THEORY of ANGULAR MOMENTUM in QUANTUM MECHANICS
The Algebraic Method
Lecture Notes by W.B. von Schlippe

1 Introductory remarks

In these notes it is assumed that the reader is familiar with the quantum mechanical definition of angular momentum, based on the correspondence with classical mechanics: the angular momentum vector $\mathbf{L}$ is defined by $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, where $\mathbf{r}$ is the position vector and $\mathbf{p}$ is the momentum vector, and then the fundamental commutation relations are imposed on $\mathbf{r}$ and $\mathbf{p}$. As a result $\mathbf{r}$ and $\mathbf{p}$, and hence also $\mathbf{L}$ become operators and the components of $\mathbf{L}$ obey the commutation relations

$$[L_i, L_j] = i\hbar\varepsilon_{ijk}L_k$$

This implies that $\mathbf{L}^2$ and L_3 have common eigenfunctions. One therefore proceeds to solve the simultaneous eigenvalue problems

$$\mathbf{L}^2 Y_{\ell m} = \hbar^2\ell(\ell+1)Y_{\ell m}, \qquad L_3 Y_{\ell m} = \hbar m Y_{\ell m}$$

Representing $\mathbf{p}$, and hence also $\mathbf{L}$, as differential operators, one then solves simultaneously two differential equations. Imposing the requirement that the eigenfunctions be regular one finds the eigenvalues $\ell = 0,\ 1,\ 2,\ \ldots,\ m = -\ell,\ -\ell+1,\ \ldots\ ,\ \ell$. ℓ is the *orbital angular momentum quantum number* and m is the *projection* or *magnetic* quantum number. The eigenfunctions $Y_{\ell m}$ are the *spherical harmonics.*

In this form the theory of angular momentum is appropriate for the description of *orbital* angular momentum, but not for *spin*, which is empirically known to take on half-integer values.

2 The Algabraic Method

In the algebraic method the commutation relations

$$[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k \tag{1}$$

are taken as the definition of angular momentum. Here $J_1 = J_x$, $J_2 = J_y$, and $J_3 = J_z$ are the cartesian components of the angular momentum (vector) operator J, and no representation in terms of differential operators is used. The important consequence of this redefinition of angular momentum is that the resulting spectrum of eigenvalues includes half-odd integer values, which are directly interpreted as eigenvalues of half-odd integer spin.

As in the analytical approach one can directly deduce the following commutation relations from definition (1):

$$[J_i, J^2] = 0 \qquad (i = 1, 2, 3) \tag{2}$$

$$[J_+, J_-] = 2\hbar J_z, \tag{3}$$

$$[J_z, J_\pm] = \pm\hbar J_\pm, \tag{4}$$

and

$$[J_\pm, J^2] = 0 \tag{5}$$

where $J^2 = J_x^2 + J_y^2 + J_z^2$ and $J_\pm = J_x \pm iJ_y$.

Exercise 1: Derive the commutation relations (2)-(5) from the definition (1).

We also note that the components of angular momentum J_x, J_y and J_z are hermitian operators since they are observables. It follows then that J^2 is also hermitian and that J_+ and J_- are related by

$$J_+ = J_-^\dagger \tag{6}$$

Let us denote the eigenstates of J_z and J^2 by $u_{\lambda m}$, i.e.

$$J^2 u_{\lambda m} = \hbar^2 \lambda u_{\lambda m} \tag{7}$$

$$J_z u_{\lambda m} = \hbar m u_{\lambda m} \tag{8}$$

The eigenfunctions are orthogonal and we assume that they are normalised, i.e. that

$$(u_{\lambda m}, u_{\lambda' m'}) = \delta_{\lambda\lambda'}\delta_{mm'}$$

Now consider the state $J_+ u_{\lambda m}$. Using the CR (5) and the eigenvalue equation (7) we get

$$J^2 J_+ u_{\lambda m} = J_+ J^2 u_{\lambda m} = \hbar^2 \lambda J_+ u_{\lambda m} \tag{9}$$

which means that $J_+ u_{\lambda m}$ is an eigenfunction of J^2 with the same eigenvalue as $u_{\lambda m}$ itself. Next, using the CR (4) and the eigenvalue equation (8), we get

$$J_z J_+ u_{\lambda m} = J_+(J_z + \hbar) u_{\lambda m} = \hbar(m+1) J_+ u_{\lambda m}$$

which means that $J_+ u_{\lambda m}$ is an eigenfunction of J_z but with an eigenvalue that is raised by $\hbar$. We shall therefore denote this state (up to a normalization factor) by $u_{\lambda,m+1}$, i.e

$$J_+ u_{\lambda m} = N_+(\lambda, m) u_{\lambda,m+1} \tag{10}$$

where $N_+(\lambda, m)$ is a normalization factor. Similarly we can also deduce that

$$J_- u_{\lambda m} = N_-(\lambda, m) u_{\lambda,m-1} \tag{11}$$

Because of the properties of the operators $J_\pm$ expressed by Eqs. (10) and (11), J_+ is called *raising operator* and J_- *lowering operator*. Collectively the two operators are also called *ladder operators*.

Using the hermiticity property (6) of the ladder operators we can find a simple relationship between the normalization factors N_+ and N_-: taking the scalar product of Eq. (10) with $u_{\lambda,m+1}$ we find

$$N_+(\lambda, m) = (u_{\lambda,m+1}, J_+ u_{\lambda m})$$

and using the definition of the hermitian conjugate we have on the R.H.S.

$$(J_+^\dagger u_{\lambda,m+1}, u_{\lambda m})$$

and hence with Eqs. (6) and (11) we get

$$N_+(\lambda, m) = N_-^*(\lambda, m+1) \tag{12}$$

We can therefore drop the subscript of N and re-write Eqs. (10) and (11) as

$$J_+ u_{\lambda m} = N(\lambda, m) u_{\lambda, m+1} \quad (13)$$
$$J_- u_{\lambda m} = N^*(\lambda, m-1) u_{\lambda, m-1} \quad (14)$$

Next we shall show that m is bounded. Consider the expectation value of the operator $[J_+, J_-]$: taking account of Eq. (3) we get

$$(u_{\lambda m}, [J_+, J_-] u_{\lambda m}) = (u_{\lambda m}, 2\hbar J_z u_{\lambda m})$$

whence, with Eqs. (8), (13) and (14) we have

$$|N(\lambda, m-1)|^2 - |N(\lambda, m)|^2 = 2\hbar^2 m \quad (15)$$

This equation is a *difference* equation. It is called that because the two terms on the L.H.S. are two values of the function N taken at *different* values of the argument m. The difference equation is similar to the more familiar *differential* equation

$$\frac{d}{dm} |N(\lambda, m)|^2 = 2\hbar^2 m$$

which has the obvious solution $|N(\lambda, m)|^2 = \hbar^2 m^2 + \text{const}$. With this in mind it is not surprising that the solution of Eq. (14) has a similar form, namely

$$|N(\lambda, m)|^2 = c - \hbar^2 m(m+1) \quad (16)$$

where c is an arbitrary constant.

Exercise 2: verify that the function (16) is the general solution of Eq. (15)

Thus, observing that obviously $|N|^2 \geq 0$, we find that

$$\hbar^2 m(m+1) \leq c \quad (17)$$

and therefore, for a fixed value of c, m has a greatest value. Let us denote the maximum value of m by j: $\max(m) = j$. Now consider Eq. (13) for $m = j$:

$$J_+ u_{\lambda j} = N(\lambda, j) u_{\lambda, j+1}$$

But for j to be the maximum value of m the R.H.S. must vanish, and this is achieved by demanding that

$$N(\lambda, j) = 0$$

Thus, putting $m = j$ in Eq. (16), we get

$$c = \hbar^2 j(j+1)$$

and substituting this into Eq. (16) for arbitrary values of m we get

$$|N(\lambda, m)|^2 = \hbar^2 [j(j+1) - m(m+1)].$$

Thus $N(\lambda, m)$ is determined up to a phase factor which we choose following the universally adopted convention of Condon and Shortley, [1] i.e. we write

$$N(\lambda, m) = \hbar\sqrt{j(j+1) - m(m+1)}. \tag{18}$$

Now Eq. (17) implies not only that there is a *maximum* value of m but also that m has a *minimum*. Denote the minimum of m by j'. Then, considering Eq. (14) together with Eq. (18) and putting $m = j'$ we get

$$J_- u_{\lambda j'} = \hbar\sqrt{j(j+1) - j'(j'-1)}\, u_{\lambda, j'-1}$$

which is consistent with the requirement that j' be the least value of m only if

$$j(j+1) - j'(j'-1) = 0.$$

This quadratic equation in j' has the two roots

$$j' = -j, \qquad \text{and} \quad j' = j+1$$

The second root must be rejected since by definition $j' \leq j$, leaving

$$j' = \min(\mathrm{m}) = -j. \tag{19}$$

We are now in a position to construct all eigenstates of J_z by successive application of J_- to the states $u_{\lambda m}$, starting from $u_{\lambda j}$:

$$J_- u_{\lambda j} = \hbar\sqrt{j(j+1) - j(j-1)}\, u_{\lambda, j-1}$$

$$J_- u_{\lambda, j-1} = \hbar\sqrt{j(j+1) - (j-1)(j-2)}\, u_{\lambda, j-2}$$

etc. until

$$J_- u_{\lambda, -j+1} = \hbar\sqrt{j(j+1) - j(j-1)}\, u_{\lambda, -j}$$

Thus the possible values of m are $m = j, j-1, j-2, \ldots, -j$, and this sequence shows that

$$\max(m) - \min(m) = 2j = \text{integer}$$

and since by definition $j \geq 0$ we get

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots \tag{20}$$

There remains the task of relating the eigenvalue λ to j. Considering the physical significance of $\hbar j$ as the maximum value of the projection of the angular momentum vector onto the z axis one can expect a simple relationship between these quantities (classically one has obviously $j = \sqrt{\lambda}$).

The required result is established immediately by noting the identity

$$J^2 = J_- J_+ + J_z^2 + \hbar J_z \tag{21}$$

[1] E.U. Condon and G.H. Shortley, *The Theory of Atomic Spectra*, Cambridge, 1935

Operating on the state $u_{\lambda j}$ we get

$$\lambda = j(j+1) \tag{22}$$

Exercise 3: Prove the identity (21) and hence deduce Eq. (22)

We note that the quantum number λ has played a purely auxiliary role. Equation (22) allows us to discard it. We shall therefore from now on label the eigenstates of the angular momentum operators by j and m instead of λ and m. The eigenvalue equations shall be written in the form

$$J^2 u_{jm} = \hbar^2 j(j+1) u_{jm} \tag{23}$$
$$J_z u_{jm} = \hbar m u_{jm} \tag{24}$$

3 Matrix Representation of the Angular Momentum Operators; Pauli Matrices

If we take the expectation values of the angular momentum operators in eigenstates of J^2 and J_z at a fixed quantum number j we get

$$(u_{jm}, J^2 u_{jm'}) = \hbar^2 j(j+1)\delta_{mm'} \tag{25}$$
$$(u_{jm}, J_z u_{jm'}) = \hbar\delta_{mm'} \tag{26}$$
$$(u_{jm}, J_\pm u_{jm'}) = \hbar\sqrt{j(j+1) - m(m\pm 1)}\delta_{m,m'\pm 1} \tag{27}$$

and hence with $J_\pm = J_x \pm iJ_y$

$$\begin{aligned}
(u_{jm}, J_x u_{jm'}) &= \frac{\hbar}{2}[\sqrt{j(j+1) - m(m-1)}\delta_{m,m'+1} + \sqrt{j(j+1) - m(m+1)}\delta_{m,m'-1}] \\
(u_{jm}, J_y u_{jm'}) &= \frac{\hbar}{2i}[\sqrt{j(j+1) - m(m-1)}\delta_{m,m'+1} - \sqrt{j(j+1) - m(m+1)}\delta_{m,m'-1}]
\end{aligned} \tag{28}$$

Let us consider the particular case of $j = 1/2$. This is an important case because of the importance of fermions, i.e. electrons and protons etc., in physical systems of interest. Then m takes on the values $-1/2$ and $1/2$. There will be four expectation values for each angular momentum operator which we can usefully arrange in the form of a matrix, e.g.

$$\begin{pmatrix} (u_{1/2\,1/2}, J_z u_{1/2\,1/2}) & (u_{1/2\,1/2}, J_z u_{1/2\,-1/2}) \\ (u_{1/2\,-1/2}, J_z u_{1/2\,1/2}) & (u_{1/2\,-1/2}, J_z u_{1/2\,-1/2}) \end{pmatrix} \tag{29}$$

With the explicit values from Eqs. (26) and (27) we get the following matrix representations of the spin-1/2 operators:

$$J_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad J_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad J_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{30}$$

and

$$J_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad J_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad J^2 = \hbar^2 \frac{3}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{31}$$

Because of the importance of the spin-1/2 matrices one also defines a set of dimensionless matrices omitting the factors of $\hbar/2$ from the matrices J_x, J_y and J_z. These matrices are called Pauli matrices; they are usually denoted by σ_x, σ_y and σ_z. Thus

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{32}$$

If we also denote the 2×2 unit matrix by σ_0, then we can easily find the following properties of the Pauli matrices:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \sigma_0 \tag{33}$$

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k \tag{34}$$

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}\sigma_0 \tag{35}$$

where we have introduced the anticommutator $\{a, b\} = ab + ba$. Thus, by adding equations (34) and (35), we get

$$\sigma_i\sigma_j = \delta_{ij}\sigma_0 + i\varepsilon_{ijk}\sigma_k \tag{36}$$

Another important property of the Pauli matrices is that, together with the unit matrix, they form a complete set of linearly independent 2×2 matrices. The linear independence is proved by showing that the equation

$$a_0\sigma_0 + a_1\sigma_x + a_2\sigma_y + a_3\sigma_z = 0 \tag{37}$$

holds iff $a_0 = a_1 = a_2 = a_3 = 0$. For the proof we note that the three Pauli matrices are traceless, i.e. $\mathrm{Tr}\,\sigma_x = \mathrm{Tr}\,\sigma_y = \mathrm{Tr}\,\sigma_z = 0$, and that the trace of the unit matrix σ_0 is 2. Thus taking the trace of Eq. (37) we get $2a_0 = 0$ and hence $a_0 = 0$. If we now multiply Eq. (37) by σ_x, use Eq. (33) and again take the trace we find $a_1 = 0$, and similarly also $a_2 = 0$ and $a_3 = 0$.

The completeness of the Pauli matrices means that an arbitrary 2×2 matrix A can be represented by a linear superposition of the Pauli matrices together with the unit matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_0\sigma_0 + a_1\sigma_x + a_2\sigma_y + a_3\sigma_z \tag{38}$$

The expansion coefficients a_0 etc. are found by a procedure similar to the one used above. Taking the trace of Eq. (38) yields $a_0 = \frac{1}{2}(a_{11} + a_{22})$, and we get similar equations for the other expansion coefficients by multiplying Eq. (38) in turn by σ_x, σ_y and σ_z and taking the traces. Thus $a_1 = \frac{1}{2}\,\mathrm{Tr}\,\sigma_x A$, $a_2 = \frac{1}{2}\,\mathrm{Tr}\,\sigma_y A$ and $a_3 = \frac{1}{2}\,\mathrm{Tr}\,\sigma_z A$.

From the commutation relations of the Pauli matrices one can also derive the following identity, which is frequently useful:

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = \vec{a} \cdot \vec{b} + i\vec{\sigma} \cdot \left(\vec{a} \times \vec{b}\right)$$

where $\vec{a}$ and $\vec{b}$ are arbitrary vectors. In particular, if we put $\vec{b} = \vec{a}$, then we get

$$(\vec{a} \cdot \vec{\sigma})^2 = a^2$$

Less frequently than the spin-1/2 matrices one uses explicit matrix representations of operators for higher spins. So it is more for curiosity than for practical usage that we can write down the matrix representation of the spin-1 operators which are 3×3 matrices since m can take on three values, -1, 0 and +1. Generalising the rule explained by Eq. (29) we find the following expressions for the spin-1 matrices:

$$J_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_y = \frac{i\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$