frac{d}{dx} \tag{17}$$

Well, we continue in this fashion and just change the basis. In particular, we'd like to consider

$$\hat{p}|\psi\rangle \tag{18}$$

where ψ is an arbitrary wavefunction. **This is an important technique.** Whenever you want to prove something general for an operator, thinking about its action on an arbitrary state usually gets you the result.

$$\begin{aligned}
\hat{p}|\psi\rangle &= \hat{1}\hat{p}\hat{1}|\psi\rangle \\
&= \int dxdy\, |x\rangle\langle x|\hat{p}|y\rangle\, \langle y|\psi\rangle \\
&= \int dxdy\, |x\rangle\langle x|\hat{p}\hat{1}|y\rangle\, \langle y|\psi\rangle \\
&= \int dxdydp\, |x\rangle\langle x|p\rangle p\langle p|y\rangle\, \langle y|\psi\rangle
\end{aligned} \tag{19}$$

So we have to know the overlap $\langle x|p\rangle$. To do so, consider

$$\begin{aligned}
\hat{1} &= \hat{1}\hat{1}\hat{1} \\
&= \int dxdy\, |x\rangle\langle x|\hat{1}|y\rangle\langle y| \\
&= \int dxdydp\, |x\rangle\langle x|p\rangle\langle p|y\rangle\langle y|
\end{aligned} \tag{20}$$

This must be equal to $\hat{1} = \int dx\, |x\rangle\langle x|$ so we must demand that

$$\int dp\, \langle x|p\rangle\langle p|y\rangle = \delta(x-y) \tag{21}$$

There can be many functions that can satisfy this relation. So what can we say more about it? This is where physics enters. We know that the **Heisenberg** principle demands that the uncertainties in the position when the momentum is specified is infinite. In other words,

$$|\langle x|p\rangle| = \text{constant} \tag{22}$$

That is, $\langle x|p\rangle$ must be a pure phase. The only way these two conditions can be simultaneously satisfied is to have

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\lambda}}\, \exp(ixp/\lambda) \tag{23}$$

with some unspecified real constant λ which has the demansion of xp or an action. And in view of the Heisenberg principle, this λ should better be proportional to $\hbar$. In fact, let's identify λ with it.

Incidently, note that the appearance of the imaginary number i here. In the current formulation of QM, imaginary number is therefore essential. You

can, of course, not ever talk about imaginary numbers since any complex number can be represented by a 2-d vector $z = (x, y)$ with somewhat weird multiplication rules. However, having an imaginary number is much more natural way of thinking about QM.

Armed with this information, we can now proceed as follows

$$\begin{aligned}
\langle x|\hat{p}|\psi\rangle &= \langle x| \int dx' dy dp \, |x'\rangle\langle x'|p\rangle p\langle p|y\rangle \, \langle y|\psi\rangle \\
&= \int dx' dy dp \, \delta(x - x')\langle x'|p\rangle p\langle p|y\rangle \, \langle y|\psi\rangle \\
&= \int dy dp \, \frac{1}{2\pi\hbar} p \, \exp\left(i(x - y)p/\hbar\right) \langle y|\psi\rangle \\
&= \frac{\hbar}{i}\frac{d}{dx} \int dy dp \, \frac{1}{2\pi\hbar} \exp\left(i(x - y)p/\hbar\right) \langle y|\psi\rangle \\
&= \frac{\hbar}{i}\frac{d}{dx} \int dy \delta(x - y) \, \langle y|\psi\rangle \\
&= \frac{\hbar}{i}\frac{d}{dx} \, \langle x|\psi\rangle
\end{aligned} \tag{24}$$

Now take a step back and breathe. What did we just do? We assumed that the **Heisenberg** principle, state vectors and operators and the reasonable norm for the state vectors and came up with the **momentum** operator! The key step was to demand that for a momentum fixed state, there cannot be any sort of position information. This led us to the **wave**

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \, e^{ixp/\hbar} \tag{25}$$

and the identification that the wavevector be related to the momentum as

$$k\hbar = p \tag{26}$$

Now you should remember that the commutation relation

$$[\hat{x}, \hat{p}] = i\hbar \tag{27}$$

leads directly to the uncertainty principle

$$\Delta x \Delta p \geq \hbar/2 \tag{28}$$

So in a way, what we have just shown is that the statements

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\, e^{ixp/\hbar} \tag{29}$$

$$[\hat{x}, \hat{p}] = i\hbar \tag{30}$$

$$\Delta x \Delta p \geq \hbar/2 \tag{31}$$

are all intimately related to each other.

Let me put it another way. Suppose that instead of assuming the uncertainty, I assume that the momentum operator $\hat{p}$ satisfies the commutation relation. Can we still get to

$$\hat{p}_x = -i\hbar\partial_x \quad ? \tag{32}$$

Let's see if we can do it. So we start with

$$\hat{p} \;=\; \int dx\, dy\, dp\, |x\rangle\langle x|p\rangle p\langle p|y\rangle\langle y| \tag{33}$$

and demand that

$$[\hat{x}, \hat{p}] = i\hbar \tag{34}$$

$$\begin{aligned} i\hbar &= [\hat{x}, \hat{p}] \\ &= \hat{x}\hat{p} - \hat{p}\hat{x} \\ &= \int dx\, dy\, dp\, x\, |x\rangle\langle x|p\rangle p\langle p|y\rangle\langle y| \\ &\quad - \int dx\, dy\, dp\, |x\rangle\langle x|p\rangle p\langle p|y\rangle\langle y|\, y \\ &= \int dx\, dy\, dp\, (x-y)\, |x\rangle\langle x|p\rangle p\langle p|y\rangle\langle y| \end{aligned} \tag{35}$$

If this is to be an identity, we must have

$$(x-y)\int dp\, \langle x|p\rangle p\langle p|y\rangle = i\hbar\delta(x-y) \tag{36}$$

which should be valid not necessarily as a functional identity but as a distributions **inside** an integral. What can we say more? Well, the right hand side is a function of $x-y$. So the product

$$\langle x|p\rangle\langle p|y\rangle \tag{37}$$

should be a function of $x-y$. The only way this can happen and the above equation satisfied is if

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\exp(ipx/\hbar) \tag{38}$$

and the same argument as before tells us that

$$\hat{p} = -i\hbar\frac{d}{dx} \tag{39}$$

With it,

$$\begin{aligned}
(x-y)\int dp\,\langle x|p\rangle p\langle p|y\rangle &= (x-y)\int dp\,p\,\frac{1}{2\pi\hbar}e^{ip(x-y)/\hbar} \\
&= \int dp\,p\frac{\hbar}{i}\frac{d}{dp}\,\frac{1}{2\pi\hbar}e^{ip(x-y)/\hbar} \\
&= -i\int\frac{dp}{2\pi}\,p\frac{d}{dp}\,e^{ip(x-y)/\hbar} \\
&= i\int\frac{dp}{2\pi}\,e^{ip(x-y)/\hbar} \\
&= i\hbar\delta(x-y)
\end{aligned} \tag{40}$$

So the wave nature of microscopic world, the commutation relation and the uncertainty principle are all very intimitly related. What dynamics provide is **how the energy depend on x and p.** For instance, if you a given a potential well, then that in a sense determines Δx and in turn determines the size of Δp and so on.

Having said that, let's see what we need to do to generalize what we leared in 1-D to 3-D. The phase space variables in 3-D are x, y, z and p_x, p_y, p_z.

The first thing we should think about is what things commute and what don't for that is at the heart of the generalizing classical to quantum. Let's think about the positions first. Question is: Is it possible to measure x and y at the same time? The answer is, yes, it should be. But does it? Well, yeah, as far as we know. So let's treat

$$[\hat{x}_i, \hat{x}_j] = 0 \tag{41}$$

$x_1 = x, x_2 = y, x_3 = z$, as given. Why is this important? Well, if they commute, then they can be indepedently measured. And if that's the case, we can specify a quantum state with them. That is, we can consider

$$|\mathbf{x}\rangle \tag{42}$$

and can define

$$\langle \mathbf{x_1}|\mathbf{x_2}\rangle = \delta^3(\mathbf{x_1}-\mathbf{x_2}) \equiv \delta(x_1-x_2)\delta(y_1-y_2)\delta(z_1-z_2) \tag{43}$$

In that case, all we have to do is repeat previous analysis 3 times with 3 name changes x, y, z and get the 3-D momentum operator

$$\hat{\mathbf{p}} = \frac{\hbar}{i}\nabla \tag{44}$$

Now I will introduce a short hand

$$\partial_x = \frac{\partial}{\partial x} \tag{45}$$

$$\partial_t = \frac{\partial}{\partial t} \tag{46}$$

and sometime write things like

$$\hat{p}_x = -i\hbar\partial_x \tag{47}$$

Note the sign. How do you determine this sign? Well, you have to fix a convention. Our convention is that

$$\langle \mathbf{x}|\mathbf{p}\rangle = \frac{1}{(2\pi\hbar)^{3/2}}\, e^{i\mathbf{x}\cdot\mathbf{p}/\hbar} \tag{48}$$

and the sign in the exponent fixes the sign in $\hat{p}_x$.

One can repeat the same argument for

$$[\hat{p}_i, \hat{p}_j] = 0 \tag{49}$$

Then how about

$$[\hat{x}_i, \hat{p}_j] = ? \tag{50}$$

To be consistent with the 1-D case, we should of course demand that

$$[\hat{x}_i, \hat{p}_i] = i\hbar \tag{51}$$

That is, $[\hat{x}, \hat{p}_x] = i\hbar$, $[\hat{y}, \hat{p}_y] = i\hbar$ and $[\hat{z}, \hat{p}_z] = i\hbar$. But how about things like $[\hat{x}, \hat{p}_y]$? Should they communte? Well can they measured simultaneously? There is no reason why this is not the case. Hence these must be zero. All of these can be conveniently expressed by using the Kronecker-δ symbol:

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \tag{52}$$

so that

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij} \tag{53}$$

2 Momentum as the generator of translations

Since we have the position eigenstate now, let's think about what exactly it means to have position and momentum. Well, classically, you know what that means. But how much of your intuitions carry thru?

First of all, consider the **translation operators** defined by

$$\hat{T}(a_x\mathbf{i})|x,y,z\rangle = |x+a_x,y,z\rangle \tag{54}$$
$$\hat{T}(a_y\mathbf{j})|x,y,z\rangle = |x,y+a_y,z\rangle \tag{55}$$
$$\hat{T}(a_z\mathbf{k})|x,y,z\rangle = |x,y,z+a_z\rangle \tag{56}$$

Now define

$$\hat{T}(a_x\mathbf{i}) = e^{-ia_x\hat{\pi}_x/\hbar} \tag{57}$$

etc. and

$$\hat{T}(\mathbf{a}) = \hat{T}(a_x\mathbf{i})\,\hat{T}(a_y\mathbf{j})\,\hat{T}(a_z\mathbf{k}) \tag{58}$$

These $\hat{\pi}_x$, etc. are called 'the generator' of translation for obvious reasons. What more can we say about $\hat{\pi}$?

Consider the Hermitian conjugation:

$$\begin{aligned}\Big(\hat{T}(a_x\mathbf{i})|x,y,z\rangle\Big)^\dagger &= \langle x,y,z|\hat{T}^\dagger(a_x\mathbf{i}) \\ &= \langle x+a_x,y,z| \end{aligned}\tag{59}$$

Now consider this.

$$\begin{aligned}\hat{T}^\dagger(a_x\mathbf{i})\,\hat{T}(a_x\mathbf{i}) &= \int d^3x \int d^3y\,|\mathbf{x}\rangle\langle\mathbf{x}|\hat{T}^\dagger(a_x\mathbf{i})\,\hat{T}(a_x\mathbf{i})|\mathbf{y}\rangle\langle\mathbf{y}| \\ &= \int d^3x \int d^3y\,|\mathbf{x}\rangle\,\langle\mathbf{x}+a_x\mathbf{i}|\,\mathbf{y}+a_x\mathbf{i}\rangle\,\langle\mathbf{y}| \\ &= \int d^3x \int d^3y\,|\mathbf{x}\rangle\delta(\mathbf{x}+a_x\mathbf{i}-(\mathbf{y}+a_x\mathbf{i})\langle\mathbf{y}| \\ &= \int d^3x \int d^3y\,|\mathbf{x}\rangle\delta(\mathbf{x}-\mathbf{y})\langle\mathbf{y}| \\ &= \int d^3x\,|\mathbf{x}\rangle\langle\mathbf{x}| = \hat{1}\end{aligned}\tag{60}$$

This can be immediately generalized to

$$\hat{T}^\dagger(\mathbf{a})\,\hat{T}(\mathbf{a}) = \hat{T}(-\mathbf{a})\hat{T}(\mathbf{a}) = 1 \tag{61}$$

This means that $\hat{T}$ is a unitary operator. That means that $\hat{\pi}$ is a Hermitian operator that satisfies

$$\hat{\pi}_x^\dagger = \hat{\pi}_x \tag{62}$$

so it is an observable.

Also note that the operation of $\hat{T}(\mathbf{a})$ on $|\mathbf{x}\rangle$

$$\hat{T}(\mathbf{a})|\mathbf{x}\rangle = |\mathbf{x}+\mathbf{a}\rangle \tag{63}$$

is different than its action on a state that is projected onto $|x\rangle$:

$$\langle\mathbf{x}|\hat{T}(\mathbf{a})|\psi\rangle = \langle\mathbf{x}-\mathbf{a}|\psi\rangle \tag{64}$$

Note also that

$$\langle\psi|\hat{T}(\mathbf{a})|\mathbf{x}\rangle = \langle\psi|\mathbf{x}+\mathbf{a}\rangle \tag{65}$$

How about commutation relationship among them? Well, for this, we need physics input. We demand that if you go 10 meters south and then 10 meters west, you end up at the same position as if you went 10 meters west and then 10 meters south. In other words, the translations must commute:

$$\hat{T}(a_x\mathbf{i})\hat{T}(a_y\mathbf{j}) = \hat{T}(a_y\mathbf{j})\hat{T}(a_x\mathbf{i}) \tag{66}$$

If a_x and a_y are small, we can Taylor-expand:

$$\begin{aligned}\hat{T}(a_x\mathbf{i})\hat{T}(a_y\mathbf{j}) &= \Big(1-(i/\hbar)(a_x\hat{\pi}_x)+O(a_x^2)\Big)\Big(1-(i/\hbar)(a_y\hat{\pi}_y)+O(a_y^2)\Big)\\ &= 1-(i/\hbar)(a_x\hat{\pi}_x+a_y\hat{\pi}_y)-(1/\hbar^2)a_xa_y\hat{\pi}_x\hat{\pi}_y+O(a_x^2,a_y^2)\end{aligned} \tag{67}$$

Comparing this with

$$\hat{T}(a_y\mathbf{j})\hat{T}(a_x\mathbf{i}) \tag{68}$$

we can easily get

$$\pi_x\pi_y = \pi_y\pi_x \tag{69}$$

or

$$[\pi_x,\pi_y] = 0 \tag{70}$$

Since they commute, we don't have to worry about the order in $\hat{T}(\mathbf{a})$ and simply write

$$\hat{T}(\mathbf{a}) = \exp\left(-i\hbar\mathbf{a}\cdot\boldsymbol{\pi}\right) \tag{71}$$

So far so good. Now what about commutation with x?

$$\begin{aligned}
\langle x|\left(\hat{x}\hat{T}(a) - \hat{T}(a)\hat{x}\right)|\psi\rangle &= \langle x|\hat{x}\hat{T}(a)|\psi\rangle - \langle x|\hat{T}(a)\hat{x}|\psi\rangle \\
&= x\langle x-a|\psi\rangle - (x-a)\langle x-a|\psi\rangle \\
&= a\langle x-a|\psi\rangle \\
&= a\langle x|\psi\rangle + O(a^2)
\end{aligned} \tag{72}$$

On the other hand

$$\begin{aligned}
\hat{x}\hat{T}(a) - \hat{T}(a)\hat{x} &= \hat{x}(1 - i\hat{\pi}a/\hbar) - (1 - i\hat{\pi}a/\hbar)\hat{x} \\
&= \frac{a}{i\hbar}(\hat{x}\hat{\pi} - \hat{\pi}\hat{x})
\end{aligned} \tag{73}$$

which implies

$$[\hat{x}, \hat{\pi}] = i\hbar \tag{74}$$

This is the same commutation relationship satisfied by $\hat{x}$ and $\hat{p}_x$.

Now let's see if we can figure out the x representation of $\hat{\pi}_x$. If we have an infinitesimal a,

$$\begin{aligned}
\langle x|\hat{T}(a_x)|\psi\rangle &= \langle x|(1 - ia\hat{\pi}/\hbar)\int dx'\,|x'\rangle\langle x|\psi\rangle \\
&= \langle x-a|\psi\rangle
\end{aligned} \tag{75}$$

This yields

$$\begin{aligned}
\langle x|\psi\rangle - (ia/\hbar)\int dx'\,\langle x|\hat{\pi}|x'\rangle\langle x'|\psi\rangle &= \langle x|\psi\rangle - (ia/\hbar)\langle x|\hat{\pi}|\psi\rangle \\
&= \langle x|\psi\rangle - a\partial_x\langle x|\psi\rangle
\end{aligned} \tag{76}$$

so

$$\langle x|\hat{\pi}|\psi\rangle = -i\hbar\partial_x\langle x|\psi\rangle \tag{77}$$

Fine, dandy. But what does this have to with **momentum**?

Well, there are many answers one can give to this question. One way is to invoke that on average, the Ehrenfest theorem

$$\frac{d\langle x\rangle}{dt} = \frac{\langle p\rangle}{m} \tag{78}$$

holds true. This you've seen in section 6. So that having a finite p means that the particle's position is changing.

Another way to think about it is to consider what equivalent 'generator' is in classical mechanics. In clasical mechanics, If a canonical pair p and q transforms according to

$$q \to \bar{q} = q + \epsilon\frac{\partial g}{\partial p} = q + \epsilon\left\{q, g\right\}_{\mathrm{PB}} = q - \epsilon\left\{g, q\right\}_{\mathrm{PB}} \tag{79}$$

$$p \to \bar{p} = p - \epsilon\frac{\partial g}{\partial q} = p + \epsilon\left\{p, g\right\}_{\mathrm{PB}} = p - \epsilon\left\{g, q\right\}_{\mathrm{PB}} \tag{80}$$

then g is called the **generator of transformation**. In the case of translation, this is given by

$$g = p \tag{81}$$

The hamiltonian changes by

$$\begin{aligned} \delta H &= \frac{\partial H}{\partial q}\delta q + \frac{\partial H}{\partial p}\delta p \\ &= \epsilon\frac{\partial H}{\partial q}\frac{\partial g}{\partial p} - \epsilon\frac{\partial H}{\partial p}\frac{\partial g}{\partial q} \\ &= \epsilon\left\{H, g\right\}_{\mathrm{PB}} \end{aligned} \tag{82}$$

Another way to see that the translation operator is the momentum operator, suppose that the time evolution of a quantum state is described by

$$i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle \tag{83}$$

where $\hat{H}$ is the generator of the time evolution. Now apply the translation operator before and after:

$$\begin{aligned} \langle x|\hat{T}(\mathbf{a})\hat{H}|\psi\rangle &= \int dx'\,\langle x|\hat{T}(\mathbf{a})|x'\rangle\langle x'|\hat{H}|\psi\rangle \\ &= \int dx'\,\delta(x - x' - a)\langle x'|\hat{H}|\psi\rangle \\ &= \langle x - a|\hat{H}|\psi\rangle \end{aligned} \tag{84}$$

and

$$\begin{aligned}\langle x|\hat{H}\hat{T}(\mathbf{a})|\psi\rangle &= \int dx' \langle x|\hat{H}|x'\rangle\langle x'|\hat{T}(\mathbf{a})|\psi\rangle \\ &= \int dx' \langle x|\hat{H}|x'\rangle\langle x'|\hat{T}(\mathbf{a})|\psi\rangle \\ &= \int dx' \langle x|\hat{H}|x'\rangle\langle x'-a|\psi\rangle \end{aligned} \tag{85}$$

Now if the space is translationally invariant, the energy of a system must not depend on the fact that whether the translation has taken place before or after the energy measurement. Therefore,

$$\hat{T}(\mathbf{a})\hat{H} - \hat{H}\hat{T}(\mathbf{a}) = 0 \tag{86}$$

In partcular, for infinitesimal $\mathbf{a}$, we should have (in coordiate representation)

$$[\nabla, \hat{H}] = 0 \tag{87}$$

This shows that the quantity representated by ∇ is a conserved quantity in the case when the space is translationally invariant. Classically, Noether Theorem tells us that that is what the momentum is. So, ∇ is **at least proportional to the momentum operator**. The proportionality constant in

$$\hat{p} = -i\hbar\nabla \tag{88}$$

can be fixed again by appealing to the correpondence principle.

3 QM vs. Classical

In optics, the propagation of wavefunction

$$\Phi = e^{A(\mathbf{x})+ik_0(L(\mathbf{x})-ct)} \tag{89}$$

can be representated by the eikonal approximation

$$(\nabla L)^2 = n^2 \tag{90}$$

where n is the index of refraction in the limit $k_0 \to \infty$ or $\lambda_0 \to 0$. This equation defines a ray.

In the $\lambda_0 \to 0$ limit, the only ray that makes contribution is the one makes $L - ct$ stationary. [1] Therefore $(\nabla L)^2 = n^2$ is equivalent to the requirement that $L - ct$ is stationary in the large k_0 limit.

On the other hand, we have the Jacobi-Hamilton equation

$$(\nabla W)^2 = 2m(E - V) \tag{92}$$

which follows from

$$\mathbf{p} = \nabla W \tag{93}$$

So if the classical equation of motion is to be an analogue of the ray optics for the underlying wave, we should have

$$\Psi = e^{A(x)+(i/\lambda)(W-Et)} \tag{94}$$

where $W - Et = S$ is the classical action so that extremizing it will give us the classical equation of motion. This should be valid in the limit $\lambda \to 0$. However, just as there are corrections to the geometrical optics, there are corrections to the above form. Here the constant λ is to be determiend by some experimental input. Classical mechanics should be recovered in the limit $\lambda \to 0$. This, of course, we know from Planck and Einstein to be $\hbar$. So let us accept that in the classical limit, the wavefunction is given by

$$\Psi = A(x)e^{iS/\hbar} \tag{95}$$

where the amplitude $A(x)$ is a slowly varying function.

Just as the eikonal approximation alone does not tell us about the Maxwell equation, this classical limit alone cannot tell us what exactly is the quantum wavefunction. However, with this relationship between the underlying wave and the manifest classical mechanics established, we can start thinking about building quantum mechanics.

Now consider time evolution. By the principle of superposition, we are restricted to consider linear equation. Therefore

$$i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle \tag{96}$$

[1] Now if ds is the line element that follows ∇L, this means that

$$\delta \int ds|\nabla L| = \delta \int ds\, n = 0 \tag{91}$$

Now unitarity demands

$$\frac{d}{dt}\int dx\,\langle\psi|\psi\rangle = 0 \tag{97}$$

or

$$\begin{aligned}\frac{d}{dt}\int dx\,\langle\psi|\psi\rangle &= \int dx\,\partial_t(\langle\psi|\psi\rangle)\\ &= \int dx\,\partial_t(\langle\psi|\psi\rangle)\\ &= \int dx\,(\partial_t\langle\psi|x\rangle)\langle x|\psi\rangle + \int dx\,\langle\psi|x\rangle(\partial_t\langle x|\psi\rangle)\\ &= \int dx\,\langle\psi|(\hat{H}-\hat{H}^\dagger)|\psi\rangle = 0\end{aligned} \tag{98}$$

Since ψ is arbitrary,

$$\hat{H} = \hat{H}^\dagger \tag{99}$$

which means that it is hermitian.

Limiting case says

$$i\hbar\partial_t\psi = -\partial_t S\psi \tag{100}$$

where $-\partial_t S$ is the classical Hamiltonian.

4 Translational invariance and conservation of linear momentum

In classical mechanics, we know that in the absence of external force, the total momentum is conseved. In particular, we know that if there if no external force on a **system** of particles, the **total momentum** of the system is conserved even if the momentum of individual particles are changing under the influence of mutual interactions.

How do we show that? Well, first of all construct the total momentum:

$$\mathbf{P}_{\text{total}} = \sum_{i=1}^{N}\mathbf{p}_i \tag{101}$$

and then take the time derivative

$$\frac{d}{dt}\mathbf{P}_{\text{total}} = \sum_{i=1}^{N} \frac{d}{dt}\mathbf{p}_i \tag{102}$$

but we know that

$$\frac{d}{dt}\mathbf{p}_i = \mathbf{F}_i \tag{103}$$

so that

$$\frac{d}{dt}\mathbf{P}_{\text{total}} = \sum_{i=1}^{N} \mathbf{F}_i = \mathbf{F}_{\text{total}} = 0 \tag{104}$$

by definition. That proves the assertion.

What about quantum mechanics system? Does this apply? Well, to answer this question, we need to know a few new things. One, we should know how to construct a wavefunction for **manybody system**. Then we need to know how such a wavefunction evolves in time.

Let's start from the Hamiltonian. In the case of a single particle, we have

$$\hat{H} = \hat{T} + \hat{V} \tag{105}$$

where $\hat{T} = \frac{\hat{p}^2}{2m}$ is the kinetic energy operator and $\hat{V}$ is the potential energy operator whose form is fixed by classical analogy.

What about 2 particle system? Well, continuting classical analogy, the classical Hamiltonian is given by

$$H_{\text{total}} = H_1 + H_2 + V_{12} \tag{106}$$

where H_1 is the Hamiltonian (energy) for particle 1 and H_2 is that for particle 2 and V_{12} is the interaction potential which is non-zero only in the presence of the second particle.

How do you prove that the total momentum is conserved in this classical case? Well, we can do the above force argument. But we can be fancy and use the Poisson bracket defined by

$$\{a, b\}_{PB} = \sum_{i=1}^{N} \left(\frac{\partial a}{\partial \mathbf{q}_i} \cdot \frac{\partial b}{\partial \mathbf{p}_i} - \frac{\partial a}{\partial \mathbf{p}_i} \cdot \frac{\partial b}{\partial \mathbf{q}_i} \right) \tag{107}$$

A quantity is conserved in time if

$$\frac{dA}{dt} = \{A, H_{\text{total}}\} \tag{108}$$

vanishes.

So consider the x component of $\mathbf{P}_{\text{total}}$

$$\begin{aligned} \frac{dP_x}{dt} &= \{P_x, H_{\text{total}}\} \\ &= \{(p_1)_x + (p_2)_x, H_{\text{total}}\} \\ &= -\frac{\partial H_{\text{total}}}{\partial (q_1)_x} - \frac{\partial H_{\text{total}}}{\partial (q_2)_x} \end{aligned} \tag{109}$$

since

$$\frac{d}{dt}p = -\frac{\partial H}{\partial q} \tag{110}$$

(this one of the Hamilton's equation. The other being $\frac{d}{dt}x = \frac{\partial H}{\partial p}$) that just goes back to the force argument. But what's useful to see in this is that the potential term V_{12} better satisfy

$$\partial_1 V_{12} + \partial_2 V_{12} = 0 \tag{111}$$

which implies that V_{12} must be a function of $|\mathbf{r}_1 - \mathbf{r}_2|$ only.

What if the function is not just a function of the differnce only? In that case, we can change variable to

$$\mathbf{r} = \mathbf{x}_1 - \mathbf{x}_2 \tag{112}$$

$$\mathbf{R} = \frac{m_1\mathbf{x}_1 + m_2\mathbf{x}_2}{m_1 + m_2} \tag{113}$$

whose canonical conjugates are

$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 \tag{114}$$

and

$$\mathbf{p} = \frac{m_2\mathbf{p}_1 - m_1\mathbf{p}_2}{m_1 + m_2} \tag{115}$$

show that

$$\frac{d}{dt}\mathbf{P} = -\frac{\partial H_{\text{total}}}{\partial \mathbf{R}} \neq 0 \tag{116}$$

Canonical conjugates

$$\begin{aligned} \mathbf{R} - \frac{m_1}{M}\mathbf{r} &= \frac{m_1\mathbf{x}_1}{M} + \frac{m_2\mathbf{x}_2}{M} - \frac{m_1\mathbf{x}_1}{M} + \frac{m_1\mathbf{x}_2}{M} \\ &= \mathbf{x}_2 \end{aligned} \tag{117}$$

$$\begin{aligned} \mathbf{R} + \frac{m_2}{M}\mathbf{r} &= \frac{m_1\mathbf{x}_1}{M} + \frac{m_2\mathbf{x}_2}{M} + \frac{m_2\mathbf{x}_1}{M} - \frac{m_2\mathbf{x}_2}{M} \\ &= \mathbf{x}_1 \end{aligned} \tag{118}$$

$$\begin{aligned} L &= m_1\frac{\mathbf{v}_1^2}{2} + m_2\frac{\mathbf{v}_2^2}{2} - V(\mathbf{x}_1 - \mathbf{x}_2) \\ &= \frac{m_1}{2}\left(\mathbf{V} + \frac{m_2}{M}\mathbf{v}\right)^2 + \frac{m_2}{2}\left(\mathbf{V} - \frac{m_1}{M}\mathbf{v}\right)^2 - V(\mathbf{r}) \\ &= \frac{M}{2}\mathbf{V}^2 + \frac{m_1 m_2}{2M}\mathbf{v}^2 - V(\mathbf{r}) \end{aligned} \tag{119}$$

$$\mathbf{P} = \frac{\partial L}{\partial \mathbf{V}} = M\mathbf{V} \tag{120}$$

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m_1 m_2}{M}\mathbf{v} \tag{121}$$

A consequence of this is that if the total momentum is to be conserved, then the potentials must necessarily of the form

$$V_{ij} = V(|\mathbf{x}_i - \mathbf{x}_j|) \tag{122}$$

Now what's good for this? Well, consider Earth orbiting the Sun. The Hamiltonian is

$$H = \frac{P_E^2}{2m_E} + \frac{P_S^2}{2m_S} - \frac{Gm_E m_S}{|\mathbf{x}_E - \mathbf{x}_S|} \tag{123}$$

Now do the canonical change of variable as above. We have

$$P_S = \frac{m_S}{m_S + m_E} P + p \tag{124}$$

$$P_E = \frac{m_E}{m_S + m_E} P - p \tag{125}$$

so that

$$H = \frac{P^2}{2M_{\text{total}}} + \frac{p^2}{2\mu} - \frac{Gm_E m_S}{r} \tag{126}$$

where $M_{\text{total}} = m_E + m_S$ and

$$\mu = \frac{m_E m_S}{m_S + m_E} \tag{127}$$

which means that we have reduced the 2-body problem to an effective 1 body problem involving the reduced mass and the distance between them only. Again this comes from

$$\frac{d}{dt}\mathbf{P} = -\frac{\partial H_{\text{total}}}{\partial \mathbf{R}} = 0 \tag{128}$$

How about quantum mechanics? Does this still work? In analogy, a 2 body problem in quantum mechanics can be written as

$$\hat{H} = \frac{\hat{p}_1^2}{2m_1} + \frac{\hat{p}_2^2}{2m_2} + \hat{V}(x_1 - x_2) \tag{129}$$

We need to ask 2 questions. One, does this mean that the total momentum is conserved? If yes, can we write down something similar to the reduced Hamiltonian as in the classical case?

Let's examine them one by one. First, the total momentum. To consider this sort of things, we need to know the position eigenstate of the 2 particles. For a single particle, we know that we can write

$$|\mathbf{x_1}\rangle \quad \text{and} \quad |\mathbf{x_2}\rangle \tag{130}$$

but what about the position space of 2 particles? Well, first thing to ask is if we can say

$$[\mathbf{x}_1, \mathbf{x}_2] = 0 \tag{131}$$

In other words, does measuring the position of particle 1 influence the subsequent measurement of the position of particle 2? The answer is, no. Think of an electron on Mars and another one on Earth. Why should the position of an electron on Mars influence an election on Earth? So the above is a reasonable thing to do. In that case, we can say that the two variables can be simultaneously measurable and we should be able to find a state which is the eigenstate of the both operators. Call it

$$|\mathbf{x}_1, \mathbf{x}_2\rangle \tag{132}$$

The same argument goes for the momentua and the cross commutators of momenta and position. Namely,

$$[\mathbf{p}_1, \mathbf{p}_2] = 0 \tag{133}$$
$$[\mathbf{p}_1, \mathbf{x}_2] = 0 \tag{134}$$
$$[\mathbf{x}_1, \mathbf{p}_2] = 0 \tag{135}$$

The translation operators for the particles are

$$\hat{T}_1(\mathbf{a})|\mathbf{r}_1, \mathbf{r}_2\rangle = |\mathbf{r}_1 + \mathbf{a}, \mathbf{r}_2\rangle \tag{136}$$

and

$$\hat{T}_2(\mathbf{a})|\mathbf{r}_1, \mathbf{r}_2\rangle = |\mathbf{r}_1, \mathbf{r}_2 + \mathbf{a}\rangle \tag{137}$$

Since p_1 and p_2 commute, we can say

$$\hat{T}_{\text{total}}(\mathbf{a}) = \hat{T}_1(\mathbf{a})\hat{T}_2(\mathbf{a}) = e^{-i\hat{\mathbf{p}}_1\cdot\mathbf{a}/\hbar}e^{-i\hat{\mathbf{p}}_2\cdot\mathbf{a}/\hbar} = e^{-i\hbar\mathbf{P}\cdot\mathbf{a}/\hbar} \tag{138}$$

Now it is obvious that the kinetic term in the total Hamiltonian commutes with $\hat{T}_{\text{total}}$. But what about the potential term? Well, since

$$\begin{aligned} \hat{T}(\mathbf{a})\hat{V}(\mathbf{r_1} - \mathbf{r_2})|\mathbf{r}_1, \mathbf{r}_2\rangle &= \hat{T}(\mathbf{a})V(\mathbf{r_1} - \mathbf{r_2})|\mathbf{r}_1, \mathbf{r}_2\rangle \\ &= V(\mathbf{r_1} - \mathbf{r_2})\hat{T}(\mathbf{a})|\mathbf{r}_1, \mathbf{r}_2\rangle \\ &= V(\mathbf{r_1} - \mathbf{r_2})|\mathbf{r}_1 + \mathbf{a}, \mathbf{r}_2 + \mathbf{a}\rangle \end{aligned} \tag{139}$$

and

$$\begin{aligned}\hat{V}(\mathbf{r_1}-\mathbf{r_2})\hat{T}(\mathbf{a})|\mathbf{r}_1,\mathbf{r}_2\rangle &= \hat{V}(\mathbf{r_1}-\mathbf{r_2})|\mathbf{r}_1+\mathbf{a},\mathbf{r}_2+\mathbf{a}\rangle \\ &= V(\mathbf{r_1}+\mathbf{a}-\mathbf{r_2}-\mathbf{a})|\mathbf{r}_1+\mathbf{a},\mathbf{r}_2+\mathbf{a}\rangle \\ &= V(\mathbf{r_1}-\mathbf{r_2})|\mathbf{r}_1+\mathbf{a},\mathbf{r}_2+\mathbf{a}\rangle \end{aligned} \tag{140}$$

we have

$$[\hat{T}(\mathbf{a}),\hat{H}_{\text{total}}]=0 \tag{141}$$

which means that

$$\frac{d}{dt}\langle P\rangle = \langle\psi|[\hat{P},\hat{H}_{\text{total}}]|\psi\rangle = 0 \tag{142}$$

for any 2 body state ψ with $V=V(\mathbf{r_1}-\mathbf{r_2})$.

So the total momentum seems to be a conserved quantity. Written this way, it is a little bit confusing. Is the total momentum exactly conserved or is it conserved in the sense that the average is conserved? In the case of quantum mechanics, the above actually implies that the total momentum is exactly conserved in any quantum process.

One can easily extend this argument and show that if the potential is of the form

$$V=\sum_{ij}V(\mathbf{r}_i-\mathbf{r}_j) \tag{143}$$

then the total momentum is conserved, just like in the classical case.

O.K. What about the hamiltonian? Can we reduce it?

Again, consider the above canonical transformation. Remember that in the classical sense, canonical transformation means that the Poisson bracket is preserved, or with new P and Q, we still have

$$\{Q_i,P_j\}=\delta_{ij} \tag{144}$$

In quantum mechanics, the analogous statement is that the commutator is preserved, or

$$[\hat{Q}_i,\hat{P}_j]=i\hbar\delta_{ij} \tag{145}$$

It turns out that if a canonical transformation preserves the Poisson bracket, then it also preserves the communtator. So

$$[\hat{r},\hat{P}]=0 \tag{146}$$

and

$$[\hat{r}, \hat{p}] = i\hbar\delta \tag{147}$$

and

$$[\hat{R}, \hat{P}] = i\hbar\delta \tag{148}$$

How do we see that the quantum Hamiltonian also transforms under this just like the classical Hamiltonian?

Well, chain-rule. That's all. Use

$$\begin{aligned} \frac{\partial}{\partial \mathbf{r}_1} &= \frac{\partial \mathbf{r}}{\partial \mathbf{r}_1}\frac{\partial}{\partial \mathbf{r}} + \frac{\partial \mathbf{R}}{\partial \mathbf{r}_1}\frac{\partial}{\partial \mathbf{R}} \\ &= \frac{\partial}{\partial \mathbf{r}} + \frac{m_1}{m_1 + m_2}\frac{\partial}{\partial \mathbf{R}} \end{aligned} \tag{149}$$

Compare this with the classical expression. Upon identifying

$$p_1 = -i\hbar\partial_1 \tag{150}$$

$$p = -i\hbar\partial_r \tag{151}$$

$$P = -i\hbar\partial_R \tag{152}$$

they are identical. Therefore, in QM, one can also write

$$\hat{H}_{\text{total}} = \frac{\hat{P}^2}{2M_{\text{total}}} + \frac{\hat{p}^2}{2\mu} + \hat{V}(r) \tag{153}$$

Since the total kinetic energy

$$\hat{K}_{\text{total}} = \frac{\hat{P}^2}{2M_{\text{total}}} \tag{154}$$

and the Hamiltonian of the relative motion

$$\hat{H}_{\text{reduced}} = \frac{\hat{p}^2}{2\mu} + \hat{V}(r) \tag{155}$$

commute, we can specify the eigenvalues of the two operators simultaneously. By changing the frame, one can always change the total momentum to zero. But the relative motion, of course is invariant.

5 Estimate of E_0

First of all, let's see if we can have the order of magnitude of the energy without complicated calculations given the parameters of the system, namely, m_e, e and also basic constant c and $\hbar$. First note that the only energy unit we can construct from these constants is $m_e c^2 = 0.511$ MeV.

Next, if we think of a process where two electric current is exchaging a photon, the amplitude is proportional to e^2 because each current couples to the photon with e. To get a probability we have to square the amplitude. Hence, the probability should be proportional to e^4 or to the fine structure constant squared: α^2. From this consideration, then roughly, one would expect

$$E \sim m_e c^2 \alpha^2 = O(10 \text{ eV}) \tag{156}$$

The ground state of the hydrogen atom has an energy of (minus) 13.6 eV and this is usually referred to as the Rydberg constant. or

$$E_n = -\frac{Z^2 R}{n^2} \tag{157}$$

To have an idea about the length scale we note that

$$E \sim e^2/r \sim \frac{\alpha}{r} \sim R \tag{158}$$

or

$$r \sim \alpha/R \sim 10^{-10}\text{m} = 1A \tag{159}$$

This is of course, nothing but the Bohr radius

$$a_0 = \frac{\hbar}{m_e c^2 \alpha} = 0.529A \tag{160}$$

Consider a system consisting of an electron and a proton. In classical mechanics, the ground state of such a system is when the electron is right on top of the proton. In that case, the potential energy is enormous since the size of proton and the electron is practially nothing. Bad news is, if this is the case, then no atom can exist! This puzzle was finally resolved with the introduction of quantum mechanics and the uncertainty principle. The hamiltonian is

$$H = \frac{p^2}{2\mu} - \frac{e^2}{r} \tag{161}$$

So if the electron's position is known to be **on top of the proton** at a certain moment, that means that the uncertainty in the momentum is

$$\Delta p \sim \hbar / r_{\text{proton}} \tag{162}$$

Now the size of the proton is about 1 fm and that means that the momentum is about 200 MeV. This is much larger than the binding coulomb energy

$$e^2/r \approx \frac{1}{137} \cdot 200\,\text{MeV} \sim 1\,\text{MeV} \tag{163}$$

so that at the next moment, the electron is nowhere in the vicinity of the proton. It also means that the average kinetic energy should be of the order

$$(100\,\text{MeV})^2/(0.5\,\text{MeV}) \sim 10^4\text{MeV} \tag{164}$$

Way too much for a bound electron.

So quantum mechanics with its uncertainty principle prevents the electron to go anywhere near the proton and stay there. To have a bound state, then, there must be a balance between the confinement and the momentum uncertainty arising from the confinement.

One way of estimating the energy scale is to think about the speed of the electron. The basic constants given in the problem are the reduced mass of the electron μ, the charge of the electron e and the Planck constant $\hbar$. We will supply c as needed, but since this is a non-relativistic problem, its appearance is not really needed.

We note that

$$\frac{e^2}{\hbar c} = \alpha = \frac{1}{137} \tag{165}$$

Therefore the combination

$$v = \frac{e^2}{\hbar} \tag{166}$$

has the dimension of speed and this is the only combination we can cook up given these constants. Then the kinetic energy is

$$\frac{p^2}{2m} = \frac{mv^2}{2} \approx m \left(\frac{e^2}{\hbar}\right)^2 /2 \tag{167}$$

Supplying enough factors of c's as needed we get

$$\frac{p^2}{2m} = \frac{mv^2}{2} \approx \frac{mc^2}{2}\left(\frac{e^2}{\hbar c}\right)^2 = \frac{mc^2\alpha^2}{2} \tag{168}$$

Now classically, for coulomb problem,

$$E = \frac{mv^2}{2} - \frac{Ze^2}{r} \tag{169}$$

Consider a circular orbit. In that case, $r = R$ is a constant. The centrifugal force is given by

$$F_{\text{cent}} = \frac{mv^2}{R} \tag{170}$$

which must balance the coulomb force

$$F_{\text{coulomb}} = \frac{Ze^2}{R^2} \tag{171}$$

This yields

$$mv^2 = \frac{Ze^2}{R} \tag{172}$$

or

$$K = \frac{mv^2}{2} = \frac{1}{2}|V(R)| \tag{173}$$

So we may guess that even for the quantum case

$$\begin{aligned} E &= \left\langle \frac{p^2}{2m} \right\rangle + \langle V \rangle \approx \left\langle \frac{p^2}{2m} \right\rangle - 2\left\langle \frac{p^2}{2m} \right\rangle \\ &= -\left\langle \frac{p^2}{2m} \right\rangle \\ &= -\frac{mc^2Z^2\alpha^2}{2} \approx -13.6Z^2\,\text{eV} \end{aligned} \tag{174}$$

which is actually exact. But don't expect this to happen everytime we or you make an estimate. All we can conclude from this is that the energy unit is

$$|E_{\text{unit}}| = mc^2\alpha^2 \tag{175}$$

The fact that we got the factor of $1/2$ right is purely accidental. For instance, the actual speed could be $1/2$ of $e^2/\hbar$ or twice that. We don't know at this point.

The book has somewhat different approach: Suppose the size of the ground state hydrogen atom is a. Then the size of the potential energy is

$$V \sim \frac{e^2}{a} \tag{176}$$

The size of the momentum is

$$p \sim \frac{\hbar}{a} \tag{177}$$

and the kinetic energy

$$K = \frac{p^2}{2\mu} \sim \frac{\hbar^2}{2\mu a^2} \tag{178}$$

So the total energy is

$$H \sim \frac{\hbar^2}{2\mu a^2} - \frac{e^2}{a} \tag{179}$$

This, as a function of a has a minimum at

$$a = \frac{\hbar^2}{m_e e^2} \tag{180}$$

and

$$E_1 \sim -\frac{m_e e^4}{2\hbar^2} \tag{181}$$

How do we estimate these? Well remember these:

$$\hbar c = 197.3\,\mathrm{MeV\,fm} = 197.3\,\mathrm{eV\,nm} \tag{182}$$

and

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137} \tag{183}$$

so

$$\begin{aligned} a &= \frac{\hbar^2}{m_e e^2} \\ &= \frac{\hbar c}{m_e c^2}\frac{\hbar c}{e^2} \\ &\approx \frac{200\,\mathrm{MeVfm}}{0.5\,\mathrm{MeV}} \times 137 \\ &= 5.5 \times 10^4 \mathrm{fm} = 0.55 \times 10^{-10} m = 0.55 \mathring{A} \end{aligned} \tag{184}$$

which is nothing but the Bohr radius and

$$\begin{aligned} E_1 &\sim -\frac{m_e e^4}{2\hbar^2} \\ &= -\frac{1}{2} m_e c^2 \left(\frac{e^2}{\hbar c}\right)^2 \\ &= -0.5 \times 0.511\,\mathrm{MeV}\,\frac{1}{137^2} \\ &= -1.36 \times 10^{-5} \mathrm{MeV} = -13.6\,\mathrm{eV} \end{aligned} \tag{185}$$

which happened to be the exact ground state energy.

All atomic phenomena happens within this energy scale and the length scale and the reason is exactly what's given above. The ground state of an atom with a few electrons with a electrostatic interaction strength of a few electon charge, must have this characteristic length and energy scale.

What happened here? Can we always do this sort of things? Well, yes and no. Yes, you should almost always be able to at least guesstimate the size of your system and the energy. No, because you don't really know whether

$$\Delta p \approx \frac{\hbar}{a} \tag{186}$$

or

$$\Delta p \approx 2\frac{\hbar}{a} \tag{187}$$

of course you know that something like

$$\Delta p \approx \frac{\hbar}{10 \times a} \tag{188}$$

can't happen because

$$\Delta p \Delta x \geq \frac{\hbar}{2} \tag{189}$$

But still, there is a big room to manuver. The fact that we got the ground state energy exactly right is just a luck than anything else. We could have as easily chosen

$$\Delta p \approx 2\frac{\hbar}{a} \tag{190}$$

and get completely different answer. But that doesn't mean that this estimate is wrong. Remember we are estimating. A factor of 2 error is certainly permitted. Just that don't be surprised that your **numerical** answer turns out to be off by a certain factor.

However, what's important is how the energy and the size depend on the parameters of the system. For instance, a is inversely proprotional to m_e and e^2. **That** should be right even if the coefficient in front may be wrong. Also for the energy, first of all we got the minus sign. That means **bound state**. So that must be right. And the m_e and e dependence should be also right. In a sense, there is no other **natural** combination of these parameters that gives you these answers.

Another thing to remember is that without a constant **conversion factor**

$$\hbar c \approx 200\,\text{eV}\,\text{nm} = 2000\,\text{eV}\,\text{Å} \tag{191}$$

all these estimate would be impossible. We will use this many many times. So, better memorize the above formula. Note that these two constants are the corner stones of modern physics. With c, time and space are unified. With $\hbar$ added, one may say that the energy and the length (space) are unified.

We can summarize these as follows:

- Energy:
  - Has the dimension of mv^2.
  - Measured in J or more conveniently in QM, eV, keV, MeV or GeV.
- Momentum:
  - Has a unit of mv or Energy/speed.

  - Natuarl unit is then eV/c, keV/c, MeV/c or GeV/c. Often times, people omit /c and just say "the value of momentum is 10 eV".

- Mass:
  - Has a unit of m or Energy/speed^2.
  - Natuarl unit is then eV/c^2, keV/c^2, MeV/c^2 or GeV/c^2. Often times, people omit /c^2 and just say "the mass of an electron is 0.511 MeV".

- Length:
  - Can be converted to momentum scale by using $\hbar c = 200\,\text{eVnm}$. Or the length "has the dimension of the inverse momentum" which manifest itself in the uncertainty principle

    $$\Delta x \Delta p \sim \hbar \tag{192}$$

    But we usually just say that the length "has the dimension of the inverse energy" since the only difference is a constant factor of c.
  - 200 nm is then equal to

    $$200\text{nm} = \frac{\hbar c}{1\text{eV}} \tag{193}$$

  - We can also write $\hbar c = 200\,\text{MeVfm} = 0.2\,\text{GeVfm}$. Therefore the size of a proton (about 1 femtometer) is equal to

    $$1\,\text{fm} = \frac{\hbar c}{200\,\text{MeV}} \tag{194}$$

  - People often just say that "1 fm is 200 MeV".

- Time:
  - Can be converted to the length scale by multiplying by c.
  - So the natural unit is **not** second but m/c, cm/c, mm/c, μm/c, nm/c, fm/c.
  - People often times omit /c and just say something like: "it takes 1 fm time for the light to go across the proton".

- Can be converted to energy scale by using $\hbar c = 200\,\mathrm{eVnm}$.
- 200 nm/c is then equal to

$$200\mathrm{nm}/c = \frac{\hbar}{1\mathrm{eV}} \tag{195}$$

- We can also write $\hbar c = 200\,\mathrm{MeVfm} = 0.2\,\mathrm{GeVfm}$. Therefore 1fm/c is equal to

$$1\,\mathrm{fm}/c = \frac{\hbar}{200\,\mathrm{MeV}} \tag{196}$$

- People often just say that "1 fm is 200 MeV".

- One more thing to remember is

$$\frac{e^2}{\hbar c} = \frac{1}{137} \tag{197}$$

  is a pure number.

This may look somewhat complicated. But this is actually very liberating. Trick is to forget about factors of $\hbar$ and c and use the energy scale for everything. And then supply the necessary factors of $\hbar$ and c at the end of calculations.

For an example, let me repeat the estimate of the Bohr radius and the ground state energy.

We have

$$H = \frac{p^2}{2m_e} - \frac{\alpha}{r} \tag{198}$$

Note that we have substituted e^2 with α. We can do that since we are ignoring $\hbar c$. Suppose the characteristic length scale is a. The uncertainty principle says

$$\Delta p a \sim 1 \tag{199}$$

so that

$$p^2 \sim 1/a^2 \tag{200}$$

Then

$$H \sim \frac{1}{2m_e a^2} - \frac{\alpha}{a} \tag{201}$$

Minimizing yields

$$a = \frac{1}{m_e \alpha} \tag{202}$$

The ground state energy is

$$E_0 \sim -\frac{m_e \alpha^2}{2} \tag{203}$$

Much simpler, isn't it? Now supply necessary $\hbar$ and c. The Bohr radius is

$$a = \frac{\hbar c}{m_e c^2 \alpha} \approx \frac{200\,\text{MeVfm}}{0.5\text{MeV}}(137) \approx 0.055\,\text{nm} \tag{204}$$

The ground state energy is

$$E_0 \sim -\frac{m_e c^2 \alpha^2}{2} \approx -\frac{0.25\text{MeV}}{137^2} \approx -13\,\text{eV} \tag{205}$$

6 Rotational Invariance and conservation of angular momentum

We saw that just as in classical mechanics, translational invariance of the system expressed as

$$[\hat{H}, \hat{T}_{\text{total}}] = 0 \tag{206}$$

or equivalently

$$[\hat{H}, \hat{P}_{\text{total}}] = 0 \tag{207}$$

leads to the conservation of total momentum.

Classically we know that rotational invariance leads to the conservation of angular momentum. Let's see if we can figure out what this means quantum mechanically.

First of all, what is rotation? Well, intuitively it's quite clear what I mean by rotation. More precisely, you can say that rotation is whatever oparation on $\mathbf{r}$ that leaves the length

$$r^2 = \mathbf{x}\cdot\mathbf{x} = x^2 + y^2 + z^2 \tag{208}$$

unchanged. For instance, if you rotate $\mathbf{x}$ around the z axis, the rotated vector is given by

$$\mathbf{x}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{209}$$

where ϕ is the angle of rotation and it is done in the counter-clock wise. Due to the identity $\cos^2 + \sin^2 = 1$, this preserves the length of the vector.

Quantum mechanically, we can define the rotation operator (around z axis) $\hat{R}_z(\phi)$ by

$$\hat{R}_z(\phi)|x, y, z\rangle = |x\cos\phi - y\sin\phi, y\cos\phi + x\sin\phi, z\rangle \tag{210}$$

To get the **generator** of the rotation $\hat{L}_z$ given by

$$\hat{R}_z = e^{-i\phi\hat{L}_z/\hbar} \tag{211}$$

consider an infinitesimal rotation with $d\phi$. In that case,

$$\mathbf{x}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & -d\phi & 0 \\ d\phi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - yd\phi \\ y + xd\phi \\ z \end{pmatrix} \tag{212}$$

Quantum mechanically, this is

$$\hat{R}_z(d\phi)|x, y, z\rangle \;=\; |x - yd\phi, y + xd\phi, z\rangle \tag{213}$$

Now we note that

$$\hat{T}_x(\epsilon)|x, y, z\rangle = |x + \epsilon, y, z\rangle = [1 - \epsilon(i/\hbar)\hat{p}_x]\,|x, y, z\rangle \tag{214}$$

so that we can write

$$\begin{aligned} \hat{R}_z(d\phi)|x, y, z\rangle &= |x - yd\phi, y + xd\phi, z\rangle \\ &= [1 + (yd\phi)(i/\hbar)\hat{p}_x]\,[1 - (xd\phi)(i/\hbar)\hat{p}_y]\,|x, y, z\rangle \\ &= \left[1 - d\phi(i/\hbar)\hat{L}_z\right]|x, y, z\rangle \end{aligned} \tag{215}$$

which leads to

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = (\mathbf{r} \times \mathbf{p})_z \tag{216}$$

We would like to show that this is a conserved quantity if the space is isotropic. That is, one direction is as good as any other. To do so, we need to show that

$$[\hat{L}_z, \hat{H}] = 0 \tag{217}$$

Now consider the kinetic term. It is a function of $\mathbf{p}^2$ only. Therefore, classically, it is certainly rotationally invariant. What about in QM? Consider first the following commutators.

$$[L_z, p_x] = [xp_y - yp_x, p_x] = [xp_y, p_x] = [x, p_x]p_y = i\hbar p_y \tag{218}$$

where I used the "product rule"

$$[AB, C] = [A, C]B + A[B, C] \tag{219}$$

This is analogous to the differentiation rule. But in this case, one must be careful keeping the order of operators right.

By the same token

$$[L_z, p_y] = -i\hbar p_x \tag{220}$$

and

$$[L_z, p_z] = 0 \tag{221}$$

since z does not appear in L_z.

Then

$$\begin{aligned} [L_z, p^2] &= [L_z, p_x^2 + p_y^2 + p_z^2] \\ &= [L_z, p_x^2] + [L_z, p_y^2] \\ &= [L_z, p_x]p_x + p_x[L_z, p_x] + [L_z, p_y]p_y + p_y[L_z, p_y] \\ &= 2i\hbar p_y p_x - 2i\hbar p_x p_y = 0 \end{aligned} \tag{222}$$

Also

$$[L_z, x] = [xp_y - yp_x, x] = -y[p_x, x] = i\hbar y \tag{223}$$

$$[L_z, y] = [xp_y - yp_x, y] = x[p_y, y] = -i\hbar x \tag{224}$$

$$[L_z, z] = 0 \tag{225}$$

so that

$$\begin{aligned}[L_z, x^2+y^2+z^2] &= [L_z, x^2] + [l_z, y^2] \\ &= x[L_z, x] + [L_z, x]x + y[L_z, y] + [L_z, x]y \\ &= 2i\hbar xy - 2i\hbar xy = 0 \end{aligned} \tag{226}$$

Therefore if the Hamiltonian is of the form

$$H = \frac{p^2}{2m} + V(|\mathbf{r}|) \tag{227}$$

then

$$[H, L_z] = 0 \tag{228}$$

which establishes

$$\frac{d}{dt}\langle L_z\rangle = (i/\hbar)\langle\psi|[H, L_z]|\psi\rangle = 0 \tag{229}$$

Now did you catch a pattern? We have

$$\begin{aligned}[L_z, p_x] &= i\hbar p_y \\ [L_z, p_y] &= -i\hbar p_x \\ [L_z, p_z] &= 0 \end{aligned} \tag{230}$$

and

$$\begin{aligned}[L_z, x] &= i\hbar y \\ [L_z, y] &= -i\hbar x \\ [L_z, z] &= 0 \end{aligned} \tag{231}$$

In general if $\hat{\mathbf{V}}$ is a vector operator,

$$\begin{aligned}[L_z, V_x] &= i\hbar V_y \\ [L_z, V_y] &= -i\hbar V_x \\ [L_z, V_z] &= 0 \end{aligned} \tag{232}$$

which actually comes from the fact that rotation operator must rotate not only the poistion vector but also **any** vector in the same manner.

- Two roles of $\hat{\mathbf{L}}$: Generator and Angular momentum
- The commutator relation

$$[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z \tag{233}$$

  is not a pure quantum effect, unlike $[\hat{x}, \hat{p}] = i\hbar$. Even in CM, the rotations do not commute.

- For an vector operator, we have the **definition**

$$(\hat{R}^\dagger \hat{\mathbf{V}} \hat{R})_i = R_{ij} \hat{V}_j \tag{234}$$

  where R_{ij} is the usual c-number rotation matrix. Note a classical definition of a vector is, 'a quantity that rotates just like $\mathbf{x}$.'

- Rotation operator with respect to a general direction $\boldsymbol{\phi}$:

$$\hat{R}(\boldsymbol{\phi}) = \exp(-i\boldsymbol{\phi}\cdot\hat{\mathbf{L}}/\hbar) \tag{235}$$

  The rotational axis is in the direction of $\boldsymbol{\phi}$ and the angle or rotation is $\phi = |\boldsymbol{\phi}|$. The convention is that this is rotation done in the counter-clock wise direction.

7 A complete set of commuting observables

Now if the space is rotationally invariant, there is nothing particular about the z axis. We might as well have chosen the x axis or y axis.

So the relation

$$L_z = xp_y - yp_x \tag{236}$$

can be immediately generalized to

$$L_x = yp_z - zp_y \tag{237}$$
$$L_y = zp_x - xp_z \tag{238}$$

A way to remember this is to use the 'cyclic' property. Look at the first term. It always is in the order that follows the cyclic permutation of xyz, that is: xyz, yzx, zxy. Then you remember that Classically

$$L = r \times p \tag{239}$$

and guess the second term to be anti-symmetric version of the first. This can be also written as

$$\hat{L_i} = i\hbar\epsilon_{ijk}x_j p_k \tag{240}$$

where ϵ_{ijk} is the completely anti-symmetric tensor with the value

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{321} = 1 \tag{241}$$

$$\epsilon_{ijk} = -\epsilon_{jik} \tag{242}$$

A very useful identity is

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{lj} \tag{243}$$

Now I said last time that for **any** vector, the action of the rotation operator has to be

$$[L_z, V_x] = i\hbar V_y \tag{244}$$

$$[L_z, V_y] = -i\hbar V_x \tag{245}$$

$$[L_z, V_z] = 0 \tag{246}$$

Now the angular momentum itself is a vector (pseudo-vector, to be precise). So it must also obey

$$[L_z, L_x] = i\hbar L_y \tag{247}$$

$$[L_z, L_y] = -i\hbar L_x \tag{248}$$

or in a cyclic (xyz) form,

$$[L_z, L_x] = i\hbar L_y \tag{249}$$

$$[L_y, L_z] = i\hbar L_x \tag{250}$$

There is nothing special about the z direction. So we must also have

$$[L_x, L_y] = i\hbar L_z \tag{251}$$

Since the commutator is anti-symmetric, we can summarize this as

$$[L_i, L_j] = i\hbar\epsilon_{ijk}L_k \tag{252}$$

Can we prove this from the operator expression

$$L_i = \epsilon_{ijk} x_j p_k \quad ? \tag{253}$$

Well, let's see if we can.

$$\begin{aligned}
[L_i, L_j] &= [\epsilon_{imn} x_m p_n, \epsilon_{jkl} x_k p_l] \\
&= \epsilon_{imn}\epsilon_{jkl} [x_m p_n, x_k p_l] \\
&= \epsilon_{imn}\epsilon_{jkl} ([x_m p_n, x_k] p_l + x_k [x_m p_n, p_l]) \\
&= \epsilon_{imn}\epsilon_{jkl} (x_m [p_n, x_k] p_l + x_k [x_m, p_l] p_n) \\
&= \epsilon_{imn}\epsilon_{jkl} (-i\hbar\delta_{nk} x_m p_l + i\hbar\delta_{ml} x_k p_n) \\
&= -i\hbar\epsilon_{imk}\epsilon_{jkl} x_m p_l + i\hbar\epsilon_{iln}\epsilon_{jkl} x_k p_n \\
&= i\hbar\epsilon_{kim}\epsilon_{kjl} x_m p_l - i\hbar\epsilon_{lin}\epsilon_{ljk} x_k p_n \\
&= i\hbar (\delta_{ij}\delta_{ml} - \delta_{il}\delta_{mj}) x_m p_l - i\hbar (\delta_{ij}\delta_{nk} - \delta_{ik}\delta_{nj}) x_k p_n \\
&= -i\hbar x_j p_i + i\hbar x_i p_j
\end{aligned} \tag{254}$$

On the other hand,

$$\begin{aligned}
& i\epsilon_{ijk}\hbar L_k \\
&\qquad = i\epsilon_{ijk}\epsilon_{klm} x_l p_m \\
&\qquad = i(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) x_l p_m \\
&\qquad = i x_i p_j - i x_j p_i
\end{aligned} \tag{255}$$

Done. For instance,

$$[L_x, L_y] = i\hbar x p_y - i\hbar y p_x = i\hbar L_z \tag{256}$$

It is also then given that the magnitude of the angular momentum

$$\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2 \tag{257}$$

is invariant under the rotation. Furthermore, since

$$[H, L_i] = 0 \tag{258}$$

for each L_i, we have

$$[H, L^2] = 0 \tag{259}$$

as well.

In this way, we have 3 mutually commuting indepedent **hermitian** operators for the rotationally invariant hamiltionian

$$\hat{H}, \hat{L^2}, \hat{L}_z \tag{260}$$

Hence any quantum states can be labelled by the eigenvalues of these three:

$$\hat{H}|E,l,m\rangle = E|E,l,m\rangle \tag{261}$$
$$\hat{L^2}|E,l,m\rangle = l(l+1)\hbar^2|E,l,m\rangle \tag{262}$$
$$\hat{L}_z|E,l,m\rangle = m\hbar|E,l,m\rangle \tag{263}$$

So what are the eigenvalues of L^2 and L_z? To do that, we first define the lowering and the raising operators

$$L_+ = L_x + iL_y \tag{264}$$
$$L_- = L_x - iL_y \tag{265}$$

One can easily show that

$$[L^2, L_\pm] = 0 \tag{266}$$

and

$$\begin{aligned} [L_\pm, L_z] &= [L_x, L_z] \pm i[L_y, L_z] \\ &= -i\hbar L_y \pm i(i\hbar L_x) \\ &= \mp\hbar\left(L_x \pm iL_y\right) = \mp\hbar L_\pm \end{aligned} \tag{267}$$

Why are we doing this? Well, suppose that we have an eigenstate of L_z with the eigenvalue given by $m\hbar$:

$$L_z|m\rangle = m\hbar|m\rangle \tag{268}$$

Now we apply $L_\pm$:

$$L_\pm L_z|m\rangle = m\hbar L_\pm|m\rangle \tag{269}$$

Suppose we apply $L_\pm$ first:

$$\begin{aligned} L_z L_\pm|m\rangle &= L_\pm L_z|m\rangle + [L_z, L_\pm]|m\rangle \\ &= m\hbar L_\pm|m\rangle \pm \hbar L_\pm|m\rangle \\ &= (m \pm 1)\hbar L_\pm|m\rangle \end{aligned} \tag{270}$$

Therefore apart from possible normalization constant α, we can say that

$$L_{\pm}|m\rangle = \alpha|m \pm 1\rangle \tag{271}$$

The only restriction that the commutator relationships impose on eigenvalues is

$$\lambda = \pm m \ \text{ with } \Delta m = 1 \tag{272}$$

This is possible in two cases:

$$\lambda = 0, \pm 1, \pm 2, ..., \pm n, ... \tag{273}$$

and

$$\lambda = \pm 1/2, \pm 3/2, \pm 5/2, ..., \pm(2n+1)/2, ... \tag{274}$$

Now we know that since $L_z = xp_y - yp_x$ involves derivative,

$$\langle x|m = 0\rangle = \text{constant} \tag{275}$$

is an eigenstate of L_z with $m = 0$. So in this case we should have the integer series.

Combined with the above relation, this implies that the possible values of m are $0, \pm 1, \pm 2, ...$

O.K. What about L^2? Denote

$$L^2|l, m\rangle = \lambda\hbar^2|l, m\rangle \tag{276}$$

First of all, we see that since

$$\langle l, m|(L_x^2 + L_y^2)|l, m\rangle \geq 0 \tag{277}$$

we should have

$$\langle l, m|(L^2 - L_z^2)|l, m\rangle = (\lambda - m^2)\hbar^2\langle l, m|l, m\rangle \geq 0 \tag{278}$$

or

$$\lambda \geq m^2 \tag{279}$$

This implies that for given l, there is a maximum value of m. Now let that maximum value be λ. Then since it is the maximum, we can't raise it any more. Or

$$L_+|l,\lambda\rangle = 0 \tag{280}$$

On the other hand, note that

$$\begin{aligned} L_-L_+ &= (L_x - iL_y)(L_x + iL_y) \\ &= L_x^2 + L_y^2 + i[L_x, L_y] \\ &= L_x^2 + L_y^2 - \hbar L_z \end{aligned} \tag{281}$$

so that

$$L_-L_+ = L^2 - L_z^2 - \hbar L_z \tag{282}$$

Then

$$\begin{aligned} 0 &= L_-L_+|l,\lambda\rangle \\ &= (L^2 - L_z^2 - \hbar L_z)|l,\lambda\rangle \\ &= (L^2 - \lambda^2 - \lambda\hbar)|l,\lambda\rangle \end{aligned} \tag{283}$$

that is,

$$L^2|l,\lambda\rangle = \lambda(1+\lambda)\hbar|l,\lambda\rangle \tag{284}$$

one can apply similar argument to the minimum value of m and ascertain that the minimum value of m is $-\lambda$. and

$$L^2|l,m\rangle = \lambda(\lambda+1)\hbar^2|l,m\rangle \tag{285}$$

for any m. So we might as well identify

$$l = \lambda \tag{286}$$

Now let's figure out the normalization. Note that

$$(L_\pm)^\dagger = L_\mp \tag{287}$$

Hence

$$\langle l,m|L_-L_+|l,m\rangle = \langle l,m+1|l,m+1\rangle c^2 \tag{288}$$

where we define c to be

$$L_+|l,m\rangle = c_+|l,m\rangle \tag{289}$$

Using

$$L_-L_+ = L^2 - L_z^2 - \hbar L_z \tag{290}$$

we get

$$\begin{aligned} \langle l,m|L_-L_+|l,m\rangle &= \langle l,m|(l(l+1)-m(m+1))\hbar^2|l,m\rangle \\ &= c_+^2 \end{aligned} \tag{291}$$

so that

$$L_+|l,m\rangle = \sqrt{l(l+1)-m(m+1)}\hbar|l,m+1\rangle \tag{292}$$

Similarly using

$$L_+L_- = L_-L_+ = L^2 - L_z^2 + \hbar L_z \tag{293}$$

we get

$$L_-|l,m\rangle = \sqrt{l(l+1)-m(m-1)}\hbar|l,m-1\rangle \tag{294}$$

Now take a breath and ask some questions.

1. Why 3 mutually commuting operators? Why not more? Why not less?
2. What has uncertainty principle has to with the quantization of the angular momentum?
3. Why can't we have $L_z^2 = L^2$?

The first question is easy to answer. Classically, a point particle state is completely specified by specifying the phase space variables $(\mathbf{x}, \mathbf{p})$. So the number of degree of freedom is 6. In quantum mechanics, uncertainty principle prevents us from specifying both the position and the momentum. Therefore, all we can specify are **half** of the classical phase space variable. So in 3-D, we have 3 independent variables. This could be (x, y, z), (p_x, p_y, p_z) of course. But they have not so nice property of changing with time. Just as in

classical mechanics, it is much better to characterize each state with the **constants** of motion. In classical mechanics, this could be arbitrary combination of initial conditions, energy, poistion or momentum. In quantum mechanics, we can't do that. If we want to specify a state with some quantum numbers, they better be measurable simultanesously. That is, the operators should commute with each other. Since H, L^2, L_z commute and they are constants of motion, they are the ideal labels in the case of rotational invariance.

The second question is harder to answer. Consider L_z. In spherical coordinate system where

$$x = r\sin\theta\cos\phi \tag{295}$$

$$y = r\sin\theta\sin\phi \tag{296}$$

$$z = r\cos\theta \tag{297}$$

one can show that

$$L_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi} \tag{298}$$

That means

$$[\phi, L_z] = i\hbar \tag{299}$$

and there is a corresponding uncertainty

$$\Delta\phi\Delta L_z \geq \frac{\hbar}{2} \tag{300}$$

Trouble is that in a sense, this represents a **confined** system. If you don't care about the actual circle, you might as well say that this is like a particle confined in a infinite potential well where the length of the potential well is 2π. In that case, we know that the momentum is quantized since a trapped wave can only be stable if it is a standing wave.

The same happens here. In a way, the above represent a particle trapped in a potential well with a **periodic** boundary condition and the possible momenta are an integer multiple of $\hbar/L$. Well, in this case, it is the angular momentum L_z that is quantized. So the possible values of L_z are

$$L_z = \pm n\hbar \tag{301}$$

with n being the integer.

It is perhaps instructive here to review how this quantization came about. We showed already that the uncertaity principle

$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{302}$$

implies that

$$\langle x|p\rangle = e^{ixp/\hbar} \tag{303}$$

This was because if one specifies the position, then there is no information whatsoever about the value of the momentum, or

$$P_p = |\langle p|x\rangle|^2 = \text{constant} \tag{304}$$

then the unitarity

$$1 = \int dx\, |x\rangle\langle x| = \int dp\, |p\rangle\langle p| \tag{305}$$

demanded that

$$\int dp\, \langle x|p\rangle\langle p|y\rangle = \delta(x-y) \tag{306}$$

Of course the only way this is possible is to have

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ixp/\hbar} \tag{307}$$

Or another way of saying it, if one specifies the momentum, the wave function is the plane wave. So if the space is infinite, we can have any value of p.

Now suppose we have a box of size L. At the walls, the waves reflect. In 1-D, we might as well say that we have a finite size string. Then we know that the only stable modes of vibration are the standing waves for which the length of the string is an integer multiple of the wavelength.

Why is that? Well, this is because interference. When the wave is reflected from the wall, unless a certain condition is met, the phase of the incoming wave and the reflected wave are random. If you have enough random phases pile up, the wave cancel itself. The standing waves are the waves where the phases of the incoming wave and the reflected wave satisfies a certain matching conditions so that the wave does not cancel itself. So even if you have many different wave mode at the beginning, after a short while,

the only ones that are still there are the standing waves. That's how all the musical instruments works.

So the lesson is, if you want to confine a wave, you can do so only if you have standing waves. That means that the wavelength and the size of the box must satisfy a certain relationship involving integers (usually linear relationship). And that means quantization of the wavelength.

So the uncertainty strikes again. The quantization of angular momentum can be ultimately traced back to the fact that a circle represents a bound state and therefore, quantization.

The third question is also due to the uncertainty. For linear momentum, one component of the vector can surely take on the whole size of the momentum. Namely,

$$\mathbf{p} = (p, 0, 0) \tag{308}$$

is possible. But this is only possible because the components of $\mathbf{p}$ commute:

$$[p_i, p_j] = 0 \tag{309}$$

But for the angular momentum, this is not the case. We have

$$[L_i, L_j] = i\hbar\epsilon_{ijk}L_k \tag{310}$$

That means that we have the uncertainty relation, for instance,

$$\Delta L_z \Delta L_x \geq \frac{\hbar}{2}|\langle L_y \rangle| \tag{311}$$

Therefore, one cannot specify all three components. Suppose we **know** that the angular momentum vector is aligned in the direction of z axis. In that case, we should have

$$\mathbf{L} = (0, 0, l) \tag{312}$$

But that means that we know all 3 components of $\mathbf{L}$ simultaneously. That can't happen. Therefore, we can't know the direction of the $\mathbf{L}$ vector. That, in turn, means that L_z can never be the as large as $|\mathbf{L}|$.

Now, we are considering a particle that's moving in a rotationally invariant radial potential. That's like Earth moving around the Sun, althogh that is much too much of a classical picture. Let's for the moment use that analogy though. When something is going around something else, the kinetic

energy can be separated into two parts, the linear kinetic energy and the rotational kinetic energy. For instance, think about a particle rotating in a circle. The angular momentum is given by

$$|L| = |r \times p| = mrv \tag{313}$$

where v is the linear speed. Therefore, the kinetic energy can be written as

$$\frac{mv^2}{2} = \frac{L^2}{2I} \tag{314}$$

with $I = mr^2$. In general, one can decompose the momentum into two directions one in the direction of $\mathbf{r}$ and one perpendicular to it.

$$\mathbf{p} = p_r \hat{\mathbf{r}} + \mathbf{p}_l \tag{315}$$

so that

$$\mathbf{p}^2 = p_r^2 + \mathbf{p}_l^2 \tag{316}$$

Now

$$\mathbf{p}_l = \mathbf{p} - \hat{\mathbf{r}}(\hat{\mathbf{r}} \cdot \mathbf{p}) \tag{317}$$

so that

$$\mathbf{p}_l^2 = \mathbf{p}^2 - (\hat{\mathbf{r}} \cdot \mathbf{p})^2 \tag{318}$$

But this is the same as

$$(\hat{r} \times \mathbf{p})^2 = \frac{\mathbf{L}^2}{r^2} \tag{319}$$

Therefore

$$\mathbf{p}^2 = p_r^2 + \frac{L^2}{r^2} \tag{320}$$

and the kinetic energy is

$$\frac{p^2}{2m} = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} \tag{321}$$

and if L is a constant of motion

$$\{L, H\}_{PB} = 0 \tag{322}$$

then the second term is a simple function of r only. Furthermore, in this case, the equation of motion becomes a 1-D equation only in r.

Does this sort of thing also happen in QM? Can we in some sense also decompose the kinetic energy into the radial part and the angular momentum part? The answer is yes.

First of all, what should the radial momentum correspond to? Classically we have

$$p_r = \frac{\mathbf{r}}{r} \cdot \mathbf{p} \tag{323}$$

But quantum mechanically, is it $\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}$ or is it $\hat{\mathbf{p}}\cdot\hat{\mathbf{r}}$? Well, if this is an observable, the operator better be Hermitian. So the natural choice is

$$\hat{p}_r = \frac{1}{2}\left(\frac{\hat{\mathbf{r}}}{r}\cdot\hat{\mathbf{p}} + \hat{\mathbf{p}}\cdot\frac{\hat{\mathbf{r}}}{r}\right) \tag{324}$$

We know

$$\frac{\mathbf{r}}{r}\cdot\hat{p} = \frac{\hbar}{i}\frac{1}{r}(x\partial_x + y\partial_y + z\partial_z) \tag{325}$$

and

$$\begin{aligned}\left(\frac{i}{\hbar}\right)\hat{\mathbf{p}}\cdot\frac{\hat{\mathbf{r}}}{r} &= \partial_x(x/r) + \partial_y(y/r) + \partial_z(z/r) + \partial_r \\ &= \frac{1}{r} - \frac{x^2}{r^3} + \frac{1}{r} - \frac{y^2}{r^3} + \frac{1}{r} - \frac{z^2}{r^3} + \partial_r \\ &= \frac{2}{r} + \partial_r \end{aligned} \tag{326}$$

Therefore

$$p_r = \frac{\hbar}{i}\left(\partial_r + \frac{1}{r}\right) \tag{327}$$

In spherical coordinates,

$$x = r\cos\phi\sin\theta \tag{328}$$

$$y = r\sin\phi\sin\theta \tag{329}$$

$$z = r\cos\theta \tag{330}$$

Consider

$$\begin{aligned}\partial_r &= \frac{\partial x}{\partial_r}\partial_x + \frac{\partial y}{\partial_r}\partial_y + \frac{\partial z}{\partial_r}\partial_z \\ &= \frac{x}{r}\partial_x + \frac{y}{r}\partial_y + \frac{z}{r}\partial_z \end{aligned} \tag{331}$$

Therefore

$$p_r = \frac{\hbar}{i}\frac{1}{r}\frac{\partial}{\partial r}r \tag{332}$$

We also get

$$[r, p_r] = i\hbar \tag{333}$$

and

$$\mathbf{x}\cdot\mathbf{p} = rp_r - \frac{\hbar}{i} \tag{334}$$

and

$$\begin{aligned}(\mathbf{x}\cdot\mathbf{p})^2 &= \left(rp_r - \frac{\hbar}{i}\right)\left(rp_r - \frac{\hbar}{i}\right) \\ &= rp_rrp_r - 2\frac{\hbar}{i}rp_r - \hbar^2 \\ &= r^2p_r^2 + \frac{\hbar}{i}rp_r - 2\frac{\hbar}{i}rp_r - \hbar^2 \\ &= r^2p_r^2 - \frac{\hbar}{i}rp_r - \hbar^2 \end{aligned} \tag{335}$$

Great. Now what? Well, we can do what we did for the classical case. But we must take care that the order is kept.

$$\hat{\mathbf{p}} = \frac{\mathbf{r}}{r}\hat{p}_r + \hat{\mathbf{p}}_l \tag{336}$$

So how do we figure out $\hat{\mathbf{p}}_l$? Now it is clear that

$$\mathbf{r}\times\hat{\mathbf{p}} = \hat{\mathbf{L}} = \mathbf{r}\times\hat{\mathbf{p}}_l \tag{337}$$

So we might as well start with $\mathbf{L}$.

Consider this:

$$\begin{aligned}
\mathbf{L}^2 &= r^2p^2 + i\hbar\left(rp_r - \frac{\hbar}{i}\right) - \left(rp_r - \frac{\hbar}{i}\right)^2 \\
&= r^2p^2 + i\hbar\left(rp_r - \frac{\hbar}{i}\right) - \left(r^2p_r^2 - \frac{\hbar}{i}rp_r - \hbar^2\right) \\
&= r^2p^2 + i\hbar\left(rp_r - \frac{\hbar}{i}\right) - \left(r^2p_r^2 + i\hbar rp_r - \hbar^2\right) \\
&= r^2p^2 - r^2p_r^2 \qquad (338)
\end{aligned}$$

Therefore

$$\begin{aligned}
\langle\mathbf{r}|\hat{H}|\psi\rangle &= \langle\mathbf{r}|\frac{\hat{\mathbf{p}}^2}{2\mu}|\psi\rangle + \langle\mathbf{r}|V(|\hat{\mathbf{r}}|)|\psi\rangle \\
&= -\frac{\hbar^2}{2\mu}\left(\partial_r^2 + \frac{2}{r}\partial_r\right)\langle\mathbf{r}|\psi\rangle + \frac{\langle\mathbf{r}|\hat{\mathbf{L}}^2|\psi\rangle}{2\mu r^2} + V(r)\langle\mathbf{r}|\psi\rangle \\
&= \left(\frac{p_r^2}{2\mu} + \frac{L^2}{2\mu r^2}\right)\langle\mathbf{r}|\psi\rangle + V(r)\langle\mathbf{r}|\psi\rangle \qquad (339)
\end{aligned}$$

Now if we choose ψ to be

$$|\psi\rangle = |E, l, m\rangle \qquad (340)$$

then

$$\begin{aligned}
\langle\mathbf{r}|\hat{H}|E,l,m\rangle &= E\langle\mathbf{r}|E,l,m\rangle \\
&= -\frac{\hbar^2}{2\mu}\left(\partial_r^2 + \frac{2}{r}\partial_r\right)\langle\mathbf{r}|E,l,m\rangle \\
&\quad + \hbar^2\frac{l(l+1)}{\mu r^2}\langle\mathbf{r}|E,l,m\rangle \\
&\quad + V(r)\langle\mathbf{r}|E,l,m\rangle \qquad (341)
\end{aligned}$$

and the equation of of motion becomes that of r only. To be more specific, we let

$$\langle\mathbf{r}|E,l,m\rangle = R(r)\Theta(\theta)\Phi(\phi) \qquad (342)$$

and get

$$-\frac{\hbar^2}{2\mu}\left(\partial_r^2 + \frac{2}{r}\partial_r\right)R(r) + \hbar^2\frac{l(l+1)}{\mu r^2}R(r) + V(r)R(r) = ER(r) \qquad (343)$$

Noting that

$$\begin{aligned} p_r &= \frac{\hbar}{i}\left(\partial_r + \frac{1}{r}\right) \\ &= \frac{\hbar}{i}\left(\frac{1}{r}\partial_r r\right) \end{aligned} \tag{344}$$

Define

$$R(r) = \frac{u(r)}{r} \tag{345}$$

then

$$p_r R(r) = \frac{\hbar}{i}\frac{1}{r}\partial_r(rR) = \frac{\hbar}{i}\frac{1}{r}\partial_r u \tag{346}$$

$$p_r p_r R(r) = -\hbar^2\frac{1}{r}\partial_r\left(r\frac{1}{r}\partial_r u\right) = -\hbar^2\frac{1}{r}\partial_r^2 u \tag{347}$$

so the above equation becomes

$$\left(-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + \hbar^2\frac{l(l+1)}{\mu r^2} + V(r)\right)u(r) = Eu(r) \tag{348}$$

whichis the same form as a 1-D Schrödinger equation with

$$V_{\text{eff}}(r) = \hbar^2\frac{l(l+1)}{\mu r^2} + V(r) \tag{349}$$

8 Vibrations and rotations of a diatomic molecules

What we have so far is enough to consider some interesting real-life phenomena. Consider diatomic molecules such as HCl.

The full Hamiltonian of this system consists of terms for 18 electrons and 2 nuclei. This is complicated enough that it is certain we are not going to get any exact solutions. A great simplification can be made, however, if we consider the fact that the electrons in an atom moves very fast. Remember that in Bohr atom, the momentum is given by (We use the trick of ignoring $\hbar$ and c first)

$$p \sim \frac{1}{a} \tag{350}$$

where $a = 1/(m_e\alpha)$. That means that

$$v = p/m_e \sim \frac{m_e\alpha}{m_e} = \alpha \tag{351}$$

Now this is speed, we better multiply this by c to get

$$v_e = \alpha c \tag{352}$$

That's about 1 percent of speed of light or about a million meters per second.

On the other hand, the nuclei are heavy and slower. Therefore, if we just concentrate on the motions of the nuclei, we may write the effective hamiltionan in the center of mass frame,

$$H_{\rm eff} = \frac{p_r^2}{2\mu} + \frac{L^2}{2mr^2} + V_{\rm eff}(r) \tag{353}$$

where $V_{\rm eff}(r)$ contains the Coulomb force between the two nuclei **and** the averaged effect of the electrons.

What should be the rough form of $V_{\rm eff}(r)$? Well, we know that it should have a minimum so that a stable bound state can be formed. We also know that at large distances the force between two neutral atoms is of the van der Waals type $V \sim 1/r^6$. We also know that we can't bring two nuclei arbitrarily close together. So roughly, we get Figure 9.4.

Now if we are interested only the states near the ground state, we can consider small vibrations around the minimum. In that case, we can expand the potential around the minimum and get the harmonic oscillator potential

$$V(r) \;=\; V(r_0) + \frac{1}{2}(r-r_0)^2 \left.\frac{d^2V}{dr^2}\right|_{r=r_0} + O((r-r_0)^3) \tag{354}$$

which should be valid for small $|r-r_0|/R$ where R is the size of the pocket. Note that since r_0 gives the minimum, the first derivative term is absent.

If this is harmonic oscillator, we would identify

$$V = \frac{1}{2}\xi^2 \left.\frac{d^2V}{dr^2}\right|_{r=r_0} = \frac{\xi^2\mu\omega^2}{2} \tag{355}$$

or

$$\omega^2 = \frac{1}{\mu} \left.\frac{d^2V}{dr^2}\right|_{r=r_0} \tag{356}$$

Now remember that the scale of atomic interaction is set by the Bohr radius (again, ignore $\hbar$ and c first for the estimates)

$$a = \frac{1}{m_e \alpha} \tag{357}$$

so the potential should be of the order

$$V \sim \frac{\alpha}{a} \tag{358}$$

and the second derivative should be of the order

$$V'' \sim \frac{\alpha}{a^3} \tag{359}$$

so

$$\omega^2 \sim \frac{1}{\mu}\frac{\alpha}{a^3} = \frac{m_e^3}{\mu}\alpha^4 \tag{360}$$

But then, the vibrational frequency is the energy so the energy levels are

$$E_n = (n + 1/2)\hbar\omega \tag{361}$$

with

$$\hbar\omega \sim \sqrt{\frac{m_e}{\mu}} m_e c^2 \alpha^2 \tag{362}$$

where we have supplied the necessary factors of $\hbar$ and c^2. Note that the Rydberg constant is $m_e^2\alpha^2/2$. So the above says that the vibrational energy differences of a diatomic molecule is about a factor of $\sqrt{m_e/\mu} \sim 1/40$ smaller than the atomic energy differences. Since the difference is what determines the radiation (or absorption) spectrum, this means that the wavelength of photons coming out of the transition between the vibrational levels are typically of the order

$$\frac{1}{40} 10\text{eV} \approx 0.3\text{eV} \tag{363}$$

Huh? Well, we have to convert this to the length scale. The above means that

$$\hbar c k = pc \sim 0.3\text{eV} \tag{364}$$

NOw

$$k = \frac{2\pi}{\lambda} \tag{365}$$

so

$$\lambda \sim \frac{2\pi\hbar c}{0.3\text{eV}} \approx \frac{6 \times 200\text{eVnm}}{0.3\text{eV}} \approx 4000\text{nm} \tag{366}$$

Optical wavelength are between 400 nm and 700 nm. So this is clearly in the IR range.

To estimate the size of the vibration, we may set the energy equal to the potential energy

$$\omega \sim \frac{\mu\omega^2}{2}\xi^2 \tag{367}$$

so that

$$\begin{aligned} \xi &\sim \sqrt{\frac{1}{\mu\omega}} \sim \sqrt{\frac{1}{\mu(m_e/\mu)^{1/2}m_e\alpha^2}} \sim \sqrt{\frac{1}{(\mu/m_e)^{1/2}m_e^2\alpha^2}} \\ &\sim \left(\frac{m_e}{\mu}\right)^{1/4} \frac{\hbar c}{m_e c^2 \alpha} \end{aligned} \tag{368}$$

where again in the last line, we supplied necessary conversion factors. Since

$$m_e/\mu \sim 1/2000 \tag{369}$$

this is about 10 % of $a = \hbar c/m_e c^2\alpha$.

Now the diatomic molecule can also rotate around the center of mass. Since we are representing the potential between the two nuclei as a harmoinic potential, we can think of the nuclei connected by a spring. Therefore the moment of inertia $I = mr^2$ is not a constant because r^2 can very. But only a small amount. So as a first approximation. But the spring in our case is quite rigid. It only stretches and contracts about 10 % and hence changes I by that amount. In that case, for a first approximation, we may say that the length is fixed and substitute

$$\mu r^2 \to I_0 = \mu r_0^2 \tag{370}$$

In that case, the energy associated with such a motion is

$$H = \frac{L^2}{2I_0} \tag{371}$$

so that the total hamiltonian becomes

$$H = \frac{p_r^2}{2\mu} + V(r) + \frac{L^2}{2I_0} \tag{372}$$

This is an approximation in the sense that the last term is independent of r. In reality it is a function of r so that this clean separation of the vibration and the rotation is not possible.

Given the level of approximation, though, we can now say that the full energy spectrum is something like

$$E_{n,l} = \hbar\omega(n + 1/2) + \frac{l(l+1)\hbar^2}{I_0} \tag{373}$$

The transition between adjacent levels of l gives

$$E_{n,l} - E_{n,l-1} = \frac{\hbar^2(l^2 + l - l^2 + l)}{2I_0} = \hbar^2 \frac{l}{I_0} \tag{374}$$

The energy scale is therefore set by

$$\begin{aligned} \frac{1}{I_0} \sim \frac{1}{\mu a^2} &= \frac{1}{\mu(1/m_e\alpha)^2} \\ &= \alpha^2 m_e \left(\frac{m_e}{\mu}\right) \end{aligned} \tag{375}$$

which is in turn a factor of $(m_e/\mu)^{1/2} \sim 1/40$ smaller than that between the vibrational energy levels. Note that if we can measure

$$E_{n,l} - E_{n,l-1} = \hbar^2 \frac{l}{I_0} \tag{376}$$

then we can deduce I_0 and in turn deduce r_0. This is one way of determining the size of a diatomic molecule.

9 L in the spherical coordinate system

We had no problem specifying a quantum state with the position of a particle so that we can have

$$|\mathbf{r}\rangle \tag{377}$$

because we know that $[x_i, x_j] = 0$.

Formulation of quantum mechanics is easiest in the cartesian coordinates mainly because there is no boundary. Many problems, however, possess symmetries such as the spherical symmetry or the cylindrical symmetry and for those problems, **Schrödinger equation itself is easiest to solve in some curvilinear coordinate system.**

Question is what to do? One can try to directly formulate quantum theory in the new coordinates. However, this presents a problem when the coordinate system is not related to a continuous canonical transformation of the Cartesian coordinates. Potential energy which usually does not involve any derivative is o.k. But the kinetic energy term is a headache. In those cases, it is best to consider the kinetic term as the **Laplacian** and use the appropriate forms of Laplacian in each coordinate system.

How do we do this in general? Well, the basic quantities are the unit vectors. In cartesian coordinate system the unit vectors are

$$e_x = (1, 0, 0) \tag{378}$$
$$e_y = (0, 1, 0) \tag{379}$$
$$e_z = (0, 0, 1) \tag{380}$$

Note that these can be obtained by

$$e_x = \partial_x \mathbf{r} \tag{381}$$
$$e_y = \partial_y \mathbf{r} \tag{382}$$
$$e_z = \partial_z \mathbf{r} \tag{383}$$

The same sort of reasoning applies to any curvilinear coordinate system. For instance, if we have r, θ, ϕ, we get

$$\begin{aligned} \mathbf{g}_r &= \partial_r \mathbf{r} = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta) \\ \mathbf{g}_\theta &= \partial_\theta \mathbf{r} = r(\cos\phi \cos\theta, \sin\phi \cos\theta, -\sin\theta) \\ \mathbf{g}_\phi &= \partial_\phi \mathbf{r} = r(-\sin\phi \sin\theta, \cos\phi \sin\theta, 0) \end{aligned} \tag{384}$$

But these are not normalized. All the difficulties of curvilinear coordinate system stems from this fact that the derivates of parameters are not unit vectors. Note that the derivates themselves communte but the derivatives associated with the normalized unit vectors

$$\mathbf{e}_r = \partial_r \mathbf{r} = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta)$$

$$\begin{aligned}\mathbf{e}_\theta &= \frac{1}{r}\partial_\theta \mathbf{r} = (\cos\phi\cos\theta, \sin\phi\cos\theta, -\sin\theta) \\ \mathbf{e}_\phi &= \frac{1}{r\sin\theta}\partial_\phi \mathbf{r} = (-\sin\phi, \cos\phi, 0)\end{aligned} \tag{385}$$

That is, $\partial_r, \partial_\theta/r, \partial_\phi/(r\sin\theta)$ **do not** commute with each other. That's where all the difficulties come from.

Now consider the gradient. What is a gradient? Well, gradient is something that gives

$$\mathbf{g_i}\cdot\nabla = \partial_i \tag{386}$$

For the spherical coordinates, this means simply

$$\begin{aligned}\nabla &= \frac{\mathbf{g}_r}{|\mathbf{g}_r|^2}\partial_r + \frac{\mathbf{g}_\theta}{|\mathbf{g}_\theta|^2}\partial_\theta + \frac{\mathbf{g}_\phi}{|\mathbf{g}_\phi|^2}\partial_\phi \\ &= \mathbf{e}_r\partial_r + \mathbf{e}_\theta\frac{1}{r}\partial_\theta + \mathbf{e}_\phi\frac{1}{r\sin\theta}\partial_\phi\end{aligned} \tag{387}$$

The angular momentum operator is

$$\mathbf{L} = \mathbf{r}\times\mathbf{p} \tag{388}$$

which involves

$$r\mathbf{e}_r\times\nabla = \mathbf{e}_r\times\left(\mathbf{e}_r\partial_r + \mathbf{e}_\theta\frac{1}{r}\partial_\theta + \mathbf{e}_\phi\frac{1}{r\sin\theta}\partial_\phi\right) \tag{389}$$

r, θ, ϕ satisfies the right hand rule in that order. So

$$r\mathbf{e}_r\times\nabla = \left(\mathbf{e}_\phi\partial_\theta - \mathbf{e}_\theta\frac{1}{\sin\theta}\partial_\phi\right) \tag{390}$$

which is now totally devoid of any factors of r or ∂_r.

Taking x and y component yields

$$L_x = \frac{\hbar}{i}\left(-\sin\phi\partial_\theta - \cot\theta\cos\phi\partial_\phi\right) \tag{391}$$

$$L_y = \frac{\hbar}{i}\left(\cos\phi\partial_\theta - \cot\theta\sin\phi\partial_\phi\right) \tag{392}$$

and

$$L^2 = -\hbar^2\left[\frac{1}{\sin\theta}\partial_\theta(\sin\theta\partial_\theta) + \frac{1}{\sin^2\theta}\partial_\phi^2\right] \tag{393}$$

10 Orbital angular momentum eigenfunctions

The reason that we went thru the above exercise is to write the form of the kinetic term explicitly. Remember that the wavefunction squared

$$P(\mathbf{r}) = |\langle \mathbf{r}|\psi\rangle|^2 \tag{394}$$

has the interpretation of the probability density for the position $\mathbf{r}$. Therefore, knowing the eigenvalues of each operators is not enough. We must figure out the form of the wavefunction to completely solve the problem.

The short term goal of this course is to solve the hydrogen atom problem. In those central force problems, the space is isotropic and hence the angular momentum is conserved. Also it naturally permits the separation of variables

$$\psi(x) = \langle x|\psi\rangle = R(r)Y_{l,m}(\theta,\phi) \tag{395}$$

The form of L^2 above further permits the separation

$$Y_{l,m}(\theta,\phi) = \Theta_{l,m}(\theta)\Phi_m(\phi) \tag{396}$$

The normalization condition

$$\int d^3r|\psi|^2 = 1 \tag{397}$$

is then separated into

$$\int_0^\infty drr^2\,|R(r)|^2 = 1 \tag{398}$$

and

$$\int d\Omega\,|Y_{l,m}(\theta,\phi)|^2 = 1 \tag{399}$$

Figuring out the form of $Y_{l,m}$ is easiest done by using the rasing and lowering operators

$$L_\pm = L_x \pm iL_y = \frac{\hbar}{i}e^{\pm i\phi}\left(\pm i\partial_\theta - \cot\theta\partial_\phi\right) \tag{400}$$

We know that

$$L_+Y_{l,l} = 0 \tag{401}$$

and $Y_{l,l} = e^{il\phi}\Theta_{l,l}(\theta)$ or

$$\begin{aligned} 0 &= (i\partial_\theta - \cot\theta\partial_\phi)\, Y_{l,l} \\ &= e^{il\phi} i\,(\partial_\theta\Theta - l\cot\theta\Theta) \end{aligned} \tag{402}$$

which is solved by

$$\Theta_{l,l}(\theta) = A_{l,l}\sin^l\theta \tag{403}$$

so

$$Y_{l,l} = A_{l,l}e^{il\phi}\sin^l\theta \tag{404}$$

All other $Y_{l,m}$ can be now obtained by applying L_- and remembering that

$$L_-|l,m\rangle = \sqrt{l(l+1) - m(m-1)}\hbar|l,m-1\rangle \tag{405}$$

and

$$\int d\Omega|Y_{l,m}| = 1 \tag{406}$$

A few of these are

$$\begin{aligned} Y_{0,0} &= \sqrt{\frac{1}{4\pi}} \\ Y_{1,\pm1} &= \mp\sqrt{\frac{3}{8\pi}}e^{\pm i\phi}\sin\theta \propto (x \pm iy)/r \\ Y_{1,0} &= \sqrt{\frac{3}{4\pi}}\cos\theta \propto z/r \\ Y_{2,\pm2} &= \sqrt{\frac{15}{32\pi}}e^{\pm 2i\phi}\sin^2\theta \propto (x \pm iy)^2/r^2 \\ Y_{2,\pm1} &= \mp\sqrt{\frac{15}{8\pi}}e^{\pm i\phi}\sin\theta\cos\theta = (x \pm iy)z/r^2 \\ Y_{2,0} &= \sqrt{\frac{5}{16\pi}}(3\cos^2\theta - 1) \propto (3z^2 - r^2)/r^2 = (2z^2 - x^2 - y^2)/r^2 \end{aligned} \tag{407}$$

How do we visualize something like this? For instance, consider $|Y_{1,0}(\phi,\theta)|^2 \propto \cos^2\theta$. This is probability density for the solid angle pointing in the direction

of ϕ, θ. Therefore, the probability is maximum when $\theta = 0, \pi$ and minimum when $\theta = \pi/2$. So one gets kind of a dumbell shape shown in Fig.9.11 and so on.

O.K. Now its time to take a breath and ask, what the hell do all these math mean? Why do these shapes make sense? First, consider $Y_{0,0}$. This is constant. So the probability to find the electron in any direction is the same. This make sense. If there is no angular momentum, there is no preferred direction.

What about $Y_{l,l}$ state? What do we expect? Well, this state has maximum L_z component. That means it has the maximum momentum of inertia $I = m\langle r_\perp^2 \rangle$ associated with it. Since the mass is fixed, the only way to achieve this is to have largest $\langle r_\perp^2 \rangle = \langle x^2 + y^2 \rangle$. The best way to achieve that is either to be a dumbell or a doughnut. We can't have dumbell due to rotational invariance. So the doughnut it is and doughnut it gets.

Now to reduce m, we have to reduce $\langle x^2 + y^2 \rangle$. Since the volume is fixed by the normalization condition, the only way to achieve this is to spread in the z direction. So the doughnut splits into two and then separates in z direction. This process continues. Now consider the $m = 0$ part. This is always given by the Legendre polynomial $P_l(\cos\theta)$ which always have maximum at $\theta = 0, \pi$,

$$|P_l(\pm 1)| = 1 \tag{408}$$

So $Y_{l,0}$ has a dumbell part. So this happens: Whenever m is reduced by 1, one more doughnut appears and then whole thing spreads in z direction some more until finally $Y_{l,0}$ has $l - 1$ doughnuts and 2 lobes and it is spread thinly along the z axis to minimize $x^2 + y^2$.

Note that for $l = 1$ state, we can write

$$Y_{1,\pm 1} = \mp\sqrt{\frac{3}{3\pi}}(x \pm iy)/r \tag{409}$$

so that we can have a linear combination

$$(Y_{1,-1} - Y_{1,1})/\sqrt{2} = \sqrt{\frac{3}{4\pi}}x/r \tag{410}$$

$$i(Y_{1,-1} + Y_{1,1})/\sqrt{2} = \sqrt{\frac{3}{4\pi}}y/r \tag{411}$$

together with

$$Y_{1,0} = \sqrt{\frac{3}{4\pi}}z/r \tag{412}$$

For the reason of symmetry, this, of course, must happen.

This means that when electrons fill the p state, or $l = 1$ state, they arrange themselves along some imaginary Cartesian axis. For instance, consider water H_2O. Oxygen has 8 electrons. Now, the inner shell $1s$ state gets 2 electrons. And then $2s$ state gets 2 electrons. The rest of 4 electrons goes to the $2p$ state and spread among the 6 lobes. This presents imbalance. Naturally, the system would like to fill it with 2 more electrons. Now 2 hydrogen comes along. H has only 1 electron in the $1s$ orbit and it can either lose it or gain another. When it encounters oxygen, it is easier for the electrons in the hydrogens to migrate over to the oxygen then 2 of the electrons in the oxygen to come over to the hydrogen which would further frustrate the symmetry. So the electrons in the hydrogen goes to the remaining 2 $2p$ states. Now these are along the Cartesian axes. Therefore the oxygen and the two hydrogen must be make an angle of about 90 degree. Coulomb force between the protons tends to push things a little and eventually ends up makeing 105 degree. A lot of chemistry can be explained in this line of argument.

Another easy example is Ammonia, NH_3. The nitrogen has 7 electrons. So it can take 3 hydrogen atoms which should all make 90 degree angles. Again, Coulomb pushes them apart a little but the basic shape remains.

10.1 Aside – Uncertainty principle in an absolutely confined space

The uncertainty principle states

$$\langle \Delta x \rangle \langle \Delta p \rangle \geq \hbar/2 \tag{413}$$

Consider a potential well with infintely high walls. In that case, the space is defined by the interval

$$x = (0, L) \tag{414}$$

Consider the ground state of this system:

$$\psi_1(x) = A \sin(\pi x/L) \tag{415}$$

Calculate

$$\langle p \rangle = \frac{A^2 \int_0^L dx \, \sin(\pi x/L)(-i\hbar)(\pi/L)\cos(\pi x/L)}{A^2 \int_0^L dx \, \sin^2(\pi x/L)} = 0 \tag{416}$$

$$\langle p^2 \rangle = \hbar^2 \frac{\pi^2}{L^2} \tag{417}$$

11 Townsend Chapter 10 – Bound state

Previously, we saw that if you have a central force potential $V(r)$ which depends only on the radius r, then the Schroedinger equation can be written in the following form

$$\langle \mathbf{r}|E, l, m\rangle = R_{E,l}(r)Y_{lm}(\theta, \phi) \tag{418}$$

and the radial part of the Schrodinger equation becomes, with $R = u/r$,

$$\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + \frac{l(l+1)\hbar^2}{2\mu r^2} + V(r)\right] u_{E,l}(r) = E u_{E,l}(r) \tag{419}$$

with the normalization condition

$$\int_0^\infty dr\, r^2\, R^*_{E,l}(r)R_{E,l}(r) = \int_0^\infty dr\, u^*_{E,l}(r)u_{E,l}(r) = 1 \tag{420}$$

The above differential equation has regular singularity at $r = 0$ if $V(r)$ does not diverge faster than $1/r^2$ at the origin. In this case, it is guaranteed that a series solution of the form

$$u = r^s \sum_{n=0}^{\infty} a_n r^n \tag{421}$$

exists. However not all solutions obtained in this way will be **physical**. By that we mean that first of all, the normalization

$$\int dr u^* u = 1 \tag{422}$$

is possible and second that the wave function never blows up anywhere.

So let's substitute this into the radial equation and see what kind of conditions we can work out. Use

$$u' = \sum_{n=0}^{\infty} a_n (n+s) r^{n+s-1} \tag{423}$$

$$u'' = \sum_{n=0}^{\infty} a_n (n+s)(n+s-1) r^{n+s-2} \tag{424}$$

then

$$-\frac{\hbar^2}{2\mu}\sum_{n=0}^{\infty}a_n(n+s)(n+s-1)r^{n+s-2}+\frac{l(l+1)\hbar^2}{2\mu}\sum_{n=0}^{\infty}a_n r^{n+s-2}+V(r)\sum_{n=0}^{\infty}a_n r^{n+s}=E\sum_{n=0}^{\infty}a_n r^{n+s} \tag{425}$$

We are suppose to match the coefficients of each power of r. Now if $V(r)$ is less singular than $1/r^2$ so that

$$\lim_{r\to 0} r^2 V(r) = 0 \tag{426}$$

then the lowest power of r is r^{s-2}. Matching the coefficients gives the condition called **indicial equation**

$$[s(s-1)-l(l+1)]a_0 = 0 \tag{427}$$

Therefore either $a_0 = 0$ or

$$s = l+1 \tag{428}$$

$$s = -l \tag{429}$$

So we have two forms of solutions

$$u_+ = r^{l+1}\sum_{n=0}^{\infty}a_n r^n \tag{430}$$

$$u_- = r^{-l}\sum_{n=0}^{\infty}a'_n r^n \tag{431}$$

with the corresponding **wavefunction**

$$R_+ = r^{l}\sum_{n=0}^{\infty}a_n r^n \tag{432}$$

$$R_- = r^{-l-1}\sum_{n=0}^{\infty}a'_n r^n \tag{433}$$

Obviously, even for $l = 0$, R_{-l} blows up. That's bad. So the physical solution must be

$$u_+ = r^{l+1}\sum_{n=0}^{\infty}a_n r^n \tag{434}$$

and corresponding

$$R_+ = r^l \sum_{n=0}^{\infty} a_n r^n \tag{435}$$

So as $r \to 0$

$$u_{E,l} \to r^{l+1} \to 0 \tag{436}$$

and

$$R_{E,l} \to r^l \to \begin{cases} 0 & \text{for } l \neq 0 \\ 1 & \text{for } l = 0 \end{cases} \tag{437}$$

Note that as l increases, there probability to find the particle near origin becomes smaller and smaller. This can be understood in two ways. Formally, take a look at the effective potential

$$V_{\text{eff}} = \frac{l(l+1)\hbar^2}{2\mu r^2} + V(r) \tag{438}$$

As l increases, the strength of **the centrifugal barrier** increases. So the classical turning point gets further away from the origin (Fig 10.1). Physically, this is what one should expect. Angular momentum can increase either by increasing the angular speed or the moment of inertia. To increase the moment of inertia, the average distance between the particle and the origin must increase. In particular, it should not come near the origin where I vanishes.

12 Hydrogen like atom

Before we charge ahead and solve this problem, it is very instructive to consider the original Bohr atom. Bohr, inspired by the de Broglie wave, postulated that due to the wave nature of the electron, only a certain orbits are allowed. These orbits has to represent **standing waves** or the circumference of the orbit must be an integer multiple of the electron's wavelength.

Now we know from the relation

$$\hbar k = \frac{\hbar 2\pi}{\lambda} = p = mv \tag{439}$$

$$\lambda = \frac{h}{mv} \tag{440}$$

Assuming a circular orbit, the circumference of the orbit is

$$2\pi R \tag{441}$$

But this is related to the speed by the centrifugal force equation

$$\frac{mv^2}{R} = \frac{Ze^2}{R^2} \tag{442}$$

or

$$R = \frac{Ze^2}{mv^2} \tag{443}$$

so

$$\frac{2\pi R}{\lambda} = n \tag{444}$$

yields

$$\frac{Ze^2 2\pi}{mv^2}\frac{mv}{h} = \frac{Ze^2}{v\hbar} = n \tag{445}$$

or

$$v_n = \frac{Ze^2}{n\hbar} \tag{446}$$

Now the total energy of this orbit is given by

$$\begin{aligned} E_n &= \frac{mv_n^2}{2} - \frac{Ze^2}{R_n} \\ &= \frac{mv_n^2}{2} - mv_n^2 \\ &= -\frac{mv_n^2}{2} \end{aligned} \tag{447}$$

again using the centrifugal force equation. Hence

$$E_n = -m\frac{Z^2e^4}{2\hbar^2 n^2} = -(mc^2)\frac{Z^2\alpha^2}{2\,n^2} \tag{448}$$

This happens to be exact.

What have we learned here? Of course, this is a happy accident. But often times, this sort of semi classical analysis is very useful to figure out the size of the effect. In this case, we learn that:

1. The dicrete energy levels are again related to the fact that the electrons are confined. Hence, some sort of **stationary** wave must be established. This in turn implies that only a certain discrete values of momentum and energy are allowed.

2. The size of the atomic energy should be

$$E \sim \alpha^2 mc^2 = O(10\,\mathrm{eV}) \tag{449}$$

3. The size of the atomic length should be

$$a_0 \sim R = \frac{e^2}{E} \sim \frac{\hbar c}{m\alpha} \tag{450}$$

What we **don't learn** are

1. Is $n = 0$ allowed?

2. What about elliptic orbits?

3. where exactly is the electron?

Some of these questions were taken up by Sommerfeld in 1916 still without the help of full quantum mechanics. It goes like this.

Let's generalize what Bohr said about the relation between the momentum and the circumference. Bohr has

$$\frac{2\pi R}{\lambda} = n \tag{451}$$

But since

$$p = \hbar k = \frac{h}{\lambda} \tag{452}$$

or

$$\lambda = \frac{h}{p} \tag{453}$$

we can also say

$$\frac{2\pi Rp}{h} = n \tag{454}$$

In other words,

$$2\pi Rp = nh \tag{455}$$

Now since we are talking about circular orbit, there is no radial motion and the momentum above is actually better represented by the angular momentum

$$L = |r \times p| = Rp \tag{456}$$

so that we can say

$$J_\phi \equiv \oint L d\phi = nh \tag{457}$$

If we have radial motion, however, the radial momentum $p_r = m\dot{r}$ is non-zero. And if electron is really represented by a wave, there must be a wave associated with this motion, too. But what's the analogous rule for the **standing wave** in this case? First of all, the orbit must be periodic. That means that when the particle comes around to the same spot, the phase of the wave must be the same as before (that is, only $2n\pi$ changes are allowed).

The phase of a wave is again determined by the wave form $e^{\frac{ipx}{\hbar}}$. If p is not constant, we should generalize this to

$$\psi \sim \exp\left(i \int p dx/\hbar\right) \tag{458}$$

the requirement that when the electron must have the same phase at the same spot regardless of the time and the position of the spot then reads to the condition

$$\oint \mathbf{p}\cdot d\mathbf{x}/\hbar = 2n\pi \tag{459}$$

or

$$\oint \mathbf{p}\cdot d\mathbf{x} = nh \tag{460}$$

In the spherical coordinate

$$\mathbf{p}\cdot d\mathbf{x} = p_r dr + p_\phi d\phi + p_\theta d\theta \tag{461}$$

In the case of central force potential, we can eliminate one angle and get

$$\mathbf{p}\cdot d\mathbf{x} = p_r dr + L d\phi \tag{462}$$

where L is a conserved angular momentum. So the single valued-ness condition leads to

$$J_r \equiv \oint p_r dr = n_r h \tag{463}$$

$$J_\phi \equiv \oint L d\phi = n_\phi h = 2L\pi \tag{464}$$

Well, there is no rigorous reason that the n_r and n_ϕ above both are integers. All we require is that $n_r + n_\phi = n$. However, it is unnatural to do so and it doesn't reproduce experimental result either.

These are the celebrated Bohr-Sommerfeld quantization rules of old quantum mechanics. To evaluate J_r requires some knowledge of classical Kepler problem. Here, I'll just list the solution.

$$r = \frac{L^2/me^2}{1+\epsilon\cos\phi} \tag{465}$$

with

$$\epsilon^2 = 1 + \frac{2L^2}{me^4}E \tag{466}$$

To use it we write

$$\begin{aligned} p_r dr &= m\frac{dr}{dt}\frac{dr}{d\phi}d\phi \\ &= m\frac{dr}{d\phi}\dot{\phi}\frac{dr}{d\phi}d\phi \\ &= m\left(\frac{dr}{d\phi}\right)^2\frac{L}{mr^2}d\phi \\ &= L\left(\frac{1}{r}\frac{dr}{d\phi}\right)^2 d\phi \end{aligned} \tag{467}$$

$$\begin{aligned} \frac{1}{r}\frac{dr}{d\phi} = \frac{d\ln r}{d\phi} &= -\frac{d}{d\phi}\ln(1+\epsilon\cos\phi) \\ &= \frac{\epsilon\sin\phi}{1+\epsilon\cos\phi} \end{aligned} \tag{468}$$

so that

$$\begin{aligned}\oint p_r dr &= L\int_0^{2\pi}\left(\frac{1}{r}\frac{dr}{d\phi}\right)^2 d\phi \\ &= L\int_0^{2\pi}\frac{\epsilon^2\sin^2\phi}{(1+\epsilon\cos\phi)^2}d\phi \end{aligned} \tag{469}$$

Let

$$u = \epsilon\sin\phi \tag{470}$$

$$dv = d\left(\frac{1}{1+\epsilon\cos\phi}\right) = \frac{\epsilon\sin\phi}{(1+\epsilon\cos\phi)^2}d\phi \tag{471}$$

so that

$$\begin{aligned}\oint p_r dr &= L\left(\frac{\epsilon\sin\phi}{1+\epsilon\cos\phi}\bigg|_0^{2\pi} - \int_0^{2\pi} d\phi\,\frac{\epsilon\cos\phi}{1+\epsilon\cos\phi}\right) \\ &= -L\int_0^{2\pi} d\phi\,\frac{\epsilon\cos\phi}{1+\epsilon\cos\phi} \\ &= -L\int_0^{2\pi} d\phi\,\frac{\epsilon\cos\phi+1-1}{1+\epsilon\cos\phi} \\ &= -L\left(2\pi - \int_0^{2\pi} d\phi\,\frac{1}{1+\epsilon\cos\phi}\right) \end{aligned} \tag{472}$$

Let

$$I = \int_0^{2\pi} d\phi\,\frac{1}{1+\epsilon\cos\phi} \tag{473}$$

Use

$$z = e^{i\phi} \tag{474}$$

to get

$$\begin{aligned} I &= \oint\frac{dz}{iz}\frac{1}{1+\epsilon(z+1/z)/2} \\ &= 2\pi\oint\frac{dz}{2\pi i}\frac{1}{(z+\epsilon(z^2+1)/2)} \\ &= \frac{4\pi}{\epsilon}\oint\frac{dz}{2\pi i}\frac{1}{(z^2+2z/\epsilon+1)} \\ &= \frac{4\pi}{\epsilon}\oint\frac{dz}{2\pi i}\frac{1}{\left[z-(-1-\sqrt{1-\epsilon})/\epsilon\right]\left[z-(-1+\sqrt{1-\epsilon})/\epsilon\right]} \end{aligned} \tag{475}$$

The integrand has poles at

$$z = -\frac{1}{\epsilon} \pm \frac{1}{\epsilon}\sqrt{1-\epsilon^2} \tag{476}$$

For $\epsilon < 1$, only the pole

$$z = -(1-\sqrt{1-\epsilon^2})/\epsilon \tag{477}$$

is in the contour. So

$$\begin{aligned} I &= \frac{4\pi}{\epsilon}\frac{1}{2\sqrt{1-\epsilon^2}/\epsilon} \\ &= \frac{2\pi}{\sqrt{1-\epsilon^2}} \end{aligned} \tag{478}$$

So all together

$$\begin{aligned} \oint p_r dr &= -L\left(2\pi - \int_0^{2\pi} d\phi\, \frac{1}{1+\epsilon\cos\phi}\right) \\ &= 2\pi L\left(\frac{1}{\sqrt{1-\epsilon^2}} - 1\right) = n_r h \end{aligned} \tag{479}$$

Remember that

$$1-\epsilon^2 = -\frac{2L^2}{me^4}E = \frac{2L^2}{me^4}|E| \tag{480}$$

so

$$n_r h = 2\pi L\left(\sqrt{\frac{me^4}{2L^2|E|}} - 1\right) \tag{481}$$

$$n_r h + 2\pi L = 2\pi\sqrt{\frac{me^4}{2|E|}} \tag{482}$$

$$E = -|E| = -\frac{me^4}{2\hbar^2(n_r+n_\phi)^2} \tag{483}$$

Now we are finally ready to tackle the Hydrogen atom. The Hamiltonian for the electron is

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2\mu} - \frac{Ze^2}{|\hat{\mathbf{r}}|} \tag{484}$$

where

$$\mu = \frac{mM}{m+M} \approx m \tag{485}$$

is the reduced mass.

We have put a factor Z in the potential term to deal with ions such as He^+ and Li^{++} as well.

The radial equation is

$$\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + \frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{Ze^2}{r}\right] u_{E,l}(r) = E u_{E,l}(r) \tag{486}$$

Now since we are interested in the bound state, it is convenient to put

$$E = -|E| \tag{487}$$

so that

$$\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + \frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{Ze^2}{r}\right] u_{E,l}(r) = -|E| u_{E,l}(r) \tag{488}$$

We would like to reduce this equation to one of the known differential equations. Now carrying around the dimensionful quantities $\hbar$ and μ and e and E then becomes nuisance.

To write the above in a more managable form, we use the fact that $\hbar c$ converts energy to length and length to energy. Let's convert everything to energy units then. To do so, we first multiply and divide by c^2 and get

$$\left[-\frac{1}{2\mu c^2}\frac{d^2}{dx^2} + \frac{l(l+1)}{2\mu c^2 x^2} - \frac{Z\alpha}{x}\right] u_{E,l}(x) = -|E| u_{E,l}(x) \tag{489}$$

where $x = r/\hbar c$ which has the dimension of 1/energy and we used $\alpha = e^2/\hbar c$. Now everything is energy. To further reduce, we divide the equation by $|E|$ to get

$$\left[-\frac{1}{2\mu c^2|E|}\frac{d^2}{dx^2} + \frac{l(l+1)}{2\mu c^2|E| x^2} - \frac{Z\alpha}{x|E|}\right] u_{E,l}(x) = -u_{E,l}(x) \tag{490}$$

The combinations $\mu c^2|E|x^2$ and $|E|x$ are now dimensionless. So this suggest that we should change variable to

$$\eta = \sqrt{\mu c^2|E|}x \tag{491}$$

The change $\eta = |E|x$ should also work but we prefer the constant term to have a pure number. This leads to

$$\left[-\frac{d^2}{d\eta^2} + \frac{l(l+1)}{\eta^2} - 2\sqrt{\frac{\mu c^2}{|E|}}\frac{Z\alpha}{\eta}\right] u_{E,l}(\eta) = -2u_{E,l}(\eta) \tag{492}$$

We are alomst there now. Everything you see now in this equation is dimensionless. You can now look up the table of differential equations and find the solution for this equation. The book does one more thing to make the equation to look even nicer. Note that as $\eta \to \infty$, the only terms that matter are the derivative term and the energy term so that from above we get as $\eta \to \infty$,

$$\frac{d^2u}{d\eta^2} \approx 2u \tag{493}$$

which has the solutions

$$u = Ae^{\sqrt{2}\eta} + Be^{-\sqrt{2}\eta} \tag{494}$$

Nothing wrong with that, but if we keep this around then we are going to have to deal with appearance of $\sqrt{2}$ in the equations. Again, nothing wrong with that, but yech, it's ugly. Also, the above tells you that we should discard the positive exponential solution because it is not normalizable.

So, let's make it nicer. An obvious thing to do is to redefine

$$\xi = \sqrt{2}\eta \tag{495}$$

to get

$$\left[-\frac{d^2}{d\xi^2} + \frac{l(l+1)}{\xi^2} - \sqrt{\frac{2\mu c^2}{|E|}}\frac{Z\alpha}{\xi}\right] u_{E,l}(\eta) = -u_{E,l}(\eta) \tag{496}$$

And in fact, many textbook uses this form of radial equation.

In our book, we redefine

$$\rho = 2\xi \tag{497}$$

so that we finally have

$$\frac{d^2u}{d\rho^2} - \frac{l(l+1)}{\rho^2}u + \left(\frac{\lambda}{\rho} - \frac{1}{4}\right)u = 0 \tag{498}$$

with $\lambda = Z\alpha\sqrt{\frac{\mu c^2}{2|E|}}$ and

$$\rho = \sqrt{\frac{8\mu|E|}{\hbar^2}}r \tag{499}$$

Rederivation Starts

Goal: To make everything dimensionless and $O(1)$ except the energy.

$$\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + \frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{Ze^2}{r}\right]u_{E,l}(r) = -|E|u_{E,l}(r) \tag{500}$$

We note that

$$\left[\frac{\hbar^2}{2\mu}\right] = [E][L]^2 \tag{501}$$

and

$$\left[Ze^2\right] = [E][L] \tag{502}$$

Hence

$$\begin{aligned} \xi_0 &= \frac{(\hbar^2/2\mu)}{Ze^2} \\ &= \frac{\hbar(\hbar c)}{2\mu Ze^2 c} \\ &= \frac{\hbar}{2\mu c Z\alpha_{EM}} \end{aligned} \tag{503}$$

has a length unit. We can use it to make everything dimensionless. To do that, we first divide by $\hbar^2/2\mu$ to get

$$\left[-\frac{d^2}{dr^2}+\frac{l(l+1)}{r^2}-\frac{1}{\xi_0\, r}\right]u=-\frac{2\mu|E|}{\hbar^2}u \tag{504}$$

Then multiply by ξ_0 to get

$$\left[-\frac{d^2}{d\eta^2}+\frac{l(l+1)}{\eta^2}-\frac{1}{\eta}\right]u=-\epsilon u \tag{505}$$

where $\epsilon \equiv \frac{2\mu|E|\xi_0^2}{\hbar^2}$ and $\eta \equiv r/\xi_0$ is now a dimensionless variable.

Now the LHS is all $O(1)$ and dimensionless. So $\epsilon = O(1)$ as well. This also indicates that ξ_0 is a characteristic length scale.

That means

$$|E| \sim \frac{\hbar^2}{\mu\xi_0^2} \sim \frac{\hbar^2}{\mu}\frac{Z^2\alpha_{EM}^2\mu^2c^2}{\hbar^2} \sim Z^2\alpha_{EM}^2\mu c^2 \tag{506}$$

as before.

One can now use this form to calculate u as a function of η. However, our book goes one step further and divides the whole thing by 4ϵ to get

$$\frac{d^2u}{d\rho^2}-\frac{l(l+1)}{\rho^2}u+\left(\frac{\lambda}{\rho}-\frac{1}{4}\right)u=0 \tag{507}$$

with $\rho = 2\sqrt{\epsilon}\eta$ and $\lambda = 1/2\sqrt{\epsilon}$.

Rederivation Ends

Why this form? Well, as far as I can see the only advantage is to have the square of the wavefunction

$$u^*u \propto e^{-\rho} \tag{508}$$

instead of $e^{-2\xi}$.

O.K. So we have a nice looking equation and we know that

$$u_{E,l} \to \rho^{l+1} \quad \text{as} \quad \rho \to 0 \tag{509}$$

$$u_{E,l} \to e^{-\rho/2} \quad \text{as} \quad \rho \to \infty \tag{510}$$

This suggests that we should try

$$u_{E,l} = e^{-\rho/2}\rho^{l+1}F(\rho) \tag{511}$$

where

$$F(\rho) = \sum_{n=0}^{\infty} c_k \rho^k \tag{512}$$

We have

$$u' = -\frac{1}{2}e^{-\rho/2}\rho^{l+1}F + (l+1)e^{-\rho/2}\rho^l F + e^{-\rho/2}\rho^{l+1}F' \tag{513}$$

and

$$\begin{aligned} u'' &= \frac{1}{4}e^{-\rho/2}\rho^{l+1}F - \frac{1}{2}(l+1)e^{-\rho/2}\rho^l F - \frac{1}{2}e^{-\rho/2}\rho^{l+1}F' \\ &\quad - \frac{1}{2}(l+1)e^{-\rho/2}\rho^l F + (l+1)le^{-\rho/2}\rho^{l-1}F + (l+1)e^{-\rho/2}\rho^l F' \\ &\quad - \frac{1}{2}e^{-\rho/2}\rho^{l+1}F' + (l+1)e^{-\rho/2}\rho^l F' + e^{-\rho/2}\rho^{l+1}F'' \\ &= \frac{1}{4}u + \frac{l(l+1)}{\rho^2}u - (l+1)e^{-\rho/2}\rho^l F \\ &\quad - e^{-\rho/2}\rho^{l+1}F' + 2(l+1)e^{-\rho/2}\rho^l F' \\ &\quad + e^{-\rho/2}\rho^{l+1}F'' \end{aligned} \tag{514}$$

Or

$$\begin{aligned} u'' - \frac{1}{4}u - \frac{l(l+1)}{\rho^2}u &= -(l+1)e^{-\rho/2}\rho^l F - e^{-\rho/2}\rho^{l+1}F' \\ &\quad + 2(l+1)e^{-\rho/2}\rho^l F' + e^{-\rho/2}\rho^{l+1}F'' \\ &= -\frac{\lambda}{\rho}e^{-\rho/2}\rho^{l+1}F \end{aligned} \tag{515}$$

Collecting, we get an equation for F

$$0 = -\frac{(l+1)}{\rho}F - F' + \frac{2(l+1)}{\rho}F' + F'' + \frac{\lambda}{\rho}F \tag{516}$$

$$\rho F'' + [(2l+2) - \rho]F' - [(l+1) - \lambda]F = 0 \tag{517}$$

This is in the form of hypergeometric equation

$$\rho F'' + [\beta - \rho] F' - \alpha F = 0 \tag{518}$$

Let

$$F = \sum_{k=0}^{\infty} c_k \rho^k \tag{519}$$

and substitute to get

$$\begin{aligned} 0 &= \rho \sum_{k=2}^{\infty} c_k k(k-1)\rho^{k-2} + \beta \sum_{k=1}^{\infty} c_k k \rho^{k-1} - \sum_{k=1}^{\infty} c_k k \rho^k - \alpha \sum_{k=0}^{\infty} c_k \rho^k \\ &= \sum_{k=1}^{\infty} c_{k+1}(k+1)k\rho^k + \beta \sum_{k=0}^{\infty} c_{k+1}(k+1)\rho^k - \sum_{k=1}^{\infty} c_k k \rho^k - \alpha \sum_{k=0}^{\infty} c_k \rho^k \end{aligned} \tag{520}$$

For $k = 0$, we have

$$\beta c_1 - \alpha c_0 = 0 \tag{521}$$

For all others we have

$$c_{k+1}k(k+1) + \beta c_{k+1}(k+1) - c_k k - \alpha c_k = 0 \tag{522}$$

or

$$c_{k+1} = c_k \frac{(k+\alpha)}{(k+\beta)(k+1)} \tag{523}$$

So

$$\begin{aligned} c_1 &= \frac{\alpha}{\beta} c_0 \\ c_2 &= \frac{(1+\alpha)}{(1+\beta)2}\frac{\alpha}{\beta} c_0 \\ c_3 &= \frac{(2+\alpha)}{(2+\beta)3}\frac{(1+\alpha)}{(1+\beta)2}\frac{\alpha}{\beta} c_0 = \frac{\alpha(\alpha+1)(\alpha+2)}{\beta(\beta+1)(\beta+2)3!} c_0 \\ &\vdots \\ c_n &= \frac{\alpha(\alpha+1)\cdots(\alpha+n-1)}{n!\,\beta(\beta+1)\cdots(\beta+n-1)} c_0 \end{aligned} \tag{524}$$

For large n, we have

$$\frac{c_{n+1}}{c_n} = \frac{(\alpha+n)}{n(\beta+n)} \to \frac{1}{n} \tag{525}$$

which is how the coeffcients of the exponential behaves. So for large n, the series behave like

$$F \sim e^{\rho} \tag{526}$$

This is not O.K. since in this case u is not normalizable. If this is not to happen, then the series must terminate at a finite k. This can happen only if

$$\alpha = -n_r \tag{527}$$

where n_r is a non-negative integer. In our case this reads

$$(l+1) - \lambda = -n_r \tag{528}$$

or

$$\lambda = n_r + (l+1) = n \tag{529}$$

where $n = 1, 2, ...$ Now since

$$\lambda = Z\alpha\sqrt{\frac{\mu c^2}{2|E|}} \tag{530}$$

this means

$$n^2 = Z^2\alpha^2\frac{\mu c^2}{2|E|} \tag{531}$$

or

$$|E| = \frac{Z^2\alpha^2\mu c^2}{2n^2} \tag{532}$$

so

$$E_n = -\frac{Z^2\alpha^2\mu c^2}{2n^2} = -\frac{13.6\,\mathrm{eV}}{n^2}Z^2 \tag{533}$$

This n is defined as the **principal quantum number**.

> More: Remember that the principal quantum number n is made up of n_r and l or $n = n_r+l+1$. So the same energy can result from no radial excitation $n_r = 0$ and all angular excitation $l = n-1$ or all radial excitation $n_r = n-1$ and no angular excitation $l = 0$. The former correspond to a circular orbit. In this case, the radius is constant. So there is no momenum component attached to the changing of the radius. Also, the electron cannot go near the center.
>
> The latter correspond to the maximally elliptic orbit. A highly elliptic orbit has big $r_{\max}/r_{\min}$ ratio. That means the radius changes a lot. Maximally elliptic orbit has no angular momentum. Therefore, this orbit is the only one that is allowed to go near the nucleus. This is called the S-state. Classically, this is NOT allowed because this would a straight line orbit that goes thru the nucleus.

When a hydrogen atom makes a transition between two energy levels, the energy difference is

$$\hbar\omega = hf = E_{n_i} - E_{n_f} = \frac{\mu c^2\alpha^2}{2}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \tag{534}$$

and the wave length

$$\begin{aligned}\frac{1}{\lambda} &= \frac{\mu c\alpha^2}{2h}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \\ &= R_H\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)\end{aligned} \tag{535}$$

where R_H is known as the Rydberg constant. One of the early triumph of the quantum mechanics was that the above expression for the Rydberg constant agreed completely with the experimentally measured value.

How big is R_H? For this we forget $\hbar$ and c for a while but remember that $h = 2\pi\hbar$ should be treated as 2π:

$$\begin{aligned} [R_H] &= \frac{\mu\alpha^2}{4\pi} \\ &= (0.511\,\mathrm{MeV})\frac{1}{137^2\,4\pi} \\ &= (2.17\,\mathrm{eV}) \end{aligned} \tag{536}$$

Now to get the length scale,

$$\begin{aligned} R_H &= \frac{(2.17\,\mathrm{eV})}{\hbar c} \\ &= \frac{2.17\,\mathrm{eV}}{197.3\,\mathrm{eVnm}} \\ &= \frac{1}{91\,\mathrm{nm}} \end{aligned} \tag{537}$$

Visible lights are from about 400nmto 700nm. So the Balmer series with the final $n_f = 2$ should give out visible light. If $n_f = 1$, then $\lambda \sim 100$nm which is UV range, known as Lyman series. If $n_f = 3$, then $\lambda \sim 1000\,\mathrm{nm}$ which is IR range, known as Paschen series.

Note that μ is somewhat sensitive to the ratio m_e/M. For duetron M actually doubles. Therefore, if the experiment is senstivie enough (1 part in 1000), then we can see that two isotope means two separate Balmer series. And so it is how deutron was discovered.

Now that we have discovered that Quantum Mechanics can naturally explain the hydrogen atom, let's see if we can figure out what a hydrogen atom actually looks like. For this, we need to know the wavefunction. **These are Laguerre functions.**

First of all, let's look at the defintion of our length parameter ρ:

$$\begin{aligned} \rho &= \sqrt{\frac{8\mu|E|}{\hbar^2}}r \\ &= \sqrt{\frac{8\mu}{\hbar^2}}\sqrt{\frac{Z^2\alpha^2\mu c^2}{2n^2}}\,r \\ &= \frac{2Z\alpha\mu c^2}{n\,\hbar c}\,r \end{aligned} \tag{538}$$

So defining the Bohr radius,

$$a_0 = \frac{\hbar c}{\mu c^2 \alpha} \tag{539}$$

we get

$$\rho = \frac{2Z}{n}\left(\frac{r}{a_0}\right) \tag{540}$$

Now ground state is when $n = n_r + 1 + l = 1$. So that means $n_r = l = 0$. And that means

$$u_{1,0}(\rho) = \rho e^{-\rho/2} c_0 \tag{541}$$

or

$$R_{1,0}(r) = 2\left(\frac{Z}{a_0}\right)^{3/2} e^{-Zr/a_0} \tag{542}$$

where the coefficients are determined by the normalization constant.

Going higher, think about

$$n = n_r + l + 1 = 2 \tag{543}$$

this can happen if

$$n_r = 0, l = 1 \tag{544}$$

$$n_r = 1, l = 0 \tag{545}$$

and the wavefunctions are

$$R_{2,0} = 2\left(\frac{Z}{2a_0}\right)^{3/2}\left(1 - \frac{Zr}{2a_0}\right) e^{-Zr/2a_0} \tag{546}$$

$$R_{2,1} = \frac{1}{\sqrt{3}}\left(\frac{Z}{2a_0}\right)^{3/2} \frac{Zr}{2a_0} e^{-Zr/2a_0} \tag{547}$$

If $n = 3$,

$$n = n_r + l + 1 = 3 \tag{548}$$

this can happen if

$$n_r = 0, l = 2 \tag{549}$$

$$n_r = 1, l = 1 \tag{550}$$

$$n_r = 2, l = 0 \tag{551}$$

and the wavefunctions are

$$R_{3,0} = 2\left(\frac{Z}{2a_0}\right)^{3/2}\left[1 - \frac{2Zr}{3a_0} - \frac{2(Zr)^2}{27a_0^2}\right]e^{-Zr/3a_0} \tag{552}$$

$$R_{3,1} = \frac{4\sqrt{2}}{9}\left(\frac{Z}{2a_0}\right)^{3/2}\frac{Zr}{a_0}\left[1 - \frac{Zr}{6a_0}\right]e^{-Zr/3a_0} \tag{553}$$

$$R_{3,2} = \frac{2\sqrt{2}}{27\sqrt{5}}\left(\frac{Z}{2a_0}\right)^{3/2}\left(\frac{Zr}{a_0}\right)^2 e^{-Zr/3a_0} \tag{554}$$

Instead of labelling the eigenstate as $|E, l, m\rangle$, we can then use

$$\psi_{n,l,m}(\mathbf{r}) = \langle \mathbf{r}|n, l, m\rangle \tag{555}$$

Note that

$$R_{n,n-1} \tag{556}$$

has one peak. Now as far as the probability goes, the probabiltiy density in r is really r^2R^2. So the location of this peak for the probability is at

$$\begin{aligned}\frac{d}{dr}(r^2R^2) &\propto \frac{d}{dr}r^{2+2(n-1)}e^{-2Zr/na_0} \\ &= 2nr^{2n-1}e^{-Zr/na_0} - (2Z/na_0)r^{2n}e^{-Zr/na_0}\end{aligned} \tag{557}$$

or

$$r = \frac{n^2a_0}{Z} \tag{558}$$

So larger n means larger orbit (in the classical sense). Could be have expected this?

Let's first consider classical mechanics. For $l = n - 1$, the effective potential is

$$[V_{\text{eff}}] = \frac{l(l+1)}{2\mu r^2} - \frac{Ze^2}{r} = \frac{n(n-1)}{2\mu r^2} - \frac{Ze^2}{r} \tag{559}$$

The minimum of this potential is at

$$\frac{d[V_{\rm eff}]}{dr} = -2\frac{n(n-1)}{2\mu r^3} + \frac{Ze^2}{r^2} = 0 \tag{560}$$

or

$$\frac{n(n-1)}{\mu} = Ze^2 r \tag{561}$$

$$r = \frac{n(n-1)}{\mu Ze^2} \tag{562}$$

Now $a_0 = 1/\mu\alpha$ so that

$$r = n(n-1)\frac{a_0}{Z} \tag{563}$$

For large n, this is almost the same as above. This is as it should be for the **correspondence principle** asserts that for large quantum number, quantum mechanics must go over to the classical mechanics.

Consider $n_r = -\alpha = 0$ ($(l = n-1)$). In that case, all c_k are zero except c_0. So your wavefunction is

$$u_{n,n-1}(r) \propto r^n\, e^{-Zr/na_0} \tag{564}$$

Consider $n_r = -\alpha = 1$ ($l = n-2$). In that case,

$$\begin{aligned} c_1 &= -\frac{1}{\beta}c_0 \\ c_2 &= 0 \\ &\cdots \end{aligned} \tag{565}$$

So your wave function is proportional to

$$\begin{aligned} u_{n,n-2}(r) &\propto r^{n-1}\left(1 - \frac{1}{2(l+1)}\frac{2Zr}{na_0}\right) e^{-Zr/na_0} \\ &\propto r^{n-1}\left(1 - \frac{1}{(n-1)}\frac{Zr}{na_0}\right) e^{-Zr/na_0} \end{aligned} \tag{566}$$

The two maxima of this function satisfies

$$r_{\rm min} r_{\rm max} = n^2(n-1)^2 a_0^2/Z^2 \tag{567}$$

Classically, an elliptic orbit with $E = Z^2\alpha^2\mu/2n^2$ satisfies

$$r_{\min} r_{\max} = n^2(l+1/2)^2 a_0^2/Z^2 \tag{568}$$

with the replacement of $l(l+1) \to (l+1/2)^2$.

For $l = n-2$ and large n, these two formulas coincide.

Or

$$\begin{aligned} r_{\max} &= a_0 n^2 \left(1 + \sqrt{1-(l+1/2)^2/n^2}\right) \\ r_{\min} &= a_0 n^2 \left(1 - \sqrt{1-(l+1/2)^2/n^2}\right) \end{aligned} \tag{569}$$

Classically

$$\begin{aligned} r_{\max} &= \frac{\alpha}{1-\epsilon} \\ &= \frac{\alpha(1+\epsilon)}{1-\epsilon^2} \\ &= \frac{\alpha(1+\epsilon)}{2|E|\alpha/\alpha_{EM}} \\ &= \alpha_{EM}\frac{(1+\epsilon)}{2|E|} \\ &= a_0 n^2(1+\epsilon) \end{aligned} \tag{570}$$

Likewise

$$\begin{aligned} r_{\min} &= \frac{\alpha}{1+\epsilon} \\ &= \frac{\alpha(1-\epsilon)}{1-\epsilon^2} \\ &= \frac{\alpha(1-\epsilon)}{2|E|\alpha/\alpha_{EM}} \\ &= \alpha_{EM}\frac{(1-\epsilon)}{2|E|} \\ &= a_0 n^2(1-\epsilon) \end{aligned} \tag{571}$$

Using $E = -\alpha_{EM}/2a_0 n^2$,

Now

$$\begin{aligned} \epsilon &= \sqrt{1-2|E|\alpha/\alpha_{EM}} \\ &= \sqrt{1-L^2/n^2} \end{aligned} \tag{572}$$

since

$$\begin{aligned} |E|\alpha/\alpha_{EM} &= \frac{\alpha_{EM}}{2a_0n^2}\frac{a_0L^2}{\alpha_{EM}} \\ &= \frac{L^2}{2n^2} \end{aligned} \tag{573}$$

Now consider uncertainty principle. We have

$$[r, p_r] = i\hbar \tag{574}$$

where

$$p_r = \frac{\hbar}{i}\left(\partial_r + \frac{1}{r}\right) = \frac{\hbar}{i}\frac{1}{r}\partial_r r \tag{575}$$

So that the uncertainty principle is

$$\Delta r \Delta p_r \geq \hbar/2 \tag{576}$$

The hamiltonian is

$$H = \frac{p_r^2}{2\mu} + \frac{l(l+1)}{2\mu r^2} - \frac{Ze^2}{r} \tag{577}$$

So let's first calculate the average radial momentum

$$\begin{aligned} \langle n,l,m|\hbar p_r|n,l,m\rangle &= \frac{\hbar}{i}\int_0^\infty dr\, r^2\, R^*_{n,l}(r)\frac{1}{r}\partial_r r R_{n,l}(r) \\ &= \frac{\hbar}{i}\int_0^\infty dr\, u^*_{n,l}(r)\partial_r u_{n,l}(r) \end{aligned} \tag{578}$$

using $R = u/r$. Note that $u_{n,l}(r)$ is real. Therefore

$$\begin{aligned} \langle n,l,m|\hat{p_r}|n,l,m\rangle &= \frac{\hbar}{2i}\int_0^\infty dr\, \frac{d}{dr}u^2_{n,l}(r) \\ &= \frac{\hbar}{2i}\, u^2_{n,l}(r)\Big|_0^\infty = 0 \end{aligned} \tag{579}$$

This is as it should be given the real nature of R and the form of p_r which involves i. Hence

$$\begin{aligned}\Delta p_r^2 &= \langle n,l,m|p_r^2|n,l,m\rangle \\ &= 2\mu\langle n,l,m|\left(H-\frac{l(l+1)}{2\mu r^2}+\frac{Ze^2}{r}\right)|n,l,m\rangle \\ &\sim \mu E \sim \frac{\mu^2 Z^2\alpha^2}{n^2}\end{aligned} \tag{580}$$

Now for a given n, there are different $l=0,1,...,n-1$ that has the same energy. The wavefunctions for $l=0,1,...,n-2$ all have extra bumps and zeros. Normally, this would mean the energy is larger but then that doesn't take account the orbit being larger. These extra zeros and bumps, though are important how the electrons fill up the energy levels and hence chemistry.

Now take a deep breath and think. What have we done here? Well, on the surface, we have worked out the full wavefunction of a hydrogen atom and we found that

$$E_n = -\frac{13.6}{n^2} \tag{581}$$

And then we said that this explains the hydrogen atom radiation pattern

$$\frac{1}{\lambda} = R_H\left(\frac{1}{n_f^2}-\frac{1}{n_i^2}\right) \tag{582}$$

We have cheated a little bit here. Can you see where? Well, the cheat is something like this. The energy levels we just worked out are supposed to be **eigenstates** of hamiltonian. Now, the thing about the eigenstates of hamiltonian is one, they are **stable** and two, each state is **orthogonal** with each other.

So in this sense, a hydrogen atom with its electron in the 2S state is just as stable as a hydrogen atom with its electron in the ground state 1S. So why does it change state by radiating a photon? Shouldn't it be **stable**?

This confusion arises from the fact that we have considered the hydrogen atom as an **isolated system**. In reality, a hydrogen atom is never really isolated. The full Hamiltonian then must also contain the energy of the light quanta and the interaction between the light (photon) and the hydrogen atom. This will come later.

We started from the two body Hamiltonian

$$\hat{H}_{\text{total}} = \frac{p_A^2}{2M} + \frac{p_e^2}{2m} - \frac{Ze^2}{|r_A - r_e|} \tag{583}$$

Then reduced it to a one-body problem

$$\hat{H}_{\text{red}} = \frac{p^2}{2\mu} - \frac{Ze^2}{r} \tag{584}$$

where r refers to the relative distance between the core nucleus and the electron and $\mu = mM/(m+M)$ is the reduced mass. Upon using

$$\hat{p} = \frac{\hbar}{i}\nabla \tag{585}$$

we then get the Schrodinger equation

$$\hat{H}_{\text{red}}\psi = \left(-\frac{\hbar^2\nabla^2}{2\mu} - \frac{Ze^2}{r}\right)\psi = E\psi \tag{586}$$

By writing $\psi_{nlm} = R_{nl}(r)Y_{lm}(\theta,\phi)$ we further reduce it to

$$\frac{1}{\rho^2}\frac{d}{d\rho}\left(\rho^2\frac{dR}{d\rho}\right) + \left[\frac{\lambda}{\rho} - \frac{1}{4} - \frac{l(l+1)}{\rho^2}\right]R = 0 \tag{587}$$

where $\rho = kr$ with $k^2 = 8\mu|E|/\hbar^2$ and

$$\lambda = \frac{Ze^2}{\hbar}\left(\frac{\mu}{2|E|}\right)^{1/2} \tag{588}$$

Setting

$$R = e^{-\rho/2}\rho^l F(\rho) \tag{589}$$

we then arrive at the hypergeometric equation

$$\rho F'' + [\beta - \rho]F' - \alpha F = 0 \tag{590}$$

with

$$\beta = 2(l+1) \tag{591}$$

$$\alpha = l + 1 - \lambda \tag{592}$$

Then we figured out that the solution of this equation behaves well only if α is some a non-negative integer, say, n_r.

This implies

$$\lambda = n_r + (l+1) = n \tag{593}$$

is an integer, too. And using the definition of λ, we get

$$E_{nl} = -\frac{\mu Z^2 e^4}{2\hbar^2 n^2} \tag{594}$$

We start from here.

Now notice something here. The equation we started with was explicitly dependent upon l. The energy we got show no sign of l. This means that as long as we can choose an appropriate n_r, different l can have the same energy. Well, for a given n and l, a non-negative n_r can be choosen only if $0 \leq l \leq n-1$. or there are n l's that satisfy

$$n - (l+1) \geq 0 \tag{595}$$

We also note that there are no magnetic quantum number m (eigenvalue of the L_z) appearing anywhere. Remember that different l, m means different quantum states and for a given l, possible m values are

$$-l \leq m \leq l \tag{596}$$

Therefore, for a given n, there can be

$$g_n = \sum_{l=0}^{n-1} \sum_{m=-l}^{l} = \sum_{l=0}^{n-1} (2l+1) = 2\frac{(n-1)n}{2} + n = n^2 \tag{597}$$

quantum states of a hydronge-like atom all having the same energy. (Actually $2n^2$ due to the electron spin.)

This is what we call degeneracy. That simply means that there are more than one quantum states corresponding to a given energy level.

Could we have expected this? In a sense, yes. Remember that the magnetic quantum number corresponds to the orientation of the angular momentum vector. Now the space is rotationally symmetric because the coulomb force is central. That means one orientation is as good as any and hence

there is no distinction between them. This tells us that at least all the different magnetic quantum numbers has to corresponds to the same energy level given l.

Why there are n degenerate different l state is somewhat harder to explain. This is sometimes called 'accidental degeneracy' because this only happens in coulomb-like potential. Briefly, this is related to the fact that in a coulomb like potential, one can have a colsed orbit. What this means is that other than the angular momentum and the energy, there is an additional conserved quantity – that of the direction of the long axis of the ellipse. This quantity is called Runge-Lenz vector.

$$\mathbf{M} = \frac{1}{2\mu}(\mathbf{p} \times \mathbf{L} - \mathbf{L} \times \mathbf{p}) - \frac{\kappa}{r}\mathbf{r} \tag{598}$$

Whenever one has a conserved quantity in physics, there usually is an associated symmetry. This time is no exception. The symmetry of the hydrogen atom hamiltonian is actually larger than the rotational symmetry implied by the form of the potential. Now symmetry means that for some quantities rotating, or changing things around in a particular way won't make any difference in the dynamics. For the symmetry associated with the Runge-Lenz vector, this happens to be direction of the major axis of the elliptical orbit.

What do we mean by degeneracy? We say there is degeneracy when one can associate many different states to a given energy. For instance, for the classical Kepler problem, we know that the direction of the angular momentum has nothing to do with the energy state. Therefore there is a degeneracy associate with the direction of the angular momentum. In this case, a 'state' means a unit vector $\mathbf{L}/L$ that points to the direction of $\mathbf{L}$.

The constancy of the Runge-Lenz vector $\mathbf{M}$ indicatest that there is another source of degeneracy – The direction of the major axis of the ellipse. Given $\mathbf{L}$, the plane of $\mathbf{M}$ is given, but it can point to any direction in that plane. An orbital 'state' is then completely specified by the 6 numbers (it has to be $6 = 3 + 3$.)

$$E, \mathbf{L}, \hat{\mathbf{M}}_\perp \tag{599}$$

where $\hat{\mathbf{M}}_\perp$ is a unit 2-D vector living in the orbital plane. All these 6 numbers can be arbitrarily chosen except that for the bound state, we must have

$$E = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} - \frac{\kappa}{r} < 0 \tag{600}$$

so that L cannot be too large compared to $|E|$.

Now the degeneracy due to the direction of $\mathbf{L}$ can be read off from this expressin of E – The energy depends only on L^2. To say something similar for $\mathbf{M}$, we need to compute

$$M^2 = \frac{1}{\mu^2}(\mathbf{p}\times\mathbf{L})^2 + G^2 - \frac{2G}{\mu}(\mathbf{p}\times\mathbf{L})\cdot\frac{\mathbf{r}}{r} \tag{601}$$

we have

$$(\mathbf{p}\times\mathbf{L})^2 = p^2L^2 - (\mathbf{p}\cdot\mathbf{L})^2 = p^2L^2 \tag{602}$$

and we use

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \epsilon_{ijk}A_iB_jC_k = \epsilon_{kij}C_kA_iB_j = \mathbf{C}\cdot(\mathbf{B}\times\mathbf{A}) \tag{603}$$

to get

$$\mathbf{r}\cdot(\mathbf{p}\times\mathbf{L}) = \mathbf{L}\cdot(\mathbf{r}\times\mathbf{p}) = L^2 \tag{604}$$

so

$$M^2 = \frac{p^2L^2}{\mu^2} + G^2 - \frac{2GL^2}{\mu r} \tag{605}$$

Now we know that

$$\frac{p^2}{2\mu} - \frac{G}{r} = -|E| \tag{606}$$

so

$$\begin{aligned} M^2 &= \frac{2L^2}{\mu}\left(\frac{G}{r} - |E|\right) + G^2 - \frac{2GL^2}{\mu r} \\ &= G^2 - \frac{2|E|L^2}{\mu} \end{aligned} \tag{607}$$

In this expression, we see that $|E|$ depends only on M^2 and L^2. Therefore, there must be degeneracy corresponding to the orientation of M^2. Now this is only possible because we know that $\mathbf{M}$ is conserved therefore constant. We can't say the same thing, for instance for p_r in the Hamiltonian because it is not a conserved quantity.

13 Time-Independent Perturbation

We begin by writing a Hamiltonian.

$$H_{\text{total}} = H_0 + \lambda V \tag{608}$$

Here, H_0 is a known and solvable hamiltonian. For instance, it could be Hydrogen-like. The perturbation Hamiltonian V can be anything from external electro-magnetic field to the average interaction with other atoms and particles.

We assume that λ is small. Then intuitively, it is reasonable to guess that the eigenstates and the eigen-energies of the total hamiltonian is not much different from those of H_0. The effect of having λV will be then just a small correction to those.

Perturbation theory is a way of formalizing this intuition and it is a very powerful calculational tool. Without perturbation theory, the success of modern physics may have not been possible. For instance, the incredibly precise calculation of the electron magnetic moment (known as $g-2$) is done by a series of perturbative calculation.

$$a_e^{\text{th}} = 1\ 159\ 652\ 359\ (282) \times 10^{-12} \tag{609}$$
$$a_e^{\text{exp}} = 1\ 159\ 652\ 410\ (200) \times 10^{-12} \tag{610}$$

Now as I said, the smallness of λ indicate to us that the eigenstate and the eigen-energy of the total hamiltonian is not much different from H_0. Since the difference between H_{tot} and H_0 is $O(\lambda)$, we expect that the difference between the exact eigen state and the eigenstate of H_0 is also $O(\lambda)$. So, let's write

$$|N\rangle = |n\rangle + \lambda|N_1\rangle + \lambda^2|N_2\rangle + \cdots \tag{611}$$

and

$$E_N = E_n + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \tag{612}$$

These should be the exact eigenstate of H_{tot} or

$$H|N\rangle = E_N|N\rangle \tag{613}$$

In other words,

$$(H_0 + \lambda V)\left(|n\rangle + \lambda|N_1\rangle + \lambda^2|N_2\rangle + \cdots\right) = \left(E_n + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots\right) \times \left(|n\rangle + \lambda|N_1\rangle + \lambda^2|N_2\rangle + \cdots\right) \tag{614}$$

Now let's collect the zero-th order terms.

$$H_0|n\rangle = E_n|n\rangle \tag{615}$$

That's what we expected. The coefficients of λ is more interesting:

$$H_0|N_1\rangle + V|n\rangle = E_n|N_1\rangle + E_n^{(1)}|n\rangle \tag{616}$$

The goal here is to find the normalized $|N_1\rangle$ and the energy correctin $E_n^{(1)}$. First, let's apply $\bar{n}$ The LHS gives

$$\langle n|H_0|N_1\rangle + \langle n|V|n\rangle = E_n\langle n|N_1\rangle + E_n^{(1)}\langle n|n\rangle \tag{617}$$

The first terms cancel because $\langle n|H_0 = \langle n|E_n$. Therefore,

$$E_n^{(1)} = \frac{\langle n|V|n\rangle}{\langle n|n\rangle} = \langle n|V|n\rangle \tag{618}$$

if you have already normalized $|n\rangle$. We'll assume so from now on.

What this tells us is that the first order approximation to the n-the energy level is given by

$$E_N = E_n + \lambda E_n^{(1)} + O(\lambda^2) \tag{619}$$

and by doing so we are making $O(\lambda^2)$ error.

To get the first correction to the state vector, let's apply $\langle m|$ but with $m \neq n$.

$$LHS = \langle m|H_0|N_1\rangle + \langle m|V|n\rangle = E_m\langle m|N_1\rangle + \langle m|V|n\rangle \tag{620}$$

$$RHS = E_n\langle m|N_1\rangle + E_n^{(1)}\langle m|n\rangle = E_n\langle m|N_1\rangle \tag{621}$$

This means

$$\langle m|N_1\rangle = \frac{\langle m|V|n\rangle}{E_n - E_m} \qquad m \neq n \tag{622}$$

Hence

$$|N_1\rangle = \sum_m |m\rangle\langle m|N_1\rangle = \sum_{m\neq n} |m\rangle \frac{\langle m|V|n\rangle}{E_n - E_m} + |n\rangle\langle n|N_1\rangle \tag{623}$$

Now the fact that this procedure doesn't fix $\langle n|N_1\rangle$ is not actually a problem. It is actually an asset. Since we are essentially free to choose $\langle n|N_1\rangle$, we will do so in a way that the normalization of $|N\rangle$ is ensured. To see this, first, square $|N\rangle$ to get

$$\begin{aligned} \langle N|N\rangle &= \left(\langle n| + \lambda\langle N_1| + \lambda^2\langle N_2| + \cdots\right)\left(|n\rangle + \lambda|N_1\rangle + \lambda^2|N_2\rangle + \cdots\right) \\ &= 1 + \lambda\left(\langle n|N_1\rangle + \langle N_1|n\rangle\right) + O(\lambda^2) \end{aligned} \tag{624}$$

So if we choose

$$\mathrm{Re}\langle n|N_1\rangle = 0 \tag{625}$$

then we are only making λ^2 error in the normalization of $|N\rangle$. So that's o.k. Now, the imaginary part of $\langle n|N_1\rangle$ is still undetermined. For the sake of simplicity, we'll simply set it to 0. So with this choice made,

$$|N\rangle = |n\rangle + \lambda \sum_{m\neq n} |m\rangle \frac{\langle m|V|n\rangle}{E_n - E_m} + O(\lambda^2) \tag{626}$$

We can continue now to the second order. Collecting the coefficient of λ^2 term, we get

$$H_0|N_2\rangle + V|N_1\rangle = E_n|N_2\rangle + E_n^{(1)}|N_1\rangle + E_n^{(2)}|n\rangle \tag{627}$$

We proceed as before. Apply $\langle n|$ to get

$$\langle n|V|N_1\rangle = E_n^{(2)} \tag{628}$$

As usual, the first terms cancel and $|N_2\rangle$ dependence drop out and things can be expressed with known quantities. Note also we used $\langle n|N_1\rangle = 0$. The second order correction to the energy is therefore,

$$\begin{aligned} E_n^{(2)} &= \langle n|V|N_1\rangle = \sum_{m\neq n} \frac{\langle n|V|m\rangle\langle m|V|n\rangle}{E_n - E_m} \\ &= = \sum_{m\neq n} \frac{|\langle n|V|m\rangle|^2}{E_n - E_m} \end{aligned} \tag{629}$$

To calculate the eigen state up to the second order, we apply $\langle k|$ with $k \neq n$.

$$E_k\langle k|N_2\rangle + \langle k|V|N_1\rangle = E_n\langle k|N_2\rangle + E_n^{(1)}\langle k|N_1\rangle \tag{630}$$

So

$$\begin{aligned}\langle k|N_2\rangle &= \frac{\langle k|V|N_1\rangle}{E_n - E_k} - \frac{E_n^{(1)}\langle k|N_1\rangle}{E_n - E_k} \\ &= \sum_{m\neq n} \frac{\langle k|V|m\rangle\langle m|V|n\rangle}{(E_n - E_m)(E_n - E_k)} - \frac{\langle n|V|n\rangle\langle k|V|n\rangle}{(E_n - E_k)^2}\end{aligned} \tag{631}$$

Again, $\langle n|N_2\rangle$ is undetermined. We proceed as before and square $|N\rangle$.

$$\langle N|N\rangle = 1 + \lambda(\langle n|N_1\rangle + \langle N_1|n\rangle) + \lambda^2(\langle N_2|n\rangle + \langle n|N_2\rangle + \langle N_1|N_1\rangle) + O(\lambda^3) \tag{632}$$

The second term is zero by our previous choice. To get rid of the third term, we set

$$\langle n|N_2\rangle = \langle N_2|n\rangle = \frac{1}{2}\langle N_1|N_1\rangle = \frac{1}{2}\sum_{m\neq n}\frac{|\langle m|V|n\rangle|^2}{(E_n - E_m)^2} \tag{633}$$

In this way, $|N\rangle$ is normalized up to $O(\lambda^3)$ error. So

$$\begin{aligned}|N_2\rangle &= \sum_k |k\rangle\langle k|N_2\rangle \\ &= \sum_{k\neq n}|k\rangle\left[\sum_{m\neq n}\frac{\langle k|V|m\rangle\langle m|V|n\rangle}{(E_n - E_m)(E_n - E_k)} - \frac{\langle n|V|n\rangle\langle k|V|n\rangle}{(E_n - E_k)^2}\right] \\ &\quad + |n\rangle\frac{1}{2}\sum_{m\neq n}\frac{|\langle m|V|n\rangle|^2}{(E_n - E_m)^2}\end{aligned} \tag{634}$$

One rarely goes beyond the second order.

14 Examples – Anharmonic Oscillator

Suppose we start with a hamiltonian,

$$H = \frac{p^2}{2m^2} + m\omega^2\frac{x^2}{2} + \lambda\epsilon\omega^4 x^4 \tag{635}$$

We know the eigenstates and the eigenenergies of the simple harmonic oscillator. This is most easily done using the creation and the annihilation operators

$$a = \frac{1}{\sqrt{2\hbar}}\left(\sqrt{\omega m}x + i\frac{p}{\sqrt{\omega m}}\right) \tag{636}$$

$$a^\dagger = \frac{1}{\sqrt{2\hbar}}\left(\sqrt{\omega m}x - i\frac{p}{\sqrt{\omega m}}\right) \tag{637}$$

They satisfy

$$[a, a^\dagger] = \frac{1}{2\hbar}\left(-i[x,p] + i[p,x]\right) = \frac{1}{2\hbar}2\hbar = 1 \tag{638}$$

and it can be easily shown that

$$a^\dagger a = \frac{1}{2\hbar}\left(\omega m x^2 + \frac{p^2}{\omega m} - \hbar\right) \tag{639}$$

In other words,

$$H_0 = \hbar\omega a^\dagger a + \frac{\hbar\omega}{2} \tag{640}$$

Hence the eigenstates of H_0 is the eigenstates of the number operator $N = a^\dagger a$ and the energy is simply given by

$$E_n = (n + 1/2)\hbar\omega \tag{641}$$

We also know that

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{642}$$

$$a|n\rangle = \sqrt{n}|n-1\rangle \tag{643}$$

$$(a^\dagger)^n|0\rangle = (a^\dagger)^{n-1}\sqrt{1}|1\rangle = (a^\dagger)^{n-2}\sqrt{1}\sqrt{2}|2\rangle = ... = \sqrt{n!}|n\rangle \tag{644}$$

A way to remember these factors is to think about the ground state. If you operator the creation op to $|0\rangle$, you should get $|1\rangle$. So the coefficient better be $n+1$. Obviously, it cannot be n or $n-1$. Also, remember

$$\langle n|a^\dagger a|n\rangle = (a|n\rangle)^\dagger(a|n\rangle) = n \tag{645}$$

so that coefficient has to be n.

Now we are ready to do the perturbation theory. First, we make the following observation:

$$x = \sqrt{\frac{\hbar}{2\omega m}}(a + a^\dagger) \tag{646}$$

Hence

$$V = \epsilon\omega^4 x^4 = \epsilon\omega^4 \left(\frac{\hbar}{2\omega m}\right)^2 (a + a^\dagger)^4 = \epsilon \left(\frac{\hbar\omega}{2m}\right)^2 (a + a^\dagger)^4 \tag{647}$$

Now if these are c-numbers, we can simply write

$$(a + a^\dagger)^4 = a^4 + 4a^3a^\dagger + 6a^2(a^\dagger)^2 + 4a(a^\dagger)^3 + (a^\dagger)^4 \tag{648}$$

But obviously, we can't do that. For instance, instead of $6a^2(a^\dagger)^2$, we will have

$$H_{22} = aaa^\dagger a^\dagger + a^\dagger a^\dagger aa + aa^\dagger aa^\dagger + a^\dagger aaa^\dagger + a^\dagger aa^\dagger a + aa^\dagger a^\dagger a \tag{649}$$

Now consider the first order expression

$$E_N^{(1)} = \langle n|V|n\rangle \tag{650}$$

Since we are sandwitching the potential with the same bra and ket, the only term that contributes to this expression is in fact the above. All other terms have mismatched a and $a^\dagger$. O.K. Once we know that, now we can start to rearrange H_{22} to a more convenient form. Remember $[a, a^\dagger] = 1$ or $aa^\dagger = N + 1$. And also

$$[N, a] = -a \qquad [N, a^\dagger] = a^\dagger \tag{651}$$

or

$$Na - aN = -a \tag{652}$$

or

$$Na = a(N - 1) \qquad aN = (N + 1)a \tag{653}$$

So

$$\begin{aligned} H_{22} &= aaa^\dagger a^\dagger + a^\dagger a^\dagger aa + aa^\dagger aa^\dagger + a^\dagger aaa^\dagger + a^\dagger aa^\dagger a + aa^\dagger a^\dagger a \\ &= a(N+1)a^\dagger + a^\dagger Na + aNa^\dagger + N(N+1) + N^2 + (N+1)N \\ &= (N+1)(N+1) + (N+1) + N(N-1) + (N+1)(N+1) \\ &\quad + N(N+1) + N^2 + (N+1)N \\ &= 6N^2 + 6N + 3 \end{aligned} \tag{654}$$

There is a factor of $\epsilon(\hbar^2\omega^2/4m^2)$ in front. So combined, we get

$$\langle n|V|n\rangle = \epsilon\frac{3\hbar^2\omega^2}{4m^2}\left(2n^2+2n+1\right) \tag{655}$$

The correction to the wave function is given by'

$$N = |n\rangle + \lambda\sum_{m\neq n}|m\rangle\frac{\langle m|V|n\rangle}{E_n - E_m} \tag{656}$$

Now, one thing is obvious here. First, since x^4 involves maximum four factors of creation and the annihilation operators, it is apparent that the sum involves on $n-4 \leq m \leq n-1$ and $n+1 \leq m \leq n+4$. That is, it involves at most 8 states.

Let's see if we can do a reasonable job for the ground state. We need to figure out these matrix elements:

$$\langle m|V|0\rangle = \langle m|x^4|0\rangle \tag{657}$$

where $m \neq 0$. Now, let's consider $m = 1$. Now, for this to have a non-zero matrix element with the ground state, there should be 1 more $a^\dagger$ than a. Otherwise, you can't change $|0\rangle$ to $|1\rangle$. However, with 4 a and $a^\dagger$, this is impossible. Therefore,

$$\langle 1|x^4|0\rangle = 0 \tag{658}$$

Now, consider $m = 2$. For this state to have non-zero matrix element with the ground state, there should be 2 more $a^\dagger$ and a. That is, terms with 3 $a^\dagger$ and 1 a will do.

$$\begin{aligned} \langle 2|a^\dagger a^\dagger aa^\dagger|0\rangle &= (a|2\rangle)^\dagger N(a^\dagger|0\rangle) \\ &= \sqrt{2}\langle 1|N\langle 1| = \sqrt{2} \end{aligned} \tag{659}$$

Also

$$\begin{aligned}\langle 2|a^\dagger a a^\dagger a^\dagger|0\rangle &= \langle 2|N\sqrt{1\times 2}|2\rangle \\ &= \sqrt{2}\langle 2|N|2\rangle = 2\sqrt{2}\end{aligned} \tag{660}$$

and

$$\begin{aligned}\langle 2|aa^\dagger a^\dagger a^\dagger|0\rangle &= (a^\dagger|2\rangle)^\dagger a^\dagger a^\dagger a^\dagger|0\rangle \\ &= \langle 3|\sqrt{3}\sqrt{1\times 2\times 3}|3\rangle \\ &= 3\sqrt{2}\end{aligned} \tag{661}$$

but

$$\langle 2|a^\dagger a^\dagger a^\dagger a|0\rangle = 0 \tag{662}$$

So, altogether,

$$\langle 2|V|0\rangle = \epsilon\left(\frac{\hbar^2\omega^2}{4m^2}\right)6\sqrt{2} = \epsilon\left(\frac{3\hbar^2\omega^2}{\sqrt{2}m^2}\right) \tag{663}$$

Now for $m=3$, there should be 3 more $a^\dagger$'s that a. That's impossible for x^4. So

$$0 = \langle 3|x^4|0\rangle \tag{664}$$

For $m=4$, only $(a^\dagger)^4$ can contribute.

$$\langle 4|(a^\dagger)^4|0\rangle = \langle 4|\sqrt{4!}|4\rangle = \sqrt{24} = 2\sqrt{6} \tag{665}$$

or

$$\langle 4|V|0\rangle = \omega\left(\frac{\hbar^2\omega^2}{4m^2}\right)2\sqrt{6} = \omega\left(\frac{\sqrt{3}\hbar^2\omega^2}{\sqrt{2}m^2}\right) \tag{666}$$

So

$$\begin{aligned}|0\rangle_{\rm tot} &= |0\rangle + \sum_{m\neq n}\lambda|m\rangle\frac{\langle m|V|n\rangle}{E_n - E_m} \\ &= |0\rangle - \lambda\epsilon\left(\frac{3\hbar^2\omega^2}{\sqrt{2}m^2}\right)\frac{|2\rangle}{2\hbar\omega} - \lambda\epsilon\left(\frac{\sqrt{3}\hbar^2\omega^2}{\sqrt{2}m^2}\right)\frac{|4\rangle}{4\hbar\omega}\end{aligned} \tag{667}$$

Let's now calculate the second order energy correction to the ground state. Our formula is

$$\begin{aligned} E_N^{(2)} &= \sum_{m\neq 0} \frac{|\langle m|V|0\rangle|^2}{E_0 - E_m} \\ &= \frac{|\langle 2|V|0\rangle|^2}{E_0 - E_2} + \frac{|\langle 4|V|0\rangle|^2}{E_0 - E_4} \\ &= -\frac{\epsilon^2 \left(\frac{3\hbar^2\omega^2}{\sqrt{2}m^2}\right)^2}{2\hbar\omega} - \frac{\epsilon^2 \left(\frac{\sqrt{3}\hbar^2\omega^2}{\sqrt{2}m^2}\right)^2}{4\hbar\omega} \\ &= -\epsilon^2 \frac{9\hbar^3\omega^3}{4m^4} - \epsilon^2 \frac{3\hbar^3\omega^3}{8m^4} \\ &= -\epsilon^2 \frac{21\hbar^3\omega^3}{8m^4} \end{aligned} \tag{668}$$

15 Degenerate Case

Now so far we have been considering the case when all energy levels are non-degenerate. When some of them *are* degenerate, then the formulations we had before breaks down and we need a new formulation of the problem. O.K. What is our problem? We again start by looking at

$$H = H_0 + \lambda V \tag{669}$$

Now, consider N-th energy level. Before, we could just write

$$|N\rangle = |n\rangle + \lambda|N^{(1)}\rangle + \lambda^2|N^{(2)}\rangle + \cdots \tag{670}$$

and expect it to be a good expansion. However, when there is a degeneracy, there is no guarantee that the degenerate energy levels would not mix. That is, in this case, a general form of $|N\rangle$ should be

$$|N\rangle = \sum_{\alpha=1}^{g} a_N^\alpha |n_\alpha\rangle + \lambda|N^{(1)}\rangle + \cdots \tag{671}$$

Here α labels each quantum states in the energy level n. For instance, if we are dealing with the hydrogen atom, α will correspond to different l and m. The point here is that the coefficient a_N^α are not necessarily small, i.e. $O(\lambda)$.

O.K. Given this, what can we work out? Consider the first order correction. Now the Schroedinger equation still says

$$(H_0 + \lambda V)(\sum_\alpha a_N^\alpha |n_\alpha\rangle + \lambda |N_1\rangle + \cdots) = (E_n + \lambda E_N^{(1)} + \cdots)(\sum_\alpha a_N^\alpha |n_\alpha\rangle + \lambda |N_1\rangle + \cdots) \quad (672)$$

Assembling the first order, we get

$$V \sum_\alpha a_N^\alpha |n_\alpha\rangle + H_0 |N_1\rangle = E_N^{(1)} \sum_\alpha a_N^\alpha |n_\alpha\rangle + E_n |N_1\rangle \quad (673)$$

Applying $\langle n_\beta|$, we get

$$\sum_\alpha a_N^\alpha \langle n_\beta|V|n_\alpha\rangle = E_N^{(1)} \sum_\alpha a_N^\alpha \{\langle n_\beta|n_\alpha\rangle \quad (674)$$

or

$$\sum_\alpha a_N^\alpha \langle n_\beta|V|n_\alpha\rangle - E_N^{(1)} \sum_\alpha a_N^\alpha \delta_{\alpha\beta} = 0 \quad (675)$$

This can be thought as a matrix equation

$$\begin{pmatrix} V_{11} - E_N^{(1)} & V_{12} & \cdots & V_{1g} \\ V_{21} & V_{22} - E_N^{(1)} & \cdots & V_{2g} \\ \vdots & \vdots & \vdots & \vdots \\ V_{g1} & \cdots & \cdots & V_{gg} - E_N^{(1)} \end{pmatrix} \begin{pmatrix} a_N^1 \\ a_N^2 \\ \vdots \\ a_N^g \end{pmatrix} = 0 \quad (676)$$

This equation has a solution only if the determinant vanishes:

$$\mathrm{Det}(V_{\alpha\beta} - E_N^{(1)} \delta_{\alpha\beta}) = 0 \quad (677)$$

In other words, only if $E_N^{(1)}$ is an eigenvalue of the matrix $V_{\beta\alpha} = \langle n_\beta|V|n_\alpha\rangle$. Now since V is a $g \times g$ hermitian matrix, the eigenvalues are real and there are g of them. And these are the first order corrections to the g degenerate $|n_\alpha\rangle$ states.

Now having said that, it seems pretty mysterious and blackmagic-like. But this is not really so. Suppose that H_0 is given by

$$H_0 = B\hat{L}^2 \quad (678)$$

and

$$\lambda V = \lambda a \hat{L}_z \quad (679)$$

then we know right away that the right eigenstate of the total hamiltonian is $|m, l\rangle$ and the energy is given by

$$E_{lm} = B\hbar^2 l(l+1) + \lambda a\hbar m \tag{680}$$

In this case, the first order perturbation theory is exact and there is not need to go further to consider wavefunction corrections or second order corrections.

Now, suppose instead we are given

$$\lambda V = \lambda a\hat{L}_x \tag{681}$$

Of course, we know again right away that the eigenstate of the total hamiltonian is given by the simultaneous eigenstate of L^2 and L_x, $|l, m_x\rangle$. Now in this simple case, we know how to solve it exactly. Just make a right combination of m, l states so that it will points to x direction instead of z direction. That is, we just need to diagonalize the perturbation part of our total hamiltionian within the degenerate energy states! The unperturbed part of our hamiltonian H_0 is already diagonal anyway. In that way, the total hamiltonian is diagonal in that particular subspace made up of degenerate energy eigenstates of H_0.

However, suppose we don't know that. How do we proceed? Well, the perturbation theory says, diagonalize the matrix

$$V_{m_1,m_2} = \langle l, m_1|a\hat{L}_x|l, m_2\rangle \tag{682}$$

For instance, this matrix element between $l = 1$ and different m states are given in the table (11.63) in the textbook:

$$V = a\frac{\hbar}{2}\begin{pmatrix} 0 & \sqrt{2} & 0 \\ \sqrt{2} & 0 & \sqrt{2} \\ 0 & \sqrt{2} & 0 \end{pmatrix} \tag{683}$$

The determinant condition:

$$\mathrm{Det}(V - E\mathbf{1}) = 0 \tag{684}$$

yields

$$E = 0, \pm a\hbar \tag{685}$$

and the eigenvectors are just eigenvectors of $\hat{L_z}$.

$$|l=1, m_x=0\rangle = \frac{1}{\sqrt{2}}\left(|l=1, m_z=1\rangle - |l=1, m_z=-1\rangle\right) \tag{686}$$

$$|l=1, m_x=-1\rangle = \frac{1}{2}\Big(|l=1, m_z=1\rangle - \sqrt{2}|l=1, m_z=0\rangle + |l=1, m_z=-1\rangle\Big) \tag{687}$$

$$|l=1, m_x=1\rangle = \frac{1}{2}\Big(|l=1, m_z=1\rangle + \sqrt{2}|l=1, m_z=0\rangle + |l=1, m_z=-1\rangle\Big) \tag{688}$$

So by diagonalizing the matrix elements between the degenerate states, all we did is just to rotate the axis from z to x.

16 Example : Fine Structure of H Atom

Relativistic Correction

The relativistic kinetic energy of an electron is given by

$$K_{\rm Rel} = \sqrt{m^2c^4 + \mathbf{p}^2c^2} - mc^2 \tag{689}$$

Assuming $\mathbf{p}^2$ is smaller than m^2c^2, we can expand the square root. Recall

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + O(x^3) \tag{690}$$

Hence

$$\begin{aligned} K_{\rm Rel} &= mc^2\sqrt{1+\mathbf{p}^2/m^2c^2} - mc^2 \\ &= mc^2\left(1 + \frac{1}{2}\frac{\mathbf{p}^2}{m^2c^2} - \frac{1}{8}\frac{\mathbf{p}^4}{m^4c^4}\right) - mc^2 + O(p^6) \\ &\approx \frac{\mathbf{p}^2}{2m} - \frac{\mathbf{p}^4}{8m^3c^2} = K_0 + H_{rc} \end{aligned} \tag{691}$$

We know that

$$K_0 \sim \alpha^2 mc^2 \tag{692}$$

So

$$H_{rc} \sim K_0^2/(mc^2) \sim \alpha^4 mc^2 \ll K_0 \tag{693}$$

Therefore it is a small perturbation.

Another small perturbation of the same order in α is given by the spin-orbit coupling. This is caused by the electron spin interacting with the magnetic field generated by the electron's motion through the nucleus' electric field.

Suppose we are in the electron rest-frame. Then what we see is instead of the electron moving, we see the proton moving. To simplify, let's consider a circular orbit. Then in effect, it's like we are sitting at the center of a current loop or a single loop solenoid.

Now we are considering a circular orbit. And the relevant quantity is the magnetic field at the center of that circle. This is because the electron has a magnetic moment due to it's spin. Therefore, electron energy should contain the magnetic interaction term

$$H_{\text{mag}} = -\boldsymbol{\mu}_e \cdot \mathbf{B} \tag{694}$$

where $\boldsymbol{\mu}$ is the electron magnetic moment. We know what electron magnetic moment is. What we are trying to figure out here is then the size of the magnetic field experienced by the electon by sitting at the center of that single loop solenoid.

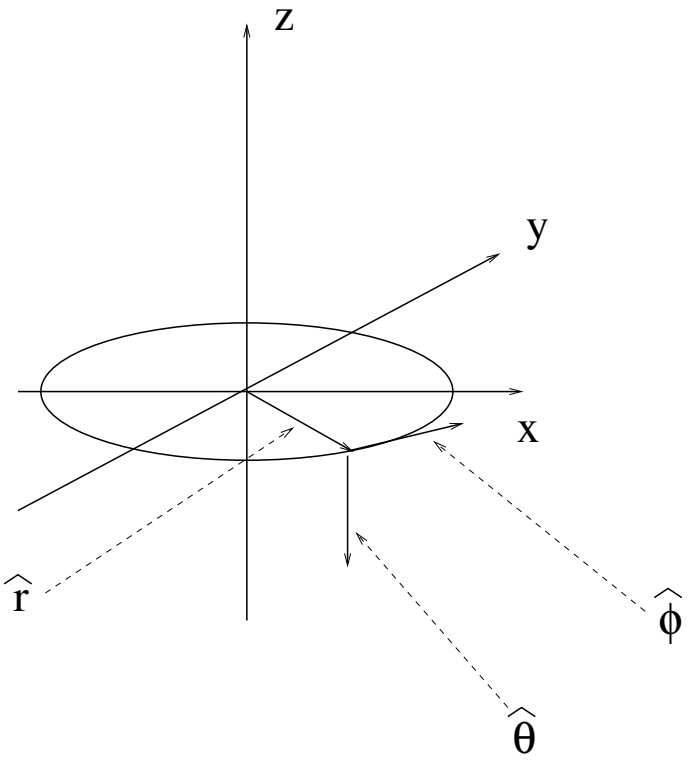
The relevant maxwell's equation reads:

$$\begin{aligned} \nabla \times \mathbf{B} &= \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J} \\ &= \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{J} \end{aligned} \tag{695}$$

In our case, the electric field is static. So we can use the Biot and Savart law. The solution of the above Maxwell's equation is given by

$$\begin{aligned} d\mathbf{B}(\mathbf{r}_2) &= -\frac{\mu_0}{4\pi} I_1 \frac{d\mathbf{l}_1 \times (\mathbf{r}_2 - \mathbf{r}_1)}{|\mathbf{r}_2 - \mathbf{r}_1|^3} \\ &= -\frac{1}{c} I_1 \frac{d\mathbf{l}_1 \times (\mathbf{r}_2 - \mathbf{r}_1)}{|\mathbf{r}_2 - \mathbf{r}_1|^3} \end{aligned} \tag{696}$$

In the case of our orbiting proton, the direction of the current is the direction of the proton. We'll take it to be in the direction of the azimuthal unit vector $\hat{\phi}$. The vector $\mathbf{r}_2 - \mathbf{r}_1$ is of course proportional to the radial unit vector $\hat{r}$.



Therefore, the magnetic field is in the $\hat{\theta}$ direction. If we choose the orbit to lie on the $x - y$ plane, this corresponds to the $-\hat{z}$ direction. We note

$$d\mathbf{l} = r d\phi \, \hat{\phi} \tag{697}$$

so

$$\begin{aligned} d\mathbf{B}(\mathbf{r}_2) &= -\frac{\mu_0}{4\pi} I_1 \frac{d\phi}{r}(-\hat{z}) \\ &= -\frac{1}{c} I_1 \frac{d\phi}{r}(-\hat{z}) \end{aligned} \tag{698}$$

The size of the magnetic field at the center can be easily obtained by a simple integration.

$$|\mathbf{B}| = \frac{\mu_0}{4\pi} I_1 \frac{2\pi}{r}$$

$$= \frac{1}{c} I_1 \frac{2\pi}{r} \tag{699}$$

Now we need to figure out what the current is. The definition of the current is

$$I = \frac{dQ}{dt} = \frac{dQ}{dr}\frac{dr}{dt} = \frac{e}{2\pi r} v \tag{700}$$

where $e = |e|$. Since we know that the angular momentum is conserved we use

$$\mathbf{L} = \mathbf{r} \times m\mathbf{v} \tag{701}$$

and get

$$v = \frac{|\mathbf{L}|}{rm} \tag{702}$$

or

$$I = \frac{e}{2\pi r} v = \frac{e\,|\mathbf{L}|}{2\pi m r^2} \tag{703}$$

So assembling everthing, we get

$$\begin{aligned} \mathbf{B} &= \frac{\mu_0}{4\pi}\frac{2\pi}{r}\frac{e\,|\mathbf{L}|}{2\pi m r^2}\hat{z} \\ &= \frac{1}{c}\frac{2\pi}{r}\frac{e\,|\mathbf{L}|}{2\pi m r^2}\hat{z} \end{aligned} \tag{704}$$

Now we know that the angular momentum is in $+\hat{z}$ direction. Hence, finally,

$$\begin{aligned} \mathbf{B} &= \frac{\mu_0}{4\pi}\frac{e}{mr^3}\mathbf{L} \\ &= \frac{1}{c}\frac{e}{mr^3}\mathbf{L} \end{aligned} \tag{705}$$

This is the magnetic field due to the *orbit*. Hence the spin-orbit coupling is given by

$$\begin{aligned} H_{\rm LS} &= -\boldsymbol{\mu}\cdot\mathbf{B} \\ &= -\left(\frac{-e}{m}\mathbf{S}\right)\cdot\frac{\mu_0}{4\pi}\frac{e}{mr^3}\mathbf{L} \quad \text{(MKS)} \end{aligned}$$

$$
\begin{aligned}
&= -\left(\frac{-e}{mc}\mathbf{S}\right)\cdot\frac{1}{c}\frac{e}{mr^3}\mathbf{L} \quad \text{(cgs)} \\
&= \frac{\mu_0}{4\pi}\frac{e^2}{m^2r^3}\mathbf{S}\cdot\mathbf{L} \quad \text{(MKS)} \\
&= \frac{e^2}{m^2c^2r^3}\mathbf{S}\cdot\mathbf{L} \quad \text{(cgs)} \\
&= \hbar\frac{\alpha}{m^2r^3c}\mathbf{S}\cdot\mathbf{L} \quad \text{(cgs, MKS)} \\
&= \frac{\mu_0\epsilon_0\hbar c}{4\pi\epsilon_0\hbar c}\frac{e^2}{m^2r^3}\mathbf{S}\cdot\mathbf{L} \\
&= \alpha\frac{\mu_0\epsilon_0\hbar c}{m^2r^3}\mathbf{S}\cdot\mathbf{L} \\
&= \alpha\frac{\hbar}{cm^2r^3}\mathbf{S}\cdot\mathbf{L} \\
&= \alpha\frac{\hbar^3c^3}{c^4m^2r^3}\left(\frac{\mathbf{S}\cdot\mathbf{L}}{\hbar^2}\right) \\
&= \frac{[\text{MeV.fm}]^3}{[\text{MeV}]^2[\text{fm}]^3} = [\text{MeV}] \qquad (706)
\end{aligned}
$$

Now let's see how big this interaction is. We use:

$$
a_0 = \frac{\hbar c}{\alpha m_e c^2} = \frac{\hbar}{\alpha m_e c} \qquad (707)
$$

So

$$
\begin{aligned}
|H_{\text{mag}}| &\sim \alpha\frac{(\hbar c)^3}{(m_ec^2)^2a_0^3} \\
&= \alpha\frac{(\hbar c)^3}{(m_ec^2)^2(\hbar c)^3/(\alpha m_ec^2)^3} \\
&= \alpha^4 m_ec^2 = \alpha^2(\alpha^2 m_ec^2) \qquad (708)
\end{aligned}
$$

This is however, not fully correct. A fully relativistic derivation yields

$$
H_{\text{LS}} = \frac{\alpha\hbar}{2m^2r^3c}\mathbf{S}\cdot\mathbf{L} \quad \text{(cgs, MKS)} \qquad (709)
$$

Fine Structure Calc

First, let's consider the relativistic correction. We notice that

$$\frac{p^2}{2m} = (H_0 + \frac{e^2}{r}) \tag{710}$$

Hence

$$\begin{aligned} H_{\rm rel} = -\frac{p^4}{8m^3c^2} &= -\frac{1}{2mc^2}\left(\frac{p^2}{2m}\right)^2 \\ &= -\frac{1}{2mc^2}\left(H_0 + \frac{e^2}{r}\right)^2 \end{aligned} \tag{711}$$

Hence

$$\begin{aligned} \langle nlm|H_{\rm rel}|nlm\rangle &= -\frac{1}{2mc^2}\langle nlm|\left(H_0^2 + H_0\frac{e^2}{r} + \frac{e^2}{r}H_0 + \left(\frac{e^2}{r}\right)^2\right)|nlm\rangle \\ &= -\frac{1}{2mc^2}\left(E_n^2 + 2E_n e^2\langle nlm|\frac{1}{r}|nlm\rangle + \left(e^2\right)^2\langle nlm|\frac{1}{r^2}|nlm\rangle\right) \\ &= -\frac{1}{2mc^2}\left(E_n^2 + 2E_n\alpha\hbar c\langle nlm|\frac{1}{r}|nlm\rangle + \alpha^2(\hbar c)^2\langle nlm|\frac{1}{r^2}|nlm\rangle\right) \end{aligned} \tag{712}$$

Now we know

$$0 = \frac{s}{4}\left[(2l+1)^2 - s^2\right]a_0^2\langle r^{s-2}\rangle - (2s+1)a_0\langle r^{s-1}\rangle + \frac{s+1}{n^2}\langle r^s\rangle \tag{713}$$

With $s = 0$, we get

$$0 = -a_0\langle r^{-1}\rangle + \frac{1}{n^2}\langle 1\rangle \tag{714}$$

or

$$\langle nlm|\frac{1}{r}|nlm\rangle = \frac{1}{n^2 a_0} \tag{715}$$

Problem is, this formula is useless to get $\langle 1/r^2\rangle$.

From our previous calculation, we also know that

$$\langle nlm|\frac{1}{r^2}|nlm\rangle = \frac{1}{n^3(l+1/2)a_0^2} \tag{716}$$

Understanding:
Kepler's 2nd law:

$$L = mr^2\dot{\phi} \tag{717}$$

is a constant
Kepler's 3rd law:

$$T^2 \propto a^3 \tag{718}$$

For cicular orbit,
Force balance:

$$\frac{e^2}{R^2} = \frac{mv^2}{R} \tag{719}$$

and

$$L = |r \times p| = Rmv \tag{720}$$

$$\begin{aligned} T &= \frac{2\pi}{\dot{\phi}} \\ &= \frac{2\pi m R^2}{L} \end{aligned} \tag{721}$$

or

$$\begin{aligned} T^2 &= \frac{4\pi^2 m^2 R^4}{L^2} \\ &= \frac{4\pi^2 m^2 R^4}{R^2 m^2 v^2} \\ &= \frac{4\pi^2 R^2}{v^2} \\ &= \frac{4\pi^2 m R^2}{m v^2} \\ &= \frac{4\pi^2 m R^3}{e^2} \end{aligned} \tag{722}$$

or

$$\frac{T}{2\pi} = \sqrt{\frac{mR^3}{e^2}} \tag{723}$$

or

$$\frac{2\pi}{T} = \sqrt{\frac{e^2}{mR^3}} \tag{724}$$

Use

$$a_0 = \frac{\hbar c}{\alpha m c^2} \tag{725}$$

$$\begin{aligned} \frac{2\pi}{T} &= \sqrt{\frac{\alpha\hbar c^3}{mc^2R^3}} \\ &= \sqrt{\frac{\alpha^2 a_0 c^2}{R^3}} \\ &= \alpha c\sqrt{\frac{a_0}{R^3}} \end{aligned} \tag{726}$$

Semi-Classical average:

$$\begin{aligned} \langle\frac{1}{r^2}\rangle_{\rm cl} &= \frac{1}{T}\int_0^T dt\,\frac{1}{r^2} \\ &= \frac{m}{T}\int_0^{2\pi} d\phi\,\frac{1}{mr^2\dot{\phi}} \\ &= \frac{m2\pi}{LT} \\ &= \frac{mc}{L}\alpha c\sqrt{\frac{a_0}{R^3}} \\ &= \frac{1}{L}\frac{1}{a_0}\sqrt{\frac{a_0}{R^3}} \end{aligned} \tag{727}$$

with

$$\frac{1}{R} = \frac{1}{n^2 a_0} \tag{728}$$

we have

$$\langle\frac{1}{r^2}\rangle_{\rm cl} = \frac{1}{n^3 L a_0^2} \tag{729}$$

The 3rd law says

$$\frac{1}{T} \propto \frac{1}{R^{3/2}} \propto \frac{1}{(n^2)^{2/3}} \propto \frac{1}{n^3} \tag{730}$$

So on dimensinal ground,

$$\langle \frac{1}{r^2} \rangle = \frac{A}{n^3(l+B)a_0^2} \tag{731}$$

Aside

So setting $s = -1$,

$$0 = \frac{-1}{4}\left[(2l+1)^2 - 1\right] a_0^2 \langle r^{-3} \rangle - (-1) a_0 \langle r^{-2} \rangle \tag{732}$$

Or

$$\langle nlm|\frac{1}{r^3}|nlm\rangle = \frac{1}{l(l+1)}\frac{1}{a_0}\langle nlm|\frac{1}{r^2}|nlm\rangle = \frac{1}{a_0^3}\frac{1}{n^3 l(l+1)(l+1/2)} \tag{733}$$

Using these and $E_n = -\alpha^2 mc^2/2$,

$$\begin{aligned}
\langle nlm|H_{\rm rel}|nlm\rangle &= -\frac{1}{2mc^2}\left(E_n^2 + 2E_n e^2 \langle nlm|\frac{1}{r}|nlm\rangle + \left(e^2\right)^2 \langle nlm|\frac{1}{r^2}|nlm\rangle\right) \\
&= -\frac{1}{2mc^2}\left(E_n^2 + 2E_n e^2 \frac{1}{a_0 n^2} + \left(e^2\right)^2 \frac{1}{a_0^2 n^3 (l+1/2)}\right) \\
&= -\frac{1}{2mc^2}\left(-3E_n^2 + 4E_n^2 \frac{n}{l+1/2}\right) \\
&= -\frac{1}{2mc^2}\frac{\alpha^4 m^2 c^4}{4n^4}\left(-3 + \frac{4n}{l+1/2}\right) \\
&= -\frac{\alpha^4 mc^2}{8n^4}\left(-3 + \frac{4n}{l+1/2}\right)
\end{aligned} \tag{734}$$

$$\begin{aligned}
H_{\rm LS} &= \frac{\alpha\hbar}{2m^2 r^3 c}\mathbf{S}\cdot\mathbf{L} \\
&= \frac{1}{2}\alpha\frac{\hbar^3 c^3}{m^2 c^4 a_0^3}\frac{a_0^3}{r^3}\frac{\mathbf{S}\cdot\mathbf{L}}{\hbar^2} \\
&= \alpha^4 (mc^2)\frac{a_0^3}{2r^3}\frac{\mathbf{S}\cdot\mathbf{L}}{\hbar^2}
\end{aligned} \tag{735}$$

$$\langle H_{\rm LS}\rangle_{njm} = \frac{\alpha^4 mc^2 a_0^3}{2}\langle 1/r^3\rangle_{nlm}(j(j+1)-l(l+1)-s(s+1))/2 \quad (736)$$

For each l we have $j = l \pm 1/2$. If $j = l + 1/2$

$$(l+1/2)(l+3/2) - l^2 - l - 3/4 = l \quad (737)$$

If $j = l - 1/2$

$$(l-1/2)(l+1/2) - l^2 - l - 3/4 = -l - 1 \quad (738)$$

So if $j = l + 1/2$,

$$\begin{aligned}\langle H_{\rm LS}\rangle_{njm} &= \frac{\alpha^4 mc^2}{4} l \frac{1}{n^3 l(l+1)(l+1/2)} \\ &= \frac{\alpha^4 mc^2}{4}\frac{1}{n^3(l+1)(l+1/2)} \\ &= \frac{\alpha^4 mc^2}{2n^3}\left(\frac{1}{l+1/2} - \frac{1}{l+1}\right) \end{aligned} \quad (739)$$

and if $j = l - 1/2$

$$\begin{aligned}\langle H_{\rm LS}\rangle_{njm} &= -\frac{\alpha^4 mc^2}{4}(l+1)\frac{1}{n^3 l(l+1)(l+1/2)} \\ &= -\frac{\alpha^4 mc^2}{4}\frac{1}{n^3 l(l+1/2)} \\ &= \frac{\alpha^4 mc^2}{2n^3}\left(\frac{1}{l+1/2} - \frac{1}{l}\right) \end{aligned} \quad (740)$$

So adding together, we get

$$\begin{aligned}\langle H_{\rm rel} + H_{LS}\rangle &= -\frac{\alpha^4 mc^2}{8n^4}\left(-3 + \frac{4n}{l+1/2}\right) + \frac{\alpha^4 mc^2}{2n^3}\left(\frac{1}{l+1/2} - \frac{1}{j+1/2}\right) \\ &= -\frac{\alpha^4 mc^2}{8n^4}\left(-3 + \frac{4n}{j+1/2}\right) \\ &= -|E_n|\frac{\alpha^2}{4n^2}\left(\frac{4n}{j+1/2} - 3\right) \end{aligned} \quad (741)$$

17 Stark Effect

As another example of using the degenerate perturbation theory, let's consider the Stark effect. This is the effect of a constant electric field on an atom, in our case, the hydrogen atom.

First, let's see what kind of potential energy this corresponds to. We know that

$$\mathbf{E} = -\nabla\Phi \tag{742}$$

where Φ is the electric potential. Since we have a constant electric field, say in the z direction, we can simply say

$$\Phi = |\mathbf{E}|\, z \tag{743}$$

Suppose we want to calculate the effect of this perturbation energy on the $n = 2$ energy level. In this level, we have $n^2 = 4$ degeneracy (we'll forget electron spin for a while). These are

$$|nlm\rangle = |2S_0\rangle, |2P_0\rangle, |2P_1\rangle, |2P_{-1}\rangle \tag{744}$$

We notice first that

$$\Phi = |\mathbf{E}|\, z = |\mathbf{E}|\, r\cos\theta = |\mathbf{E}|\, r\sqrt{\frac{4\pi}{3}}Y_{10}(\theta,\phi) \tag{745}$$

and it contains no ϕ dependence. Hence, the matrix element between states with different m vanishes. This means only the following and their hermitian conjugates can be non-zero:

$$\langle 1S_0|z|1S_0\rangle \tag{746}$$
$$\langle 1S_0|z|2S_0\rangle \tag{747}$$
$$\langle 2S_0|z|2S_0\rangle \tag{748}$$
$$\langle 2S_1|z|2S_1\rangle \tag{749}$$
$$\langle 2S_{-1}|z|2S_{-1}\rangle \tag{750}$$
$$\tag{751}$$

Now consider the classical parity operation,

$$\mathbf{x} \to -\mathbf{x} \tag{752}$$

The quantum version of this reads

$$\Pi^\dagger \mathbf{x} \Pi = -\mathbf{x} \tag{753}$$

where the operator Π is the partity operator. In polar coordinate, this means $\theta \to \theta + \pi$.

Now we know that the θ part of Y_{lm} is the Legendre polynomial P_l^m. Its definition is

$$P_l^m(\cos\theta) \;=\; \frac{1}{2^l l!}\,(-1)^m(1-\mu^2)^{m/2}\frac{d^{m+l}}{d\mu^{m+l}}(\mu^2-1)^l\bigg|_{\mu\to\cos\theta} \tag{754}$$

What's clear from this definition is that P_l^m is l-th order polynomial in $\sin\theta$ and $\cos\theta$ and furthermore, all terms should have l factors of sin and cos. In other words, Y_l^m can be represented by some combination of $x^{l-a-b}y^a z^b/r^l$. This implies that the effect of parity operation on Y_{lm} is

$$\Pi Y_{lm} = (-1)^l Y_{lm} \tag{755}$$

since under parity, $(x,y,z) \to -(x,y,z)$ but $r = \sqrt{x^2+y^2+z^2} \to r$.

Another way of seeing this: We have

$$L_\pm = L_x \pm iL_y = \frac{\hbar}{i}e^{\pm i\phi}\left(\pm i\partial_\theta - \cot\theta\partial_\phi\right) \tag{756}$$

and

$$\Theta_{l,l}(\theta) = A_{l,l}\sin^l\theta \tag{757}$$

so

$$Y_{l,l} = A_{l,l}e^{il\phi}\sin^l\theta \tag{758}$$

So

$$Y_{lm} = (\hat{L}_-)^{l-m}Y_{ll} \tag{759}$$

has l factors of $\sin\theta$ and $\cos\theta$ no matter what the value of m is.

Let's then consider

$$I = \langle nlm|z|nl'm'\rangle \tag{760}$$

We know that $z = -\Pi^\dagger z \Pi$. So,

$$I = -\langle nlm|\Pi^\dagger z \Pi|nl'm'\rangle \tag{761}$$

We also know $\Pi|nlm\rangle = (-1)^l|nlm\rangle$. So,

$$I = \langle nlm|z|nl'm'\rangle = -(-1)^{l+l'}\langle nlm|z|nl'm'\rangle \tag{762}$$

Hence, unless $l + l'$ is an odd integer, the matrix element is 0. This is called the 'dipole selection rule'. In our case, this means that all matrix element between $l = 2$ states vanish and $\langle 2S_0|z|2S_0\rangle$ vanish. In other words, the only non-vanishing matrix element is $\langle 2S_0|z|2P_0\rangle$ and its hermitian conjugate.

So, our matrix looks like

$$V_{\alpha\beta} = \begin{pmatrix} 0 & \Delta & 0 & 0 \\ \Delta^* & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{763}$$

where $\Delta = e\,|\mathbf{E}|\,\langle 2S_0|z|2P_0\rangle$. Here $\alpha = 1$ is $2S_0$ and $\alpha = 2$ is $2P_0$. The energy correction of course is given by

$$E^{(1)} = \pm\,|\Delta| \tag{764}$$

O.K. Let's then calculate Δ. We know that

$$\langle x|2S_0\rangle = \frac{2}{(2a_0)^{3/2}}(1 - r/2a_0)e^{-r/2a_0}Y_{00} = \frac{2}{(2a_0)^{3/2}}(1 - r/2a_0)e^{-r/2a_0}\sqrt{\frac{1}{4\pi}} \tag{765}$$

and

$$\langle x|2P_0\rangle = \frac{1}{\sqrt{3}(2a_0)^{3/2}}\frac{r}{a_0}e^{-r/2a_0}Y_{10} = \frac{1}{\sqrt{3}(2a_0)^{3/2}}\frac{r}{a_0}e^{-r/2a_0}\sqrt{\frac{3}{4\pi}}\cos\theta \tag{766}$$

So

$$\begin{aligned}
\langle 2S_0|z|2P_0\rangle &= \langle 2S_0|r\cos\theta|2P_0\rangle \\
&= \int dr\, r^3\,(1 - r/2a_0)(r/a_0)\,e^{-r/a_0}\frac{2}{\sqrt{3}(2a_0)^3}\frac{\sqrt{3}}{4\pi}\int_0^{2\pi} d\phi \int_{-1}^{1} d\cos\theta\cos^2\theta \\
&= \int dr\, r^3\,(1 - r/2a_0)(r/a_0)\,e^{-r/a_0}\frac{2}{\sqrt{3}(2a_0)^3}\frac{\sqrt{3}}{4\pi}(2\pi)\frac{2}{3} \\
&= \frac{a_0}{12}\left(4! - \frac{5!}{2}\right) \\
&= a_0\,(2 - 5) \\
&= -3a_0
\end{aligned} \tag{767}$$

Hence, as far as the matrix elements between $2S_0$ and $2P_0$ are concerned,

$$V_{\alpha\beta} = -3e\,|\mathbf{E}|\,a_0\sigma_x \tag{768}$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{769}$$

and the eigenvectors of this matrix is of course

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{770}$$

which corresponds to

$$|E_2 \pm \Delta\rangle = \frac{1}{\sqrt{2}}\left(|2S_0\rangle \mp |2P_0\rangle\right) \tag{771}$$

Now let's see if we can make a sense out of this result. First of all, why does the diagonal terms all vanish? Well, as I said, the perturbation energy here corresponds to non-vanishing electric dipole moment. Consider the $2S_0$ state for instance. It's radially symmetric state. Hence, naturally the dipole moment associated with the state vanishes. What about the $2P_0$ state? Well, the state is not radially symmetric. However, the density distribution of electron looks like this: Fig.10.17 in the book. Now, the value of wavefunction in the positive z direction is plus and in the negative z direction is minus. However, to get the probability, we need to square it. That means the probability for the electron to be on the positive z axis and the probability to be on the negative z axis is the same. Hence, no dipole moment.

No consider summing the two wavefunctions. The emphasis here is on the 'wavefunction', not the probability. If we sum the probability, there will be no difference between $+z$ and $-z$ again. However, if we sum the wavefunctions, there *is* a difference. Since we are subtracting something that's proportional to z from a radially symmetric function, the result is a lobsided funciton in the z direction. This is a fine example of quantum interference. Notice again that if we are only allowed to add probabilities, this will never happen. The sum of two symmetric (in the z direction) probability is again symmetric. This lobsided-ness only happens since we are only allowed to add the *wavefunctions* or *quantum amplitudes* not the probabilities.

18 Definition of Angular Momentum

Defining relation : Angular momentum operators satisfy

$$[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k \tag{772}$$

Or if you define

$$\mathcal{J} = \hat{\mathbf{J}}/\hbar \tag{773}$$

this becomes

$$[\mathcal{J}_i, \mathcal{J}_j] = i\epsilon_{ijk}\mathcal{J}_k \tag{774}$$

NOTATION ALERT

From now on, I'll identify $\mathcal{J}$ and $\hat{\mathbf{J}}$. That means I am not going to write $\hbar$ from now on except to make a few points. You should remember to supply them at the end. That is, just change $\hat{\mathbf{J}} \to \hat{\mathbf{J}}/\hbar$.

It doesn't matter what kind of angular momentum $\hat{\mathbf{J}}$ represents. It could be the orbital angular momentum, spin angular momentum, or any combination of them. If it has to do with an angular momentum, it must satisfy Eq.(772) or equivalently Eq.(774).

$\hat{\mathbf{J}}$ commutes with $\hat{\mathbf{J}}^2$: That is $[\hat{\mathbf{J}}^2, \hat{J}_i] = 0$ for all i:

$$\begin{aligned}
[\hat{\mathbf{J}}^2, \hat{J}_z] &= [\hat{J}_x^2, \hat{J}_z] + [\hat{J}_y^2, \hat{J}_z] \\
&= \hat{J}_x[\hat{J}_x, \hat{J}_z] + [\hat{J}_x, \hat{J}_z]\hat{J}_x + \hat{J}_y[\hat{J}_y, \hat{J}_z] + [\hat{J}_y, \hat{J}_z]\hat{J}_y \\
&= \hat{J}_x(-i\hat{J}_y) + (-i\hat{J}_y)\hat{J}_x + \hat{J}_y(i\hat{J}_x) + (i\hat{J}_x)\hat{J}_y = 0
\end{aligned} \tag{775}$$

Now there is nothing particular about z-direction. So it must be true that $[\hat{\mathbf{J}}^2, \hat{J}_i] = 0$ for all i.

Raising and Lowering operators :

It should be familiar by now:

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y \tag{776}$$

From the last point,

$$[\hat{\mathbf{J}}^2, \hat{J}_\pm] = 0 \tag{777}$$

is obvious. Also

$$\begin{aligned}
[\hat{J}_z, \hat{J}_\pm] &= [\hat{J}_z, \hat{J}_x \pm i\hat{J}_y] \\
&= [\hat{J}_z, \hat{J}_x] \pm i[\hat{J}_z, \hat{J}_y] \\
&= i\hat{J}_y \pm i(-i\hat{J}_x) \\
&= \pm(\hat{J}_x \pm i\hat{J}_y) \\
&= \pm\hat{J}_\pm
\end{aligned} \tag{778}$$

Possible eigenvalues of $\hat{J}_z$:

First of all, since $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ commutes, we can specify the eigenvalues of both. Let $|j, m\rangle$ be the corresponding eigenstates with

$$\hat{J}_z|j, m\rangle = m|j, m\rangle \tag{779}$$

We'll get j later.

Now consider

$$\begin{aligned}
\hat{J}_z\hat{J}_\pm|j, m\rangle &= \hat{J}_\pm\hat{J}_z|j, m\rangle \pm \hat{J}_\pm|j, m\rangle \\
&= (m \pm 1)\hat{J}_\pm|j, m\rangle
\end{aligned} \tag{780}$$

Therefore $\hat{J}_\pm|j, m\rangle$ is also an eigenstate of $\hat{J}_z$ and it has the eigenvalue of $m \pm 1$. Since $[\hat{\mathbf{J}}^2, \hat{J}_\pm] = 0$, the value of j does not change.

Now redefine

$$\hat{J}'_x = \hat{J}_x \tag{781}$$

$$\hat{J}'_y = -\hat{J}_y \tag{782}$$

$$\hat{J}'_z = -\hat{J}_z \tag{783}$$

Then it is easy to check that

$$\hat{\mathbf{J}}'^2 = \hat{\mathbf{J}}^2 \tag{784}$$

and

$$[\hat{J}'_x, \hat{J}'_y] = i\hat{J}'_z \tag{785}$$

$$[\hat{J}'_y, \hat{J}'_z] = i\hat{J}'_x \tag{786}$$

$$[\hat{J}'_z, \hat{J}'_x] = i\hat{J}'_y \tag{787}$$

However, the previous state $|j, m\rangle$ now satisfies

$$\hat{J}'_z|j, m\rangle = -mj, \hat{m} \tag{788}$$

Now there is nothing particular about the coordinate system (x, y, z) and $(x, -y, -z)$ as long as we maintain the right-handedness (and we did). So if m is an eigenvalue of $\hat{J}_z$, then so should $-m$.

So combined with the fact that all m's are related by an integer, this means

$$m = -m_{\max}, -m_{\max} + 1, -m_{\max} + 2, \cdots, m_{\max} - 2, m_{\max} - 1, m_{\max} \tag{789}$$

where $m_{\max}$ is the maximum value of m.

So what are the possible values of $m_{\max}$? Well, so far the only condition we can say is

$$2m_{\max} = \text{positive integer} \tag{790}$$

so

$$m_{\max} = \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \cdots \tag{791}$$

Note the appearance of half the integers. There is nothing in our argument so far that prevents them and there will be none to come. So they must be allowed. This is a pure mathematical statement coming from the commutation relationship Eq.(774). Amazing thing is, the nature *does* allow them.

Possible values of $\hat{\mathbf{J}}^2$: Note first

$$\begin{aligned} [\hat{J}_+, \hat{J}_-] &= [\hat{J}_x + i\hat{J}_y, \hat{J}_x - i\hat{J}_y] \\ &= -i[\hat{J}_x, \hat{J}_y] + i[\hat{J}_y, \hat{J}_x] \\ &= -ii\hat{J}_z + i(-i\hat{J}_z) \\ &= 2\hat{J}_z \end{aligned} \tag{792}$$

and

$$\begin{aligned}\hat{J}_+\hat{J}_- &= (\hat{J}_x + i\hat{J}_y)(\hat{J}_x - i\hat{J}_y) \\ &= \hat{J}_x^2 + \hat{J}_y^2 - i\hat{J}_x\hat{J}_y + i\hat{J}_y\hat{J}_x \\ &= \hat{J}_x^2 + \hat{J}_y^2 - i[\hat{J}_x, \hat{J}_y] \\ &= \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z \end{aligned} \tag{793}$$

also

$$\begin{aligned}\hat{J}_-\hat{J}_+ &= (\hat{J}_x - i\hat{J}_y)(\hat{J}_x + i\hat{J}_y) \\ &= \hat{J}_x^2 + \hat{J}_y^2 + i\hat{J}_x\hat{J}_y - i\hat{J}_y\hat{J}_x \\ &= \hat{J}_x^2 + \hat{J}_y^2 + i[\hat{J}_x, \hat{J}_y] \\ &= \hat{J}_x^2 + \hat{J}_y^2 - \hat{J}_z \end{aligned} \tag{794}$$

Hence

$$\hat{\mathbf{J}}^2 = \hat{J}_-\hat{J}_+ + \hat{J}_z^2 + \hat{J}_z \tag{795}$$

Now consider the maximum m state $|j, m_{\max}\rangle$. We know that we can't raise its m value any more. Therefore

$$\hat{\mathbf{J}}^2|j, m_{\max}\rangle = m_{\max}(m_{\max}+1)|j, m_{\max}\rangle \tag{796}$$

Since $\hat{J}_z$ and $\hat{\mathbf{J}}^2$ commute, all other m state must also give

$$\hat{\mathbf{J}}^2|j, m\rangle = m_{\max}(m_{\max}+1)|j, m\rangle \tag{797}$$

So we may use $m_{\max}$ to label the possible eigenvalues of $\hat{\mathbf{J}}^2$.

So the possible values of $\hat{\mathbf{J}}^2$ are:

$$\begin{aligned}\frac{3}{4} &= \left(\frac{1}{2}\right)\left(1+\frac{1}{2}\right) \\ 2 &= 1(1+1) \\ \frac{15}{4} &= \frac{3}{2}\left(1+\frac{3}{2}\right) \\ 5 &= 2(2+1) \\ &\vdots \end{aligned} \tag{798}$$

Normalization of $\hat{J}_\pm|j,m\rangle$: Use

$$\hat{J}_-\hat{J}_+ = \hat{\mathbf{J}}^2 - \hat{J}_z^2 - \hat{J}_z \tag{799}$$

so let

$$\hat{J}_+|j,m\rangle = c_+|j,m+1\rangle \tag{800}$$

$$\begin{aligned}\langle j,m|\hat{J}_-\hat{J}_+|j,m\rangle &= |c_+|^2 \\ &= j(j+1) - m(m+1)\end{aligned} \tag{801}$$

where we have used $(\hat{J}_-)^\dagger = \hat{J}_+$. Hence (similar argument works for $\hat{J}_-|j,m\rangle$)

$$\hat{J}_\pm|j,m\rangle = \sqrt{j(j+1) - m(m\pm 1)}|j,m\pm 1\rangle \tag{802}$$

Again, just remember that you can't raze $m = j$ any more and can't lower $m = -j$ any more.

19 Matrix Representation of J for Fixed j

$j = 1/2$

We abbreviate: $|m\rangle \equiv |j,m\rangle$. For the z component

$$S^z_{mn} = \langle m|\hat{S}_z|n\rangle = m\delta_{mn} \tag{803}$$

or the matrix S^z is

$$S^z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{804}$$

For the matrix S^x, use

$$\hat{S}_x = (\hat{S}_+ + \hat{S}_-)/2 \tag{805}$$

so that

$$\begin{aligned}\langle +1/2|\hat{S}_x|+1/2\rangle &= 0 \\ \langle +1/2|\hat{S}_x|-1/2\rangle &= \langle +1/2|\hat{S}_+|-1/2\rangle/2 \\ &= 1/2 \\ \langle -1/2|\hat{S}_x|+1/2\rangle &= \langle -1/2|\hat{S}_-|+1/2\rangle/2 \\ &= 1/2 \\ \langle -1/2|\hat{S}_x|-1/2\rangle &= 0\end{aligned} \tag{806}$$

using

$$\sqrt{j(j+1)-m(m-1)} = \sqrt{3/4+(1/2)(1/2)} = 1 \tag{807}$$

for $m = 1/2$ and

$$\sqrt{j(j+1)-m(m+1)} = \sqrt{3/4+(1/2)(1/2)} = 1 \tag{808}$$

for $m = -1/2$. So the matrix S^x is

$$S^x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{809}$$

For the matrix S^y, use

$$\hat{S}_y = (\hat{S}_+ - \hat{S}_-)/2i \tag{810}$$

so that

$$\begin{aligned} \langle +1/2|\hat{S}_y|+1/2\rangle &= 0 \\ \langle +1/2|\hat{S}_y|-1/2\rangle &= (-i/2)\langle +1/2|\hat{S}_-|-1/2\rangle \\ &= -i/2 \\ \langle -1/2|\hat{S}_y|+1/2\rangle &= (i/2)\langle -1/2|\hat{S}_-|+1/2\rangle \\ &= i/2 \\ \langle -1/2|\hat{S}_y|-1/2\rangle &= 0 \end{aligned} \tag{811}$$

So the matrix S^y is

$$S^y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{812}$$

Or

$$S^i = \frac{1}{2}\sigma_i \tag{813}$$

where σ_i are the Pauli matrices.

It is easy to check that the matrix themselve satisfy

$$[S_i, S_j] = S_iS_j - S_jS_i = i\epsilon_{ijk}S_k \tag{814}$$

$j = 1$

Obviously,

$$S_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{815}$$

Let's get $S_x = (S_+ + S_-)/2$. Let's do it this way. For S_+, non-zero elements are

$$\begin{aligned} \langle 1|\hat{S}_+|0\rangle &= \sqrt{1(1+1) - 0(0+1)} = \sqrt{2} && (816) \\ \langle 0|\hat{S}_+|-1\rangle &= \sqrt{1(1+1) - (-1)((-1)+1)} = \sqrt{2} && (817) \end{aligned}$$

or

$$S_+ = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \tag{818}$$

For S_-, non-zero elements are

$$\begin{aligned} \langle 0|\hat{S}_-|1\rangle &= \sqrt{1(1+1) - 1(1-1)} = \sqrt{2} && (819) \\ \langle -1|\hat{S}_-|0\rangle &= \sqrt{1(1+1) - 0(0-1)} = \sqrt{2} && (820) \end{aligned}$$

or

$$S_- = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \tag{821}$$

So

$$\begin{aligned} S_x &= (S_+ + S_-)/2 \\ &= \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} && (822) \end{aligned}$$

and

$$\begin{aligned} S_y &= (S_+ - S_-)/2i \\ &= \begin{pmatrix} 0 & -\frac{i}{\sqrt{2}} & 0 \\ \frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ 0 & \frac{i}{\sqrt{2}} & 0 \end{pmatrix} \end{aligned} \tag{823}$$

Again, it is easy to check that these satisfy the angular momentum commutation relation.

20 Adding Angular Momenta

Quantum Sum : If we have two angular momenta $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$,

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}} \tag{824}$$

or

$$\hat{J}_i = \hat{L}_i + \hat{S}_i \tag{825}$$

just like the classical case.

Operators : The most important fact is that measuring one angular momentum does not interfere with measuring another:

$$[\hat{L}_i, \hat{S}_j] = 0 \tag{826}$$

for all i and j.

This means that we can specify both l, m_l and s, m_s at the same time.

We write the eigenstate as

$$|l, m_l; s, m_s\rangle \equiv |l, m_l\rangle \otimes |s, m_s\rangle \tag{827}$$

where the symbol $\otimes$ is the 'direct product'. It is used to join two independent vector space. In this case one spanned by the eigenstates of $\hat{\mathbf{L}}$ with the eigenvalue of $\hat{\mathbf{L}}^2$ fixed at $l(l+1)$ and another spanned by the eigenstates of $\hat{\mathbf{S}}$ with the eigenvalue of $\hat{\mathbf{S}}^2$ fixed at $s(s+1)$. Then the operator $\hat{L}_i$ acts only on $|l, m_l\rangle$ part and the operator $\hat{S}_i$ acts only on $|s, m_s\rangle$ part.

Commutation Relation : Since they do not interfere with each other, the sum also satisfies

$$[\hat{J}_i, \hat{J}_j] = i\epsilon_{ijk}\hat{J}_k \tag{828}$$

That means we can define

$$\hat{\mathbf{J}}^2 = (\hat{\mathbf{L}} + \hat{\mathbf{S}})^2 \tag{829}$$

$$\hat{J}_z = \hat{L}_z + \hat{S}_z \tag{830}$$

and

$$\hat{J}_\pm = \hat{L}_\pm + \hat{S}_\pm \tag{831}$$

and all the things we have shown in the previous section are valid.

Consider commutation between $\hat{\mathbf{J}}^2$ and $\hat{\mathbf{L}}^2$ and $\hat{\mathbf{S}}^2$. First of all,

$$[\hat{\mathbf{L}}^2, \hat{\mathbf{S}}^2] = 0 \tag{832}$$

Now

$$\hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}}\cdot\hat{\mathbf{S}} \tag{833}$$

Then since $[\hat{\mathbf{L}}, \hat{\mathbf{L}}^2] = 0$,

$$[\hat{\mathbf{J}}^2, \hat{\mathbf{L}}^2] = 0 \tag{834}$$

Similarly,

$$[\hat{\mathbf{J}}^2, \hat{\mathbf{S}}^2] = 0 \tag{835}$$

Also it is easy to see that

$$[\hat{\mathbf{L}}^2, \hat{J}_z] = 0 \tag{836}$$

and

$$[\hat{\mathbf{S}}^2, \hat{J}_z] = 0 \tag{837}$$

since $[\hat{\mathbf{L}}^2, \hat{L}_z] = [\hat{\mathbf{S}}^2, \hat{L}_z] = 0$ and $[\hat{\mathbf{L}}^2, \hat{S}_z] = [\hat{\mathbf{S}}^2, \hat{S}_z] = 0$.

So instead of labeling the eigenstate with $lm_l; sm_s$ we can use

$$|l, s; j, m_j\rangle_j \tag{838}$$

where m_j is the eigenvalue of $\hat{J}_z$. I am going to put the subscript j (think of it as 'joint') on states like this to avoid confusion with $|l, m_l; s, m_s\rangle$ states.

Possible eigenvalues of $\hat{J}_z$: The answer is simple. Since $\hat{J}_z = \hat{L}_z + \hat{S}_z$, we must get all possible $m_l + m_s$.

Highest eigenvalue of $\hat{\mathbf{J}}^2$: Consider the $m_l = l$ and $m_s = s$ state:

$$|l, l; s, s\rangle \tag{839}$$

Apply $\hat{J}_z$ to get

$$\hat{J}_z|l, l; s, s\rangle = (l + s)|l, l; s, s\rangle \tag{840}$$

Apply $\hat{\mathbf{J}}^2$ to get

$$\begin{aligned} \hat{\mathbf{J}}^2|l, l; s, s\rangle &= (\hat{J}_-\hat{J}_+ + \hat{J}_z^2 + \hat{J}_z)|l, l; s, s\rangle \\ &= (l + s)(l + s + 1)|l, l; s, s\rangle \end{aligned} \tag{841}$$

The last line follows because $\hat{J}_+ = \hat{L}_+ + \hat{S}_+$ annihilates $|l, l; s, s\rangle$. Therefore

$$|l, l; s, s\rangle = |l, s; l + s, l + s\rangle_j \tag{842}$$

that is, $j = l + s$ and $m_j = l + s$. Obviously, this is the highest j state and highest m_j state that can be formed from $|l, m_l\rangle$ and $|s, m_s\rangle$.

All other m_j states with this $j = l + s$ can be then formed by applying $\hat{J}_-$. For instance

$$\begin{aligned} |l, s; l+s, l+s-1\rangle_j &= \frac{1}{\sqrt{(l + s)(l + s + 1) - (l + s)(l + s - 1)}}\hat{J}_-|l, l\rangle \otimes |s, s\rangle \\ &= \frac{1}{\sqrt{2(l + s)}}\left((\hat{L}_-|l, l\rangle) \otimes |s, s\rangle + |l, l\rangle \otimes (\hat{S}_-|s, s\rangle)\right) \\ &= \frac{1}{\sqrt{2(l + s)}}\left(\sqrt{2l}|l, l-1\rangle \otimes |s, s\rangle + \sqrt{2s}|l, l\rangle \otimes |s, s-1\rangle\right) \end{aligned} \tag{843}$$

Next highest : O.K. Fine. But what about the combination

$$|\psi\rangle = \frac{1}{\sqrt{2(l + s)}}\left(\sqrt{2s}|l, l-1\rangle \otimes |s, s\rangle - \sqrt{2l}|l, l\rangle \otimes |s, s-1\rangle\right) \tag{844}$$

which can be easily shown to be orthogonal to $|l, s; l + s, l + s - 1\rangle_j$?

Remember, $\langle l, m_l|$ only reacts with $|l', m_l'\rangle$ and $\langle s, m_s|$ only with $|s', m_s'\rangle$.

Obviously,

$$\hat{J}_z|\psi\rangle = (l + s - 1)|\psi\rangle \tag{845}$$

But also

$$\begin{aligned}
\hat{J}_+|\psi\rangle &= (\hat{L}_+ + \hat{S}_+)\frac{1}{\sqrt{2(l+s)}}\left(\sqrt{2s}|l, l{-}1\rangle \otimes |s, s\rangle - \sqrt{2l}|l, l\rangle \otimes |s, s{-}1\rangle\right) \\
&= \frac{1}{\sqrt{2(l+s)}}\Bigg(\sqrt{2s}\sqrt{l(l+1) - (l-1)l}|l, l\rangle \otimes |s, s\rangle \\
&\qquad - \sqrt{2l}\sqrt{s(s+1) - (s-1)s}|l, l\rangle \otimes |s, s\rangle\Bigg) \\
&= 0
\end{aligned} \tag{846}$$

That means

$$\hat{\mathbf{J}}^2|\psi\rangle = (\hat{\mathbf{J}}_-\hat{\mathbf{J}}_+ + \hat{J}_z^2 + \hat{J}_z)|\psi\rangle = (l + s - 1)(l + s)|\psi\rangle \tag{847}$$

which means

$$|\psi\rangle = |l, s; l{+}s{-}1, l{+}s{-}1\rangle_j \tag{848}$$

and all other $|l, s; l{+}s{-}1, m_j\rangle_j$ can be obtained by applying $\hat{J}_-$.

$$\begin{aligned}
&\hat{J}_-\hat{J}_+|l, m_l; s, m_s\rangle = \\
&\quad \hat{J}_-\Big(\sqrt{l(l+1) - m_l(m_l+1)}|l, m_l+1; s, m_s\rangle + \sqrt{s(s+1) - m_s(m_s+1)}|l, m_l; s, m_s+1\rangle\Big) \\
&= \Big[\sqrt{l(l+1) - m_l(m_l+1)}\sqrt{l(l+1) - (m_l+1)m_l}|l, m_l; s, m_s\rangle \\
&\qquad + \sqrt{l(l+1) - m_l(m_l+1)}\sqrt{s(s+1) - m_s(m_s-1)}|l, m_l+1; s, m_s-1\rangle \\
&\qquad + \sqrt{s(s+1) - m_s(m_s+1)}\sqrt{l(l-1) - m_l(m_l-1)}|l, m_l-1; s, m_s+1\rangle \\
&\qquad + \sqrt{s(s+1) - m_s(m_s+1)}\sqrt{s(s+1) - (m_s+1)m_s}|l, m_l; s, m_s\rangle\Big]
\end{aligned} \tag{849}$$

Degrees of Freedom : There are $2l+1$ $|l, m_l\rangle$ states and $2s+1$ $|s, m_s\rangle$ states. So there must be $(2l+1)(2s+1)$ number of states in the joint states of $|l, s; j, m_j\rangle$. Let's check. We have stated that the possible states are $j = l+s, \cdots, |l-s|$. So

$$\begin{aligned}
\text{d.o.f.} &= \sum_{j=|l-s|}^{l+s} \sum_{m_j=-j}^{j} 1 \\
&= \sum_{j=|l-s|}^{l+s} (2j+1) \\
&= 2\left(\frac{(l+s)(l+s+1)}{2} - \frac{(|l-s|-1)|l-s|}{2}\right) + l + s - |l-s| + 1 \\
&= (2l+1)(2s+1) \qquad (850)
\end{aligned}$$

regardless of the sign of $l-s$.

Lesson I : The highest j is $l+s$ and the next highest is $l+s-1$. This process can be continued to show that all values from $l+s$ and $|l-s|$ are possible.

Clebsch-Gordon Coefficients : The overlap function

$$\langle l, m_l; s, m_s | l, s; j, m_j \rangle_j \qquad (851)$$

is the **Clebsch-Gordon Coefficient**.

Let's have an example. Consider adding 2 spin 1/2.

$$\begin{aligned}
|1/2, 1/2; 1/2, 1/2\rangle &= |1/2, 1/2\rangle \otimes |1/2, 1/2\rangle, \\
|1/2, 1/2; 1/2, -1/2\rangle &= |1/2, 1/2\rangle \otimes |1/2, -1/2\rangle, \\
|1/2, -1/2; 1/2, 1/2\rangle &= |1/2, -1/2\rangle \otimes |1/2, 1/2\rangle, \\
|1/2, -1/2; 1/2, -1/2\rangle &= |1/2, -1/2\rangle \otimes |1/2, -1/2\rangle, \qquad (852)
\end{aligned}$$

The total spin is

$$\hat{\mathbf{S}}_{\text{total}} = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2 \qquad (853)$$

The heighest j and m_j state is

$$|1/2, 1/2; 1, 1\rangle = |1/2, 1/2; 1/2, 1/2\rangle \qquad (854)$$

Taking $\hat{S}_- = \hat{S}^1_- + \hat{S}^2_-$ yields

$$\begin{aligned} |1/2,1/2;1,0\rangle &= \frac{1}{\sqrt{2-1(1-1)}}\hat{S}_-|1,1\rangle \\ &= \frac{1}{\sqrt{2}}\Big(|1/2,-1/2;1/2,1/2\rangle + |1/2,1/2;1/2,-1/2\rangle\Big) \end{aligned} \tag{855}$$

Taking it once more

$$\begin{aligned} |1/2,1/2;1,-1\rangle &= \frac{1}{\sqrt{2-0(1-1)}}\hat{S}_-|1,0\rangle \\ &= |1/2,-1/2;1/2,-1/2\rangle \end{aligned} \tag{856}$$

The state

$$|1/2,1/2;0,0\rangle = \frac{1}{\sqrt{2}}\Big(|1/2,-1/2;1/2,1/2\rangle - |1/2,1/2;1/2,-1/2\rangle\Big) \tag{857}$$

is orthogonal to all the above combination and it is easy to show that it has $j=0$.

So the non-zero Clebsch-Gordan coefficients are

$$\begin{aligned} \langle 1/2,1/2;1/2,1/2|1/2,1/2;1,1\rangle &= 1 \\ \langle 1/2,1/2;1/2,-1/2|1/2,1/2;1,0\rangle &= 1/\sqrt{2} \\ \langle 1/2,1/2;-1/2,1/2|1/2,1/2;1,0\rangle &= 1/\sqrt{2} \\ \langle 1/2,1/2;1/2,-1/2|1/2,1/2;0,0\rangle &= 1/\sqrt{2} \\ \langle 1/2,1/2;-1/2,1/2|1/2,1/2;0,0\rangle &= -1/\sqrt{2} \end{aligned} \tag{858}$$

Let's do one more example. Consider adding 1/2 to 1. Then,

$$\begin{aligned} |1,1/2;3/2,3/2\rangle &= |1,1;1/2,1/2\rangle \\ \hat{J}_-|1,1/2;3/2,1/2\rangle &= \frac{1}{\sqrt{(3/2)(1+3/2)-3/2(3/2-1)}}\hat{J}_-|1,1/2;3/2,3/2\rangle \\ &= \frac{1}{\sqrt{3}}\left(\sqrt{1(1+1)-1(1-1)}|1,0;1/2,1/2\rangle + \sqrt{(1/2)(1+1/2)-(1/2)(1/2-\right. \\ &= \frac{1}{\sqrt{3}}\Big(\sqrt{2}|1,0;1/2,1/2\rangle + |1,1;1/2,-1/2\rangle\Big) \end{aligned}$$

$$
\begin{aligned}
|1,1/2;3/2,-1/2\rangle &= \frac{1}{\sqrt{(3/2)(1+3/2)-1/2(1/2-1)}}\hat{J}_-|1,1/2;3/2,1/2\rangle \\
&= \frac{1}{2\sqrt{3}}\left(\sqrt{2}\sqrt{2}|1,-1;1/2,1/2\rangle + 2\sqrt{2}|1,0;1/2,-1/2\rangle\right) \\
&= \frac{1}{\sqrt{3}}\left(|1,-1;1/2,1/2\rangle + \sqrt{2}|1,0;1/2,-1/2\rangle\right) \\
|1,1/2;3/2,-3/2\rangle &= |1,-1;1/2,-1/2\rangle
\end{aligned}
$$

The other j state is $j = 1/2$ state. They are

$$
|1,1/2;1/2,1/2\rangle = \frac{1}{\sqrt{3}}\left(|1,0;1/2,1/2\rangle - \sqrt{2}|1,1;1/2,-1/2\rangle\right) \tag{860}
$$

$$
\begin{aligned}
|1,1/2;1/2,-1/2\rangle &= \hat{J}_-|1,1/2;1/2,1/2\rangle \\
&= \frac{1}{\sqrt{3}}\left(\sqrt{2}|1,-1;1/2,1/2\rangle + |1,0;1/2,-1/2\rangle - 2|1,0;1/2,-1/2\rangle\right) \\
&= \frac{1}{\sqrt{3}}\left(\sqrt{2}|1,-1;1/2,1/2\rangle - |1,0;1/2,-1/2\rangle\right) \qquad (861)
\end{aligned}
$$

Clebsch-Gordan coefficients can be easily taken from these expresions.

21 Zeeman Effect

Stark effect is due to the fact that electon-proton system can be polarized by Electric field. Therefore, what's important is actually the induced electric dipole moment. Remember that

$$
\hat{H}_{\text{pert}} = e|\mathbf{E}|\hat{z} \tag{862}
$$

Correct the factor e**.** and we found that the zero-th order wavefunctions should be organized as

$$
|2,1,1\rangle, |2,1,-1\rangle, \tag{863}
$$

$$
|\psi_-\rangle = \frac{1}{\sqrt{2}}\left(|2,0,0\rangle - |2,1,0\rangle\right) \tag{864}
$$

$$
|\psi_+\rangle = \frac{1}{\sqrt{2}}\left(|2,0,0\rangle + |2,1,0\rangle\right) \tag{865}
$$

with the corresponding corrections

$$E_2^{(1)} = 0, 0, 3e|E|a_0, -3e|E|a_0 \tag{866}$$

That means that

$$\langle\psi_-|z|\psi_+\rangle = 3a_0 \tag{867}$$
$$\langle\psi_+|z|\psi_-\rangle = -3a_0 \tag{868}$$

Remember that the dipole moment is defined to be

$$\mathbf{d} = |q|\mathbf{r} \tag{869}$$

where the vector $\mathbf{r}$ is defined to be the position of the *positive charge* from the *negative charge*. Therefore, in this case, $|\psi_-\rangle$ state corresponds to

$$\mathbf{d} = -3a_0 e\mathbf{k} \tag{870}$$

and $|\psi_+\rangle$ state corresponds to

$$\mathbf{d} = 3a_0 e\mathbf{k} \tag{871}$$

This is an induced electric dipole moment.

Now consider having Magnetic field. You don't have induce anything for an atom to interact with Magnetic field. A charge particle with a non-zero angular momentum, by definition constitutes a current loop and therefore a magnetic dipole moment. For the orbital angular momentum the relationship is

$$\boldsymbol{\mu}_l = \frac{q}{2mc}\mathbf{L} \tag{872}$$

For the spin, this is modified

$$\boldsymbol{\mu}_s = g\frac{q}{2mc}\mathbf{S} \tag{873}$$

the modification factor is 2 for the electron and about 2.8 for the proton.

So the unperturbed Hamiltonian

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} - \frac{e^2}{r} \tag{874}$$

is modified by

$$\begin{aligned}\hat{H}_{\text{Zeeman}} &= -\left(\hat{\boldsymbol{\mu}}_l + \hat{\boldsymbol{\mu}}_s\right)\cdot\mathbf{B} \\ &= \left(\hat{L}_z + 2\hat{S}_z\right)\frac{eB}{2m_e c}\end{aligned} \tag{875}$$

where I have chosen

$$\mathbf{B} = B\mathbf{e}_z \tag{876}$$

and used the fact that $q = -e$ for the electron. This, of course, we can write as

$$\hat{H}_{\text{Zeeman}} = \left(\hat{J}_z + \hat{S}_z\right)\frac{eB}{2m_e c} \tag{877}$$

Now ask youself: Is

$$|ls; jm_j\rangle \tag{878}$$

an eigenstate of this perturbing hamiltonian?

Answer is no. This is because we have choosen our mutually commuting operators to be $\mathbf{L}^2, \mathbf{S}^2, \mathbf{J}^2$ and $\mathbf{J}^z$. $\hat{S}_z$ does not commute with $\mathbf{J}^2$ due to the $\mathbf{L}\cdot\mathbf{S}$ term.

Is that O.K.? Well, sure. It just means that we gonna have to use perturbation theory.

But wait a minute. Is it really impossible to diagonalize both the unperturbed and the perturbing Hamiltonians?

Well, question is whether

$$[\hat{H}_0, \hat{H}_{\text{Zeeman}}] = 0 \tag{879}$$

Since B is constant, this is certainly true. That is, we should be able to find common eigenstates. This can be done by specifying the values of

$$\hat{L}^2, \hat{L}_z, \hat{S}^2, \hat{S}_z \tag{880}$$

instead of l, s, j, m_j.

So why bother?

Well, this is because our choice of unperturbed Hamiltonian $\hat{H}_0$ above is a little too naive.

Why did we bother with l, s, j, m_j any way? Well, that because we wanted to calculate the spin-orbit coupling which went like

$$\mathbf{L} \cdot \mathbf{S} = (\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2)/2 \tag{881}$$

So specifying j, l, s, m_j has its advantage.

Question is: Which effect is bigger? If the Zeeman effect is smaller than the spin-orbit coupling (and also of course the relativistic correction), then we must conisider the unperturbed Hamiltonian to be

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} - \frac{Ze^2}{r} - \frac{\hat{\mathbf{p}}^4}{8m^3c^2} + \frac{Ze^2}{2m^2c^2r^3}\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \tag{882}$$

Since most of the energy is determined by j, l, s, m_j, expressing the unperturbed eigenstate with l, m, s, m_s is possible but awkward and cumbersome. So that's why we are going to do *perturbation* theory calculation when it looks as though we can calculate the exact values.

But this should also gives you the following point: If the magnetic field is big enough to overcome the spin-orbit coupling, then it becomes advantageous to use l, m, s, m_s basis instead of j, l, s, m basis. In that case, the spin orbit coupling becomes perturbation.

When can this happen? How big the magnetic field should be? Well, let's estimate.

We know that

$$E_{s.o} \sim mc^2\alpha^4 \tag{883}$$

and

$$E_{\text{Zeeman}} \sim \frac{eB\hbar}{mc} \tag{884}$$

so the two becomes of a similar strength when

$$\begin{aligned} B &\sim \frac{m^2c^3\alpha^4}{e\hbar} \\ &= \frac{m^2c^4\alpha^4}{e(\hbar c)} \\ &= \frac{(mc^2\alpha^2)^2}{\hbar c}\frac{1}{e} \\ &= \frac{(10\,\text{eV})^2}{200\,\text{nm}\,\text{eV}}\frac{1}{e} \\ &\sim \frac{\text{V}}{\text{nm}} \end{aligned} \tag{885}$$

Magnetic field is measured by the Tesla in MKS unit

$$1\text{Tesla} = \frac{\text{V sec}}{\text{m}^2} \tag{886}$$

so we need to multiply the avove with $1/c$ to get it in terms of Tesla

$$B \sim \frac{V}{10^{-9}\,\text{m}}\frac{1}{3\times 10^8 \text{m/s}} \sim 1\text{T} - 10\text{T} \tag{887}$$

So if $B \ll 1\text{T}$, we take $\hat{H}_{\text{Zeeman}}$ to be the perturbation and calculate

$$\langle ls; jm_j|\hat{H}_{\text{Zeeman}}|ls; jm_j\rangle \;\; = \;\; m_j\frac{eB\hbar}{2m_e c} + \langle ls; jm_j|\hat{S}_z|ls; jm_j\rangle\frac{eB}{2m_e c} \tag{888}$$

Now what should $|ls; jm_j\rangle$ be?

Well, since $s = 1/2$, there can be only 4 relevant combinations:

$$|l, m\rangle \otimes |+1/2\rangle, |l, m+1\rangle \otimes |-1/2\rangle \tag{889}$$

both gives $m_j = m + 1/2$.

$$|l, m\rangle \otimes |-1/2\rangle, |l, m-1\rangle \otimes |+1/2\rangle \tag{890}$$

both gives $m_j = m - 1/2$.

So consider the linear combinations

$$|\psi_+\rangle = \alpha_+|l, m\rangle \otimes |+1/2\rangle + \beta_+|l, m+1\rangle \otimes |-1/2\rangle \tag{891}$$

and

$$|\psi_-\rangle = \alpha_-|l, m\rangle \otimes |-1/2\rangle + \beta_-|l, m-1\rangle \otimes |+1/2\rangle \tag{892}$$

Applying $\hat{J}_-\hat{J}_+$ we get

$$\begin{aligned}
\hat{J}_-\hat{J}_+|\psi_+\rangle \;\; &= \;\; \hat{J}_-\Big[\alpha_+\hat{J}_+|l, m\rangle \otimes |+1/2\rangle + \beta_+\hat{J}_+|l, m+1\rangle \otimes |-1/2\rangle\Big] \\
&= \;\; \hat{J}_-\Big[\alpha_+\sqrt{l(l+1) - m(m+1)}|l, m+1\rangle \otimes |+1/2\rangle \\
&\qquad + \beta_+\sqrt{l(l+1) - (m+1)(m+2)}|l, m+2\rangle \otimes |-1/2\rangle
\end{aligned}$$

$$
\begin{aligned}
&\quad + \beta_+ |l, m+1\rangle \otimes |+1/2\rangle \Big] \\
&= \Big[\alpha_+ (l(l+1) - m(m+1)) |l, m\rangle \otimes |+1/2\rangle \\
&\quad \alpha_+ \sqrt{l(l+1) - m(m+1)} |l, m+1\rangle \otimes |-1/2\rangle \\
&\quad + \beta_+ (l(l+1) - (m+1)(m+2)) |l, m+1\rangle \otimes |-1/2\rangle \\
&\quad + \beta_+ \sqrt{l(l+1) - (m+1)m} |l, m\rangle \otimes |+1/2\rangle \\
&\quad + \beta_+ |l, m+1\rangle \otimes |-1/2\rangle \Big] \\
&= \Big[\alpha_+ (l(l+1) - m(m+1)) |l, m\rangle \otimes |+1/2\rangle \\
&\quad + \beta_+ \sqrt{l(l+1) - (m+1)m} |l, m\rangle \otimes |+1/2\rangle \\
&\quad \alpha_+ \sqrt{l(l+1) - m(m+1)} |l, m+1\rangle \otimes |-1/2\rangle \\
&\quad + \beta_+ (l(l+1) - (m+1)(m+2)) |l, m+1\rangle \otimes |-1/2\rangle \\
&\quad + \beta_+ |l, m+1\rangle \otimes |-1/2\rangle \Big] \qquad (893)
\end{aligned}
$$

And

$$
\begin{aligned}
\left(\hat{J}_z^2 + \hat{J}_z \right) |\psi_+\rangle &= (m + 1/2)(m + 3/2) |\psi_+\rangle \\
&= m^2 + 2m + 3/4 \qquad (894)
\end{aligned}
$$

If this is to be an eigenvalue equation,

$$
\lambda \alpha_+ = \alpha_+ (l(l+1) - m(m+1)) + \beta_+ \sqrt{l(l+1) - (m+1)m} \qquad (895)
$$

and

$$
\begin{aligned}
\lambda \beta_+ &= \alpha_+ \sqrt{l(l+1) - m(m+1)} \\
&\quad + \beta_+ (l(l+1) - (m+1)(m+2)) \\
&\quad + \beta_+ \qquad (896)
\end{aligned}
$$

Equating gives

$$
\begin{aligned}
(\beta_+ / \alpha_+) \sqrt{l(l+1) - (m+1)m} &= (\alpha_+ / \beta_+) \sqrt{l(l+1) - m(m+1)} \\
&\quad - 2m - 1 \qquad (897)
\end{aligned}
$$

Let $x = \alpha/\beta$. Then

$$x^2(\alpha_+/\beta_+)\sqrt{l(l+1)-m(m+1)} - x(2m+1) - \sqrt{l(l+1)-m(m+1)} = 0 \quad (898)$$

or

$$\begin{aligned} x &= \frac{(2m+1) \pm \sqrt{(2m+1)^2 + 4(l(l+1)-m(m+1))}}{2\sqrt{l(l+1)-m(m+1)}} \\ &= \frac{(2m+1) \pm 2(l+1/2)}{2\sqrt{l(l+1)-m(m+1)}} \end{aligned} \quad (899)$$

We note

$$\begin{aligned} l(l+1)-m(m+1) &= (l+1/2)^2 - (m+1/2)^2 \\ &= (l+m+1)(l-m) \end{aligned} \quad (900)$$

and the numerator

$$(2m+1) \pm 2(l+1/2) = \begin{cases} 2(m+l+1) \\ -2(l-m) \end{cases} \quad (901)$$

so

$$(\alpha/\beta) = \begin{cases} \sqrt{\dfrac{m+l+1}{l-m}} \\ -\sqrt{\dfrac{l-m}{m+l+1}} \end{cases} \quad (902)$$

and

$$\begin{aligned} \lambda_1 &= (l^2+l-m^2-3m-1) + (\alpha/\beta)\sqrt{l(l+1)-m(m+1)} \\ &= l^2+l-m^2-3m-1+(l+m+1) \\ &= l^2+2l-m^2-2m \end{aligned} \quad (903)$$

so

$$\begin{aligned} j(j+1) &= l^2+2l-m^2-2m+m^2+2m+3/4 \\ &= l^2+2l+3/4 \\ &= (l+1/2)(l+3/4) \end{aligned} \quad (904)$$

and

$$\begin{aligned}\lambda_2 &= (l^2+l-m^2-3m-1)+(\alpha/\beta)\sqrt{l(l+1)-m(m+1)} \\ &= (l^2+l-m^2-3m-1)-(l-m) \\ &= l^2-(m^2+2m+1) \end{aligned} \tag{905}$$

so

$$\begin{aligned} j(j+1) &= l^2-m^2-2m-1+m^2+2m+3/4 \\ &= l^2-1/4 \\ &= (l-1/2)(l+1/2) \end{aligned} \tag{906}$$

$$(l-1/2)(l+1/2) = l^2-1/4 \tag{907}$$

$$(l+1/2)(l+3/2) = l^2+2l+3/4 \tag{908}$$

Note

$$l+m+1+l-m = 2l+1 \tag{909}$$

so

$$|l,1/2;l+1/2,m+1/2\rangle_j = \sqrt{\frac{l+m+1}{2l+1}}|l,m\rangle\otimes|+1/2\rangle+\sqrt{\frac{l-m}{2l+1}}|l,m+1\rangle\otimes|-1/2\rangle \tag{910}$$

with $j = l+1/2$ and

$$|l,1/2;l-1/2,m+1/2\rangle_j = \sqrt{\frac{l-m}{2l+1}}|l,m\rangle\otimes|+1/2\rangle-\sqrt{\frac{l+m+1}{2l+1}}|l,m+1\rangle\otimes|-1/2\rangle \tag{911}$$

with $l = l-1/2$

Next we need to ask whether we need to consider the degenerate perturbation theory since the spin-orbit plus relativistic correction only depend on j.

For $\hat{J}_z$ term, the answer is certainly no because if m_j are different, it gives 0. What about $\hat{S}_z$? Can it couple different m_j states? Well, let's look at the structure given above. You can easily see that there is no overlap between the components of different m_j state at all. Therefore, even though the energy levels are degenerate, the matrix $\langle j, m_j|\hat{H}_{\text{Zeeman}}|j, m'_j\rangle$ is already diagonal. In that case, we don't need the additional step of diagonalizing the matrix.

So let's compute

$$\begin{aligned}
\langle l+1/2, m+1/2|\hat{S}_z|l+1/2, m+1/2\rangle &= \left(\sqrt{\frac{l+m+1}{2l+1}}\langle l, m|\otimes\langle +1/2| + \sqrt{\frac{l-m}{2l+1}}\langle l, m+1|\otimes\langle -\right. \\
&\quad \times\left(\frac{1}{2}\sqrt{\frac{l+m+1}{2l+1}}|l, m\rangle\otimes|+1/2\rangle - \frac{1}{2}\sqrt{\frac{l-m}{2l+1}}|l, m+1\right. \\
&= \frac{1}{2}\left(\frac{l+m+1}{2l+1} - \frac{l-m}{2l+1}\right) \\
&= \frac{1}{2}\left(\frac{2m+1}{2l+1}\right)
\end{aligned}$$

so

$$j_z + s_z = m_j + m_j\frac{1}{2l+1} \tag{913}$$

and

$$\begin{aligned}
\langle l-1/2, m-1/2|\hat{S}_z|l-1/2, m-1/2\rangle &= \left(\sqrt{\frac{l-m}{2l+1}}\langle l, m|\otimes\langle +1/2| - \sqrt{\frac{l+m+1}{2l+1}}\langle l, m+1|\otimes\langle -\right. \\
&\quad \times\left(\frac{1}{2}\sqrt{\frac{l-m}{2l+1}}|l, m\rangle\otimes|+1/2\rangle + \frac{1}{2}\sqrt{\frac{l+m+1}{2l+1}}|l, m+1\right. \\
&= \frac{1}{2}\left(\frac{l-m}{2l+1} - \frac{l+m+1}{2l+1}\right) \\
&= -\frac{1}{2}\left(\frac{2m+1}{2l+1}\right)
\end{aligned}$$

so

$$j_z + s_z = m_j - m_j\frac{1}{2l+1} \tag{915}$$

so

$$E^{(1)}_{\text{Zeeman}} = \frac{e\hbar B}{2mc} m_j \left(1 \pm \frac{1}{2l+1}\right) \tag{916}$$

depending on whether $j = l \pm 1/2$.

This last factor

$$g = \left(1 \pm \frac{1}{2l+1}\right) \tag{917}$$

is called the Lande g factor.

22 The remaining corrections to the Hydrogen

So far without external force, we have (we'll take $\mu \approx m_e$)

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} - \frac{Ze^2}{r} - \frac{\hat{\mathbf{p}}^4}{8m^3c^2} + \frac{mc^2\alpha^4}{2}\frac{a_0^3}{r^3}\frac{\hat{\mathbf{S}}\cdot\hat{\mathbf{L}}}{\hbar^2} \tag{918}$$

which gave us

$$E_{nj} = \frac{mc^2(Z\alpha)^2}{2n^2} - \frac{mc^2(Z\alpha)^4}{2n^3}\left(\frac{1}{j+\frac{1}{2}} - \frac{3}{4n}\right) \tag{919}$$

Darwin Term

23 Many body problem

Starting point: N body hamilotonian

$$\hat{H}_N = \sum_{i=1}^{N} \frac{\hat{\mathbf{p}}_i^2}{2m} + \sum_{i=1}^{N} v_{\text{ext}}(\mathbf{r}_i) + \frac{1}{2}\sum_{i\neq j} v_{\text{int}}(\mathbf{r}_i - \mathbf{r}_j) \tag{920}$$

N-body state:

$$|\Psi_N\rangle \tag{921}$$

Here Ψ_N denote a quantum state. Fermionic means, for instance,

$$\langle \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N | \, \Psi_N \rangle = - \langle \mathbf{r}_2, \mathbf{r}_1, \cdots, \mathbf{r}_N | \, \Psi_N \rangle \tag{922}$$

We assume:

$$|\Psi_N\rangle = \frac{1}{\sqrt{N!}} \sum_\sigma (-1)^\sigma |\psi_{\sigma(1)}, \psi_{\sigma(2)}, \cdots, \psi_{\sigma(N)}\rangle \tag{923}$$

where σ's are all permutations of $(1, 2, \cdots, N)$ and $(-1)^\sigma$ are the signature of σ and

$$|\psi_{\sigma(1)}, \psi_{\sigma(2)}, \cdots, \psi_{\sigma(N)}\rangle = |\psi_{\sigma(1)}\rangle \otimes |\psi_{\sigma(2)}\rangle \otimes \cdots \otimes |\psi_{\sigma(N)}\rangle \tag{924}$$

Check the normalization:

$$\langle \Psi_N | \Psi_N \rangle = \frac{1}{N!} \sum_{\sigma,\tau} (-1)^\sigma (-1)^\tau \, \langle \psi_{\sigma_1}, \cdots, \psi_{\sigma_N} | \, \psi_{\tau_1}, \cdots, \psi_{\tau_N} \rangle \tag{925}$$

Since all ψ_i must be different (assume they are orthonormal), this is non-zero only when $\tau = \sigma$. So the sum is 1.

Now calculate the expectation value:

$$\langle \psi_N | \hat{H}_N | \psi_N \rangle \tag{926}$$

Let's do this term by term. First the kinetic energy term

$$\begin{aligned}
&\langle \psi_N | \hat{K}_N | \psi_N \rangle \\
&= \sum_{i=1}^N \int \prod_{j=1}^N d^3 r_j \, \langle \psi_N | \, \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N \rangle \, \frac{\hat{\mathbf{p}}_i^2}{2m} \, \langle \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N | \, \psi_N \rangle \\
&= \frac{1}{N!} \sum_{i=1}^N \int \prod_{j=1}^N d^3 r_j \, \sum_{\sigma,\tau} (-1)^\sigma (-1)^\tau \prod_{k=1}^N \psi^*_{\sigma_k}(\mathbf{r}_k) \frac{\hat{\mathbf{p}}_i^2}{2m} \prod_{l=1}^N \psi_{\tau_l}(\mathbf{r}_l) \\
&= \frac{1}{N!} \int \prod_{j=1}^N d^3 r_j \, \sum_{\sigma,\tau} (-1)^\sigma (-1)^\tau \prod_{k=1}^N \psi^*_{\sigma_k}(\mathbf{r}_k) \sum_{i=1}^N \prod_{\substack{l=1 \\ l \neq i}}^N \psi_{\tau_l}(\mathbf{r}_l) \frac{\hat{\mathbf{p}}_i^2}{2m} \psi_{\tau_i}(\mathbf{r}_i) \\
&= \frac{1}{N!} \sum_{\sigma,\tau} (-1)^\sigma (-1)^\tau \sum_i \int d^3 r_i \, \left(\psi^*_{\sigma_i}(\mathbf{r}_i) \frac{\hat{\mathbf{p}}_i^2}{2m} \psi_{\tau_i}(\mathbf{r}_i) \right) \prod_{\substack{j=1 \\ j \neq i}}^N \int d^3 r_j \, \psi_{\sigma_j}(\mathbf{r}_j) \psi_{\tau_j}(\mathbf{r}_j) \\
&= \sum_{i=1}^N \int d^3 r_i \, \psi^*_i(\mathbf{r}_i) \frac{\hat{\mathbf{p}}_i^2}{2m} \psi_i(\mathbf{r}_i)
\end{aligned} \tag{927}$$

since only $\sigma = \tau$ contributes.

Exactly the same line of reasoning gets you

$$\langle\psi_N|\hat{V}_{\rm ext}|\psi_N\rangle = \sum_{i=1}^{N}\int d^3r_i\,\psi_i^*(\mathbf{r}_i)v_{\rm ext}(\mathbf{r}_i)\psi_i(\mathbf{r}_i) \tag{928}$$

The interaction term is another matter altogether:

$$\begin{aligned}
&\langle\psi_N|\hat{V}_{\rm int}|\psi_N\rangle \\
&= \frac{1}{2}\sum_{i\neq j}\int\prod_{k=1}^{N}d^3r_k\,\langle\psi_N|\,\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N\rangle\,\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\,\langle\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N|\,\psi_N\rangle \\
&= \frac{1}{2}\sum_{i\neq j}\frac{1}{N!}\sum_{\sigma,\tau}(-1)^\sigma(-1)^\tau\int\prod_{k=1}^{N}d^3r_k\prod_{l=1}^{N}\psi^*_{\sigma_l}(\mathbf{r}_l)\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\prod_{m=1}^{N}\psi^*_{\tau_m}(\mathbf{r}_m)
\end{aligned} \tag{929}$$

The product is non-zero only in two cases. (i) $\sigma = \tau$ (ii) $\tau = \hat{P}_{ij}\sigma$ where $\hat{P}_{ij}$ is the exchange operator. One exchange operator difference implies that $(-1)^\sigma(-1)^\tau = -1$. So I get

$$\begin{aligned}
&\langle\psi_N|\hat{V}_{\rm int}|\psi_N\rangle \\
&= \frac{1}{2}\sum_{i\neq j}\frac{1}{N!}\sum_{\sigma}\int d^3r_id^3r_j\;|\psi_{\sigma_i}(\mathbf{r}_i)|^2\,\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\left|\psi_{\sigma_j}(\mathbf{r}_j)\right|^2 \\
&\quad -\frac{1}{2}\sum_{i\neq j}\frac{1}{N!}\sum_{\sigma}\int d^3r_id^3r_j\,\psi^*_{\sigma_i}(\mathbf{r}_i)\psi^*_{\sigma_j}(\mathbf{r}_j)\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\psi_{\tau_i}(\mathbf{r}_i)\psi_{\tau_j}(\mathbf{r}_j) \\
&= \frac{1}{2}\sum_{i\neq j}\frac{1}{N!}\sum_{\sigma}\int d^3r_id^3r_j\;|\psi_{\sigma_i}(\mathbf{r}_i)|^2\,\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\left|\psi_{\sigma_j}(\mathbf{r}_j)\right|^2 \\
&\quad -\frac{1}{2}\sum_{i\neq j}\frac{1}{N!}\sum_{\sigma}\int d^3r_id^3r_j\,\psi^*_{\sigma_i}(\mathbf{r}_i)\psi^*_{\sigma_j}(\mathbf{r}_j)\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\psi_{\sigma_j}(\mathbf{r}_i)\psi_{\sigma_i}(\mathbf{r}_j) \\
&= \frac{1}{2}\sum_{i\neq j}\int d^3r_id^3r_j\;|\psi_i(\mathbf{r}_i)|^2\,\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\,|\psi_j(\mathbf{r}_j)|^2 \\
&\quad -\frac{1}{2}\sum_{i\neq j}\int d^3r_id^3r_j\,\psi^*_i(\mathbf{r}_i)\psi^*_j(\mathbf{r}_j)\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\psi_j(\mathbf{r}_i)\psi_i(\mathbf{r}_j)
\end{aligned} \tag{930}$$

where $\tau = \hat{P}_{ij}\sigma$ is used and we renamed $\mathbf{r}_i \to \mathbf{r}_{\sigma_i}$ to sum over σ.

Taking ψ_i's themselves to be variational parameters, we should demand

$$\delta_k\left(\left\langle\,\Psi_N|\,\hat{H}_N\,|\Psi_N\,\right\rangle - \epsilon_k\,\langle\Psi_N|\Psi_N\rangle\right) = 0 \tag{931}$$

for each k. This yields

$$\begin{aligned}
\epsilon_k\psi_k(\mathbf{r}) &= \frac{\hat{\mathbf{p}}^2}{2m}\psi_k(\mathbf{r}) + v_{\rm ext}(\mathbf{r})\psi_k(\mathbf{r}) \\
&\quad + \frac{1}{2}\frac{\delta}{\delta\psi_k^*(\mathbf{r})}\sum_{i\neq j}\int d^3r_i d^3r_j \; |\psi_i(\mathbf{r}_i)|^2 \, \hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\, |\psi_j(\mathbf{r}_j)|^2 \\
&\quad - \frac{1}{2}\frac{\delta}{\delta\psi_k^*(\mathbf{r})}\sum_{i\neq j}\int d^3r_i d^3r_j \; \psi_i^*(\mathbf{r}_i)\psi_j^*(\mathbf{r}_j)\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\psi_j(\mathbf{r}_i)\psi_i(\mathbf{r}_j) \\
&= \frac{\hat{\mathbf{p}}^2}{2m}\psi_k(\mathbf{r}) + v_{\rm ext}(\mathbf{r})\psi_k(\mathbf{r}) \\
&\quad + \sum_{i\neq k}\int d^3r_i \, |\psi_i(\mathbf{r}_i)|^2 \, \hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r})\psi_k(\mathbf{r}) \\
&\quad - \frac{1}{2}\sum_{i\neq j}\int d^3r_i d^3r_j \; \delta_{ik}\delta^{(3)}(\mathbf{r}_i-\mathbf{r})\psi_j^*(\mathbf{r}_j)\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\psi_j(\mathbf{r}_i)\psi_i(\mathbf{r}_j) \\
&\quad - \frac{1}{2}\sum_{i\neq j}\int d^3r_i d^3r_j \; \psi_i^*(\mathbf{r}_i)\delta_{jk}\delta^{(3)}(\mathbf{r}_j-\mathbf{r})\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r}_j)\psi_j(\mathbf{r}_i)\psi_i(\mathbf{r}_j) \\
&= \frac{\hat{\mathbf{p}}^2}{2m}\psi_k(\mathbf{r}) + v_{\rm ext}(\mathbf{r})\psi_k(\mathbf{r}) \\
&\quad + \sum_{i\neq k}\int d^3r_i \, |\psi_i(\mathbf{r}_i)|^2 \, \hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r})\psi_k(\mathbf{r}) \\
&\quad - \frac{1}{2}\sum_{j\neq k}\int d^3r_j \; \psi_j^*(\mathbf{r}_j)\psi_k(\mathbf{r}_j)\hat{v}_{\rm int}(\mathbf{r}-\mathbf{r}_j)\psi_j(\mathbf{r}) \\
&\quad - \frac{1}{2}\sum_{i\neq k}\int d^3r_i \; \psi_i^*(\mathbf{r}_i)\psi_k(\mathbf{r}_i)\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r})\psi_i(\mathbf{r}) \\
&= \frac{\hat{\mathbf{p}}^2}{2m}\psi_k(\mathbf{r}) + v_{\rm ext}(\mathbf{r})\psi_k(\mathbf{r}) \\
&\quad + \sum_{i\neq k}\int d^3r_i \, |\psi_i(\mathbf{r}_i)|^2 \, \hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r})\psi_k(\mathbf{r}) \\
&\quad - \sum_{i\neq k}\int d^3r_i \; \psi_i^*(\mathbf{r}_i)\psi_k(\mathbf{r}_i)\hat{v}_{\rm int}(\mathbf{r}_i-\mathbf{r})\psi_i(\mathbf{r}) \qquad (932)
\end{aligned}$$

24 Lagrangian Formulation of NR QM

We start from the Maxwell's equation

$$\nabla\cdot\mathbf{E} = \rho \tag{933}$$
$$\nabla\cdot\mathbf{B} = 0 \tag{934}$$
$$\nabla\times\mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B} \tag{935}$$
$$\nabla\times\mathbf{B} = \frac{\partial}{\partial t}\mathbf{E} + \mathbf{J} \tag{936}$$
$$\tag{937}$$

The second equation says there is no magnetic monopole. Hence

$$\mathbf{B} = \nabla\times\mathbf{A} \tag{938}$$

The third then implies that

$$\nabla\times(\mathbf{E} + \partial_t\mathbf{A}) = 0 \tag{939}$$

or

$$\mathbf{E} + \partial_t\mathbf{A} = -\nabla\phi \tag{940}$$

Substituing into the first gives me

$$-\partial_t\nabla\cdot\mathbf{A} - \nabla^2\phi = \rho \tag{941}$$

into the 4-th gives me

$$\begin{aligned}\nabla\times\nabla\times\mathbf{A} &= \partial_t(-\partial_t\mathbf{A} - \nabla\phi) + \mathbf{J} \\ &= \mathbf{e}^i\epsilon_{ijk}\nabla^j\epsilon_{klm}\nabla^l A^m = -\partial_t^2\mathbf{A} - \nabla\partial_t\phi + \mathbf{J} \\ &= \mathbf{e}^i(\delta_{il}\delta jm - \delta_{im}\delta_{jl})\nabla^j\nabla^l A^m = -\partial_t^2\mathbf{A} - \nabla\partial_t\phi + \mathbf{J} \\ &= \nabla\nabla\cdot\mathbf{A} - \nabla^2\mathbf{A} = -\partial_t^2\mathbf{A} - \nabla\partial_t\phi + \mathbf{J} \end{aligned} \tag{942}$$

or

$$\left(\partial_t^2 - \nabla^2\right)\mathbf{A} = -\nabla\nabla\cdot\mathbf{A} - \nabla\partial_t\phi + \mathbf{J} \tag{943}$$

We choose a gauge condition

$$\partial_t\phi = -\nabla\cdot\mathbf{A} \tag{944}$$

to get

$$\left(\partial_t^2 - \nabla^2\right)\mathbf{A} = \mathbf{J} \tag{945}$$

Also the first equation becomes

$$\begin{aligned} \rho &= -\partial_t \nabla\cdot\mathbf{A} - \nabla^2\phi \\ &= (\partial_t^2 - \nabla^2)\phi \end{aligned} \tag{946}$$

or in the covariant notation

$$\partial^\mu \partial_\mu A_\nu = J_\nu \tag{947}$$

O.K. So we must have all the signs right.

For a particle with a charge q, the Lorentz force is

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v}\times\mathbf{B}\right) \tag{948}$$

What we want is a Lagrangian that will give us this force. For that, we do the following manipulation.

$$\begin{aligned} F_i &= q\left(-\nabla_i\phi - \partial_t A_i\right) + q\epsilon_{ijk}v_j\,\epsilon_{klm}\nabla_l A_m \\ &= q\left(-\nabla_i\phi - \partial_t A_i\right) + q(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})v_j\,\nabla_l A_m \\ &= q\left(-\nabla_i\phi - \partial_t A_i + \mathbf{v}\cdot\nabla_i\mathbf{A} - \mathbf{v}\cdot\nabla A_i\right) \\ &= q\left(-\nabla_i\phi - \partial_t A_i\right) + q(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})v_j\,\nabla_l A_m \\ &= q\left(-\nabla_i\phi - \partial_t A_i + \mathbf{v}\cdot\nabla_i\mathbf{A} - \mathbf{v}\cdot\nabla A_i\right) \\ &= -q\nabla_i\left(\phi - \mathbf{v}\cdot\mathbf{A}\right) - q\frac{d}{dt}A_i \\ =?\ & \frac{\partial}{\partial x^i}(-V) - \frac{d}{dt}\frac{\partial}{\partial v_i}(-V) \end{aligned} \tag{949}$$

Whose solution is

$$-V = -q\phi + q\mathbf{v}\cdot\mathbf{A} \tag{950}$$

So the total Lagrangian is

$$L = T - V = \frac{m\mathbf{v}^2}{2} - q\phi + q\mathbf{v}\cdot\mathbf{A} \tag{951}$$

Now the canonical momentum is

$$\mathbf{p} = \frac{\partial}{\partial \mathbf{v}} L = m\mathbf{v} + q\mathbf{A} \tag{952}$$

or the kinetic momentum is

$$m\mathbf{v} = \mathbf{p} - q\mathbf{A} \tag{953}$$

The Hamiltonian is given by

$$\begin{aligned} H &= \mathbf{p}\cdot\mathbf{v} - \mathbf{L} \\ &= \mathbf{p}\cdot(\mathbf{p} - q\mathbf{A})/m - \frac{(\mathbf{p} - q\mathbf{A})^2}{2m} + q\phi - q(\mathbf{p} - q\mathbf{A})\cdot\mathbf{A}/m \\ &= \frac{(\mathbf{p} - q\mathbf{A})^2}{2m} + q\phi \end{aligned} \tag{954}$$

This is what corresponds to the Quantum Mechanical Hamiltonian operator with the *canonical* momentum given by the gradient.

Now let's have an alternative derivation of the same result. Start from the QM hamiltonian,

$$H = \frac{\mathbf{p}^2}{2m} + q\phi \tag{955}$$

The Schrödinger equation is

$$H\psi(x,t) = i\partial_t \psi(x) \tag{956}$$

Consider the phase change $\psi(x) \rightarrow e^{iqf_0}\psi(x)$. In this case, since f_0 is a constant, nothing changes. The Schrödinger equation stays the same, the norm stays the same, the energy and momentum stays the same. In other words, the phase of a wavefunction is *not* observable.

Now allow f to be x and t dependent. Now the Schrödinger equation is modified to

$$-\hbar^2 \frac{(\nabla + iq\nabla f)^2}{2m}\psi = i(\partial_t + iq\partial_t f)\psi \tag{957}$$

However, the phase of a wavefunction is supposed to be unobservable. This now says that it is. To remedi the situation, consider a vector field $\mathbf{A}$ that under the same phase change transforms to

$$\mathbf{A} \rightarrow \mathbf{A} + \nabla f \tag{958}$$

and a scalar that changes to

$$\phi \to \phi - \partial_t f \tag{959}$$

and modify our original Schrödinger equation to

$$H\psi = \left[\frac{(\mathbf{p} - q\mathbf{A})^2}{2m} + q\phi\right]\psi = i\partial_t\psi \tag{960}$$

Then the combined transformation leaves the Schrödinger equation invariant. The above transformation of $\mathbf{A}$ and ϕ is called the 'Gauge Tranformation' and the vector potential and the electric potential are often referred to as the 'gauge field'. The above grandient-gauge coupling is called the 'minimal coupling'. The fact that the hamiltonian is intact means that the hamiltonian is 'gauge invariant'.

Notice that by postulating the gauge invariance of the hamiltonian, we somehow got back the vector and scalar potential *without* knowing anything about E & M! Or did we? Well, not quite. We have the interaction between the matter and the gauge field, but we can't derive Maxwell's equation from the above hamiltonian. This part comes from constructing again a Lagrangian made up of $\mathbf{A}$ and ϕ that is gauge invariant (and quadratic in the gradient). After that we have full (classical) electro-magnetism and quantum mechanics for the charged particles.

This way of getting E & M is called the 'gauge principle' and it is a very powerful principle. In fact, the modern day Standard Model (which encompasses E & M, Weak and the Strong interactions) is constructed just this way.

Complex $U(1)$ is a real $SO(2)$ or 2-d rotation.

All that is required is actually to generalize the wavefunction to the wavefunction multiplets

$$\psi \to \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} \tag{961}$$

and require that the rotation matrix that leaves the complex norm $\psi^\dagger\psi$ intact be unobservable. Then instead of one set of a vector and scalar gauge field, we get a multiple set of gauge field. Each set corresponds to each rotation angle.

Now, we know that the rotation in more that 2-dimensinal space *do not* commute. That is, changing the order of rotations results in different final state. In 2-d, this doesn't matter because there is only one rotation axis and whichever way you do it, it ends up at the same position. This means that the rotations don't know about each other. In terms of gauge field, this means that the gauge field do not interact with each other directly. It is generated or absorbed by matter, but a photon can't generate another photon because it has no charge itself. This is why E and M is so much simpler than the other interactions.

Other interactions such as the weak and the strong interactions and ultimately the Grand Unified Theory, all rely on many dimensional (in complex space) rotations which do not commute. That means, different rotations has to know about each other. They have to know who's on first and who's on second. In terms of gauge field, this means that they do interact with each other. Unlike photons, a gauge field particle of a non-commuting kind (usually called 'non-abelian gauge field') can generate another gauge field particle. And amazing thing is, this works. The so-called Standard Model is the most accurate theory of matter man-kind has ever devised. We keep searching, but so far no violation of the standard model has been reported yet. The story may change, though. A few weeks ago, a team at Brookhaven national lab announced their measurement of muon anomalous magnetic moment (at the 7-th digit!). And there is a compelling reason to believe that this is due to a new set of principles called the supersymmetry. However, the jury is still out on that. No definite conclusions yet.

25 Time Dependent Perturbation

So far we have considered perturbations which do not depend on time. But in many physical system, the perturbation does depend on time. For instance, Microwave oven. The perturbation in this case is the microwave in the form of $V(t) = A\sin(wt - kx)$.

Question is, how do we deal with it? Consider a hamiltonian

$$H = H_0 + \lambda V(t) \tag{962}$$

where the unperturbed hamiltonian H_0 is time independent and we know how to solve for its eigenfunctions and eigenvalues.

What we want to solve is the time dependent Schödinger equation

$$H|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle \tag{963}$$

Let's call the eigenfunctions of H_0 $|n\rangle$ and the eigenvalues E_n. And write

$$|\psi(t)\rangle = \sum_{n=0}^{\infty} c_n(t)e^{-iE_nt/\hbar}\,|n\rangle \tag{964}$$

This is possible because H_0 is a hermitian operator. Therefore, its eigenfunctions form a complete set. That is, they can span any function. Applying H to $|\psi\rangle$ yields

$$H|\psi(t)\rangle \;=\; \sum_{n=0}^{\infty} c_n(t)e^{-iE_nt/\hbar}\,(E_n + \lambda V(t))|n\rangle \tag{965}$$

Applying $i\hbar\partial_t$ yields

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle \;=\; \sum_{n=0}^{\infty} e^{-iE_nt/\hbar}\,(E_nc_n(t) + i\hbar\dot{c}_n(t))|n\rangle \tag{966}$$

Equating the two, we get

$$\sum_{n=0}^{\infty} c_n(t)e^{-iE_nt/\hbar}\,\lambda V(t)|n\rangle = \sum_{n=0}^{\infty} e^{-iE_nt/\hbar}\,i\hbar\dot{c}_n(t)|n\rangle \tag{967}$$

To get the coefficients, we apply $\langle m|e^{iE_mt/\hbar}$ to get

$$\sum_{n=0}^{\infty} c_n(t)e^{-i(E_n-E_m)t/\hbar}\,\langle m|\lambda V(t)|n\rangle = i\hbar\dot{c}_m(t) \tag{968}$$

This is an infinitely coupled set of linear differential equations for the coefficients $c_n(t)$ and in principle solvable provided that we know the matrix element $\langle m|\lambda V(t)|n\rangle$.

In practice, the usual situation goes something like this. Before $t = 0$, we have a system which is in a stationary state of H_0. At $t = 0$, we turn on the time dependent perturbation (you turn on your microwave oven) and then ask: What is the probability of finding my system in another stationary state of H_0 at t?

That is, we start from

$$c_n(0) = \delta_{ni} \tag{969}$$

where i is the initial state we start from and ask the value of the *transition probability*

$$P_{im}(t) = |\langle m|\psi(t)\rangle|^2 = |c_m(t)|^2 \tag{970}$$

Now since $\lambda V(t)$ is supposed to be a perturbation, we can safely say that $\langle m|\lambda V(t)|n\rangle$ is much smaller than any of the E_n. In other words, as time goes by, the most likely state the system will be in is still the initial state $|i\rangle$ and the *transition* probabilities $P_{m\neq i}$ should stay small at all times.

We want to use this fact to solve the above coupled equation for $c_n(t)$. Let

$$c_i(t) = 1 + \lambda f_i(t) + O(\lambda^2) \tag{971}$$

and

$$c_{m\neq i}(t) = \lambda f_m(t) + O(\lambda^2) \tag{972}$$

The from above, the equation for $m = i$ becomes

$$\begin{aligned} i\hbar\dot{c}_i(t) &= i\hbar\lambda\dot{f}_i(t) + O(\lambda^2) \\ &= \sum_{n\neq i}^{\infty} \lambda f_n(t) e^{-i(E_n-E_i)t/\hbar} \langle i|\lambda V(t)|n\rangle + (1+\lambda f_i(t))\langle i|\lambda V(t)|i\rangle + O(\lambda^2) \end{aligned} \tag{973}$$

Collecting the terms linear in λ yields

$$i\hbar\dot{f}_i(t) = \langle i|V(t)|i\rangle f_i(t) \tag{974}$$

or

$$f_i(t) = \exp\left(-i\int_0^t dt' \langle i|V(t')|i\rangle/\hbar\right) \tag{975}$$

For $m \neq i$,

$$\begin{aligned} i\hbar\dot{c}_{m\neq i}(t) &= i\hbar\lambda\dot{f}_{m\neq i}(t) + O(\lambda^2) \\ &= \sum_{n\neq i}^{\infty} \lambda f_n(t) e^{-i(E_n-E_m)t/\hbar} \langle i|\lambda V(t)|n\rangle + (1+\lambda f_i(t)) e^{-i(E_i-E_m)t/\hbar} \langle m|\lambda V(t)|i\rangle \end{aligned} \tag{976}$$

Again collecting the coefficients of λ, we get

$$i\hbar \dot{f}_{m\neq i}(t) = e^{-i(E_i-E_m)t/\hbar}\langle m|V(t)|i\rangle \tag{977}$$

or

$$f_{m\neq i}(t) = -\frac{i}{\hbar}\int_0^t dt'\, e^{-i(E_i-E_m)t'/\hbar}\langle m|V(t')|i\rangle \tag{978}$$

The leading order perturbation theory transition amplitude is then ($i \neq m$)

$$P_{im}(t) = \lambda^2 \left|\frac{1}{\hbar}\int_0^t dt'\, e^{-i(E_i-E_m)t'/\hbar}\langle m|V(t')|i\rangle\right|^2 + O(\lambda^3) \tag{979}$$

Consider the following perturbation:

$$\hat{V}(t) = 2\hat{V}_0(\hat{x},\hat{p})\,\cos(\omega t) \tag{980}$$

The transition amplitude in this case is given by ($m \neq i$)

$$\begin{aligned} P_{im}(t) &= \lambda^2 \left|\frac{1}{\hbar}\int_0^t dt'\, e^{-i(E_i-E_m)t'/\hbar}\langle m|V(t')|i\rangle\right|^2 \\ &= \lambda^2 \left|\frac{1}{\hbar}\int_0^t dt'\, \left(e^{-i(E_i-E_m+\hbar\omega)t'/\hbar} + e^{-i(E_i-E_m-\hbar\omega)t'/\hbar}\right)\langle m|V_0|i\rangle\right|^2 \end{aligned} \tag{981}$$

Remember this is $P_{im}(t)$. So we need to figure out its behavior as both the function of time and m. First of all, of course, the whole thing vanishes unless

$$\langle m|V_0|i\rangle \neq 0 \tag{982}$$

Secondly, most of the times, it turned out that we are rather more interested in the *rate* of the transition than the transition probability itself. To calculate the rate, we take the derivate. Setting $\Delta E = E_m - E_i$ for convenience,

$$\begin{aligned} W_{im}(t) &= \frac{d}{dt}P_{im}(t) \\ &= \lambda^2 \frac{d}{dt}\left|\frac{1}{\hbar}\int_0^t dt'\, \left(e^{i(\Delta E-\hbar\omega)t'/\hbar} + e^{i(\Delta E+\hbar\omega)t'/\hbar}\right)\langle m|V_0|i\rangle\right|^2 \end{aligned}$$

$$
\begin{aligned}
&= \frac{\lambda^2}{\hbar^2}\frac{d}{dt}\Big[\int_0^t dt'\,\left(e^{i(\Delta E-\hbar\omega)t'/\hbar}+e^{i(\Delta E+\hbar\omega)t'/\hbar}\right)\\
&\qquad\times\int_0^t dt''\,\left(e^{-i(\Delta E-\hbar\omega)t''/\hbar}+e^{-i(\Delta E+\hbar\omega)t''/\hbar}\right)\Big]\,|\langle m|V_0|i\rangle|^2\\
&= \frac{\lambda^2}{\hbar^2}\Big[\left(e^{i(\Delta E-\hbar\omega)t/\hbar}+e^{i(\Delta E+\hbar\omega)t/\hbar}\right)\int_0^t dt''\,\left(e^{-i(\Delta E-\hbar\omega)t''/\hbar}+e^{-i(\Delta E+\hbar\omega)t''/\hbar}\right)\\
&\quad+\left(e^{-i(\Delta E-\hbar\omega)t/\hbar}+e^{-i(\Delta E+\hbar\omega)t/\hbar}\right)\int_0^t dt'\,\left(e^{i(\Delta E-\hbar\omega)t'/\hbar}+e^{i(\Delta E+\hbar\omega)t'/\hbar}\right)\Big]\\
&\quad\times|\langle m|V_0|i\rangle|^2 \qquad (983)
\end{aligned}
$$

Changing $t'-t\to t'$ and $t''-t\to t''$, we get

$$
\begin{aligned}
W_{im}(t) &= \frac{\lambda^2}{\hbar^2}\Big[\int_{-t}^0 dt'\,e^{-i(\hbar\omega-\Delta E)t'/\hbar}+\int_{-t}^0 dt'\,e^{i(\hbar\omega-\Delta E)t'/\hbar}\\
&\quad+\int_{-t}^0 dt'\,e^{-i(\hbar\omega+\Delta E)t'/\hbar}+\int_{-t}^0 dt'\,e^{i(\hbar\omega+\Delta E)t'/\hbar}\\
&\quad+2\int_0^t dt'\,e^{i\hbar\omega(t+t')}+2\int_0^t dt'\,e^{-i\hbar\omega(t+t')}\Big]\\
&\quad\times|\langle m|V_0|i\rangle|^2\\
&= \frac{\lambda^2}{\hbar^2}\Big[\int_{-t}^t dt'\,e^{-i(\hbar\omega-\Delta E)t'/\hbar}+\int_{-t}^t dt'\,e^{-i(\hbar\omega+\Delta E)t'/\hbar}\\
&\qquad+2\int_0^t dt'\,e^{i\hbar\omega(t+t')}+2\int_0^t dt'\,e^{-i\hbar\omega(t+t')}\Big]\,|\langle m|V_0|i\rangle|^2\\
&= \frac{\lambda^2}{\hbar^2}\Big[\int_{-t}^t dt'\,e^{-i(\hbar\omega-\Delta E)t'/\hbar}+\int_{-t}^t dt'\,e^{-i(\hbar\omega+\Delta E)t'/\hbar}\\
&\qquad+2\int_{-2t}^{2t} dt'\,e^{i\hbar\omega t}-2\int_{-t}^t dt'\,e^{i\hbar\omega t'}\Big]\,|\langle m|V_0|i\rangle|^2
\end{aligned}
\qquad (984)
$$

For very short times, we can make the taylor expansion of the exponential and the first term is

$$
W_{im}(t) \approx 8t\frac{\lambda^2}{\hbar^2}\,|\langle m|V_0|i\rangle|^2 \qquad (985)
$$

or

$$
P_{im}(t) \approx 4t^2\frac{\lambda^2}{\hbar^2}\,|\langle m|V_0|i\rangle|^2 \qquad (986)
$$

This is valid when $t \ll \hbar/\left|\hbar\omega \pm \Delta E\right|$. So for very small time, the probability grows quadratically provided that $\langle m|V_0|i\rangle \neq 0$.

For large times, the behavior is quite different. We know that

$$\begin{aligned} d_T(\omega) = \int_{-T}^{T} dt\, e^{i\omega t} &= \frac{e^{i\omega T} - e^{-i\omega T}}{i\omega} \\ &= \frac{2\sin(\omega T)}{\omega} \end{aligned} \tag{987}$$

In the limit of 0 omega,

$$\lim_{\omega\to 0} d_T(\omega) = 2T \tag{988}$$

Hence the height of the central peak grows lineary with time. The width of the central peak is given roughly by

$$\Delta\omega \approx \frac{\pi}{T} \tag{989}$$

That is, the width drops as T grows. Or

$$\Delta(\hbar\omega)T \approx \pi \tag{990}$$

This is an example of time-energy Heisenberg uncertainty. As time grows, the uncertainty in the energy (frequency) falls to zero. The integral under of the function at hand is

$$\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} d_T(\omega) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi\omega}\, 2\sin(\omega T) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{dx}{x}\, \sin(x) = 1 \tag{991}$$

That is, we have a function whose peak grows indefintely as time elapses and the width becomes smaller. At the same time, the integral remains the same. In the limit of large time, such function goes to the Dirac delta-function. What we have here is actually a statement that the free particle eigenstates are complete. This is of course true because they are eigenfunctions of an Hermitian operator. In other words,

$$\int_{-\infty}^{\infty} dt\, e^{i\omega t} = 2\pi\delta(\omega) \tag{992}$$

The appearance of (2π) factors can sometimes cause a little headache. An easy way to remeber those factors is this : First of all, only the frequencies and

wavevectors get them. Or equivalently, only the energies and the momenta get them. Second, dk or $d\omega$ always appear in the combination

$$\frac{dk}{(2\pi)} \quad \text{and} \quad \frac{d\omega}{(2\pi)} \tag{993}$$

and the δ functions always appear in the combination

$$(2\pi)\delta(k) \quad \text{and} \quad (2\pi)\delta(\omega) \tag{994}$$

For example,

$$\int_{-\infty}^{\infty} \frac{d\omega}{(2\pi)} e^{-i\omega t} = \delta(t) \tag{995}$$

For another example, the phase space volume

$$d\Gamma = \frac{dx\, dp}{h} = \frac{dx\, dp}{\hbar(2\pi)} = dx\, \frac{dk}{(2\pi)} \tag{996}$$

In our case, using the above expression for the delta function gives

$$\begin{aligned} \lim_{t\to\infty} W_{im}(t) &= \lim_{t\to\infty} \frac{d}{dt} P_{im}(t) \\ &= \frac{\lambda^2}{\hbar^2}\, 2\pi \Big[\delta(\omega - \Delta E/\hbar) + \delta(\omega + \Delta E/\hbar)\Big] \, |\langle m|V_0|i\rangle|^2 \\ &= \frac{\lambda^2}{\hbar}\, 2\pi \Big[\delta(\hbar\omega - \Delta E) + \delta(\hbar\omega + \Delta E)\Big] \, |\langle m|V_0|i\rangle|^2 \end{aligned} \tag{997}$$

That is, as the time elapses, the transition probability vanishes for all m *unless* $\hbar\omega$ happens to be equal to $\pm(E_m - E_i)$. In other words, transition happens only if ω is a resonance frequency if we apply the perturbation for a long time. The first term says

$$\hbar\omega = E_m - E_i \tag{998}$$

Hence, it is the abosorption rate. If this was electromagnetic field this would be really the photon absorption rate. The second term corresponds to

$$\hbar\omega = E_i - E_m \tag{999}$$

and hence is the emission rate.

Let's consder the absorption rate. Having a δ-function is mathematically convenient, but in real life, nothing is really infinitely thin and infinitely high. So the above formula as a transition rate has little use.

A more useful expression is obtained by Fermi and is dubbed as the 'Golden Rule' by him. We call it, of course, 'Fermi's Golden Rule'. First of all, we realize that in many cases, the final state m is continuous and degenerate. For instance, it can be just a 3-momentum. Then it is of course continuous and all momenta with the same magnitude $\mathbf{k}^2$ will be degenerate as far as the kinetic energy is concerned. In this case, we really have to sum over a certain range of m.

Consider now a small interval of energy $(E_m - dE/2, E_m + dE/2)$. Let's write the number of states in that interval as

$$dN = g(E_m)dE \tag{1000}$$

Then summing over such states becomes

$$\begin{aligned} \bar{W}_{im} &= \int_{E_m-\delta}^{E_m+\delta} g(E_m)\, W_{im}(T)\, dE \\ &= \int_{E_m-\delta}^{E_m+\delta} dE\, g(E_m)\, d_T(E_m - E_i, \omega)\, \frac{2\pi}{\hbar^2}\, |\langle i|(\lambda V_0)|m\rangle|^2 \end{aligned} \tag{1001}$$

As $T \to \infty$, d_T becomes a δ-function, and we get

$$\bar{W}_{im} = \frac{2\pi}{\hbar}\, g(E_m)\, |\langle i|(\lambda V_0)|m\rangle|^2 \tag{1002}$$

This 'Fermi's Golden Rule' describes the transition rate of a state to a continuum state and is a truly useful in many physical situations.

A Supplemental – Uncertainty and Commutator

Suppose A, B and C are Hermitian and we have

$$[A, B] = iC \tag{1003}$$

Now consider

$$\Delta A^2 = \langle A^2 \rangle - \langle A \rangle^2 \tag{1004}$$
$$\Delta B^2 = \langle B^2 \rangle - \langle B \rangle^2 \tag{1005}$$

Landau-Lifshitz

$$\begin{aligned} \langle (\alpha A + iB)^\dagger (\alpha A + iB) \rangle &= \langle \alpha^2 A^2 + B^2 + i\alpha AB - i\alpha BA \rangle \\ &= \alpha^2 \langle A^2 \rangle + \langle B^2 \rangle + i\alpha \langle [A, B] \rangle \geq 0 \\ &= \alpha^2 \langle A^2 \rangle + \langle B^2 \rangle - \alpha \langle C \rangle \geq 0 \end{aligned} \tag{1006}$$

For this to be true for arbitrary α, the quadratic equation must not have any solutions. So we must have

$$b^2 - 4ac \leq 0 \tag{1007}$$

or

$$4\langle A^2 \rangle \langle B^2 \rangle \geq \langle C \rangle^2 \rightarrow \langle A^2 \rangle \langle B^2 \rangle \geq \frac{1}{4} \langle C \rangle^2 \tag{1008}$$

Now we could have easily used $\Delta \hat{A} = \hat{A} - \langle A \rangle$ and $\Delta \hat{B} = \hat{B} - \langle B \rangle$ since

$$[\Delta \hat{A}, \Delta \hat{B}] = [\hat{A}, \hat{B}] \tag{1009}$$

So we have established

$$\Delta A \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle| \tag{1010}$$

Alternative: Schwartz inequality:

$$\begin{aligned} 0 &\leq \langle (A - tB)^\dagger (A - tB) \rangle \\ &= \langle (A^\dagger - t^* B^\dagger)(A - tB) \rangle \\ &= \langle A^\dagger A \rangle + t^* t \langle B^\dagger B \rangle - t \langle A^\dagger B \rangle - t^* \langle B^\dagger A \rangle \end{aligned} \tag{1011}$$

This is valid for any complex t. In particular, if we set

$$t = \frac{\langle B^\dagger A \rangle}{\langle B^\dagger B \rangle} = \left(\frac{\langle A^\dagger B \rangle}{\langle B^\dagger B \rangle} \right)^* \tag{1012}$$

we get

$$0 \leq \langle A^\dagger A \rangle - \frac{\langle A^\dagger B \rangle \langle B^\dagger A \rangle}{\langle B^\dagger B \rangle} \tag{1013}$$

Or

$$\langle A^\dagger A \rangle \langle B^\dagger B \rangle \geq |\langle A^\dagger B \rangle|^2 \tag{1014}$$

With Hermitian operators, we have

$$\langle A^2 \rangle \langle B^2 \rangle \geq |\langle AB \rangle|^2 \tag{1015}$$

Using

$$AB = \frac{1}{2}\{A, B\} + \frac{1}{2}[A, B] \tag{1016}$$

and the fact that the anti-commutator

$$\{A, B\} = AB + AB \tag{1017}$$

is Hermitian and the commutator is anti-hermitian, we get

$$|\langle AB \rangle|^2 = \frac{1}{4}|\langle [A, B] \rangle|^2 + \frac{1}{4}|\langle \{A, B\} \rangle|^2 \tag{1018}$$

so

$$\langle A^2 \rangle \langle B^2 \rangle \geq |\langle AB \rangle|^2 \geq \frac{1}{4}|\langle [A, B] \rangle|^2 + \frac{1}{4}|\langle \{A, B\} \rangle|^2 \geq \frac{1}{4}|\langle [A, B] \rangle|^2 \tag{1019}$$

B Supplemental – Calculation of $\langle nlm|r^s|nlm\rangle$

We use the following relation. Start from

$$I(s) = \int_0^\infty d\rho\, f_s'''(\rho)\, u(\rho)\, u(\rho) \tag{1020}$$

Do integration by part:

$$\begin{aligned} I(s) &= \int_0^\infty d\rho\, f_s'''(\rho)\, u(\rho)\, u(\rho) \\ &= f_s''\, u\, u|_0^\infty - 2\int_0^\infty d\rho\, f_s'' u' u \\ &= -2f' u' u|_0^\infty + 2\int_0^\infty d\rho\, f'(u'' u + u' u') \\ &= 2\int_0^\infty d\rho\, f' u'' u + 2\int_0^\infty d\rho\, f' u' u' \\ &= 2\int_0^\infty d\rho\, f' u'' u + 2f(u'u')|_0^\infty - 4\int_0^\infty d\rho\, f u'' u' \\ &= 2\int_0^\infty d\rho\, f' u'' u - 4\int_0^\infty d\rho\, f u'' u' \end{aligned} \tag{1021}$$

Or in short,

$$\int_0^\infty d\rho\, (fu' - (f'u)/2)u'' = -\int_0^\infty d\rho\, f''' uu/4 \tag{1022}$$

Also

$$\int_0^\infty d\rho\, f' uu = fuu|_0^\infty - 2\int_0^\infty d\rho\, fuu' \tag{1023}$$

or

$$\int_0^\infty d\rho\, huu' = -\frac{1}{2}\int_0^\infty d\rho\, h' uu \tag{1024}$$

So let's start from

$$u'' - \frac{l(l+1)}{\rho^2}u + \left(\frac{\lambda}{\rho} - \frac{1}{4}\right)u = 0 \tag{1025}$$

We multiply it with $[fu' - f'u/2]$ and integrate

$$0 = \int_0^\infty d\rho \left[[fu' - f'u/2]u'' - [fu' - f'u/2]\frac{l(l+1)}{\rho^2}u + [fu' - f'u/2]\left(\frac{\lambda}{\rho} - \frac{1}{4}\right)u\right]$$

$$
\begin{aligned}
&= \int_0^\infty d\rho\,[fu' - f'u/2]u'' - \int_0^\infty d\rho\,\left[\frac{l(l+1)}{\rho^2} - \left(\frac{\lambda}{\rho} - \frac{1}{4}\right)\right] fu'u \\
&\quad + \frac{1}{2}\int_0^\infty d\rho\,\left[\frac{l(l+1)}{\rho^2} - \left(\frac{\lambda}{\rho} - \frac{1}{4}\right)\right] f'uu \\
&= -\frac{1}{4}\int_0^\infty d\rho\, f'''uu \\
&\quad + \frac{1}{2}\int_0^\infty d\rho\, uu\frac{d}{d\rho}\left(f\left[\frac{l(l+1)}{\rho^2} - \left(\frac{\lambda}{\rho} - \frac{1}{4}\right)\right]\right) \\
&\quad + \frac{1}{2}\int_0^\infty d\rho\,\left[\frac{l(l+1)}{\rho^2} - \left(\frac{\lambda}{\rho} - \frac{1}{4}\right)\right] f'uu \\
&= -\frac{1}{4}\int_0^\infty d\rho\, f'''uu \\
&\quad + \frac{1}{2}\int_0^\infty d\rho\, uuf\left[-2\frac{l(l+1)}{\rho^3} + \frac{\lambda}{\rho^2}\right] \\
&\quad + \int_0^\infty d\rho\,\left[\frac{l(l+1)}{\rho^2} - \left(\frac{\lambda}{\rho} - \frac{1}{4}\right)\right] f'uu
\end{aligned}
\tag{1026}
$$

Let $f = \rho^{s+1}$ Then with $\lambda = n$

$$
\begin{aligned}
0 &= -\frac{1}{4}(s+1)s(s-1)\langle\rho^{s-2}\rangle + \frac{1}{2}\left[-2l(l+1)\langle\rho^{s-2}\rangle + n\langle\rho^{s-1}\rangle\right] \\
&\quad + l(l+1)(s+1)\langle\rho^{s-2}\rangle - n(s+1)\langle\rho^{s-1}\rangle + \frac{1}{4}(s+1)\langle\rho^s\rangle \\
&= \frac{s}{4}\left[(2l+1)^2 - s^2\right]\langle\rho^{s-2}\rangle - n(s+1/2)\langle\rho^{s-1}\rangle + \frac{1}{4}(s+1)\langle\rho^s\rangle
\end{aligned}
\tag{1027}
$$

Use $\rho = \frac{2r}{a_n n}$ to finally get

$$
0 = \frac{s}{4}\left[(2l+1)^2 - s^2\right] a_0^2\langle r^{s-2}\rangle - (2s+1)a_0\langle r^{s-1}\rangle + \frac{s+1}{n^2}\langle r^s\rangle \tag{1028}
$$

$\langle 1/r^2\rangle$ Calc

We start from our home work problem. When we added B/r^2 to the potential, the energy was given by

$$E = -\frac{Z^2\hbar^2}{2ma_0^2}\frac{1}{(n_r + 1/2 + \sqrt{(l+1/2)^2 + 2mb})^2} \tag{1029}$$

where $b = B/\hbar^2$. Now for small mb, we can expand: Use

$$\begin{aligned}\sqrt{(l+1/2)^2 + 2mb} &= (l+1/2)\sqrt{1 + 2mb/(l+1/2)^2} \\ &= (l+1/2)\left(1 + mb/(l+1/2)^2 + \cdots\right) \\ &\approx (l+1/2) + mb/(l+1/2)\end{aligned} \tag{1030}$$

So

$$\begin{aligned}E &\approx -\frac{Z^2\hbar^2}{2ma_0^2}\frac{1}{(n_r + l + 1 + mb/(l+1/2))^2} \\ &\approx -\frac{Z^2\hbar^2}{2ma_0^2}\frac{1}{(n_r+l+1)^2} + \frac{mb}{l+1/2}\frac{Z^2\hbar^2}{ma_0^2}\frac{1}{(n_r+l+1)^3}\end{aligned} \tag{1031}$$

We compare this with

$$E^{(0)} = -\frac{Z^2\hbar^2}{2ma_0^2}\frac{1}{n^2} \tag{1032}$$

and make the identification $n = n_r + l + 1$ so that

$$E = -\frac{Z^2\hbar^2}{2ma_0^2}\frac{1}{n^2} + \frac{mb}{l+1/2}\frac{Z^2\hbar^2}{ma_0^2}\frac{1}{n^3} + O(b^2) \tag{1033}$$

Now compare this with the perturbation result

$$E = E^{(0)} + \langle n \left| \frac{B}{r^2} \right| n \rangle + O(B^2) \tag{1034}$$

Comparing terms linear in B, we get

$$\langle n \left| \frac{1}{r^2} \right| n \rangle = \frac{Z^2}{a_0^2}\frac{1}{(l+1/2)n^3} \tag{1035}$$

C Supplemental – Classical Kepler

Consider

$$H = \frac{p^2}{2m} - \frac{G}{r} \tag{1036}$$

Separate momentum into r and angles

$$H = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} - \frac{G}{r} \tag{1037}$$

We know that L is constant. So this is now 1-D problem. Hamilton's equation

$$\frac{dr}{dt} = \frac{\partial H}{\partial p_r} = \frac{p_r}{m} \tag{1038}$$

$$\frac{dp_r}{dt} = -\frac{\partial H}{\partial r} = \frac{L^2}{2mr^3} - \frac{G}{r^2} \tag{1039}$$

Hence the energy is

$$E = H = m\frac{\dot{r}^2}{2} - \frac{L^2}{2mr^2} + \frac{G}{r} \tag{1040}$$

which leads to

$$\frac{dr}{dt} = \sqrt{\frac{2E}{m} - \frac{L^2}{m^2r^2} + \frac{2G}{mr}} \tag{1041}$$

or

$$\frac{dr}{\sqrt{\frac{2E}{m} - \frac{L^2}{m^2r^2} + \frac{2G}{mr}}} = dt \tag{1042}$$

Now we consider

$$L = \text{constant} = mr^2\dot{\phi} \tag{1043}$$

therefore

$$dt = \frac{mr^2}{L}d\phi \tag{1044}$$

and

$$\frac{dr}{\sqrt{\frac{2E}{m} - \frac{L^2}{m^2 r^2} + \frac{2G}{mr}}} = \frac{mr^2}{L} d\phi \tag{1045}$$

which leads to

$$d\phi = \frac{L}{r^2} \frac{dr}{\sqrt{2Em - \frac{L^2}{r^2} + \frac{2Gm}{r}}} \tag{1046}$$

change variable to $u = 1/r$ to get

$$d\phi = -L \frac{du}{\sqrt{2Em - L^2 u^2 + 2Gmu}} \tag{1047}$$

To solve this, you need to complete the square

$$L^2 u^2 - 2Gmu = (Lu - Gm/L)^2 - (Gm/L)^2 \tag{1048}$$

so that

$$2Em - \frac{L^2}{r^2} + \frac{2Gm}{r} = 2Em + (Gm/L)^2 - (Lu - Gm/L)^2 \tag{1049}$$

Change variable to $\eta = Lu - Gm/L$ to get

$$d\phi = -\frac{d\eta}{\sqrt{2Em + (Gm/L)^2 - \eta^2}} \tag{1050}$$

change again to

$$\eta = \sqrt{2Em + (Gm/L)^2} \cos\zeta \tag{1051}$$

which yields

$$d\phi = d\zeta \tag{1052}$$

or

$$\cos\phi = \frac{1}{\sqrt{2Em + (Gm/L)^2}} (Lu - Gm/L) \tag{1053}$$

Rearraging

$$\cos\phi\sqrt{2Em + (Gm/L)^2} + (Gm/L) = L/r \tag{1054}$$

or by multiplying (L/Gm),

$$\cos\phi\sqrt{2EmL^2/G^2m^2 + 1} + 1 = \frac{L^2}{Gm}\frac{1}{r} \tag{1055}$$

or by defining

$$\epsilon^2 = 1 + 2EL^2/G^2m \tag{1056}$$

$$\alpha = \frac{L^2}{Gm} \tag{1057}$$

we get

$$\frac{\alpha}{r} = 1 + \epsilon\cos\phi \tag{1058}$$

solving for r yields

$$r = \frac{\alpha}{1 + \epsilon\cos\phi} \tag{1059}$$

The period is given by

$$\begin{aligned} T &= \int dt = \int_0^{2\pi} \frac{mr^2}{L} d\phi \\ &= \int dt = \frac{m}{L}\int_0^{2\pi} d\phi \frac{\alpha^2}{(1 + \epsilon\cos\phi)^2} \end{aligned} \tag{1060}$$

Consider

$$I(\zeta) = \int_0^{2\pi} d\phi \, \frac{1}{\zeta + \epsilon\cos\phi} \tag{1061}$$

Use

$$z = e^{i\phi} \tag{1062}$$

to get

$$\begin{aligned} I &= \oint \frac{dz}{iz} \frac{1}{\zeta + \epsilon(z+1/z)/2} \\ &= 2\pi \oint \frac{dz}{2\pi i} \frac{1}{(z\zeta + \epsilon(z^2+1)/2)} \\ &= \frac{4\pi}{\epsilon} \oint \frac{dz}{2\pi i} \frac{1}{(z^2 + 2z\zeta/\epsilon + 1)} \\ &= \frac{4\pi}{\epsilon} \oint \frac{dz}{2\pi i} \frac{1}{\left[z - (-\zeta - \sqrt{\zeta^2 - \epsilon})/\epsilon\right] \left[z - (-1 + \sqrt{\zeta^2 - \epsilon})/\epsilon\right]} \end{aligned} \tag{1063}$$

The integrand has poles at

$$z = -\frac{\zeta}{\epsilon} \pm \frac{1}{\epsilon}\sqrt{\zeta^2 - \epsilon^2} \tag{1064}$$

For $\epsilon < 1$, only the pole

$$z = -(\zeta - \sqrt{\zeta^2 - \epsilon^2})/\epsilon \tag{1065}$$

is in the contour. So

$$\begin{aligned} I &= \frac{4\pi}{\epsilon} \frac{1}{2\sqrt{\zeta^2 - \epsilon^2}/\epsilon} \\ &= \frac{2\pi}{\sqrt{\zeta^2 - \epsilon^2}} \end{aligned} \tag{1066}$$

and

$$\begin{aligned} -\frac{dI}{d\zeta} &= \int_0^{2\pi} d\phi \, \frac{1}{(\zeta + \epsilon\cos\phi)^2} \\ &= -2\pi \frac{d}{d\zeta} (\zeta^2 - \epsilon^2)^{-1/2} \\ &= -2\pi(-1/2)(\zeta^2 - \epsilon^2)^{-3/2}(2\zeta) \\ &= 2\pi\zeta \frac{1}{(\zeta^2 - \epsilon^2)^{3/2}} \end{aligned} \tag{1067}$$

also

$$\frac{d^n}{d\zeta^n} I = (2\pi)(-1)^n \, n! \int_0^{2\pi} d\phi \, \frac{1}{(\zeta + \epsilon\cos\phi)^{n+1}} \tag{1068}$$

Then

$$\begin{aligned} T &= \int dt \\ &= \frac{m}{L}\int_0^{2\pi} d\phi \frac{\alpha^2}{(1+\epsilon\cos\phi)^2} \\ &= -\frac{m\alpha^2}{L}\,\frac{d}{d\zeta}\int_0^{2\pi} d\phi \frac{\alpha^2}{(\zeta+\epsilon\cos\phi)}\bigg|_{\zeta=1} \\ &= \frac{m\alpha^2}{L} 2\pi \frac{1}{(1-\epsilon^2)^{3/2}} \end{aligned} \tag{1069}$$

Remember

$$1-\epsilon^2 = -2EL^2/G^2m = 2|E|L^2/G^2m \tag{1070}$$

and

$$\alpha = \frac{L^2}{Gm} \tag{1071}$$

So

$$\begin{aligned} T &= 2\pi m \frac{L^4}{G^2m^2L}\left(\frac{G^2m}{2L^2|E|}\right)^{3/2} \\ &= \pi G\sqrt{\frac{m}{2}}|E|^{-2/3} \\ &= \pi G\sqrt{\frac{m}{2|E|^3}} \\ &= \pi G\sqrt{\frac{m}{2}}\sqrt{\frac{2^3n^6}{m^3G^6}} \\ &= \frac{2\pi n^3}{G^2m} \end{aligned} \tag{1072}$$

Calculate

$$\begin{aligned} \frac{1}{T}\int_0^T r dt &= \frac{1}{T}\int_0^{2\pi} r\frac{dt}{d\phi}d\phi \\ &= \frac{1}{T}\int_0^{2\pi} r\frac{mr^2}{L}d\phi \\ &= \frac{1}{T}\frac{m}{L}\int_0^{2\pi} r^3 d\phi \\ &= \frac{1}{T}\frac{m}{L}\int_0^{2\pi}\left(\frac{\alpha}{1+\epsilon\cos\phi}\right)^3 d\phi \end{aligned} \tag{1073}$$

$$\begin{aligned}
\frac{1}{T}\int_0^T r^s dt &= \frac{1}{T}\int_0^{2\pi} r^s \frac{dt}{d\phi} d\phi \\
&= \frac{1}{T}\int_0^{2\pi} r^s \frac{mr^2}{L} d\phi \\
&= \frac{1}{T}\frac{m}{L}\int_0^{2\pi} r^{2+s} d\phi \\
&= \frac{1}{T}\frac{m}{L}\int_0^{2\pi} \left(\frac{\alpha}{1+\epsilon\cos\phi}\right)^{2+s} d\phi \\
&= \frac{m\alpha^{2+s}}{LT}\int_0^{2\pi} \left(\frac{1}{1+\epsilon\cos\phi}\right)^{2+s} d\phi \qquad (1074)
\end{aligned}$$

The integral is is an integral expression of the Legendre polynomials so that

$$\begin{aligned}
R_s &\equiv \frac{1}{T}\int_0^T r^s dt \\
&= \frac{m\alpha^{2+s}}{LT}\int_0^{2\pi} \left(\frac{1}{1+\epsilon\cos\phi}\right)^{2+s} d\phi \\
&= \frac{m\alpha^{2+s}}{LT} 2\pi (1-\epsilon^2)^{-(s+2)/2} P_{s+1}(1/\sqrt{1-\epsilon^2}) \qquad (1075)
\end{aligned}$$

Let $z = 1/\sqrt{1-\epsilon^2}$. Then

$$P_{s+1}(z) = \frac{LT}{2\pi m}\frac{1}{\alpha^{s+2} z^{(s+2)}} R_s \qquad (1076)$$

The Legendre Polynomials satisfy the following recursion relation:

$$(2s+1) z P_s(z) = (s+1) P_{s+1}(z) + s P_{s-1}(z) \qquad (1077)$$

So R_s's satisfy

$$0 = (s+1)\frac{LT}{2\pi m}\frac{1}{\alpha^{s+2} z^{(s+2)}} R_s - (2s+1)\frac{LT}{2\pi m}\frac{z}{\alpha^{s+1} z^{(s+1)}} R_{s-1} + s\frac{LT}{2\pi m}\frac{1}{\alpha^s z^{(s)}} R_{s-2} \qquad (1078)$$

Or

$$0 = s\alpha^2 R_{s-2} - (2s+1)\alpha R_{s-1} + (s+1)\frac{1}{z^2}R_s \tag{1079}$$

Using

$$1/z^2 = 1 - \epsilon^2 = \frac{2L^2|E|}{G^2 m} = \frac{2L^2}{G^2 m}\frac{G^2 m}{2n^2} = \frac{L^2}{n^2} \tag{1080}$$

and

$$\alpha = \frac{L^2}{Gm} \tag{1081}$$

yields

$$0 = s\frac{L^4}{G^2 m^2}R_{s-2} - (2s+1)\frac{L^2}{Gm}R_{s-1} + (s+1)\frac{L^2}{n^2}R_s \tag{1082}$$

or

$$0 = sL^2 a_0^2 R_{s-2} - (2s+1)a_0 R_{s-1} + (s+1)\frac{1}{n^2}R_s \tag{1083}$$

with $a_0 = 1/Gm$.

This should be compared with the QM result

$$0 \;=\; s\left[(l+1/2)^2 - s^2/4\right]a_0^2\langle r^{s-2}\rangle - (2s+1)a_0\langle r^{s-1}\rangle + \frac{s+1}{n^2}\langle r^s\rangle \tag{1084}$$

Consider this:

$$\bar{r^n} = \frac{1}{T}\int_0^\infty dt\, r^n \tag{1085}$$

so

$$\begin{aligned} P(r) &= \frac{1}{T}\frac{dt}{dr} \\ &= \frac{L}{m\alpha^2}\frac{(1-\epsilon^2)^{3/2}}{2\pi} \times \frac{\alpha m}{L\sqrt{\epsilon^2 - (\alpha/r - 1)^2}} \\ &= \frac{(1-\epsilon^2)^{3/2}}{2\pi\alpha\sqrt{\epsilon^2 - (\alpha/r-1)^2}} \\ &= \frac{(1-\epsilon^2)^{3/2} r}{2\pi\alpha\sqrt{\epsilon^2 r^2 - (\alpha - r)^2}} \\ &= \frac{(1-\epsilon^2)^{3/2} r}{2\pi\alpha\sqrt{(\epsilon^2-1)r^2 + 2\alpha r - \alpha^2}} \end{aligned} \tag{1086}$$

since

$$\begin{aligned}\frac{dr}{dt} &= \frac{d}{dt}\frac{\alpha}{1+\epsilon\cos\phi} \\ &= \alpha\frac{\epsilon\sin\phi}{(1+\epsilon\cos\phi)^2}\dot{\phi} \\ &= \frac{r^2}{\alpha}\sqrt{\epsilon^2-(\alpha/r-1)^2}\frac{L}{mr^2} \\ &= \frac{L}{\alpha m}\sqrt{\epsilon^2-(\alpha/r-1)^2} \qquad (1087)\end{aligned}$$

with

$$\frac{\alpha}{r} = 1+\epsilon\cos\phi \qquad (1088)$$

But this is good only between

$$r_{\min} = \frac{\alpha}{1+\epsilon} \qquad (1089)$$

$$r_{\max} = \frac{\alpha}{1-\epsilon} \qquad (1090)$$

(For bound state, $\epsilon < 1$.)

Calculate

$$\begin{aligned}\frac{1}{T}\int_0^T \frac{1}{r}dt &= \frac{1}{T}\int_0^{2\pi}\frac{1}{r}\frac{dt}{d\phi}d\phi \\ &= \frac{1}{T}\int_0^{2\pi}\frac{1}{r}\frac{mr^2}{L}d\phi \\ &= \frac{1}{T}\frac{m}{L}\int_0^{2\pi} r d\phi \\ &= \frac{1}{T}\frac{m}{L}\int_0^{2\pi}\frac{\alpha}{1+\epsilon\cos\phi}d\phi \\ &= \frac{\alpha}{T}\frac{m}{L}\frac{2\pi}{\sqrt{1-\epsilon^2}} \\ &= \frac{\alpha}{\pi G\sqrt{\frac{m}{2}}|E|^{-2/3}}\frac{m}{L}\frac{2\pi}{\sqrt{2|E|L^2/G^2m}} \\ &= 2\frac{\alpha m}{|E|^{-1}}\frac{1}{L^2}\end{aligned}$$

$$
\begin{aligned}
&= \frac{\alpha m|E|}{L^2} \\
&= \frac{L^2 m}{Gm}\frac{|E|}{L^2} \\
&= \frac{|E|}{G} \\
&= \frac{mc^2G^2}{2Gn^2} \\
&= \frac{mG}{2n^2} \qquad (1091)
\end{aligned}
$$

Note:

$$
|E| = \frac{mG^2}{2n^2} \qquad (1092)
$$

$$
\begin{aligned}
\frac{1}{T}\int_0^T \frac{1}{r^2}dt &= \frac{1}{T}\int_0^{2\pi} \frac{1}{r^2}\frac{dt}{d\phi}d\phi \\
&= \frac{1}{T}\int_0^{2\pi} \frac{1}{r^2}\frac{mr^2}{L}d\phi \\
&= \frac{1}{T}\frac{m}{L}\int_0^{2\pi} d\phi \\
&= \frac{1}{T}\frac{m}{L}2\pi \\
&= \frac{1}{\pi G\sqrt{\frac{m}{2|E|^3}}}\frac{m}{L}2\pi \\
&= \frac{1}{G\sqrt{\frac{m}{2|E|^3}}}\frac{m}{L}2 \\
&= \sqrt{\frac{2|E|^3}{m}}\frac{m}{LG}2 \\
&= \sqrt{8m|E|^3}\frac{1}{LG} \\
&= \sqrt{\frac{8m^4G^6}{2^3n^6}}\frac{1}{LG} \\
&= \frac{m^2G^2}{n^3L} \qquad (1093)
\end{aligned}
$$

$$\begin{aligned}
\frac{1}{T}\int_0^T \frac{1}{r^3}dt &= \frac{1}{T}\int_0^{2\pi} \frac{1}{r^3}\frac{dt}{d\phi}d\phi \\
&= \frac{1}{T}\int_0^{2\pi} \frac{1}{r^3}\frac{mr^2}{L}d\phi \\
&= \frac{1}{T}\frac{m}{L}\int_0^{2\pi} d\phi\frac{1}{r} \\
&= \frac{1}{T}\frac{m}{L}\int_0^{2\pi} d\phi\frac{1+\epsilon\cos\phi}{\alpha} \\
&= \frac{1}{T}\frac{m}{L}\frac{2\pi}{\alpha} \\
&= \frac{G^2m}{2\pi n^3}\frac{m}{L}\frac{2\pi Gm}{L^2} \\
&= \frac{G^3m^3}{n^3L^3} \qquad (1094)
\end{aligned}$$

In general, we have

$$\begin{aligned}
\frac{1}{T}\int_0^T \frac{1}{r^s}dt &= \frac{1}{T}\int_0^{2\pi} \frac{1}{r^s}\frac{dt}{d\phi}d\phi \\
&= \frac{1}{T}\int_0^{2\pi} \frac{1}{r^s}\frac{mr^2}{L}d\phi \\
&= \frac{1}{T}\frac{m}{L}\int_0^{2\pi} d\phi\frac{1}{r^{s-2}} \\
&= \frac{1}{T}\frac{m}{L}\int_0^{2\pi} d\phi\left(\frac{1+\epsilon\cos\phi}{\alpha}\right)^{s-2} \\
&= \frac{1}{T}\frac{m}{L\alpha^{s-2}}\int_0^{2\pi} d\phi\,(1+\epsilon\cos\phi)^{s-2} \qquad (1095)
\end{aligned}$$

The integral is is an integral expression of the Legendre polynomials so that

$$\begin{aligned}
R_{-s} &\equiv \frac{1}{T}\int_0^T \frac{1}{r^s}dt \\
&= \frac{1}{T}\frac{m}{L\alpha^{s-2}}\int_0^{2\pi} d\phi\,(1+\epsilon\cos\phi)^{s-2} \\
&= \frac{1}{T}\frac{m}{L\alpha^{s-2}}2\pi(1-\epsilon^2)^{(s-2)/2}\,P_{s-2}(1/\sqrt{1-\epsilon^2}) \qquad (1096)
\end{aligned}$$

Or

$$P_s(1/\sqrt{1-\epsilon^2}) = \frac{LT\alpha^s}{m}\frac{1}{2\pi(1-\epsilon^2)^{s/2}}\,R_{-s-2} \tag{1097}$$

Let $z = 1/\sqrt{1-\epsilon^2} \propto n$. Then

$$P_s(z) = \frac{LT}{2\pi m}\,\alpha^s z^s R_{-s-2} \tag{1098}$$

The Legendre Polynomials satisfy the following recursion relation:

$$(2s+1)zP_s(z) = (s+1)P_{s+1}(z) + sP_{s-1}(z) \tag{1099}$$

So I_s's satisfy

$$(2s+1)z\frac{LT}{2\pi m}\,\alpha^s z^s R_{-s-2} = (s+1)\frac{LT}{2\pi m}\,\alpha^{s+1}z^{s+1}R_{-s-3} + s\frac{LT}{2\pi m}\,\alpha^{s-1}z^{s-1}R_{-s-1} \tag{1100}$$

or

$$(2s+1)\alpha R_{-s-2} = (s+1)\alpha^2 R_{-s-3} + s\frac{1}{z^2}R_{-s-1} \tag{1101}$$

or

$$(2s-1)\alpha R_{-s-1} = s\alpha^2 R_{-s-2} + (s-1)\frac{1}{z^2}R_{-s} \tag{1102}$$

or leting $-s = \nu$,

$$(-2\nu-1)\alpha R_{\nu-1} = -\nu\alpha^2 R_{\nu-2} + (-\nu-1)\frac{1}{z^2}R_{\nu} \tag{1103}$$

or

$$0 = (-2\nu-1)\alpha R_{\nu-1} + \nu\alpha^2 R_{\nu-2} + (\nu+1)\frac{1}{z^2}R_{\nu} \tag{1104}$$

For positive n, we had

$$0 = s\alpha^2 R_{s-2} - (2s+1)\alpha R_{s-1} + (s+1)\frac{1}{z^2}R_s \tag{1105}$$

which then turned out to be good for both positive and negative n.

D Supplemental – Symmetry and Conservation Laws

In classical mechanics, given a Lagrangian $L(q_i)$, the equation of motion is given by

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \tag{1106}$$

Now suppose a transformation

$$\mathbf{q}_i \rightarrow \mathbf{q}_i + \boldsymbol{\epsilon} \tag{1107}$$

leaves the action $S = \int dt\, L$ unchanged. Here $\boldsymbol{\epsilon}$ is constant. That means

$$\int dt\, L(\mathbf{q_i} + \boldsymbol{\epsilon}, \dot{\mathbf{q}}_\mathbf{i}) = \int dt\, L(\mathbf{q_i}, \dot{\mathbf{q}}_\mathbf{i}) + \int dt\, \boldsymbol{\epsilon} \cdot \sum_{i=1}^{N} \frac{\partial L}{\partial \mathbf{q}_i} \tag{1108}$$

or

$$\boldsymbol{\epsilon} \cdot \sum_{i=1}^{N} \frac{\partial L}{\partial \mathbf{q_i}} = 0 \tag{1109}$$

since $\boldsymbol{\epsilon}$ is arbitrary,

$$\sum_{i=1}^{N} \frac{\partial L}{\partial \mathbf{q_i}} = \frac{d}{dt} \sum_{i=1}^{N} \mathbf{p_i} = 0 \tag{1110}$$

So the symmetry of the Lagrangian directly leads to the conservation of total momentum.

In general suppose

$$q_i \rightarrow q_i + \epsilon \phi_i(q_i) \tag{1111}$$

Then

$$\begin{aligned} \delta S &= \int dt\, L(q + \epsilon\phi, \dot{q} + \epsilon\dot{\phi}) - \int dt\, L(q, \dot{q}) \\ &= \int dt \left[\epsilon\phi \frac{\partial L}{\partial q} + \epsilon\dot{\phi} \frac{\partial L}{\partial \dot{q}} \right] \\ &= \epsilon \int dt \left[\phi \frac{\partial L}{\partial q} - \phi \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \right] + \epsilon\, \phi \frac{\partial L}{\partial \dot{q}} \Big|_{t_i}^{t_f} = 0 \end{aligned} \tag{1112}$$

But the first term in the last line vanishes upon using the equation of motion. That means that

$$\phi \frac{\partial L}{\partial \dot{q}}\bigg|_{t_i}^{t_f} = \phi \frac{\partial L}{\partial \dot{q}}\bigg|_{t_i} - \phi \frac{\partial L}{\partial \dot{q}}\bigg|_{t_f} = 0 \tag{1113}$$

or

$$\phi_i \frac{\partial L}{\partial \dot{q}_i} = \text{Constant} \tag{1114}$$

along the path of the motion.

For instance, suppose L is invariant under rotation.

$$x \to x - y\epsilon \tag{1115}$$
$$y \to y + x\epsilon \tag{1116}$$

Then

$$\phi_x = -y \tag{1117}$$
$$\phi_y = x \tag{1118}$$

it immdiately follows that

$$\phi_i \frac{\partial L}{\partial \dot{q}_i} = x \frac{\partial L}{\partial \dot{y}} - y \frac{\partial L}{\partial \dot{x}} \tag{1119}$$

But then by construction,

$$\frac{\partial L}{\partial \dot{x}} = p_x \tag{1120}$$
$$\frac{\partial L}{\partial \dot{y}} = p_y \tag{1121}$$

so that

$$\phi_i \frac{\partial L}{\partial \dot{q}_i} = x \frac{\partial L}{\partial \dot{y}} - y \frac{\partial L}{\partial \dot{x}} = x p_y - y p_x = \text{Constant}. \tag{1122}$$

E Runge-Lenz

Another way of seeing it is as follows. The principal quantum number is actually a combination of radial and angular:

$$n = n_r + l + 1 \tag{1123}$$

So that it depends on the combination only instead of something like

$$n_r l \tag{1124}$$

This happens only if the oscillation in the radial direction and the oscillation in the angle are **commensurate**. That is, in phase. If they are not in phase, the orbit can't stay in one place. It will precess. The fact that it does not precess is given by the constancy of $\mathbf{M}$ and the fact that n depends only on the combination.

> More: Why this degeneracy? Remember that the principal quantum number n is made up of n_r and l or $n = n_r + l + 1$. So the same energy can result from no radial excitation $n_r = 0$ and all angular excitation $l = n-1$ or all radial excitation $n_r = n-1$ and no angular excitation $l = 0$. The former correspond to a circular orbit. In this case, the radius is constant. So there is no momenum component attached to the changing of the radius. The latter correspond to the maximally elliptic orbit. A highly elliptic orbit has big $r_{\max}/r_{\min}$ ratio. That means the radius changes a lot. Maximally elliptic orbit has no angular momentum, that is, it is a straight line trajectory.
>
> Classically, we have
>
> $$r = \frac{\alpha}{1 + \epsilon \cos\phi} \tag{1125}$$
>
> with
>
> $$\alpha = \frac{L^2}{Ze^2 m} \tag{1126}$$
>
> $$\epsilon^2 = 1 - \frac{2|E|\alpha}{Ze^2} \tag{1127}$$

or

$$|E| = (1-\epsilon^2)\frac{Ze^2}{2\alpha} \tag{1128}$$

So given $|E|$, you can find an infinite number of α and ϵ that can satisfy this. Now the correspondence principle states that for large quantum numbers, QM goes to CM. Therefore, Quantum mechanics must somehow reflect such classical degeneracy at least in the classical limit.

Now we know what α corresponds to. It is the magnitude of the angular momentum. What about ϵ? Is there an independent quantity that gives us ϵ?

The Lenz-Lunge vector is given by

$$\mathbf{M} = \frac{1}{\mu}(\mathbf{p}\times\mathbf{L}) - \frac{G}{r}\mathbf{r} \tag{1129}$$

Squaring it yields

$$M^2 = \frac{1}{\mu^2}(\mathbf{p}\times\mathbf{L})^2 + G^2 - \frac{2G}{\mu}(\mathbf{p}\times\mathbf{L})\cdot\frac{\mathbf{r}}{r} \tag{1130}$$

we have

$$(\mathbf{p}\times\mathbf{L})^2 = p^2L^2 - (\mathbf{p}\cdot\mathbf{L})^2 = p^2L^2 \tag{1131}$$

and we use

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \epsilon_{ijk}A_iB_jC_k = \epsilon_{kij}C_kA_iB_j = \mathbf{C}\cdot(\mathbf{B}\times\mathbf{A}) \tag{1132}$$

to get

$$\mathbf{r}\cdot(\mathbf{p}\times\mathbf{L}) = \mathbf{L}\cdot(\mathbf{r}\times\mathbf{p}) = L^2 \tag{1133}$$

so

$$M^2 = \frac{p^2L^2}{\mu^2} + G^2 - \frac{2GL^2}{\mu r} \tag{1134}$$

Now we know that

$$\frac{p^2}{2\mu} - \frac{G}{r} = -|E| \tag{1135}$$

so

$$\begin{aligned} M^2 &= \frac{2L^2}{\mu}\left(\frac{G}{r} - |E|\right) + G^2 - \frac{2GL^2}{\mu r} \\ &= G^2 - \frac{2|E|L^2}{\mu} \end{aligned} \tag{1136}$$

Remember

$$\epsilon^2 = 1 - 2|E|L^2/G^2 m \tag{1137}$$

for bound states. Hence

$$M^2 = G^2\epsilon^2 \tag{1138}$$

So the magnintude of M^2 is, apart from the constant G^2, the eccentricity ϵ. But that's not a big deal since by fixing E and L we fix ϵ. What it does not fix is the directionof $\mathbf{M}$. So how do we quantify that?

Now consider an operation that changes the value of L_z. A simple rotation in along x axis will do. These operations must leave the energy intact. In classical mechanics, of course it does. In QM, changing L_z is accomplished by the operators

$$L_\pm = L_x \pm iL_y \tag{1139}$$

since

$$L \pm |n, l, m\rangle = \sqrt{l(l+1) - m(m \pm 1)}|n, l, m \pm 1\rangle \tag{1140}$$

and since

$$[\hat{H}, \hat{L}_\pm] = 0 \tag{1141}$$

this operation of changing the value of L_z with L^2 fixed nothing to the energy.

How do we represent chaning the direction of $\mathbf{M}$ quantum mechanically? Well, let's first write down the QM version of Runge-Lenz

$$\mathbf{M} = \frac{1}{2\mu}(\mathbf{p} \times \mathbf{L} - \mathbf{L} \times \mathbf{p}) - \frac{G}{r} \tag{1142}$$

Like in CM, this is conserved:

$$[H, \mathbf{M}] = 0 \tag{1143}$$

Since this is a vector, we must have the vector transformation rules:

$$[M_x, L_x] = 0 \tag{1144}$$
$$[M_x, L_y] = i\hbar M_z \tag{1145}$$
$$[M_x, L_z] = -i\hbar M_y \tag{1146}$$

and all other combinations. Self commutation is a bit more complicated:

$$[M_i, M_j] = -\frac{2i\hbar}{m} H \epsilon_{ijk} L_k \tag{1147}$$