Notes on Quantum Mechanics: Angular Momentum

Jasper Chen

1 The Angular Momentum Operators

From the classical definition of Angular Momentum

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

we can define operators for each of the spatial dimension in 3D as follows:

$$\begin{aligned} L_x &= yp_z - zp_y \\ L_y &= zp_x + xp_z \\ L_z &= xp_y - yp_x \end{aligned} \tag{2}$$

The hat notation for operators have been abbreviated but it should be made aware that the above quantities are all operators. From these definitions, we can further define an operator for total angular momentum squared:

$$L^2 = L_x^2 + L_y^2 + L_z^2 \tag{3}$$

The commutator relations are as follows:

$$\begin{aligned} [L^2, L_{any}] &= 0 \\ [L_x, L_y] &= i\hbar L_z \\ [L_y, L_z] &= i\hbar L_x \\ [L_z, L_x] &= i\hbar L_y \end{aligned}$$

The form of the commutators suggest a "ladder operator" approach to this problem, similar to the approach taken in the 1D harmonic oscillator. Specifically, we can define operators such that their commutation relations will be self-referential, and thus leading to a "ladder" structure of states. We then define the following operators:

$$L_\pm = L_x \pm iL_y \tag{4}$$

Note that $L_+ = L_-^\dagger$. The commutation relations for these operators are:

$$[L^2, L_\pm] = 0 \tag{5}$$

$$[L_z, L_\pm] = \pm\hbar L_\pm \tag{6}$$

Since L^2 and L_z commute, we assert that there is a common eigenfunction f of both operators. Furthermore, let it be labeled by the integers l and m — such that $f \equiv Y_l^m$. The l will correspond to the eigenvalues of L^2 while m will correspond to the eigenvalues of L_z. Based on this notation, we assert further that

$$L^2 Y_l^m = \hbar^2 l(l+1) Y_l^m \tag{7}$$

$$L_z Y_l^m = \hbar m Y_l^m. \tag{8}$$

Based on the commutator relationship above, these two assertions lead to the following properties:

$$L^2 L_\pm Y_l^m = \hbar^2 l(l+1)(L_\pm Y_l^m) \tag{9}$$

$$L_z L_\pm Y_l^m = \hbar(m \pm 1)(L_\pm Y_l^m) \tag{10}$$

In words, if Y_l^m is an eigenfunction of L^2 and $L_\pm$, so is $L_\pm Y_l^m$. This fact validates our definition for $L_\pm$. It implies that there is a "ladder" of angular momentum states that can be traversed via the *lowering* and *raising* operators.

2 The Ladder

Now we have noticed that the eigenvalue of L_z after a transformation by $L_\pm$ has the value of m in its eigenvalue raised or lowered by 1. We can now shorten the notation by defining:

$$L_\pm Y_l^m = Y_l^{m\pm 1}. \tag{11}$$

Does this ladder have an upper and/or lower limit? Certainly — recall that each m value in the ladder corresponds to a value of L_z — if m is unbounded, the angular momentum in the z direction can become so large that it exceeds the total momentum, which is nonsense. A similar argument applies to the down direction. So there must be a maximum and a minimum m such that any state beyond that vanishes:

$$L_+ Y_l^{m_{max}} = 0 \tag{12}$$
$$L_- Y_l^{m_{min}} = 0. \tag{13}$$

By exploiting the fact that the inner product of two 0 vectors must be 0 and the fact that L_- is the adjoint of L_+, one can arrive at the conclusions that $m_{max} = l$ and $m_{min} = -l$. Therefore, in any given ladder with l, there are a total of $2l + 1$ states.

Before we proceed, I want to derive the normalization constant needed for the raising and lowering operator. We begin by assuming that

$$L_+ f_l^m = N_m^l f_l^{m+1}$$

and noticing that

$$L_+^\dagger = L_-$$

We then use an inner product and the Hermitian-pair nature of the raising and lowering operators to derive N_l^m given that f_l^m is already normalized.

$$\langle f_m^l | L_- L_+ f_m^l \rangle = \langle L_+ f_m^l | L_+ f_m^l \rangle = |N_l^m|^2$$

Then, using the algebraic fact that $L_- L_+ = L^2 - L_z^2 - \hbar L_z$:

$$\begin{aligned}
|N_l^m|^2 &= \langle f_m^l | (L^2 - L_z^2 - \hbar L_z) f_l^m \rangle \\
&= \langle f_m^l | L^2 f_m^l \rangle - \langle f_m^l | L_z^2 f_m^l \rangle - \langle f_m^l | \hbar L_z f_m^l \rangle \\
&= \hbar^2 l(l+1) - \hbar^2 m^2 - \hbar^2 m \\
&= \hbar^2 [l(l+1) - m^2 - m] \\
&= \hbar^2 [l(l+1) - m(m+1)]
\end{aligned}$$

By symmetry:

$$N_l^m = \hbar \sqrt{l(l+1) - m(m\pm 1)}$$

Note that when $m = \pm l$, N_l^m vanishes when obtained from $L_\pm$.

3 Y_l^m, or Spherical Harmonics

Rewriting the definition of Angular Momentum in terms of momentum operators [Eq (1)], we can derive the explicit form of the angular momentum operators in spherical coordinates:

$$L_x = \frac{\hbar}{i}\left(-\sin\phi \frac{\partial}{\partial\theta} - \cos\phi\cot\theta \frac{\partial}{\partial\phi}\right) \tag{14}$$
$$L_y = \frac{\hbar}{i}\left(+\cos\phi \frac{\partial}{\partial\theta} - \sin\phi\cot\theta \frac{\partial}{\partial\phi}\right) \tag{15}$$

and of course

$$L_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi} \tag{16}$$

The form of L_z is in accordance with the form of momentum and energy operators; all are reflections of Noether's Theorem.
Now we can explicitly derive a formula for Y_l^m. The process involves sophisticated methods using separation of variables and will not be discussed in detail here. Y_l^m forms a family of functions conventionally called *Spherical Harmonics* and takes the general form of:

$$Y_l^m = Ne^{im\phi}P_l^m(\cos\theta) \tag{17}$$

where N is a normalization constant and P_l^m is the *Associated Legendre Function* defined by:

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{\frac{|m|}{2}}\left(\frac{d}{dx}\right)^{|m|+l}(x^2-1)^l \tag{18}$$

with $m \in \mathbf{Z}$ and $l \in \mathbf{N}$. Associated Legendre Functions obey orthogonality in l:

$$\int_{-1}^{1} P_k^m P_l^m = \frac{2}{(2l+1)}\frac{(l+|m|)!}{(l-|m|)!}\delta_{kl}. \tag{19}$$

4 When $l = \frac{n}{2}$ but n is odd

If l is a half-integer, that necessarily means all the m's in the ladder are also half-integer. However, when m is a half integer, the transformation $\phi \to \phi + 2\pi$ would lead to the Spherical Harmonic be equal to -1 times itself. To show this, let $m = \frac{o}{2}$, where o is an odd integer.

$$\begin{aligned}
Y_l^m(\theta, \phi+2\pi) &= Ne^{i\frac{o}{2}(\phi+2\pi)}P_l^m(cos\theta) \\
&= Ne^{i\frac{o}{2}2\pi}e^{i\frac{o}{2}\phi}P_l^m(cos\theta) \\
&= Ne^{io\pi}e^{i\frac{o}{2}\phi}P_l^m(cos\theta) \\
&= -Ne^{i\frac{o}{2}\phi}P_l^m(cos\theta) = -Y_l^m(\theta,\phi)
\end{aligned}$$

Under a rotation by 2π, the system must look exactly the same, for they correspond to the same physical state:

$$Y_l^m(\theta, \phi+2\pi) = Y_l^m(\theta,\phi) = -Y_l^m(\theta,\phi)$$

This then implies that Y_l^m is necessarily 0, which is aphysical since the system it describes does not even exist. Therefore, only integer values of m and thus l describe physical systems.

Alternatively, one could have solved the Schrödinger Equation in three dimensions with spherical coordinates and noticed that a half-integer l leads to a divergent solution, but I think this way of thinking is more physics-like.

5 What of the half-integer l's then?

It turns out that half-integer values of l describe an intrinsic property of particles that differs fundamentally from orbital angular momentum. Just like how the Earth is spinning along its axis while orbiting around the Sun, a particle like an electron also has two "types" of angular momentum. Its orbital angular momentum is described by the integer values of l and m while its intrinsic angular momentum is described by the half-integer values — this has now become to be known as the *spin* of a particle.

The algebraic formulation of spin is actually exactly the same as that of orbital angular momentum's, but we can no longer talk in terms of the Spherical Harmonics. All the results regarding operators above will still hold when describing spins in systems. Instead of L, I will use S to distinguish spin. I will adopt the notation $|sm\rangle$ (as Griffiths does) to talk about a particle that is in some half-integer ladder represented by s and in the m-th S_z state.

5.1 spin one-half

When $s = \frac{1}{2}$, m is either $\frac{1}{2}$ or $-\frac{1}{2}$ (by the discussion of ladders above). We can represent the general state with a column two-vector $\begin{pmatrix} a \\ b \end{pmatrix}$ with the basis $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$, where $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ represents the $+\frac{1}{2}$ case $\psi_\uparrow$ ("up") and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ represents the $-\frac{1}{2}$ case $\psi_\downarrow$ ("down").

Then, from previous results about operators, we know that:

$$\begin{aligned} S_z \psi_\uparrow &= \frac{\hbar}{2} \psi_\uparrow \\ S_z \psi_\downarrow &= -\frac{\hbar}{2} \psi_\downarrow \\ S^2 \psi_\uparrow = S^2 \psi_\downarrow &= \frac{3\hbar^2}{4} \psi_\uparrow \end{aligned}$$

Except now the operators are 2 by 2 matrices and it is possible to obtain them.

Let $S_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, then $a = \frac{\hbar}{2}$ and $c = 0$, doing the same on $\psi_\downarrow$ gives that $b = 0$ and $d = -\frac{\hbar}{2}$. Factoring out $\frac{\hbar}{2}$, we get that

$$S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \equiv \frac{\hbar}{2} \sigma_x$$

Additionally, $S_+ \psi_\downarrow = \psi_\uparrow$ and $S_- \psi_\uparrow = \psi_\downarrow$. After some algebra we can arrive at the conclusion that

$$\begin{aligned} S_+ &= \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ S_- &= \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

Note that $S_+ \psi_\uparrow = S_- \psi_\downarrow = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0$, which is consistent with our ladder formulation. From the expressions for S_+ and S_-, it is possible to obtain S_x and S_y via the relation $S_\pm = S_x \pm i S_y$:

$$\begin{aligned} S_x &= \frac{S_+ + S_-}{2} = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ S_y &= \frac{S_+ - S_-}{2i} = \frac{\hbar}{2i} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \frac{\hbar}{2i} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \end{aligned}$$

Many of the building blocks of our physical world have spin-$\frac{1}{2}$ — protons, electrons, neutrons, quarks ... Measuring the S_z of them at any point in time will give *either* $\frac{1}{2}$ or $-\frac{1}{2}$. We can also simply choose to measure S_x or S_y instead, and all we have to do would be just to change the basis in which we express the general state of a spin-$\frac{1}{2}$ system. This is a valid operation because all the spin matrices are hermitian, which implies that their eigenvectors necessarily span the space of column two-vectors. No matter what basis we are in though, all that we can ever measure is going to be the set of values $\left\{-\frac{1}{2}, \frac{1}{2}\right\}$. After all, calling a certain axis to be "the" z-axis was a *choice* and has nothing to do with what is going on in nature. Choice of a coordinate system should not have an impact on the intrinsic observable values.