

## Lecture 5: The Hydrogen Atom (continued).

In the previous lecture we have solved the eigenvalue problem for the hydrogen atom,

$$\hat{H}\psi(r, \theta, \varphi) = E\psi(r, \theta, \varphi) \quad (1)$$

where the Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 + V(r)$$

with the reduced mass of the atom

$$\mu = m_e m_p / (m_e + m_p)$$

and the Coulomb potential  $V(r) = -e^2/r$ .

We have also seen that the Hamiltonian with a spherically symmetric potential energy function commutes with the angular momentum operators  $\hat{L}_z$  and  $\hat{L}^2$ :

$$[\hat{L}_z, \hat{H}] = 0, \quad [\hat{L}^2, \hat{H}] = 0$$

and of course also  $[\hat{L}_z, \hat{L}^2] = 0$ , and this implies that  $\hat{H}$ ,  $\hat{L}_z$  and  $\hat{L}^2$  have common eigenfunctions, i.e. that we have also

$$\hat{L}^2\psi(r, \theta, \varphi) = \hbar^2\ell(\ell+1)\psi(r, \theta, \varphi) \quad (2)$$

$$\hat{L}_z\psi(r, \theta, \varphi) = \hbar m\psi(r, \theta, \varphi) \quad (3)$$

Then, since the Laplacian operator  $\nabla^2$  in spherical coordinates is of the form of:

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \left[ \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left( \sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right]$$

and the expression in square brackets is, up to a factor of  $-\hbar^2$ , the angular momentum operator  $\hat{L}^2$ , we could rewrite the Hamiltonian in the form of:

$$\hat{H} = -\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{\hat{L}^2}{2\mu r^2} + V(r)$$

and we could factorize the wave function into the radial wave function  $R(r)$  and the spherical harmonic  $Y(\theta, \varphi)$ :

$$\psi(r, \theta, \varphi) = R(r) Y(\theta, \varphi)$$

Then we found that the radial wave function satisfies the radial wave equation:

$$\frac{1}{\rho^2} \frac{d}{d\rho} \left( \rho^2 \frac{dR}{d\rho} \right) + \left[ \frac{\lambda}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - \frac{1}{4} \right] R = 0 \quad (4)$$

where  $\rho = \varepsilon r$  with  $\varepsilon^2 = 8\mu|E|/\hbar^2$ , and asymptotically  $R(\rho)|_{\rho \rightarrow \infty} \sim \exp(-\rho/2)$ .

The exact solution had to be of the form of

$$R(\rho) = \rho^\ell e^{-\rho/2} L(\rho)$$

where  $L(\rho)$  satisfies the differential equation

$$\rho L'' + [2(\ell + 1) - \rho] L' + (n - \ell - 1)L = 0 \quad (5)$$

The requirement of square integrability of the wave function selected solutions with eigenvalues

$$E_n = -\frac{1}{2} \alpha^2 \mu c^2 \frac{1}{n^2} \quad (6)$$

where  $\alpha$  is the fine structure constant whose numerical value of close to  $1/137$ , and  $n = 1, 2, 3, \dots$  is the principle quantum number.  $E_1$  is the ground state energy,  $E_2$  is the energy of the first excited state, *etc.*

The excitation energy can be released in transitions to lower lying levels by emission of photons of energy

$$E_{nm} = \frac{1}{2} \alpha^2 \mu c^2 \left( \frac{1}{n^2} - \frac{1}{m^2} \right)$$

and corresponding frequency  $\nu_{nm} = E_{nm}/2\pi\hbar$ , *i.e.*

$$\nu_{nm} = \frac{\alpha^2 \mu c^2}{4\pi\hbar} \left( \frac{1}{n^2} - \frac{1}{m^2} \right)$$

(Balmer formula).

### The hydrogen wave function.

The differential equation (5) is the **Laguerre equation**. Its polynomial solutions are the **associated Laguerre polynomials**  $L_{n+\ell}^{2\ell+1}(\rho)$ . Explicit expressions can be derived using the definition of  $L_n^m(\rho)$  in terms of the Laguerre polynomials  $L_n(\rho)$ <sup>1</sup>:

$$L_n^m(\rho) = \frac{d^m}{d\rho^m} L_n(\rho), \quad L_n(\rho) = e^\rho \frac{d^n}{d\rho^n} (e^{-\rho} \rho^n)$$

Tracing back the relation of the radial wave function  $R(r)$  and the associated Laguerre polynomials we find

$$R_{n\ell}(\rho) = -a^{-3/2} N_{n\ell} e^{-\rho/2} \rho^\ell L_{n+\ell}^{2\ell+1}(\rho)$$

where  $N_{n\ell}$  is a normalization factor which can be found by demanding  $R_{n\ell}(r)$  to be normalized:

$$\int_0^\infty [R_{n\ell}(r)]^2 r^2 dr = 1$$

The result of this calculation is

$$N_{n\ell} = \frac{2}{n^2} \sqrt{\frac{(n - \ell - 1)!}{[(n + \ell)!]^3}}$$

---

<sup>1</sup>Unfortunately there are different conventions used in the literature; the convention used here is that of such widely read textbooks as Blokhintsev and Schiff; the no less well known textbooks by Messiah and by Smirnov, also the mathematical reference works by Abramowicz and Stegun and by Ryzhik and Gradshtein and the algebraic computer package Mathematica define  $L_n^m(x) = (-1)^m d^m L_{n+m}(x)/dx^m$ .

The radial wave functions are orthogonal:

$$\int_0^{\infty} R_{n\ell}(r)R_{n'\ell'}r^2dr = \delta_{nn'}\delta_{\ell\ell'}$$

The constant  $a$  in the expression of  $R_{n\ell}$  is the **Bohr radius**:  $a = \hbar^2/\mu e^2 = 0.529 \times 10^{-10}$  m. The sign of  $R_{n\ell}$  is chosen such that the wave function is positive near the origin.

The hydrogen wave functions can now be written in the form of

$$\psi_{n\ell m}(r, \theta, \phi) = R_{n\ell}(r)Y_{\ell m}(\theta, \phi)$$

where

$$n = 1, 2, 3, \dots; \quad \ell = 0, 1, 2, \dots, n - 1$$

and

$$m = -\ell, -\ell + 1, -\ell + 2, \dots, \ell.$$

Thus we have  $n$  values of  $\ell$  for a fixed value  $n$  of the principal q.n., and for every value of the orbital angular momentum q.n. we have  $2\ell + 1$  values of  $m$ . Therefore the number of different states with the same  $n$  is  $n^2$ . When there is more than one wave function at a given energy eigenvalue, then that level is said to be **degenerate**. In the case of the hydrogen atom the  $n$ -th energy level is  $n^2$ -fold degenerate. Only the ground state,  $n = 1$ , is nondegenerate.

The degeneracy of the stationary states of hydrogen is related to the spherical symmetry of the potential energy. The spherical symmetry can be broken, for instance by placing the atoms in an electric or magnetic field. Then the Hamiltonian does not commute with the angular momentum operators. As a result the degeneracy is removed and the spectral lines split up into several lines. The degeneracy is also removed when relativistic effects are taken into account. This corresponds to the observed fine structure of the hydrogen spectral lines. We will discuss this quantitatively in another lecture.

### Discussion of the Hydrogen Wave Functions.

The probability that the electron lie in the volume element  $dV$  at  $(r, \theta, \phi)$  is, by general rules of quantum mechanics, equal to  $|\psi_{n\ell m}(r, \theta, \phi)|^2 dV$ . Integrating over  $\theta$  from 0 to  $\pi$  and over  $\varphi$  from 0 to  $2\pi$ , we get

$$w_{n\ell}(r) = |R_{n\ell}(r)|^2 r^2 dr$$

which is the probability of the electron to lie between  $r$  and  $r + dr$ . The function

$$D_{n\ell}(r) = r^2 R_{n\ell}^2(r) \tag{7}$$

is the **radial distribution function**. For the ground state and the first and second excited states with  $\ell = 0$  the radial distribution functions are shown in Fig. 1. The abscissa was chosen to be the radial distance  $r$  in units of the Bohr radius  $a$ . In Bohr's theory of the hydrogen atom,  $a$  is the radius of the circular electron orbit in the ground state of the hydrogen atom. This is also where the ground state distribution function  $D_{10}$  is seen to peak. We therefore conclude that the most probable distance of the electron from the nucleus is given by the Bohr radius  $a$ . For distances greater than  $a$  the distribution function drops sharply. Thus although the surface of the atom is fuzzy, but its radius is roughly equal to  $a$ .

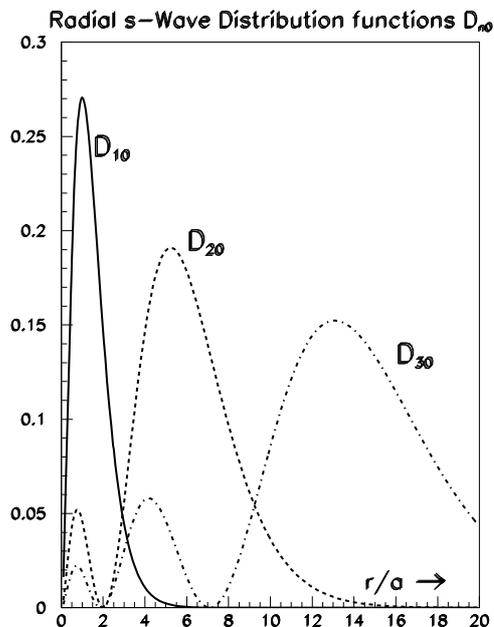


Figure 1: s-wave radial distribution functions  $D_{n0}$ ,  $n = 1, 2, 3$ .

In the first excited state,  $n = 2$ , the distribution function has a small peak near  $a$  and a large peak near  $6a$ . Thus the most likely distance of the electron from the nucleus is in this case  $r \approx 6a$ . The distribution function of the second excited state has three peaks. The most pronounced peak lies close to  $14a$ , corresponding to the most probable distance of the electron from the nucleus.

We can make the above observations completely quantitative by calculating the expectation value of  $r$  for arbitrary states  $\psi_{n\ell m}(r, \theta, \phi)$ . By the general definition of expectation values of observables, this is given by

$$\langle r \rangle_{n\ell m} = \int \psi_{n\ell m}^*(r, \theta, \phi) r \psi_{n\ell m}(r, \theta, \phi) dV$$

and substituting  $\psi_{n\ell m}(r, \theta, \phi) = R_{n\ell}(r)Y_{\ell m}(\theta, \phi)$  we see that the volume integral is the product of an integral of  $|Y_{\ell m}(\theta, \phi)|^2$  over the polar angles and an integral of  $R_{n\ell}^2(r)r^3$  over  $r$  from 0 to  $\infty$ . The former integral is unity by normalization of the spherical harmonics, and the latter integral is evaluated to give

$$\langle r \rangle_{n\ell} = n^2 a \left\{ 1 + \frac{1}{2} \left[ 1 - \frac{\ell(\ell+1)}{n^2} \right] \right\}$$

In spectroscopic terminology the ground state distribution is called **K shell**, the distributions in the 1st and 2nd excited states are called **L shell** and **M shell**, respectively. States with  $\ell = 0, 1, 2, 3$  are called *s*, *p*, *d* and *f* states, respectively, and one denotes the states by the principal quantum number  $n$  and the spectroscopic symbol. Thus the ground state is denoted 1*s* state, the two L shell states are 2*s* and 2*p*, *etc.*

### Hydrogen-like systems.

There are many systems with hydrogen-like structure. The most obvious ones are alkali atoms. They consist of a nucleus surrounded by electrons forming closed shells and a single valence

electron in a higher shell. In first approximation the nucleus together with the electrons of the closed shells are a point-like system of charge  $+e$ , and the outer electron moves in the Coulomb field of this charge.

Another example is the singly-ionized helium atom. Here the electron is moving in the field of the nucleus of charge  $2e$ . More generally we can consider ions of atomic number  $Z$  with a single electron bound to the nucleus. Then the electron is moving in the field of a charge  $Ze$ . For such ions we can repeat the calculation we have done for the hydrogen atom replacing everywhere  $e^2$  by  $Ze^2$ . The energy levels become

$$E_n = -\frac{1}{2} Z^2 \alpha^2 \mu c^2 \frac{1}{n^2} \quad (8)$$

and the expectation values of  $r$  are

$$\langle r \rangle_{n\ell} = \frac{n^2 a}{Z} \left\{ 1 + \frac{1}{2} \left[ 1 - \frac{\ell(\ell+1)}{n^2} \right] \right\}$$

The latter result shows that for sufficiently large values of  $Z$  the electron can be inside the nucleus with a non-negligible probability. This means that the assumption that the electron moves in the field of a point-like charge is not correct. The spectra of such ions are slightly different from those given by Eq. (8).

An interesting system is **positronium**, the bound state of an electron and a positron. All results found for the hydrogen atom remain valid if we take the correct reduced mass of the system, which is  $\mu = m_e/2$ . Similar to positronium is **protonium**, *i.e.* the bound state of a proton with an antiproton. On the level of nonrelativistic quantum mechanics we can again take over all results found for hydrogen, putting the reduced mass equal to a half of the proton mass. The Bohr radius of protonium is  $a_p = 2\hbar^2/m_p e^2 = 5.76 \times 10^{-14}$  m. This is still almost two orders of magnitude greater than the proton radius  $r_p \approx 0.7 \times 10^{-15}$  m. Therefore protonium can remain in this state for a fairly long time before annihilating.

Other systems studied both theoretically and experimentally are **mesic atoms**, that is atoms in which a negatively charged muon or pion is bound to the nucleus.

Consider a  $\mu$ -mesic atom. The Bohr radius of this system is  $a = \hbar^2/\mu Z e^2$  where

$$\mu = m_\mu A / (m_\mu + A)$$

$A$  is the atomic mass and  $m_\mu$  is the mass of the muon:  $m_\mu \approx 210m_e$ . The radius of the K shell of the  $\mu$ -mesic atom is about 200 times smaller than the radius of the electronic K shell. Therefore the muon lies inside the electron cloud and is therefore almost entirely under the effect of the nuclear Coulomb potential. There are small but interesting deviations between the observed spectra and the spectra calculated within the framework of the nonrelativistic theory discussed so far.

Pions are even heavier than muons: their mass is about 280 electron masses. Therefore the K shell of a pionic atom is even smaller than that of a  $\mu$ -mesic atom. A negatively charged pion, captured by an atom, has therefore an even higher probability of being inside the nucleus. Then, since the pion is a hadron, *i.e.* a particle that interacts not only electromagnetically

but also by the strong nuclear force, the spectra of pionic atoms are affected by the strong interaction.

In spite of this additional difficulty, the measurement of spectra of pionic atoms yields the most accurate determination of the pion mass. Measured are the X-rays emitted by the atom when the pion makes transitions from higher to lower energy levels. Ignoring for the moment any effect other than the above nonrelativistic calculation, we recall the formula for the spectral frequencies:

$$\nu_{nm} = \frac{\alpha^2 \mu c^2}{4\pi\hbar} \left( \frac{1}{n^2} - \frac{1}{m^2} \right)$$

where now the reduced mass is

$$\mu = m_\pi / (1 + m_\pi/m_p) \approx 122 \text{ MeV}$$

The frequency of radiation emitted in transitions between low lying energy levels lies in the X-ray range. For instance, for the transition from  $m = 2$  to  $n = 1$  we get a frequency of about  $\nu_{12} \approx 6 \times 10^{17} \text{ s}^{-1}$ , and the corresponding wavelength is  $\lambda_{12} \approx 0.5 \text{ nm}$ . Compare this with the range of wavelengths of X-rays: 1 to 0.01 nm.

The next heavier meson is the  $K$  meson with a mass of 494 MeV. The measurement of X-ray spectra of  $K$  mesonic atoms gives the most accurate determination of the kaon mass.

### The Bohr magneton.

We can now calculate the current flowing in the hydrogen atom. To do this we substitute the hydrogen wave function into the expression for the probability current density multiplied by the charge of the electron,  $-e$ :

$$\mathbf{j} = -e \frac{\hbar}{\mu} \text{Im} (\psi \nabla \psi^*)$$

where as before the reduced mass is denoted  $\mu$ .

Since we have the wave function in polar coordinates, we shall also write the  $\nabla$  operator in polar coordinates:

$$\nabla_r = \frac{\partial}{\partial r}, \quad \nabla_\theta = \frac{1}{r} \frac{\partial}{\partial \theta}, \quad \nabla_\varphi = \frac{1}{r \sin \theta} \frac{\partial}{\partial \varphi}$$

hence

$$j_r = -e \frac{\hbar}{\mu} \text{Im} \psi \frac{\partial \psi^*}{\partial r}, \quad j_\theta = -e \frac{\hbar}{\mu r} \text{Im} \psi \frac{\partial \psi^*}{\partial \theta}, \quad j_\varphi = -e \frac{\hbar}{\mu r \sin \theta} \text{Im} \psi \frac{\partial \psi^*}{\partial \varphi}$$

where

$$\psi \equiv \psi_{nlm}(r, \theta, \varphi) = N_{nl} R_{nl}(r) P_n^\ell(\cos \theta) e^{im\varphi}$$

hence  $j_r = j_\theta = 0$  and

$$j_\varphi = -e \frac{\hbar m}{\mu r \sin \theta} |\psi_{nlm}(r, \theta, \varphi)|^2$$

It is not surprising that there are no radial and no meridial currents:  
 A nonzero radial current is incompatible with the stability of the atom;  
 A nonzero meridial current would imply an accumulation of charge on one of the poles.

We put a surface element  $d\sigma$  perpendicular to the current density. The current  $dI$  through  $d\sigma$  is

$$dI = j_\varphi d\sigma$$

This current makes a closed loop which gives rise to a magnetic moment  $d\vec{M}$ . By the Biot-Savart law we have

$$d\vec{M} = \frac{\hat{n}}{c} A dI = \frac{\hat{n}}{c} A j_\varphi d\sigma$$

where  $A$  is the surface area of the circle enclosed by the current, *i.e.*  $A = \pi r^2 \sin^2 \theta$ ,  $\hat{n}$  is the unit vector perpendicular to  $A$  and  $d\sigma = r d\theta dr$ . Since  $\hat{n}$  points in the  $z$  direction, we get

$$dM_z = \frac{\pi r^2 \sin^2 \theta}{c} j_\varphi d\sigma$$

The magnetic moment of the atom is found by integrating over the entire meridial plane:

$$M_z = \frac{\pi}{c} \int r^2 \sin^2 \theta j_\varphi d\sigma$$

After substitution of  $j_\varphi$  we get

$$M_z = -\frac{e\hbar m}{2\mu c} \int |\psi|^2 2\pi r \sin \theta d\sigma$$

and we note that  $dV = 2\pi r \sin \theta d\sigma$  is the volume of an annular tube. Therefore the integral is just the integral of  $|\psi|^2$  over the whole volume, which is unity by the normalization of the wave function. Our final result is therefore

$$M_z = -\frac{e\hbar m}{2\mu c} = -M_B m$$

The constant  $M_B = e\hbar/2\mu c$  is the **Bohr magneton**. Since the  $z$  component of angular momentum (strictly: the eigenvalue of the  $z$  component of the angular momentum operator) is  $L_z = \hbar m$ , we have also

$$M_z = -\frac{e}{2\mu c} L_z$$

*i.e.* the magnetic moment is proportional to the  $z$  component of angular momentum.

### The Spin of the Electron.

#### **Introduction.**

With high resolution spectrometers one finds many spectral lines split into two. This is called **fine structure**. The first attempt at an explanation of the fine structure was made by Sommerfeld in 1916, even before the birth of quantum mechanics. Developing the Bohr theory of the hydrogen atom to a relativistic theory, Sommerfeld could derive an expression of the fine

structure that was quantitatively correct for all observed spectra. Only years later discrepancies were found which called for another explanation.

Goudsmith and Uhlenbeck proposed the hypothesis that the electron has an intrinsic angular momentum, the spin, and associated with it a magnetic moment. The spin had to be  $\hbar/2$  so that its projection on an arbitrary axis could take on two values:  $+\hbar/2$  (“spin-up”) and  $-\hbar/2$  (“spin-down”). The interaction of the electron’s magnetic moment with the magnetic field of its orbital motion gives rise to an additional potential energy which takes on two values corresponding to the two orientations of spin.

### Revisiting the theory of angular momentum.

In Lecture 4 we have based the definition of angular momentum on the classical definition  $\vec{L} = \vec{r} \times \vec{p}$ . Replacing the vectors  $\vec{r}$  and  $\vec{p}$  by the operators  $\hat{r}$  and  $\hat{p}$  we derived commutation relations of the components of angular momentum. Working in polar coordinates we were led to a differential equation whose square integrable solutions were the spherical harmonics  $Y_{\ell m}(\theta, \varphi)$ , and we found the eigenvalues of  $\hat{L}_z$  to be integer multiples of  $\hbar$  and the eigenvalues of  $\hat{L}^2$  to be  $\hbar^2 \ell(\ell + 1)$ .

But the eigenvalues of the angular momentum operators can be derived without solving a differential equation. If we base the derivation on the commutation relations, then we find that the eigenvalues of the  $z$  component of angular momentum can be  $\pm\hbar/2$ . This is just what we need to describe the spin of the electron.

### The Algebraic Method

Thus we take as the definition of angular momentum the following commutation relations:

$$[\hat{J}_i, \hat{J}_j] = i\hbar \varepsilon_{ijk} \hat{J}_k \quad (9)$$

Here  $\hat{J}_1 = \hat{J}_x$ ,  $\hat{J}_2 = \hat{J}_y$ , and  $\hat{J}_3 = \hat{J}_z$  are the cartesian components of the angular momentum (vector) operator  $\hat{J}$ , 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 (9):

$$[\hat{J}_i, \hat{J}^2] = 0 \quad (i = 1, 2, 3) \quad (10)$$

$$[\hat{J}_+, \hat{J}_-] = 2\hbar \hat{J}_z, \quad (11)$$

$$[\hat{J}_z, \hat{J}_{\pm}] = \pm\hbar \hat{J}_{\pm}, \quad (12)$$

and

$$[\hat{J}_{\pm}, \hat{J}^2] = 0 \quad (13)$$

where  $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$  and  $\hat{J}_{\pm} = \hat{J}_x \pm i\hat{J}_y$ .

**Exercise 1:** Derive the commutation relations (10)-(13) from the definition (9).

We also note that the components of angular momentum  $\hat{J}_x$ ,  $\hat{J}_y$  and  $\hat{J}_z$  are hermitian operators since they are observables. It follows then that  $\hat{J}^2$  is also hermitian and that  $\hat{J}_+$  and  $\hat{J}_-$  are related by

$$\hat{J}_+ = \hat{J}_-^\dagger \quad (14)$$

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

$$\hat{J}^2 u_{\lambda m} = \hbar^2 \lambda u_{\lambda m} \quad (15)$$

$$\hat{J}_z u_{\lambda m} = \hbar m u_{\lambda m} \quad (16)$$

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

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

Now consider the state  $\hat{J}_+ u_{\lambda m}$ . Using the CR (13) and the eigenvalue equation (15) we get

$$\begin{aligned} \hat{J}^2 \hat{J}_+ u_{\lambda m} &= \hat{J}_+ \hat{J}^2 u_{\lambda m} \\ &= \hbar^2 \lambda \hat{J}_+ u_{\lambda m} \end{aligned} \quad (17)$$

which means that  $\hat{J}_+ u_{\lambda m}$  is an eigenfunction of  $\hat{J}^2$  with the same eigenvalue as  $u_{\lambda m}$  itself. Next, using the CR (12) and the eigenvalue equation (16), we get

$$\hat{J}_z \hat{J}_+ u_{\lambda m} = \hat{J}_+ (\hat{J}_z + \hbar) u_{\lambda m} = \hbar(m+1) \hat{J}_+ u_{\lambda m}$$

which means that  $\hat{J}_+ u_{\lambda m}$  is an eigenfunction of  $\hat{J}_z$  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.*

$$\hat{J}_+ u_{\lambda m} = N_+(\lambda, m) u_{\lambda, m+1} \quad (18)$$

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

$$\hat{J}_- u_{\lambda m} = N_-(\lambda, m) u_{\lambda, m-1} \quad (19)$$

Because of the properties of the operators  $\hat{J}_\pm$  expressed by Eqs. (18) and (19),  $\hat{J}_+$  is called *raising operator* and  $\hat{J}_-$  *lowering operator*. Collectively the two operators are also called *ladder operators*.

Using the hermiticity property (14) of the ladder operators we can find a simple relationship between the normalization factors  $N_+$  and  $N_-$ : taking the scalar product of Eq. (18) with  $u_{\lambda, m+1}$  we find

$$N_+(\lambda, m) = (u_{\lambda, m+1}, \hat{J}_+ u_{\lambda m})$$

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

$$(\hat{J}_+^\dagger u_{\lambda, m+1}, u_{\lambda m})$$

and hence with Eqs. (14) and (19) we get

$$N_+(\lambda, m) = N_-^*(\lambda, m+1) \quad (20)$$

We can therefore drop the subscript of  $N$  and re-write Eqs. (18) and (19) as

$$\hat{J}_+ u_{\lambda m} = N(\lambda, m) u_{\lambda, m+1} \quad (21)$$

$$\hat{J}_- u_{\lambda m} = N^*(\lambda, m-1) u_{\lambda, m-1} \quad (22)$$

Next we shall show that  $m$  is bounded. Consider the expectation value of the operator  $[\hat{J}_+, \hat{J}_-]$ : taking account of Eq. (11) we get

$$(u_{\lambda m}, [\hat{J}_+, \hat{J}_-]u_{\lambda m}) = (u_{\lambda m}, 2\hbar\hat{J}_z u_{\lambda m})$$

whence, with Eqs. (16), (21) and (22) we have

$$|N(\lambda, m-1)|^2 - |N(\lambda, m)|^2 = 2\hbar^2 m \quad (23)$$

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. (23) has a similar form, namely

$$|N(\lambda, m)|^2 = c - \hbar^2 m(m+1) \quad (24)$$

where  $c$  is an arbitrary constant.

**Exercise 2:** verify that the function (24) is the general solution of Eq. (23)

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

$$\hbar^2 m(m+1) \leq c \quad (25)$$

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. (21) for  $m = j$ :

$$\hat{J}_+ u_{\lambda j} = N(\lambda, j) u_{\lambda, j+1}$$

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. (24), we get

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

and substituting this into Eq. (24) 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,<sup>2</sup> *i.e.* we write

$$N(\lambda, m) = \hbar \sqrt{j(j+1) - m(m+1)}. \quad (26)$$

<sup>2</sup>E.U. Condon and G.H. Shortley, *The Theory of Atomic Spectra*, Cambridge, 1935

Now Eq. (25) 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. (22) together with Eq. (26) and putting  $m = j'$  we get

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

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

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

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

$$j' = \min(m) = -j. \tag{27}$$

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

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

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

*etc.* until

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

Thus the possible values of  $m$  are

$$m = j, j-1, j-2, \dots, -j,$$

and this sequence shows that

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

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

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots \tag{28}$$

There remains the task of relating the eigenvalue  $\lambda$  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

$$\hat{J}^2 = \hat{J}_- \hat{J}_+ + \hat{J}_z^2 + \hbar \hat{J}_z \tag{29}$$

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

$$\lambda = j(j+1) \tag{30}$$

**Exercise 3:** Prove the identity (29) and hence deduce Eq. (30)

We note that the quantum number  $\lambda$  has played a purely auxiliary role. Equation (30) 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  $\lambda$  and  $m$ . The eigenvalue equations shall be written in the form

$$\hat{J}^2 u_{jm} = \hbar^2 j(j+1) u_{jm} \quad (31)$$

$$\hat{J}_z u_{jm} = \hbar m u_{jm} \quad (32)$$

### Matrix Representation of the Angular Momentum Operators; Pauli Matrices

If we take the expectation values of the angular momentum operators in eigenstates of  $\hat{J}^2$  and  $\hat{J}_z$  at a fixed quantum number  $j$  we get

$$(u_{jm}, \hat{J}^2 u_{jm'}) = \hbar^2 j(j+1) \delta_{mm'} \quad (33)$$

$$(u_{jm}, \hat{J}_z u_{jm'}) = \hbar m \delta_{mm'} \quad (34)$$

$$(u_{jm}, \hat{J}_\pm u_{jm'}) = \hbar \sqrt{j(j+1) - m(m \pm 1)} \delta_{m, m' \pm 1} \quad (35)$$

and hence with  $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$

$$\begin{aligned} (u_{jm}, \hat{J}_x u_{jm'}) &= \frac{\hbar}{2} [\sqrt{j(j+1) - m(m-1)} \delta_{m, m'+1} \\ &\quad + \sqrt{j(j+1) - m(m+1)} \delta_{m, m'-1}] \\ (u_{jm}, \hat{J}_y u_{jm'}) &= \frac{\hbar}{2i} [\sqrt{j(j+1) - m(m-1)} \delta_{m, m'+1} \\ &\quad - \sqrt{j(j+1) - m(m+1)} \delta_{m, m'-1}] \end{aligned} \quad (36)$$

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.

We shall use a special notation for the spin-1/2 operators and eigenvalues: instead of  $\hat{J}$  we write  $\hat{S}$ , the spin quantum number will be  $s$  and the projection quantum number  $m_s$ .

Thus  $m_s$  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.*

$$\begin{pmatrix} (u_{1/2 1/2}, \hat{S}_z u_{1/2 1/2}) & (u_{1/2 1/2}, \hat{S}_z u_{1/2 -1/2}) \\ (u_{1/2 -1/2}, \hat{S}_z u_{1/2 1/2}) & (u_{1/2 -1/2}, \hat{S}_z u_{1/2 -1/2}) \end{pmatrix} \quad (37)$$

With the explicit values from Eqs. (34)-(36) we get the following matrix representations of the spin-1/2 operators:

$$\hat{S}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (38)$$

and

$$\hat{S}_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \hat{S}_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}^2 = \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (39)$$

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  $\hat{S}_x$ ,  $\hat{S}_y$  and  $\hat{S}_z$ . These matrices are called **Pauli matrices**; they are usually denoted by  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$ . Thus

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (40)$$

If we also denote the  $2 \times 2$  unit matrix by  $\sigma_0$ , then we can easily find the following properties of the Pauli matrices:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \sigma_0 \quad (41)$$

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k \quad (42)$$

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}\sigma_0 \quad (43)$$

where we have introduced the anticommutator  $\{a, b\} = ab + ba$ . Thus, by adding equations (42) and (43), we get

$$\sigma_i\sigma_j = \delta_{ij}\sigma_0 + i\varepsilon_{ijk}\sigma_k \quad (44)$$

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

$$a_0\sigma_0 + a_1\sigma_x + a_2\sigma_y + a_3\sigma_z = 0 \quad (45)$$

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.  $\text{Tr } \sigma_x = \text{Tr } \sigma_y = \text{Tr } \sigma_z = 0$ , and that the trace of the unit matrix  $\sigma_0$  is 2. Thus taking the trace of Eq. (45) we get  $2a_0 = 0$  and hence  $a_0 = 0$ . If we now multiply Eq. (45) by  $\sigma_x$ , 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  $2 \times 2$  matrix  $A$  can be represented by a linear superposition of the Pauli matrices together with the unit matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_0\sigma_0 + a_1\sigma_x + a_2\sigma_y + a_3\sigma_z \quad (46)$$

The expansion coefficients  $a_0$  etc. are found by a procedure similar to the one used above. Taking the trace of Eq. (46) yields  $a_0 = \frac{1}{2}(a_{11} + a_{22})$ , and we get similar equations for the other expansion coefficients by multiplying Eq. (46) in turn by  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$  and taking the traces. Thus  $a_1 = \frac{1}{2}\text{Tr } \sigma_x A$ ,  $a_2 = \frac{1}{2}\text{Tr } \sigma_y A$  and  $a_3 = \frac{1}{2}\text{Tr } \sigma_z A$ .

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

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

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  $3 \times 3$  matrices since  $m$  can

take on three values, -1, 0 and +1. Generalising the rule explained by Eq. (37) we find the following expressions for the spin-1 matrices:

$$\begin{aligned}\hat{J}_x &= \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\ \hat{J}_y &= \frac{i\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \\ \hat{J}_z &= \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}\end{aligned}$$