PHYS294 Tutorial 2

Question 9.15 from the textbook

February 5, 2024

If you want, you can only focus on the first part. I put the second part for more extra OPTIONAL information on where the quantization of the corrected energy levels in the Zeeman effect comes from. Although it's optional, I think it will help you understand where the quantization of angular momentum comes from.

1 Question 9.15

Recall that the shift in the energy spectrum is given by the following formula once the external magentic field is applied:

$$\Delta E = m\mu_B B \tag{1}$$

where $m = -l, l+1, \ldots, l-1, l$ (quantum number), $\mu_B = \frac{e\hbar}{2m_e}$ is the Bohr magneton, and B is the external magnetic field. Since we are given $l = 0$ and $l = 3$ for the two electrons in the Helium atom, the energy correction for the $l = 3$ electron is given by

$$E = E_0 + m\mu_B B \tag{2}$$

with $m = -3, -2, -1, 0, 1, 2, 3$. So the original $2l + 1$ degenerate level split into seven levels. The separation between the adjacent levels are simply

$$\mu_B B = \left(\frac{5.79 \times 10^{-5}\text{eV}}{T}\right)(0.8T) = 4.6 \times 10^{-5}\text{eV} \tag{3}$$

We will now briefly go over why the energy correction is the way it is.

Consider a classical orbit of an electron going around a atomic nucleus in a circular motion. Since electron is a charged object, it going around in a circle will create a current loop and a magnetic field. If we now apply an external magnetic field, the magnetic field creating by the current loop will interact with the external B field (Just like a compass needle interacting with an external B field).

If the magnetic field created by the loop current aligns with the external magnetic field, the orbit is at the lowest potential. However, suppose the two magnetic fields are anti-parallel. Then, the disk associated with the electron current loop will tend to re-align itself to make sure that the two magnetic fields align. That is, there is a torque felt by the electron orbit. This toque is given by

$$\mathbf{\Gamma} = i\mathbf{A} \times \mathbf{B} = (i\mathbf{A}) \times \mathbf{B} = \mu \times \mathbf{B} \quad \Rightarrow \quad \Gamma = \mu B \sin\theta \tag{4}$$

where we have defined $\mu = i\mathbf{A}$ as the magnetic moment of the loop (with i being the current and $\mathbf{A}$ being the area-vector). Recall that the potential energy due to this two magnetic interaction can be obtained as the negative of work. That is, since work is defined as

$$W = -\int \Gamma d\theta = -\int \mu B \sin\theta d\theta = \mu B \cos\theta + C \tag{5}$$

Since the potential can be set to have any reference point, we can neglect C and write

$$U = -\mu B \cos\theta = -\mu \cdot \mathbf{B}. \tag{6}$$

When the angle between the two magnetic fields are parallel ($\theta = 0$), the potential is at its minimum $U = -\mu B$. This is because the two magnetic field orientations are already at its favoured energy configuration. However, if the two magnetic fields are anti-parallel ($\theta = \pi/2$), the potential is at its maximum, $U = \mu B$. You can imagine this being a configuration where a mass is at the top of the hill so it has greater potential energy than the one at minimum.

Now that we figured out what the potential looks like between two interacting magnetic fields, let us adjust it to our specific situation. Since we are modelling the atom having a circular orbit of electron, the velocity and current are defined as

$$v = \frac{2\pi r}{T}, \quad i = \frac{e}{T} \tag{7}$$

respectively. Here, r is the radius of the orbit, T is the period, and e is the electric charge. The current has units of charge per time so it makes sense that it is defined as $i = e/T$. The magnetic moment is given as

$$\mu = iA = i(\pi r^2). \tag{8}$$

Replacing $i = \frac{ev}{2\pi r}$, we find

$$\mu = iA = i(\pi r^2) = \frac{ev}{2\pi r}(\pi r^2) = \frac{1}{2}evr. \tag{9}$$

The above formula is very similar to the angular momentum:

$$L = m_e vr \tag{10}$$

where m_e is the electron mass. The ratio between M and L is composed of the fundamental constants and is called the gyro-magnetic ratio:

$$\frac{\mu}{L} = \frac{\frac{1}{2}evr}{m_e vr} = \frac{e}{2m_e} \quad \Rightarrow \quad \mu = \frac{e}{2m_e}L \tag{11}$$

Going to a vector notation, we can therefore write

$$\boxed{\mu = -\frac{e}{2m_e}\mathbf{L}} \tag{12}$$

The minus sign comes from the fact that the charge of an electron is negative and the current is defined as the amount of positive charge being pumped.

In QM, the above equation can be directly applied with appropriate quantization of the angular momentum. That is, we are able to write the atomic energy as

$$E = E_0 + \Delta E = E_0 - \mu \cdot \mathbf{B} = E_0 - \left(-\frac{e}{2m_e}\mathbf{L}\right) \cdot \mathbf{B} \tag{13}$$

Assuming $\mathbf{B}$ is in the z direction, we have

$$E = E_0 + \frac{e}{2m_e}L_z B \tag{14}$$

here we wrote L_z as a number but in QM it is really an operator and it's corresponding eigenvalues are quantized as $m\hbar$ with m being the quantum number. We therefore have

$$\boxed{\Delta E = \frac{e\hbar}{2m_e}mB} \tag{15}$$

2 Why the quantization?

Let us begin with the classical angular momentum, which is given by the following expression:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \tag{16}$$

where $\mathbf{r}$ and $\mathbf{p}$ are position and momentum of a particle respectively. This means, on the quantum side, we have

$$\begin{cases} L_x = yp_z - zp_y \\ L_y = zp_x - xp_z \\ L_z = xp_y - yp_x \end{cases} \Rightarrow \begin{cases} \hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y = y\left(-i\hbar\frac{\partial}{\partial z}\right) - z\left(-i\hbar\frac{\partial}{\partial y}\right) \\ \hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z = z\left(-i\hbar\frac{\partial}{\partial x}\right) - x\left(-i\hbar\frac{\partial}{\partial z}\right) \\ \hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = x\left(-i\hbar\frac{\partial}{\partial y}\right) - y\left(-i\hbar\frac{\partial}{\partial x}\right) \end{cases} \tag{17}$$

First of all, unlike the classical angular momentum which *commutes* in all three directions, the quantum angular momentum operators fails to commute with one another. You can think operators as some matrices which in general do not commute with one another. This means that the order in which you do the multiplication matters. For example, suppose we have matrices A and B. Then, in general

$$AB \neq BA \tag{18}$$

To quantify this, we define a commutator as

$$\boxed{AB - BA = [A, B]} \tag{19}$$

If the above quantity is zero, the two operators commute. If they don't there is an associated uncertainty between the two operators (which I won't go over in this note). But you probably recall that there is an associated uncertainty between position $\hat{x}$ and the momentum $\hat{p}$ of a particle. You cannot simultaneously know the position and momentum with perfect accuracy. This is because these two operators do not commute. So let us write a definition of the commutator for this scenario:

$$[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x} \tag{20}$$

Because this is yet again another operator, to figure out what this object is, we need to act on some function $f(x)$. We have

$$\begin{aligned}
[\hat{x}, \hat{p}]\, f(x) &= (\hat{x}\hat{p} - \hat{p}\hat{x})\, f(x) \\
&= x(-i\hbar\frac{df(x)}{dx}) - (-i\hbar\frac{d(xf(x))}{dx}) \\
&= -i\hbar x f'(x) + i\hbar\,(f(x) + xf'(x)) \\
&= i\hbar f(x)
\end{aligned} \tag{21}$$

So we conclude that $[\hat{x}, \hat{p}] = i\hbar$. The reason why we got this number is simple. The order in which we multiply a function x and taking the derivative d/dx matters. These to operations do not commute.

Now that we know the basics of the commutation relation, let us look at the commutation relation for the each angular momentum operators:

$$\begin{cases}
\left[\hat{L}_x, \hat{L}_y\right] = i\hbar\hat{L}_z \\
\left[\hat{L}_y, \hat{L}_z\right] = i\hbar\hat{L}_x \\
\left[\hat{L}_z, \hat{L}_x\right] = i\hbar\hat{L}_y
\end{cases} \tag{22}$$

These can be obtained by using the canonical commutation relation $[\hat{x}, \hat{p}] = i\hbar$. For instance,

$$\begin{aligned}
\left[\hat{L}_x, \hat{L}_y\right] &= [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] \\
&= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] - [\hat{y}\hat{p}_z, \hat{x}\hat{p}_z] - [\hat{z}\hat{p}_y, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z]
\end{aligned} \tag{23}$$

Note that the two commutators in the middle are zero. Since the partial derivatives does not effect the position observable. We are left with

$$\begin{aligned}
\left[\hat{L}_x, \hat{L}_y\right] &= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] \\
&= \hat{y}\hat{p}_x\,[\hat{p}_z, \hat{z}] + \hat{p}_y\hat{x}\,[\hat{z}, \hat{p}_z] \\
&= (\hat{p}_y\hat{x} - \hat{y}\hat{p}_x)\,[\hat{z}, \hat{p}_z] \\
&= i\hbar\,(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) = i\hbar\hat{L}_z
\end{aligned} \tag{24}$$

This means, just like the position and the momentum which are non-commuting observable, the angular momentum in three directions will have associated uncertainty relation. However, there is an observable that commutes with all angular momentum operators in all three directions. This is given by

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \tag{25}$$

which is called the total angular momentum. One can check that indeed this will commute with all $\hat{L}_\alpha$ with $\alpha = \{x, y, z\}$: $\left[\hat{L}^2, \hat{L}_\alpha\right] = 0$. For instance,

$$\left[\hat{L}^2, \hat{L}_x\right] = \left[\hat{L}_x^2, \hat{L}_x\right] + \left[\hat{L}_y^2, \hat{L}_x\right] + \left[\hat{L}_z^2, \hat{L}_x\right] \tag{26}$$

The first term trivially commutes. There are the same operators. The non-commutative part comes from the last two terms. Using the commutator identity $[AB, C] = A[B, C] + [A, C]B$ (you can find these on Wikipedia page on "Commutator"). We can therefore write

$$\left[\hat{L}^2, \hat{L}_x\right] = \hat{L}_y\left[\hat{L}_y, \hat{L}_x\right] + \left[\hat{L}_y, \hat{L}_x\right]\hat{L}_y + \hat{L}_z\left[\hat{L}_z, \hat{L}_x\right] + \left[\hat{L}_z, \hat{L}_x\right]\hat{L}_z \tag{27}$$

The remaining commutators are the ones we already know from Eq. (22). Using another property of the commutator ($[\hat{A}, \hat{B}] = -\left[\hat{B}, \hat{A}\right]$), we finally have

$$\begin{aligned}\left[\hat{L}^2, \hat{L}_x\right] =&\hat{L}_y\left(-i\hbar\hat{L}_z\right) + \left(-i\hbar\hat{L}_z\right)\hat{L}_y + \hat{L}_z\left(i\hbar\hat{L}_y\right) + \left(i\hbar\hat{L}_y\right)\hat{L}_z \\ =&0.\end{aligned} \tag{28}$$

Likewise $\left[\hat{L}^2, \hat{L}_y\right] = 0$ and $\left[\hat{L}^2, \hat{L}_z\right] = 0$ Focusing on the commutation relation between $\hat{L}^2$ and $\hat{L}_z$, we are then able to write the following eigenvalue equation

$$\boxed{\hat{L}^2 f = \lambda f, \quad \hat{L}_z f = \mu f} \tag{29}$$

This is due to a theorem in Linear algebra that states that there exists a common eigenvector for two commuting matrices. A very simple way of proving this is given as the following. Suppose we have a non-degenerate eigenvalue equation

$$\hat{H}\left|\psi\right\rangle = E\left|\psi\right\rangle \tag{30}$$

If we act with an operator $\hat{A}$ that commutes with $\hat{H}$, we have

$$\hat{L}\hat{H}\left|\psi\right\rangle = \hat{L}E\left|\psi\right\rangle \tag{31}$$

Since $[\hat{L}, \hat{H}] = 0$ and $\hat{L}$ trivially commutes with the number E, we can write

$$\hat{H}\hat{L}\left|\psi\right\rangle = E\hat{L}\left|\psi\right\rangle \tag{32}$$

This means some set of vectors $\hat{L}\left|\psi\right\rangle$ is also eigenvectors of $\hat{H}$ since it satisfies the following equation

$$\hat{H}\left(\hat{L}\left|\psi\right\rangle\right) = E\left(\hat{L}\left|\psi\right\rangle\right) \tag{33}$$

Since $\left|\psi\right\rangle$ are non-degenerate, it must be the case that $\hat{L}\left|\psi\right\rangle$ is some constant multiple of $\left|\psi\right\rangle$. That is,

$$\hat{L}\left|\psi\right\rangle = \gamma\left|\psi\right\rangle \tag{34}$$

So, $\left|\psi\right\rangle$ is an eigenvector of $\hat{L}$ with eigenvector γ. The two equations are therefore

$$\hat{H}\left|\psi\right\rangle = E\left|\psi\right\rangle, \quad \hat{L}\left|\psi\right\rangle = \gamma\left|\psi\right\rangle \tag{35}$$

which is what we have in Eq. (29).

Let us now solve Eq. (29). First, it is convenient to introduce the so-called raising and lowering angular momentum operators. These are defined as

$$\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y \tag{36}$$

The commutation relation between these operators and $\hat{L}_z$ is given by

$$\begin{aligned}\left[\hat{L}_z, \hat{L}_\pm\right] &= \left[\hat{L}_z, \hat{L}_x \pm i\hat{L}_y\right] \\ &= \left[\hat{L}_z, \hat{L}_x\right] \pm i\left[\hat{L}_z, \hat{L}_y\right] \\ &=i\hbar\hat{L}_y \pm i\left(-i\hbar\hat{L}_x\right) \\ &= \pm\hbar\left(\hat{L}_x \pm i\hat{L}_y\right)\end{aligned} \tag{37}$$

So we have

$$\boxed{\left[\hat{L}_z, \hat{L}_\pm\right] = \pm\hbar\hat{L}_\pm \quad \text{and} \quad \left[\hat{L}^2, \hat{L}_\pm\right] = 0} \tag{38}$$

The second commutator is obvious. Note then we can write

$$\hat{L}^2\left(\hat{L}_\pm f\right) = \hat{L}_\pm\left(\hat{L}^2 f\right) = \hat{L}_\pm(\lambda f) = \lambda\left(\hat{L}_\pm f\right) \tag{39}$$

This means $\hat{L}_\pm f$ is an eigenfunction of $\hat{L}^2$ with the same eigenvalue λ. Furthermore,

$$\begin{aligned}\hat{L}_z\left(\hat{L}_\pm f\right) &= \left(\hat{L}_z\hat{L}_\pm - \hat{L}_\pm\hat{L}_z\right) f + \hat{L}_\pm\hat{L}_z f \\ &= \pm\,\hbar\hat{L}_\pm f + \hat{L}_\pm(\mu f) \\ &= (\mu \pm \hbar)\,\hat{L}_\pm f\end{aligned} \tag{40}$$

From which we can conclude $\hat{L}_\pm f$ is an eigenfunction of $\hat{L}_z$ with new eigenvalue $\mu \pm \hbar$. The raising operator will shift the eigenvalue μ by $+\hbar$ and the lowering operator will shift the eigenvalue by $-\hbar$. This means with given λ, we will get a ladder of states with the largest eigenvalue being λ. That is, the eigenvalue equation for the largest $\hat{L}_z$ is given as

$$\hat{L}_z f_t = \hbar l f_t; \quad \hat{L}^2 f_t = \lambda f_t \tag{41}$$

here, $\hbar l$ is the maximum value one could take for $\hat{L}_z$. From here, note

$$\begin{aligned}\hat{L}_\pm\hat{L}_\mp &= \left(\hat{L}_x \pm i\hat{L}_y\right)\left(\hat{L}_x \mp i\hat{L}_y\right) \\ &= \hat{L}_x^2 + \hat{L}_y^2 \mp i\left(\hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x\right)\end{aligned} \tag{42}$$

Note that the first two terms are the total angular momentum squared minus the z component squared. The last term is nothing but a commutator between $\hat{L}_x$ and $\hat{L}_y$. We therefore have

$$\hat{L}_\pm\hat{L}_\mp = \hat{L}^2 - \hat{L}_z^2 \mp i\left(i\hbar\hat{L}_z\right) \tag{43}$$

From which we can write

$$\hat{L}^2 = \hat{L}_\pm\hat{L}_\mp + \hat{L}_z^2 \mp \hbar\hat{L}_z \tag{44}$$

Acting the above operator with the top eigenvalue f_t, we have

$$\begin{aligned}\hat{L}^2 f_t &= \left(\hat{L}_-\hat{L}_+ + \hat{L}_z^2 + \hbar\hat{L}_z\right) f_t \\ &= \left(0 + \hbar^2 l^2 + \hbar^2 l\right) f_t \\ &= \hbar^2 l(l+1) f_t\end{aligned} \tag{45}$$

From which we can conclude $\boxed{\lambda = \hbar^2 l(l+1)}$. Similarly, there is a minimum eigenvector f_b that the angular momentum cannot be lowered any further. The equation is given as $\hat{L}_- f_b = 0$. Let $\hbar\bar{l}$ be the eigenvalue for that bottom ladder. The two equations are therefore

$$\hat{L}_z f_b = \hbar\bar{l} f_b \quad \text{and} \quad \hat{L}^2 f_b = \lambda f_b \tag{46}$$

Again using $\hat{L}^2$ in terms of $\hat{L}_\pm$ and $\hat{L}_z$, we have

$$\left(\hat{L}_+\hat{L}_- + \hat{L}_z^2 - \hbar\hat{L}_z\right) f_b = \left(0 + \hbar^2\bar{l}^2 - \hbar^2\bar{l}\right) f_b = \hbar^2\left(\bar{l}(\bar{l}-1)\right) f_b \tag{47}$$

So we have $\lambda = \hbar^2\left(\bar{l}(\bar{l}-1)\right)$. Since the two lambdas must be the same, we must have

$$\hbar^2 l(l+1) = \hbar^2\bar{1}(\bar{1}-1) \tag{48}$$

The only possible answer is that $\bar{l} = -l$. This means the eigenvalues for $\hat{L}_z$ goes from $-l$ to l in integer steps. We call this m. Since there are N integer steps from $-l$ to l, we have

$$l = -l + N \quad \Rightarrow \quad l = \frac{N}{2} \tag{49}$$

So the total angular momentum is eigher integer or half-integer spin.