Angular momentum in quantum mechanics

In this article, we will quantize angular momentum and calculate its eigenvalues. From classical mechanics, you may recall:

$$\vec{L} = \vec{r} \times \vec{p} \tag{1}$$

In other words, in Cartesian coordinate:

$$L_x = yp_z - zp_y \tag{2}$$

$$L_y = zp_x - xp_z \tag{3}$$

$$L_z = xp_y - yp_x \tag{4}$$

These formulas also make sense in quantum mechanics, if we interpret L_x,L_y,L_z,x,y,z,p_x,p_y,p_z not as numbers but as operators (i.e. matrices).

It is necessary to find the commutators between the different components of angular momentum to understand the quantum version of angular momentum. Using the following facts,

$$[x, p_x] = [y, p_y] = [z, p_z] = i\hbar \tag{5}$$

$$[x, y] = [y, z] = [z, x] = [p_x, p_y] = [p_y, p_z] = [p_z, p_x] = 0 \tag{6}$$

$$[x, p_y] = [x, p_z] = [y, p_x] = [y, p_z] = [z, p_x] = [z, p_y] = 0 \tag{7}$$

we obtain:

$$\begin{aligned} [L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\ &= [yp_z - zp_y, zp_x] - [yp_z - zp_y, xp_z] \\ &= [yp_z, zp_x] - [-zp_y, xp_z] \\ &= -i\hbar yp_x + i\hbar p_y x = i\hbar L_z \end{aligned} \tag{8}$$

Similarly, we can easily obtain:

$$[L_y, L_z] = i\hbar L_x, \qquad [L_z, L_x] = i\hbar L_y \tag{9}$$

Notice that $[L_a, L_b] = i\hbar L_c$ if a, b, c is in a cyclic order. We can actually write this as

$$[L_i, L_j] = i\hbar \epsilon_{ijk} L_k \tag{10}$$

where Einstein summation convention is assumed for repeated index k, even though it doesn't appear as an upper index and a lower index. Of course, we also assume i, j, k run from 1 to 3 and $L_1 = L_x, L_2 = L_y, L_3 = L_z$.

In our earlier article on Heisenberg's uncertainty principle, we mentioned that we can have a state with definite position, because $[x, y] = [y, z] = [z, x] = 0$. The same is true for momentum. However, you see here that the same is not true for angular momentum, because its components do not commute with each other; we cannot have a state with definite L_x, definite L_y and definite L_z at the same time. So, let's just try to find a state with definite L_z only. In other words, let's try to find the eigenvalues and eigenvectors of L_z. We could have found the eigenvalues and eigenvectors of L_x or L_y instead, but it is a common practice to consider L_z. After all, it doesn't matter, because calling which components are x-component, y-component and z-component is totally arbitrary. We would have gotten the same eigenvalues if we considered L_x and L_y instead of L_z.

From (9), we can easily obtain the following relations:

$$[L_z, L_x + iL_y] = \hbar(L_x + iL_y) \tag{11}$$

$$[L_z, L_x - iL_y] = -\hbar(L_x - iL_y) \tag{12}$$

If we define as follows:

$$L_+ \equiv L_x + iL_y \tag{13}$$

$$L_- \equiv L_x - iL_y \tag{14}$$

we can rewrite the above formulas as follows:

$$[L_z, L_+] = \hbar L_+, \qquad [L_z, L_-] = -\hbar L_- \tag{15}$$

Moreover, let's define "L^2" as follows:

$$L^2 = L_x^2 + L_y^2 + L_z^2 \tag{16}$$

This corresponds to the square of the magnitude of angular momentum. Then, it is easy to check the followings:

$$[L^2, L_x] = [L^2, L_y] = [L^2, L_z] = 0 \tag{17}$$

$$[L^2, L_+] = [L^2, L_-] = 0 \tag{18}$$

$$L^2 = L_z^2 + L_z + L_- L_+ = L_z^2 - L_z + L_+ L_- \tag{19}$$

Now, let's consider a simultaneous eigenvector of L^2 and L_z. Such a vector always exists, as L^2 and L_z commute which implies that we can always find a basis in which the matrices L^2 and L_z are diagonal. Let's denote such a vector $|(\tilde{l}^2), m\rangle$, where

$$L^2|(\tilde{l}^2), m\rangle = \tilde{l}^2\hbar^2|(\tilde{l}^2), m\rangle \tag{20}$$

$$L_z|(\tilde{l}^2), m\rangle = m\hbar|(\tilde{l}^2), m\rangle \tag{21}$$

In other words, we label such a vector by the eigenvalues of L^2 and L_z. Note that $\tilde{l}^2$ and m are dimensionless, as $\hbar$ has the dimension of angular momentum.

Then, it is easy to see the followings:

$$\begin{aligned} L^2 L_+|(\tilde{l}^2), m\rangle &= L_+ L^2|(\tilde{l}^2), m\rangle \\ &= L_+ l^2\hbar^2|(\tilde{l}^2), m\rangle \\ &= \tilde{l}^2\hbar^2 L_+|(\tilde{l}^2), m\rangle \end{aligned} \tag{22}$$

Therefore, we conclude:

$$L^2(L_+|(\tilde{l}^2), m\rangle = \tilde{l}^2\hbar^2(L_+|(\tilde{l}^2), m\rangle) \tag{23}$$

In other words, $L_+|(\tilde{l}^2), m\rangle$ is an eigenvector of L^2 with eigenvalues $\tilde{l}^2\hbar^2$. We also have:

$$\begin{aligned} L_z L_+|(\tilde{l}^2), m\rangle &= (L_+ L_z + \hbar L_+)|(\tilde{l}^2), m\rangle \\ &= L_+ m\hbar|(\tilde{l}^2), m\rangle + \hbar L_+|(\tilde{l}^2), m\rangle \\ &= (m+1)\hbar L_+|(\tilde{l}^2), m\rangle \end{aligned} \tag{24}$$

Therefore, we conclude:

$$L_z(L_+|(\tilde{l}^2), m\rangle) = (m+1)\hbar(L_+|(\tilde{l}^2), m\rangle) \tag{25}$$

In other words, $L_+|(\tilde{l}^2), m\rangle$ is an eigenvector of L_z with eigenvalues $(m+1)\hbar$.

Problem 1. Similarly, show that $L_-|(\tilde{l}^2), m\rangle$ is an eigenvector of L^2 and L_z with eigenvalues of $\tilde{l}^2\hbar^2$ and $(m-1)\hbar$.

Now, given $|(\tilde{l}^2), m\rangle$, we can apply multiple numbers of L_+s or L_-s to construct an eigenvector of L^2 and L_z with the same eigenvalue of L^2, but a different eigenvalue of L_z. For example,

$$L_z\left((L_+)^n|(\tilde{l}^2), m\rangle\right) = (m+n)\hbar\left((L_+)^n|(\tilde{l}^2), m\rangle\right) \tag{26}$$

$$L_z\left((L_-)^n|(\tilde{l}^2), m\rangle\right) = (m-n)\hbar\left((L_-)^n|(\tilde{l}^2), m\rangle\right) \tag{27}$$

In other words, $(L_+)^n|(\tilde{l}^2), m\rangle$ is proportional to $|(\tilde{l}^2), m+n\rangle$ and $(L_-)^n|(\tilde{l}^2), m\rangle$ is proportional to $|(\tilde{l}^2), m-n\rangle$.

At first glance, given an eigenvector $|(\tilde{l}^2), m\rangle$, it may seem that one can construct other eigenvectors with L_z eigenvalues as high as or as low as we want, if we choose big enough n. But, this is troublesome, as the square of L_z should never be bigger than L^2 eigenvalues, as:

$$L^2 - L_z^2 = L_x^2 + L_y^2 \geq 0 \tag{28}$$

However, the formulas (26) and (27) are always satisfied. The only possibility is that $(L_+)^n|(\tilde{l}^2), m\rangle$ is zero for big ns. Similarly for $(L_-)^n|(\tilde{l}^2), m\rangle$. This implies that there exists k such that $(L_+)^k|(\tilde{l}^2), m\rangle \neq 0$, but $(L_+)^{k+1}|(\tilde{l}^2), m\rangle = 0$. Such a vector $(L_+)^k|(\tilde{l}^2), m\rangle$ is called highest weight vector. Now, let's consider a highest weight vector $|(\tilde{l}^2), l\rangle$. Then, we have:

$$L_+|(\tilde{l}^2), l\rangle = 0 \tag{29}$$

which implies

$$0 = L_- L_+ |(\tilde{l^2}), l\rangle = (L^2 - L_z^2 - \hbar L_z)|(\tilde{l^2}), l\rangle \tag{30}$$
$$= (\tilde{l^2} - l^2 - l)\hbar^2 |(\tilde{l^2}), l\rangle \tag{31}$$

As $|(\tilde{l^2}), l\rangle \neq 0$, we conclude

$$\tilde{l^2} = l^2 + l = l(l+1) \tag{32}$$

In fact, it is customary to use the notation $|l,\ m\rangle$ instead of $|(\tilde{l^2}), m\rangle$ or $|(l(l+1)), m\rangle$. From now on, we will use this notation. In other words,

$$L^2|l,\ m\rangle = l(l+1)\hbar^2|l,\ m\rangle \tag{33}$$
$$L_z|l,\ m\rangle = m\hbar|l,\ m\rangle \tag{34}$$

On the other hand, we can apply L_-s to $|l,\ l\rangle$ to construct eigenvectors with lower L_z eigenvalues. Of course, this process should terminate at some point and the vector so obtained must be zero. Let's find how lowest the eigenvalue of L_z can be. Let's say $(L_-)^a|l,\ l\rangle \neq 0$, while $(L_-)^{a+1}|l,\ l\rangle = 0$. Then, we have:

$$\begin{aligned} 0 &= L_+(L_-)^{a+1}|l,\ l\rangle = L_+ L_-(L_-)^a|l,\ l\rangle && (35)\\ &= (L^2 - L_z^2 + \hbar L_z)(L_-)^a|l,\ l\rangle && (36)\\ &= \left(l(l+1) - (k-a)^2 + (k-a)\right)\hbar^2|l,\ l\rangle && (37) \end{aligned}$$

Since $(L_-)^a|l,\ l\rangle \neq 0$, we have:

$$l(l+1) = (l-a)^2 - (l-a) \tag{38}$$

Therefore, we obtain:

$$a = 2l, -1 \tag{39}$$

As a can't be negative we conclude $a = 2l$. Now, notice that a should be an integer, as we cannot act "2.5" or "3.7" times of operator L_-. Therefore, we see that l must be non-negative half-integer as follows:

$$l = 0, 1/2, 1, 3/2, 2\cdots \tag{40}$$

Now, what is the lowest possible value for L_z given l? From (27), we have:

$$L_z\left((L_-)^{2l}|l,\ l\rangle\right) = -l\left((L_-)^{2l}|l,\ l\rangle\right) \tag{41}$$

Here, we see that $(L_-)^{2l}|l,\ l\rangle$ must be proportional to $|l,\ -l\rangle$.

Finally, we see that the following vectors

$$|l,\ l\rangle,\ |l,\ l-1\rangle,\ |l,\ l-2\rangle, \cdots, |l,\ -l+1\rangle,\ |l,\ -l\rangle \tag{42}$$

span the eigenvectors of L^2 with eigenvalue $l(l+1)$. It is also easy to see that this vector space is $(2l+1)$ dimensional.

Therefore, we conclude, in 3 space dimensions, both the square of magnitude of angular momentum and the value of the angular momentum along certain direction (for example $\hat{z}$ direction as considered in this article) are quantized.

Finally, let me conclude this article with four comments.

First, in our earlier article "Noether's theorem," we have seen that L_z generates the rotation around x-y plane. Recall that this is because

$$\{L_z, x\} = y, \qquad \{L_z, y\} = -x \tag{43}$$

Here, I want to mention that the z-component of the position doesn't change under the rotation around x-y plane, because $\{L_z, z\} = 0$, which you can check. What about the Poisson bracket between other components of angular momentum and the components of position? Remember the cylic order. We can change x to y and y to z and z to x. Thus,

$$\{L_x, y\} = z, \qquad \{L_x, z\} = -y, \qquad \{L_x, x\} = 0 \tag{44}$$

By changing x to y and y to z and z to x again, we get

$$\{L_y, z\} = x, \qquad \{L_y, x\} = -z, \qquad \{L_y, y\} = 0 \tag{45}$$

Summarizing all these, we can write

$$\{L_i, r_j\} = \epsilon_{ijk} r_k \tag{46}$$

where $r_1 = x$, $r_2 = y$, $r_3 = z$. Similarly, it is not hard to find

$$\{L_i, p_j\} = \epsilon_{ijk} p_k \tag{47}$$

Notice that (46) and (47) have the same form. In other words,

$$\{L_i, Q_j\} = \epsilon_{ijk} Q_k \tag{48}$$

where Q_i is r_i for (46) and p_i for (47). Is this a coincidence?

No. Recall that the angular momentum generates the rotation, and the rotation acts the same for all vectors. You sort of have seen this in our earlier article "Coriolis force, revisiting." For a constant vector $\vec{Q}$, you saw that the (anti-clockwise) rotation around the axis $\vec{\omega}$ is given by $\vec{\omega} \times \vec{Q}$, i.e., $\epsilon_{ijk}\omega_i Q_j$. This is easy to confirm from (48). The clockwise rotation around the axis $\vec{\omega}$ is generated by $\omega_i L_i$. Therefore, the anti-clockwise rotation around the same axis is generated by $-\omega_i L_i$. Thus,

$$\{-\omega_i L_i, Q_k\} = -\epsilon_{ikj}\omega_i Q_j = \epsilon_{ijk}\omega_i Q_j \tag{49}$$

which perfectly agrees. Then, what will happen if Q_j is the angular momentum L_j? By plugging in this (48), we obtain

$$\{L_i, L_j\} = \epsilon_{ijk} L_k \tag{50}$$

If we consider the relation between the Poisson bracket and the commutator, we see that we have derived (10) just from the property that a vector such as angular momentum must satisfy. Then, how about $[L_z, L^2]$ or $\{L_z, L^2\}$? L^2 is a scalar, as it is the dot product of two vectors, $\vec{L}$ and $\vec{L}$. A scalar doesn't transform under rotation. For example, a 30 cm ruler is 30 cm, no matter how you rotate it. Therefore, we conclude that $[L_z, L^2]$ is zero. Notice that we have derived (17) without doing any calculation.

Second, there are two types of angular momentum in quantum mechanics: orbital angular momentum and spin angular momentum (also called "spin"). Orbital angular momentum is due to the motion of particle while spin is not. Spin is an intrinsic angular momentum that has no classical counterparts. A particle has an angular momentum called spin even when it's not rotating or moving. Also, different particles have different spins. For example, Higgs boson has spin 0 (i.e., $l = 0$), electrons and quarks have spin 1/2 (i.e., $l = 1/2$), photon has spin 1 (i.e., $l = 1$) and graviton has spin 2 (i.e., $l = 2$). (Let me correct what I just said. Every particle that is not spin 0 has an angular momentum called spin even when it is not moving. Of course, if a particle is spin 0 it doesn't have spin.) As we explained in "Electron magnetic moment" we cannot attribute the spin of an electron to its motion, because the g-factor is not 1, but approximately 2. In our article "Spherical harmonics," we will give you yet another reason why we cannot attribute the spin of an electron to its motion.

In this article, we derived the commutation relations of angular momentum, namely, $[L_i, L_j] = i\hbar\epsilon_{ijk}L_k$ from the classical relation $\vec{L} = \vec{r} \times \vec{p}$ and by promoting $\vec{r}$ and $\vec{p}$ into quantum operators that satisfy certain commutation relations. As $\vec{r}$ and $\vec{p}$ concern the actual position and the actual momentum of particle, the commutation relation of angular momentum we derived only concern for orbital angular momentum, not for spin angular momentum. You are right. However, recall what we just said in the last comment. $[L_i, Q_j] = i\epsilon_{ijk}Q_k$ must be satisfied for any vector Q_i. This is also true when Q_i is the spin angular momentum, and L_i, the angular momentum concerned is the spin angular momentum. Thus, we see that $[L_i, L_j] = i\hbar\epsilon_{ijk}L_k$ must be true when L_i is the spin angular momentum. Then, the subsequent calculations and derivations will exactly follow like the ones presented in this article. Then, we will get the same conclusion as we obtained for orbital angular momentum. However, as you will see in "Spherical harmonics," for orbital angular momentum, only integer ls are allowed, while there is no same restriction for spin angular momentum. Instead, as we saw in (39), $2l$ must be a non-negative integer. In other words, for spin, we see that l must be a "half-integer." Anyhow, the concept of spin must be assumed in quantum mechanics. From the point of view of quantum mechanics, there is no reason why there should be spin, the intrinsic angular momentum. However, if you learn quantum field theory, you will see that the concept of spin comes out much more naturally.

Third, let me explain how Stern-Gerlach experiment is related to the spin of electron. As we just mentioned, electron has spin 1/2 (i.e., $l = 1/2$). It means that the eigenvalues of L_z are given by $\hbar/2$ and $-\hbar/2$, as we have $m = 1/2$ and $-1/2$. In Stern-Gerlach experiment, an external magnetic field is exerted on electron, which has a magnetic moment. Recall from our earlier article on magnetic dipole that $U = -\vec{\mu} \cdot \vec{B}$. Recall also from our earlier article on electron magnetic moment that the spin magnetic moment is given by $\vec{\mu}_s = -g(e/2m)\vec{S}$, where $\vec{S}$ is the spin angular momentum of an electron. Combining these two equations, the Hamiltonian of the electron is given by

$$H = -g\frac{e}{2m}\vec{S} \cdot \vec{B} \tag{51}$$

Without loss of generality, let's find a coordinate system in which $\vec{B}$ is aligned along z axis. In other words, let's set $\vec{B} = B\hat{z}$. Then, the above formula becomes

$$H = -g\frac{e}{2m}BS_z \tag{52}$$

where S_z is the z-component of spin angular momentum. What are the possible observed value of the Hamiltonian above? A Hamiltonian, just like any other observable, can have only its eigenvalue as its observed values. As S_z can be only $\hbar/2$ and $-\hbar/2$, we have

$$H = \mp g\frac{e}{4m}B\hbar \tag{53}$$

Thus, it can have only two values. This is the reason why only two spots are seen on the screen in Stern-Gerlach expperiment. S_z can indeed have only two values.

Fourth, let me comment that one uses mathematics behind the angular momentum to derive the area spectrum in loop quantum gravity, even though the area spectrum has nothing to do with the angular momentum itself.

Problem 2. Check the following. (Hint[1])

$$\langle l,\ m|L_x|l,\ m\rangle = \langle l,\ m|L_y|l,\ m\rangle = 0 \tag{54}$$

Problem 3. Verify that the norm of the following vector is $\hbar\sqrt{l(l+1) - m(m+1)}$:

$$L_+|l,\ m\rangle \tag{55}$$

(Hint[2])

Summary

- $[L_x, L_y] = i\hbar L_z$, $[L_y, L_z] = i\hbar L_x$, $[L_z, L_x] = i\hbar L_y$. In other words, $[L_i, L_j] = i\epsilon_{ijk}L_k$.
- $L_+ = L_x + iL_y$ raises the eigenvalue of L_z by $\hbar$ and $L_- = L_x - iL_y$ lowers the eigenvalue of L_z by $\hbar$.
- L_+ and L_- don't change the eigenvalues of $L^2 = L_x^2 + L_y^2 + L_z^2$.
- L_z and L^2 commute. Therefore, we can diagonalize them in a common basis.

[1] Express L_x and L_y in terms of L_+ and L_-. Use also the fact that L_z is a Hermitian matrix which implies that two vectors with different eigenvalues for L_z are orthogonal to each other.

[2] The solution is in the next article.

- $L^2|l,\ m\rangle = l(l+1)\hbar^2|l,\ m\rangle, \qquad L_z|l,\ m\rangle = m\hbar|l,\ m\rangle.$
- m can have values as $-l, -l+1, \cdots, l-1, l$.
- l can be only half-integers, such as $0, 1/2, 1, 3/2, 2, 5/2 \cdots$.