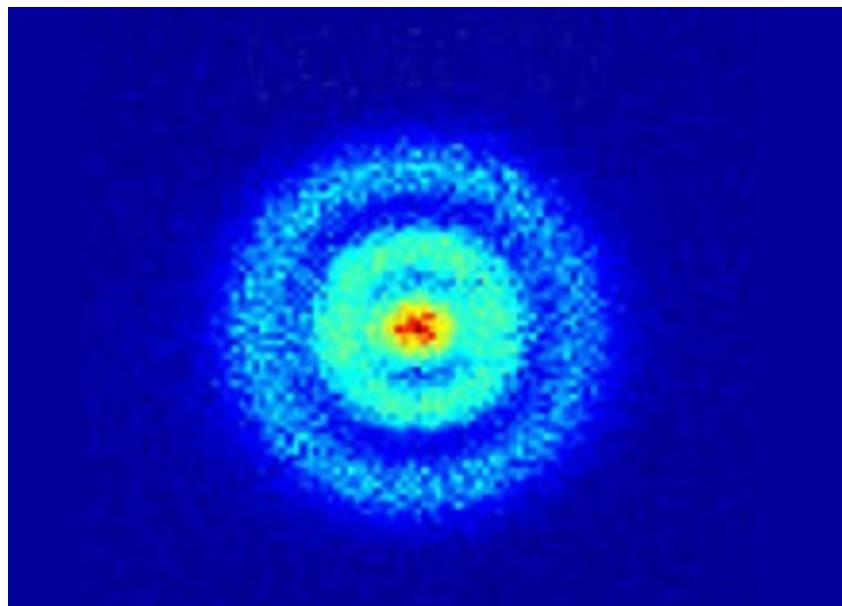


PHYS3008: Atomic Physics

Dr Debasish Banerjee
Physics and Astronomy
University of Southampton

Semester 1, 2025/26 academic year



The first direct experimental observation of the nodal structure of an electronic orbital of a hydrogen atom.

Stodolna et al., Phys. Rev. Lett. 110, 213001 (2013).

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 4 |
| 2 | Quantum Mechanics Revision | 5 |
| 2.1 | Operators | 5 |
| 2.2 | Eigenfunctions and Eigenvalues | 5 |
| 2.3 | Degeneracy | 5 |
| 2.4 | Expectation and Matrix Elements | 6 |
| 2.5 | Hermitian Operators | 6 |
| 2.6 | Orthogonality | 7 |
| 2.7 | Completeness | 7 |
| 2.8 | Commuting and Non-Commuting Operators | 7 |
| 2.9 | Schrödinger's Equation | 9 |
| 2.10 | Good Quantum Numbers | 10 |
| 2.11 | Linear Momentum | 11 |
| 2.12 | Orbital Angular Momentum | 11 |
| 2.13 | Vector space formalism of quantum mechanics | 15 |
| 2.14 | Spin of a Single Electron | 16 |
| 2.15 | Combining Spin and Wave function. | 17 |
| 2.16 | Multi-Electron Wave functions | 18 |
| 3 | The Hydrogen Atom | 20 |
| 3.1 | Solution of the Schrödinger equation | 20 |
| 3.2 | Comparison with the Bohr model (non-examinable) | 27 |
| 3.3 | Comparison with Experiment | 28 |
| 4 | Time-Independent Perturbation Theory | 32 |
| 4.1 | Non-degenerate | 32 |
| 4.2 | 1D Infinite Potential Well | 36 |
| 4.3 | 2D Infinite Potential Well | 37 |
| 4.4 | Stark Effect | 40 |
| 5 | The Fine Structure of the Hydrogen Atom | 44 |

| | | |
|----------|---|-----------|
| 5.1 | Spin-Orbit Coupling | 44 |
| 5.2 | Addition of Angular Momenta | 46 |
| 5.3 | Term Symbol | 47 |
| 5.4 | Relativistic Mass Correction | 47 |
| 5.5 | Darwin Correction | 47 |
| 5.6 | Sum of Fine Structure Corrections | 48 |
| 5.7 | Other effects | 48 |
| 6 | The Interaction of Atoms and Electromagnetic Waves | 49 |
| 6.1 | Time-Dependent Perturbation Theory | 49 |
| 6.2 | Einstein A and B Coefficients | 50 |
| 6.3 | Fermi's Golden Rule | 52 |
| 6.4 | Selection Rules | 55 |
| 6.5 | Forbidden Transitions | 57 |
| 7 | Two-Electron Atoms/Ions | 58 |
| 7.1 | Two-Electron Hamiltonian | 58 |
| 7.2 | The Wave function | 58 |
| 7.3 | Exchange Term | 60 |
| 7.4 | Comparison with Experiments | 61 |
| 7.5 | Optical transitions | 63 |
| 8 | Multi-Electron Atoms | 64 |
| 8.1 | The Hamiltonian and the Central Field Approximation | 64 |
| 8.2 | Self-Consistent Atomic Orbital Calculations | 65 |
| 8.3 | Alkali Metals | 66 |
| 8.4 | Periodic Table | 67 |
| 8.5 | Fine Structure of Multi-Electron Atoms | 71 |
| 8.6 | LS or Russell-Saunders Coupling | 74 |
| 8.7 | Hund's Rules | 76 |
| 8.8 | Optical Selection Rules | 78 |
| 8.9 | Zeeman Effect | 79 |

1 Introduction

In this course the application of quantum mechanics to atoms is studied. In the second-year course PHYS2003, the theoretical framework of quantum mechanics was introduced, together with the study of some physical systems. In PHYS3008 some of the problems that beset the first models of atoms in the early 20th century are discussed, and it is shown that quantum mechanics provides complete agreement with experiment.

The concept of a basic constituent of matter (*atom*) was theorised in several ancient cultures, the most well-known proponents being the Greek philosophers Leucippus and Democritus (in the fifth century BCE). Although there was no direct experimental evidence, several prominent physicists (e.g. Boyle, Gassendi) of the 17th century supported atomism, and Isaac Newton (in *Opticks* 1704) referred to matter as being composed of "solid, massy, hard, impenetrable, moveable particles". The first proposal based on experimental evidence that matter behaves as if it were composed of atoms was by Dalton in 1808 ("Law of multiple proportions"), developing suggestions by Proust ("Law of definite proportions") and Lavoisier in the late 1700s. Despite these advances, in the 19th century relatively few physicists considered atoms to be real entities (in contrast to chemists), and full acceptance by the scientific community only came around the year 1908 due to the agreement with experiment of Einstein's explanation of Brownian motion in terms of atoms. After the discovery of the electron in the period 1897-1899 it was clear that atoms must have a substructure, and the first atomic models were proposed in the first decade of the 20th century. The Bohr model (1913) was the first quantum model of the atom, invoking radical non-classical postulates. It had some spectacular successes, such as explaining the Balmer series of spectral lines, but could not accommodate all the (rapidly increasing) experimental results on atoms.

Many experimental data such as the spectral lines of elements (from 1853) and their fine-structure (from 1892), the Zeeman effect (1896) and the anomalous Zeeman effect (1898) etc, could not be fully explained by the early models of atoms. Quantum mechanics is needed in order to accommodate all these data in one theory, and several of these topics will be covered in the course.

These notes have been originally written by Prof. David Smith and were substantially updated by Prof. Tim Morris and Prof. Simone De Liberato. Minor modifications were introduced by Dr Pierre Thibault (year 19/20). Dr. Andrew Akeroyd made considerable revisions for the academic year 20/21, and minor revisions for the years 21/22, 22/23, 23/24 and 24/25. I have made minor revisions for the year 25/26.

2 Quantum Mechanics Revision

This section provides a brief review of some of the key concepts in PHYS2003 (Quantum Physics) that will be needed in PHYS3008. It is recommended to refer back to the notes for PHYS2003 throughout the course.

2.1 Operators

Within quantum mechanics all physically measurable quantities, called observables, can be mathematically represented by a linear Hermitian operator (more technically: *self-adjoint*). An operator is a mathematical object which acts on a function (the wave function in quantum mechanics) to produce another function. Examples of operators are *momentum* (which is the action of a differential operator, $\frac{\partial}{\partial x}$), or the *position* (the action of multiplication by x).

2.2 Eigenfunctions and Eigenvalues

For any particular operator there are certain functions which when acted on by the operator give as a result the original function scaled by a number *i.e.* if \hat{O} is an operator (denoted by a hat on an algebraic symbol), $\psi(x)$ is a function, and λ is a number then

$$\hat{O} \psi(x) = \lambda \psi(x).$$

In this case $\psi(x)$ is called an *eigenfunction* and λ is the corresponding eigenvalue.

In quantum mechanics, eigenvalues and eigenfunctions have a particular significance because the only values that an observable can take are the eigenvalues of the operator which represents it. The restriction that experimental observables should have real values requires the operators that correspond to observables to be Hermitian operators (see section 2.5), which necessarily have real eigenvalues.

2.3 Degeneracy

It is possible for there to be more than one eigenfunction with the same eigenvalue. This is called degeneracy. The degeneracy of a particular eigenvalue is the number of eigenfunctions with that eigenvalue e.g. in the following case the distinct functions $\psi_1(x)$ and $\psi_2(x)$ have the same eigenvalue λ :

$$\hat{O} \psi_1(x) = \lambda \psi_1(x), \quad \hat{O} \psi_2(x) = \lambda \psi_2(x).$$

Recall that any linear combination of the two wavefunctions $\psi_1(x)$ and $\psi_2(x)$ is also an eigenfunction of \hat{O} with the same eigenvalue.

2.4 Expectation and Matrix Elements

Whilst a single measurement of any particular observable always leads to one of the eigenvalues of the corresponding operator, the general wave function of a quantum system (e.g. before measurement) does not have to be an eigenfunction of any particular observable. In this case, measurements of the same observable on multiple identical copies of the system will not lead to the same result for the outcome of the measurement of each copy. The average value of the observable measured, which is called the expectation value, can be calculated, for 1D wave functions, using the following formula

$$\langle \hat{O} \rangle = \int \psi^*(x) \hat{O} \psi(x) dx ,$$

where $\langle \hat{O} \rangle$ is the expectation value of the operator \hat{O} , $\psi(x)$ is the wave function which describes the system before the measurement and the integral is over the whole of space. One can introduce a convenient notation for the integrals of this form:

$$\langle \psi_i | \hat{O} | \psi_j \rangle = \int \psi_i^*(x) \hat{O} \psi_j(x) dx , \quad (2.1)$$

where ψ_j is the j^{th} member of a set of wave functions (likewise for ψ_i). Taking $i = j$ gives the expectation value $\langle \hat{O} \rangle$ for a particular wave function. Each of these integrals is called a **matrix element**, because the values of all possible pairs (all i, j) of the integral for a certain set of wave functions form the elements a matrix. The resulting matrix is said to be a *representation* of \hat{O} in the basis of the set of wave functions ψ_i .

2.5 Hermitian Operators

All measurable quantities are real numbers with no imaginary component. This restricts any operator which represents an observable to be **Hermitian** (or *self-adjoint*). The mathematical definition of a Hermitian operator is that for all functions $f(x)$ and $g(x)$

$$\int f^*(x) \hat{O} g(x) dx = \int (\hat{O} f(x))^* g(x) dx .$$

From this it can be shown that all the eigenvalues of the operator are real (see PHYS2003 notes).

Operators act on a function and the order of the symbols is crucial in many cases. In the above example, in the first integral the operator is acting on the function $g(x)$, whereas in the second integral the operator is now acting on the function $f(x)$, with the result then being complex conjugated.

2.6 Orthogonality

Another extremely important property of the eigenfunctions of a Hermitian operator is that they form a set which is orthogonal *i.e.*

$$\int \psi_i^*(x) \psi_j(x) dx = \langle \psi_i | \psi_j \rangle = 0 \quad \text{unless} \quad i = j.$$

For normalised eigenfunctions the value of the matrix element with $i = j$ is equal to one and thus one can write

$$\langle \psi_i | \psi_j \rangle = \delta_{ij}.$$

This property is known as orthonormality. If there are degenerate eigenvalues then not all the eigenfunctions are orthonormal, but a linear combinations of these can be chosen to obtain an orthonormal set.

2.7 Completeness

A further important property of the eigenfunctions of a Hermitian operator representing some observable is that they form what is called a complete set. That is, any possible wave function of the system can be written down as a linear superposition of the eigenfunctions of this Hermitian operator, *i.e.*

$$\Psi(x) = \sum_i a_i \psi_i(x). \quad (2.2)$$

This result is equivalent to expanding a *vector* $\Psi(x)$ in a *basis* $\{\psi_i(x)\}$. By applying the orthonormality property of the eigenfunctions, the coefficients of the terms of this expansion can be obtained as follows (where i is a fixed index and j is an index to be summed over):

$$\langle \psi_i | \Psi \rangle = \sum_j a_j \langle \psi_i | \psi_j \rangle = \sum_j \delta_{ij} a_j = a_i \quad \text{or} \quad a_i = \langle \psi_i | \Psi \rangle. \quad (2.3)$$

2.8 Commuting and Non-Commuting Operators

As mentioned earlier, when two or more operators act on a function the order in which they are applied can give different results *e.g.* if \hat{O} and \hat{M} are both operators, then $\hat{O}\hat{M}\psi$ may not equal $\hat{M}\hat{O}\psi$. To aid the discussion of this property, the commutator $[\hat{O}, \hat{M}]$ of two operators \hat{O} and \hat{M} is defined by

$$[\hat{O}, \hat{M}] = \hat{O}\hat{M} - \hat{M}\hat{O}.$$

If the commutator of two operators is zero then they are said to commute, and the order of applying them does not matter. An important result is that if two

operators commute then all the eigenfunctions of one operator are also all the eigenfunctions of the other and *vice versa* (called *common eigenfunctions*).

Non-examinable section starts here

The proof (for the case of one operator not having degenerate eigenvalues) is as follows:

Let ψ_i^M be an eigenfunction of operator \hat{M} with eigenvalue λ_i^M . Let operator \hat{M} commute with operator \hat{O} and so

$$\hat{M}(\hat{O}\psi_i^M) = \hat{O}(\hat{M}\psi_i^M) = \lambda_i^M(\hat{O}\psi_i^M)$$

i.e. $\hat{O}\psi_i^M$ is an eigenfunction of \hat{M} with the (non-degenerate) eigenvalue λ_i^M . What is the effect of \hat{O} on ψ_i^M ? Usually one would expect a new function to be formed but then the above equation states that this new function is also an eigenfunction of \hat{M} with eigenvalue λ_i^M . Since the eigenvalue λ_i^M is not degenerate, then $\hat{O}\psi_i^M$ can only be the same eigenfunction ψ_i^M scaled by some number. If this number is denoted by λ_i^O then we have

$$\hat{O}\psi_i^M = \lambda_i^O\psi_i^M$$

i.e. ψ_i^M is an eigenfunction of \hat{O} with eigenvalue λ_i^O . In the case where λ_i^M is degenerate the proof is longer.

Non-examinable section ends here

Due to the fact that the eigenfunctions of two commuting operators are common eigenfunctions, this means that it is possible for a system to exist in a state which has a precisely defined value of both observables associated with these operators, i.e. measurements of each observable on this state give only one answer with certainty, rather than being capable of giving more than one answer with associated probabilities.

Conversely, if two operators do not commute then it is not possible to have certain values of both observables from a measurement on the state, which is Heisenberg's uncertainty principle

$$\Delta O \Delta M \geq \frac{1}{2} |\langle [\hat{O}, \hat{M}] \rangle|,$$

where ΔO and ΔM are the standard deviations of multiple measurements of the observables O and M on identical states. Note that on the right hand side of the equation there is an expectation value. This means that it the uncertainty relation depends on the state that the system is in i.e. its wave function. Taking the two operators to be \hat{p} and \hat{x} gives the well known uncertainty relation of Heisenberg

$$\Delta x \Delta p \geq \frac{1}{2} \hbar,$$

with the equality being obtained for Gaussian wave functions.

Non-examinable section starts here

In order to prove Heisenberg's uncertainty principle we start with the definition of the variance of an operator as

$$\Delta O^2 = \langle \psi | \hat{O}^2 | \psi \rangle - \langle \psi | \hat{O} | \psi \rangle^2.$$

We then introduce for a generic operator \hat{O} the state

$$|\psi_O\rangle = (\hat{O} - \langle \psi | \hat{O} | \psi \rangle) |\psi\rangle$$

such that

$$\langle \psi_O | \psi_O \rangle = \Delta O^2.$$

We thus have that

$$\Delta O^2 \Delta M^2 = \langle \psi_O | \psi_O \rangle \langle \psi_M | \psi_M \rangle \geq |\langle \psi_O | \psi_M \rangle|^2$$

where at the last step we exploited the Cauchy-Schwarz inequality. Now $\langle \psi_O | \psi_M \rangle$ is in general a complex number $\alpha + i\beta$; this means that

$$\Delta O^2 \Delta M^2 \geq |\langle \psi_O | \psi_M \rangle|^2 = \alpha^2 + \beta^2 \geq \beta^2$$

with

$$\beta = \frac{\langle \psi_O | \psi_M \rangle - \langle \psi_M | \psi_O \rangle}{2i} = \frac{1}{2i} \langle \psi | [\hat{O}, \hat{M}] | \psi \rangle.$$

Putting all this together and taking the square root we finally find

$$\Delta O \Delta M \geq \frac{1}{2} \left| \langle [\hat{O}, \hat{M}] \rangle \right|.$$

Non-examinable section ends here

2.9 Schrödinger's Equation

Schrödinger's equation describes how a wave function develops in time. It states that the time derivative of the wave function is proportional to the result of acting the total energy operator, the Hamiltonian \hat{H} , on the wave function *i.e.*

$$\hat{H}\psi = i\hbar \frac{\partial}{\partial t} \psi.$$

The standard Hamiltonian for a particle such as an electron contains two parts, the kinetic energy and its potential energy:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x).$$

The kinetic energy part is given by

$$\frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}.$$

It can be shown that the time dependence of an arbitrary wave function can be written down in a relatively simple way. If at a specific time (i.e. $t = 0$) the expansion of the arbitrary wave function in terms of the energy eigenfunctions of the system is known, then at any time later the form of this wave function is given by

$$\psi(x, t) = \sum_j \langle \psi_j^E | \psi(t=0) \rangle \psi_j^E(x) e^{-i \frac{E_j}{\hbar} t},$$

where $\psi_j^E(x)$ are the energy eigenfunctions, and E_j is the corresponding eigenvalue of energy.

2.10 Good Quantum Numbers

An eigenfunction of the Hamiltonian can be labelled by the quantum number associated with its energy e.g. the principle quantum number n for the hydrogen atom, ψ_n . This is a convenient way of referring to eigenfunctions without writing their explicit mathematical forms. However, degeneracy in energy often means that the quantum number associated with energy does not alone uniquely define the eigenfunction. In this case one needs extra labels i.e. additional quantum numbers.

For a set of operators (i.e. observables) that all commute with each other (e.g. \hat{H} , \hat{L}^2 and \hat{L}_z for the hydrogen atom, just including the Coulomb interaction and the kinetic energy term in the expression for \hat{H}) their eigenfunctions are common, as discussed in section 2.8. Hence one can label the eigenfunction of energy by the quantum numbers associated with each operator of this commuting set of operators (ψ_{nlm}). These quantum numbers are called **good quantum numbers**, because if a measurement of an observable associated with one of the operators in this set is carried out on the state then the outcome will be the associated quantum number with certainty (probability=1), rather than different results (with various probabilities) from each measurement on identical states. The latter would be an example of a *bad quantum number* (e.g. an operator that does not commute with all the others, such as \hat{L}_y for the above case of the hydrogen atom, which does not commute with \hat{L}_z).

As we will see later, there are situations that can arise where the operators we are using to provide these labels no longer commute with the (complete) Hamiltonian for a given system. This will happen when perturbations to the Hamiltonian that are present in the system that is under consideration can no longer be ignored (e.g. the internal magnetic field in the hydrogen atom, which leads to an extra term in the Hamiltonian).

2.11 Linear Momentum

It is informative to look more closely at when linear momentum is conserved. Classically, linear momentum of a body is conserved as long as a force does not act on the body. Quantum mechanically, we require that the operators representing the three components of the linear momentum, namely

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}, \quad \hat{p}_y = -i\hbar \frac{\partial}{\partial y}, \quad \hat{p}_z = -i\hbar \frac{\partial}{\partial z},$$

have to commute with the Hamiltonian. It is reasonably easy to show that the linear momentum operators commute with the kinetic energy part of the Hamiltonian. In order for them to commute with the potential energy part of the Hamiltonian, $V(\mathbf{r})$, one has to evaluate $[-i\hbar\partial/\partial x, V(\mathbf{r})]$ etc acting on an arbitrary wave function $\psi(\mathbf{r})$, where we ignore $-i\hbar$ in what follows and write $\psi(\mathbf{r}) = \psi$ and $V(\mathbf{r}) = V$:

$$\frac{\partial}{\partial x} (V\psi) - V \frac{\partial}{\partial x} (\psi) = \frac{\partial V}{\partial x} \psi + V \frac{\partial \psi}{\partial x} - V \frac{\partial \psi}{\partial x} = \frac{\partial V}{\partial x} \psi,$$

and similar for $[-i\hbar\partial/\partial y, V(\mathbf{r})]$ and $[-i\hbar\partial/\partial z, V(\mathbf{r})]$. Hence the condition that linear momentum is conserved is given by $\frac{\partial}{\partial x} V(\mathbf{r}) = 0$ everywhere (and same for \hat{p}_y and \hat{p}_z). This is (of course) equivalent to saying that there must be no forces acting on the body described by the Hamiltonian (i.e. the same as the classical condition).

Another way of looking at the fact that the derivative of the potential must be zero at all points in space, is that if we were as an observer to move by any arbitrary displacement within the system, the system would be identical around us. This is called **translational symmetry** and it turns out there is a fundamental link between symmetries of the system and conserved quantities. Basically, any global symmetry has an associated operator whose expectation value can be measured. If the dynamics of the system preserves the symmetry, then the operator is conserved. *Noether's Theorem* provides the detailed mathematical link.

2.12 Orbital Angular Momentum

The classical definition of the angular momentum \mathbf{L} (or moment of the momentum) is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}.$$

In quantum mechanics the operators representing the components of the orbital angular momentum are as follows:

$$\begin{aligned}\hat{L}_x &= \hat{y}\hat{p}_z - \hat{z}\hat{p}_y = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \\ \hat{L}_y &= \hat{z}\hat{p}_x - \hat{x}\hat{p}_z = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \\ \hat{L}_z &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right).\end{aligned}$$

In spherical polar coordinates these operators translate into:

$$\begin{aligned}\hat{L}_x &= i\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right), \\ \hat{L}_y &= i\hbar\left(-\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right), \\ \hat{L}_z &= -i\hbar\frac{\partial}{\partial\phi}.\end{aligned}$$

Finally, there is the operator corresponding to the square of the total angular momentum. This is defined as

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

and in spherical polar coordinates it becomes:

$$\hat{L}^2 = -\hbar^2 \left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta} \left(\sin\theta\frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \right].$$

It is possible to show that the different components of the angular momentum do not commute with each other, but that each commutes with the total angular momentum operator. Therefore, when discussing orbital angular momentum we will use the eigenvalues of the total angular momentum and one of the components. This is chosen to be the z component as this is the simplest to write down in spherical polar co-ordinates.

Since the angular momentum of a particle which experiences zero force (torque) is conserved, all of the operators discussed above must commute with the kinetic energy part of the Hamiltonian for a free particle (which can be shown). For the same reason, if the potential energy is uniform in space (no r, θ, ϕ dependence and so no force) then these operators will commute with the full Hamiltonian. Moreover, if the potential is spherically symmetric (i.e. $V(r, \theta, \phi) = V(r)$ with no θ, ϕ dependence), then the orbital angular momentum operators (which have $\partial/\partial\theta$ and $\partial/\partial\phi$ terms but no $\partial/\partial r$ term) commute with the Hamiltonian and orbital angular momentum will be conserved.

In this case, the eigenfunctions of the Hamiltonian are eigenfunctions of the total angular momentum operator (\hat{L}^2) and of the operator associated with the z

component of the angular momentum (\hat{L}_z) i.e. three commuting operators which necessarily have common eigenfunctions. Understanding these eigenfunctions is important. We will start by solving for the eigenfunctions of the total angular momentum operator \hat{L}^2 . The partial derivatives in \hat{L}^2 involve either θ or ϕ and thus we will try a *separation of variables* solution *i.e.* labelling $Y(\theta, \phi)$ as the eigenfunctions of \hat{L}^2 one then writes $Y(\theta, \phi) = \Theta(\theta)\Phi(\phi)$. If we insert this separable solution into the eigenvalue equation for \hat{L}^2 using Λ as the unknown eigenvalues (*i.e.* the equation $\hat{L}^2 Y(\theta, \phi) = \Lambda Y(\theta, \phi)$) then we can rearrange to get the equation

$$\frac{-\hbar^2 \sin(\theta)}{\Theta(\theta)} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta(\theta)}{\partial \theta} \right) - \Lambda \sin^2 \theta = \frac{\hbar^2}{\Phi(\phi)} \frac{\partial^2 \Phi(\phi)}{\partial \phi^2}.$$

Since the two sides of these equations do not depend on the same variable, the only way this can be true is if they are both constants. We can thus rewrite these equations using a (separation) constant b to give

$$\frac{-\hbar^2 \sin(\theta)}{\Theta(\theta)} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta(\theta)}{\partial \theta} \right) - \Lambda \sin^2 \theta = b$$

and

$$\frac{\hbar^2}{\Phi(\phi)} \frac{\partial^2 \Phi(\phi)}{\partial \phi^2} = b.$$

As the second equation is simpler we will solve this first. The solutions are given by

$$\Phi(\phi) = e^{im\phi}$$

where $b = -(\hbar^2 m^2)$ and m need not be an integer. These solutions are also the eigenfunctions of the z component of angular momentum operator \hat{L}_Z . However, in order for the solutions to be physically reasonable they are required to be single valued *i.e.*

$$\Phi(\phi) = \Phi(\phi + 2\pi).$$

This condition requires m to be an integer and the associated eigenvalue of the \hat{L}_z operator is $\hbar m$. For each of these solutions we now need to solve the above θ dependent equation (called *Laplace's equation in spherical co-ordinates*) (which we write again, but now with $b = -(\hbar^2 m^2)$)

$$\frac{-\hbar^2 \sin(\theta)}{\Theta(\theta)} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta(\theta)}{\partial \theta} \right) - \Lambda \sin^2 \theta = -\hbar^2 m^2$$

to find $\Theta(\theta)$, with the knowledge that m (which has come from the separation constant b) is now an integer.

The solutions of this equation consist of functions based upon a sum of terms involving different powers of sin and cos functions and have an associated quantum number l which gives the eigenvalue of the total angular momentum operator via $\Lambda = \hbar^2 l(l+1)$. It turns out that physically reasonable solutions of this equation only exist for $-l \leq m \leq l$.

The $Y(\theta, \phi)$ are known as spherical harmonics. Examples of these functions are given below. The letters on the left of the table are a standard way that orbitals are labelled based on their total angular momentum, where s corresponds to $l=0$, p corresponds to $l=1$ etc.

| | |
|---|---|
| s | $Y_{0,0} = \sqrt{\frac{1}{4\pi}}$ |
| p | $Y_{1,0} = \sqrt{\frac{3}{4\pi}} \cos \theta$ $Y_{1,\pm 1} = \mp \sqrt{\frac{3}{8\pi}} \sin \theta \exp(\pm i\phi)$ |
| d | $Y_{2,0} = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1)$ $Y_{2,\pm 1} = \mp \sqrt{\frac{15}{8\pi}} \cos \theta \sin \theta \exp(\pm i\phi)$ $Y_{2,\pm 2} = \sqrt{\frac{15}{32\pi}} \sin^2 \theta \exp(\pm i2\phi)$ |

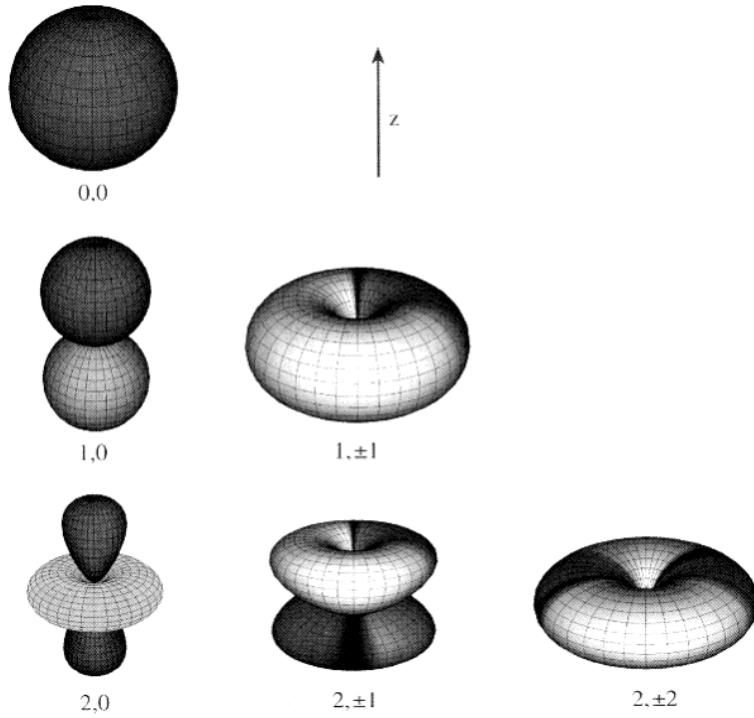


Figure 1: Representation of the shapes of the spherical harmonics $Y_{l,m}$ for $l \leq 2$. In the dark and light shaded regions the wave function have opposite signs and for $m > 0$ the phase changes by $2m\pi$ as the azimuthal angle ϕ increases from 0 to 2π .

The shapes of these spherical harmonics are shown in Figure 1 above. The value

of $|Y_{lm}|$ at a particular value of θ and ϕ is given by the distance from the origin of the spherical coordinate system to a point on the depicted surface. In this figure the dark shading represents positive amplitude for the wave function, while the light shading represents negative amplitude: e.g. in the representation of Y_{10} the upper lobe ($\theta < \pi/2$) has a positive amplitude, while the lower lobe ($\theta > \pi/2$) has a negative amplitude. All $Y_{l,m}$ with $m=0$ have azimuthal symmetry (*i.e.* no ϕ dependence). For $Y_{l,m}$ with non-zero m , the sign of the wavefunction changes m times as ϕ increases from 0 to π : this may be seen in the change in the shading in Figure 1 in the representations of $Y_{2,1}$ and $Y_{2,2}$.

2.13 Vector space formalism of quantum mechanics

Schrödinger's approach to quantum mechanics (wave mechanics, 1926) and Heisenberg's approach (matrix mechanics, 1925) looked very different when the theories were proposed. However, both theories gave the same predictions for physical systems that could be calculated (e.g. a harmonic oscillator and the hydrogen atom). Within a few years it was shown that both are different aspects of the same underlying theory that makes use of an infinite dimensional vector space (called *Hilbert space*). The notation $\langle \psi_i | \psi_j \rangle$ for the integral of the product of wave functions that was introduced in Sec 2.4 has a mathematical meaning in which $\langle \psi |$ and $|\psi \rangle$ are vectors, and are called a *bra vector* and a *ket vector* respectively. A bra vector and a ket can be represented by row vectors and column vectors respectively, and $\langle \psi |$ and $|\psi \rangle$ can be obtained from each other by the *adjoint* operation (which takes the complex conjugate and the transpose of the column or row vector). The two vectors can be combined to give a number (in analogy with the scalar product of two vectors in classical mechanics), and this is referred to as an *inner product* $\langle \psi_i | \psi_j \rangle$, also called a *dot product*. Operators can be represented by matrices.

The postulates of quantum mechanics are written in terms of $|\psi\rangle$, which is called the *state vector*. The state vector has unit *length* (which is called the *norm*), which leads to a normalised wave function. The state vector can be expanded in terms of basis vectors of the Hilbert space, and the eigenvectors of a Hermitian operator form *one such basis* for the Hilbert space. The eigenvectors of the Hamiltonian, labelled by $|E_i\rangle$, satisfy $\hat{H}|E_i\rangle = E_i|E_i\rangle$ with E_i being the eigenvalues. A state vector can be expanded in terms of these basis vectors $|E_i\rangle$ via

$$|\psi\rangle = \sum_i a_i |E_i\rangle , \quad (2.4)$$

with a_i being the (discrete) components in this basis. The eigenvectors of position $|x\rangle$ (where $\hat{x}|x\rangle = x|x\rangle$) and linear momentum $|p\rangle$ (where $\hat{p}|p\rangle = p|p\rangle$) also form a basis for the Hilbert space. Since x and p always take continuous values, the expansion of the state vector in these bases takes the following form

(with integrals replacing summations):

$$|\psi\rangle = \int \psi(x) |x\rangle dx, \quad |\psi\rangle = \int \psi(p) |p\rangle dp.$$

Here $\psi(x)$ and $\psi(p)$ are the position-space wave function and momentum-space wave functions (in one dimension) respectively i.e. the continuous components of the state vector in the bases $\{|x\rangle\}$ and $\{|p\rangle\}$, where $a_i \rightarrow a_x = \psi(x)$ etc. This is the meaning of the wave function in the vector space formalism. The associated bra vector $\langle\psi|$ is obtained by the adjoint operation, and it can be shown that the inner product for $|\psi_i\rangle$ and $|\psi_j\rangle$ (when these state vectors are both expanded in $|x\rangle$ eigenvectors) can be expressed as follows:

$$\langle\psi_i|\psi_j\rangle = \int \psi_i^*(x) \psi_j(x) dx, \quad (2.5)$$

in agreement with what was written earlier. However, $\langle\psi_i|\psi_j\rangle$ could also be evaluated in the $|p\rangle$ basis, the $|E_i\rangle$ basis, or in any other basis.

The position-space wave function of a state vector $|\psi\rangle$ in vector-space notation is given by $\langle x|\psi\rangle$. This is analogous to eq. (2.3), where in the continuous case the basis $\{|\psi_i\rangle\}$ has been changed to the basis $\{|x\rangle\}$. Eq. (2.2) is obtained from eq. (2.4) by multiplying from the left by $\langle x|$, as follows:

$$|\psi\rangle = \sum_i a_i |E_i\rangle \rightarrow \langle x|\psi\rangle = \sum_i a_i \langle x|E_i\rangle \rightarrow \psi(x) = \sum_i a_i \psi_i(x). \quad (2.6)$$

An eigenfunction $\psi_i(x)$ is the position-space wave function (i.e. the components in the $\{|x\rangle\}$ basis, denoted by $\langle x|E_i\rangle$) of an eigenvector $|E_i\rangle$ of the Hamiltonian. The vector-space approach to quantum mechanics is covered in more detail in the course PHYS6003 (Advanced Quantum Physics). In PHYS3008 it is sufficient to think of inner products like $\langle\psi_i|\psi_j\rangle$ and $\langle\psi_i|\hat{O}|\psi_j\rangle$ as being a notation for integrals over a products of wave functions, as in eq. (2.5) and eq. (2.1), which is equivalent to working in the position basis.

2.14 Spin of a Single Electron

We now discuss spin angular momentum S , which has no classical counterpart. Spin angular momentum is different from orbital angular momentum L . Spin angular momentum was discovered in the Stern-Gerlach experiment in 1922 in which particles were seen to be deflected by a magnetic field even when the particles were arranged such that $l = 0$. This deflection cannot then be attributed to orbital angular momentum, and it confirmed the existence of another type of angular momentum, which is now known as spin angular momentum.

The spin operators act on a state vector of spin $|\psi_s\rangle$, which resides in a vector space V_s . It is important to note that this vector space V_s is distinct from the Hilbert space (called V_H , say) in which the state vector $|\psi\rangle$ resides. The Hilbert

space V_H is of infinite dimension, with bases corresponding to all observables in classical mechanics (such as position $|x\rangle$, momentum $|p\rangle$, energy $|E_i\rangle$ and orbital angular momentum etc). The spin of a particle is an intrinsic property and does not depend on its spatial coordinates and motion. The vector space of the spin of an electron is of dimension two, as there are only two possible values ($\pm\hbar/2$) for a measurement of one of its components (say \hat{s}_Z), and this quantum number is labelled by m_s .

The commutation relations for the spin operators (\hat{s}_x , \hat{s}_y , \hat{s}_z) are postulated to be the same as those for orbital angular momentum, and lead to the Pauli spin matrices. Thus for each electron one can define two commuting operators, one choice being the magnitude squared of the spin, $\hat{\mathbf{s}}^2$, and the z -component of the spin, \hat{s}_z . For an electron, the magnitude squared of the spin is always measured to be $3\hbar^2/4$ and the z -component has two eigenstates with eigenvalues $m_s = \pm\hbar/2$. The common eigenvectors of these two operators are often taken as the working basis (called the *standard basis*) of the vector space of spin. These eigenvectors can be labelled/represented in a number of different ways. In this course the state with positive z -component of spin ($m_s = +\hbar/2$) is labelled by $|\uparrow\rangle$ and the state with the negative component ($m_s = -\hbar/2$) is labelled by $|\downarrow\rangle$.

2.15 Combining Spin and Wave function.

The (total) state vector of a particle (such as an electron) is given by the *tensor product* of the state vector $|\psi\rangle$ and the spin state vector $|\psi_s\rangle$, written as $|\psi\rangle \otimes |\psi_s\rangle$. This is studied in detail in PHYS6003. The tensor product operation \otimes combines two column vectors into one column vector of dimensionality equal to the product of the individual dimensionalities e.g. a 2×1 column vector and a 2×1 column vector are combined into a 4×1 column vector. In quantum mechanics $|\psi\rangle$ is of infinite dimension and $|\psi_s\rangle$ is of dimension two. It is common to write $|\psi\rangle \otimes |\psi_s\rangle$ as simply $|\psi\rangle |\psi_s\rangle$, and we will adopt this convention. The operators from the Hilbert space V_H (e.g. \hat{H} , \hat{x} , \hat{p} etc) act on $|\psi\rangle$ only, while the operators of spin act on $|\psi_s\rangle$ only. In general, $|\psi_s\rangle$ is a linear combination of $|\uparrow\rangle$ and $|\downarrow\rangle$.

In atomic physics it will be very useful to use single electron eigenvectors of energy that are labelled by a discrete index or set of indices (quantum numbers) that identifies this orbital e.g. $|n\rangle$, with n being the quantum number of energy, and the associated position-space wave function being $\psi_n(\mathbf{r})$. For hydrogen it would actually have the indices n, l, m which you met in PHYS2003 last year and will be reviewed in chapter 3. As mentioned before, we will usually be working in the position basis and so we write $|\psi\rangle |\psi_s\rangle$ as the product of the wave function and state vector of spin i.e. $\psi_n(\mathbf{r}) |\psi_s\rangle$.

For multi-electron atoms our choice will not quite be these but rather ones that correspond to the energy levels of a simplified Hamiltonian where electron-electron interactions are ignored or otherwise suppressed in some way. For electrons that are bound (rather than free at $r = \infty$), these orbitals also form a

complete set of eigenfunctions. For any bound state electron wave function one can write

$$\psi = \sum_{n,s} a_{n,s} \psi_n(\mathbf{r}) |s\rangle . \quad (2.7)$$

Here $s = 1, 2$ corresponds to $|\uparrow\rangle$ and $|\downarrow\rangle$, and $|a_{n,s}|^2$ is the probability to find the electron with spin s and occupying the single-electron orbital ψ_n .

2.16 Multi-Electron Wave functions

Quantum mechanics normally describes point particles such us a single electron. However in atoms we will in general be dealing with multi-electron systems. In this case the wave function describing the system depends on the spatial and spin co-ordinates of each electron, and on time, *i.e.* for two electrons (as in helium)

$$\psi(\mathbf{r}_1, s_1; \mathbf{r}_2, s_2; t) .$$

The form of this wave function is constrained by the fact that electrons are indistinguishable, as discussed in PHYS2003 and PHYS2024. It is not possible to stick meaningful identity labels on the individual electrons (*i.e.* specifying them uniquely as *electron 1* and *electron 2*) and then keep track of both of them as they move around their orbits, as one could in principle do with classical particles that have a trajectory. These labels of **1** and **2** on particles in quantum mechanics are just dummy labels, and no physically measurable property can be affected by an exchange of these particle labels. This means that a multi-electron wave function must have *exchange symmetry*. That is, if the labels 1 and 2 on the spatial co-ordinates and spin are exchanged then any measurable quantity, and in particular the probability density, must be unchanged, *i.e.*

$$|\psi(\mathbf{r}_1, s_1; \mathbf{r}_2, s_2; t)|^2 = |\psi(\mathbf{r}_2, s_2; \mathbf{r}_1, s_1; t)|^2 .$$

Let us consider an operator whose effect is to exchange the two particle labels ($1 \rightarrow 2$ and $2 \rightarrow 1$). As shown in PHYS2024, the requirement of exchange symmetry on a two-particle wave function requires the wave function to be either symmetric under the exchange of particle labels (*i.e.* the wave function is unchanged after the labels are swapped) or anti-symmetric (*i.e.* the wave function picks up an overall minus sign with respect to the original wave function). Electrons have anti-symmetric wave functions, which is the explanation of Pauli's exclusion principle in quantum mechanics. A two-electron wavefunction can sometimes be written as the product of single particle electron wavefunctions. The overall wave function should be anti-symmetric, which means that a symmetric spatial wave function must be paired with an anti-symmetric spin state vector (and vice versa). These two solutions have consequences for the energy levels of helium (chapter 7).

| S | M_S | ψ_{spin} | $\psi_{spatial}$ |
|-----|-------|---|---|
| 0 | 0 | $\frac{1}{\sqrt{2}}(\uparrow\rangle_1 \downarrow\rangle_2 - \downarrow\rangle_1 \uparrow\rangle_2)$ | $\frac{1}{\sqrt{2}}[u_A(\mathbf{r}_1)u_B(\mathbf{r}_2) + u_B(\mathbf{r}_1)u_A(\mathbf{r}_2)]$ |
| 1 | +1 | $ \uparrow\rangle_1 \uparrow\rangle_2$ | |
| | 0 | $\frac{1}{\sqrt{2}}(\uparrow\rangle_1 \downarrow\rangle_2 + \downarrow\rangle_1 \uparrow\rangle_2)$ | $\frac{1}{\sqrt{2}}[u_A(\mathbf{r}_1)u_B(\mathbf{r}_2) - u_B(\mathbf{r}_1)u_A(\mathbf{r}_2)]$ |
| | -1 | $ \downarrow\rangle_1 \downarrow\rangle_2$ | |

Table 1: Spin wave functions and spatial eigenfunctions for a two-electron atom with electronic configuration (quantum numbers) designated by the labels A and B . The upper row has anti-symmetric ψ_{spin} and symmetric $\psi_{spatial}$, and vice versa for the lower row.

The spatial part of a two-electron eigenfunction is given by

$$\psi_{AB}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} [u_A(\mathbf{r}_1)u_B(\mathbf{r}_2) \pm u_B(\mathbf{r}_1)u_A(\mathbf{r}_2)]. \quad (2.8)$$

Here u_A and u_B are single particle eigenfunctions (i.e. they just depend on \mathbf{r}_1 or \mathbf{r}_2 but not on both) with quantum numbers given by A and B . Under the exchange of particle labels one can see that the plus sign gives the symmetric state, while the minus sign gives the anti-symmetric state. Note that both terms are multiplied by $1/\sqrt{2}$ for normalisation, as other (normalised) choices like $1/2$ and $\sqrt{3}/2$ for the first and second term respectively would not respect the exchange symmetry. The allowed spin wave states are

$$|\uparrow\rangle_1|\uparrow\rangle_2, \quad \frac{1}{\sqrt{2}}(|\uparrow\rangle_1|\downarrow\rangle_2 + |\downarrow\rangle_1|\uparrow\rangle_2), \quad |\downarrow\rangle_1|\downarrow\rangle_2, \quad \frac{1}{\sqrt{2}}(|\uparrow\rangle_1|\downarrow\rangle_2 - |\downarrow\rangle_1|\uparrow\rangle_2).$$

It can be seen that the first three are symmetric with respect to particle exchange, while last one is anti-symmetric. Note that only these states are allowed by the exchange symmetry.

Hence in order to have an overall two-electron wave function that is anti-symmetric one has the possibilities in Table 1. The values of the total spin (S) and z -component of total spin (M_S) are also given, and are found from combining the spins of the individual electrons. There is one spin state with $S = 0$ (called a singlet) and three spin states with $S = 1$ (called a triplet).

One can see Pauli's exclusion principle following from the form of these two-particle wavefunctions in Table 1. Setting $A = B$ (i.e. n, l, m to be the same for both electrons) leads to the vanishing of the anti-symmetric $\psi_{spatial}$, and so two electrons cannot have all four quantum numbers the same. Although the symmetric $\psi_{spatial}$ does not vanish for $A = B$, the anti-symmetric ψ_{spin} necessarily has the two electrons with a different z -component of spin.

3 The Hydrogen Atom

As the simplest atom and the only atom for which analytic wave functions can be derived, hydrogen is clearly the foundation of all atomic physics. The solution of the Schrödinger equation applied to hydrogen was covered in the course PHYS2003. The presentation here will cover some of the details of the mathematical solution of the equations, and will also focus on interpreting and understanding the results. Quantum Mechanics was first applied successfully to the hydrogen atom in late 1925 by Pauli, using Heisenberg's matrix mechanics. Schrödinger obtained the same results for the energy levels in early 1926 using his wave equation, which is the approach below.

3.1 Solution of the Schrödinger equation

The time-independent Schrödinger equation for the hydrogen atom is

$$\left[\frac{-\hbar^2}{2\mu} \nabla^2 - \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}|} \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (3.1)$$

where μ is the so-called reduced mass given by $\mu = m_e m_p / (m_e + m_p)$, m_e being the mass of the electron and m_p the mass of the proton, and \mathbf{r} is the displacement vector from the proton to the electron. We will take $\mu = m_e$, as $m_p \gg m_e$. There are a number of implicit assumptions in using the time-independent Schrödinger equation e.g. the equation is non-relativistic and the proton is taken to be a point-like particle. These assumptions do not give the exact values for the allowed energy levels in the hydrogen atom (the fine structure is covered later on in the course) but provide an excellent first approximation (*gross structure*) to the true energy levels.

In order to fully define the problem we also require boundary conditions. These are based on the physically reasonable assumptions that the wave function must tend to zero as r tends to infinity and the wave function must be single valued.

One important feature of the above Schrödinger equation is that the electrostatic potential is a so-called central potential, $V(r, \theta, \phi) \rightarrow V(r)$, i.e. it depends only on the distance from the origin (r) and not on the direction in space (θ, ϕ). Another way of saying this is that the potential is spherically symmetric. These two identical statements strongly suggest that we should solve the Schrödinger equation using spherical coordinates (instead of Cartesian coordinates), which will result in simpler differential equations. In spherical coordinates the Laplace operator ∇^2 is written

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}, \quad (3.2)$$

which can be rewritten using the total orbital angular momentum operator as

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{1}{\hbar^2 r^2} \hat{L}^2. \quad (3.3)$$

The fact that the radial and angular derivatives are in separate terms in the Laplace operator suggests that we trial a separation of variables solution of the form

$$\psi(r, \theta, \phi) = R(r)Y(\theta, \phi). \quad (3.4)$$

Substituting this trial solution into the Schrödinger equation and rearranging gives

$$\frac{1}{R(r)} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{2\mu E}{\hbar^2} r^2 + \frac{2\mu e^2}{\hbar^2 4\pi \varepsilon_0} r \right] R(r) = \frac{1}{\hbar^2 Y(\theta, \phi)} \hat{L}^2 Y(\theta, \phi). \quad (3.5)$$

As the two sides of this equation do not depend on the same variables, for the equation to be true they must both equal something which does not depend on any variables i.e. both sides must equal the same constant. This constant can be taken to be $l(l+1)$, and then the right-hand side of eq. (3.5) is the eigenfunction equation for angular momentum

$$\hat{L}^2 Y(\theta, \phi) = \hbar^2 l(l+1) Y(\theta, \phi), \quad (3.6)$$

that we met in sec. 2.12. Recall that the eigenfunctions $Y(\theta, \phi)$ are the spherical harmonics

$$Y_{l,m}(\theta, \phi) = a_{l,m} P_l^m(\cos \theta) e^{im\phi},$$

where l is the total orbital angular momentum quantum number, m the magnetic quantum number, P_l^m the so-called *associated Legendre polynomials*, and $a_{l,m}$ are normalisation constants.

From eq. (3.5), it then follows that

$$\frac{1}{R(r)} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{2\mu E}{\hbar^2} r^2 + \frac{2\mu e^2}{\hbar^2 4\pi \varepsilon_0} r \right] R(r) = l(l+1). \quad (3.7)$$

It is our task now to solve this *radial equation*, with the knowledge that l is a positive integer ($l = 0, 1, 2, \dots$), which is the result of solving the θ, ϕ equation for angular momentum. By rearrangement the radial equation becomes

$$\left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{e^2}{4\pi \varepsilon_0 r} + \frac{\hbar^2 l(l+1)}{2\mu r^2} \right] R(r) = ER(r). \quad (3.8)$$

Note that $R(r)$ does not cancel out from both sides because the first term on the left-hand side is a differential operator acting on $R(r)$. If l is not equal to zero then the orbital angular momentum term is effectively a short range repulsive potential (opposite in sign to the Coulomb attractive potential). This is exactly the centrifugal potential, seen in second year Classical Mechanics.

Expanding out the derivatives using the product rule, multiplying through by $-2\mu/\hbar^2$, using the formula for the Bohr radius $a_0 = 4\pi\hbar^2\varepsilon_0/(\mu e^2)$, and replacing $\partial/\partial r$ by d/dr etc because there is only one variable, the equation becomes:

$$\left[\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} + \frac{2}{ra_0} - \frac{l(l+1)}{r^2} \right] R(r) = -\frac{2\mu E}{\hbar^2} R(r). \quad (3.9)$$

For large r , all except the first term on the left-hand side can be neglected, and one has

$$\frac{d^2 R(r)}{dr^2} \sim -\frac{2\mu E}{\hbar^2} R(r). \quad (3.10)$$

If $E > 0$ this is just the equation for a free particle, with the solutions being oscillatory functions (cosines and sines) corresponding to unbound states. Bound states (i.e. the hydrogen atom) have $E < 0$. Defining the positive quantity

$$a = \sqrt{\frac{\hbar^2}{-2\mu E}}, \quad (3.11)$$

the solutions then are $R(r) \sim \exp(\pm r/a)$. The solution with the plus sign is rejected because it does not have the necessary behaviour of $R(r) \rightarrow 0$ as $r \rightarrow \infty$. Therefore only the solution with the negative sign is kept, and one can write:

$$R(r) = e^{-r/a} u(r), \quad (3.12)$$

where $u(r)$ is some function of r to be determined in what follows.

Substituting this expression for $R(r)$ into (3.9), expanding out the derivatives and rearranging, we get:

$$\frac{d^2 u(r)}{dr^2} + \frac{2}{r} \frac{du(r)}{dr} - \frac{l(l+1)}{r^2} u(r) = 2 \left(\frac{1}{a} - \frac{1}{a_0} \right) \frac{u(r)}{r} + \frac{2}{a} \frac{du(r)}{dr}. \quad (3.13)$$

To solve this equation we try a power series in r :

$$u(r) = \sum_{s=s_0}^{\infty} c_s r^s = r^{s_0} + c_{s_0+1} r^{s_0+1} + c_{s_0+2} r^{s_0+2} + \dots \quad (3.14)$$

Here s_0 is the lowest value of the summation integer s , and s_0 will be determined shortly. Without loss of generality we choose the coefficient of r^{s_0} (i.e. c_{s_0}) to be 1 (the equation is a linear differential equation and so any scaling of the obtained solution will also be a solution - we are scaling our trial solution so that $c_{s_0} = 1$). In eq. (3.13) all terms on the right-hand side can be moved to the left-hand side in order to give zero on the right-hand side. Substituting the power series for $u(r)$ (given in eq. (3.14)) in this rearranged version of eq. (3.13), it is clear that every power of r on the left-hand side must cancel separately in order give zero on the right-hand side. After a little tidying one finds that this requirement of cancellation of each power of r gives the *recurrence relation*:

$$[s(s+1) - l(l+1)] c_s = 2 \left(\frac{s}{a} - \frac{1}{a_0} \right) c_{s-1}. \quad (3.15)$$

If we look at the lowest power $s = s_0$ then the right-hand side of this equation is missing>equals zero (i.e. there is no coefficient c_{s_0-1} , due to $s_0 - 1$ not being allowed) and one has a quadratic equation given by $[s_0(s_0 + 1) - (l + 1)]c_{s_0} = 0$, and $c_{s_0}=1$ as mentioned earlier. By inspection the solutions are $s_0 = l$ or $s_0 = -l - 1$. The latter solution must be rejected because $u(r)$ has a negative power for $l = 0, 1, 2$ etc) and so this term would diverge at the origin $r = 0$.

Hence the lowest power is $s_0 = l$, and using the recurrence relation eq. (3.15) all the higher coefficients (starting from $c_{s_0} = 1$) can be computed. However, on inspection of eq. (3.15) it can be seen that for large s , one has $s^2 c_s \sim 2s c_{s-1}/a$. Iterating this relation, we find

$$c_s \sim \frac{2}{sa} c_{s-1} \sim \frac{2^2}{s(s-1)a^2} c_{s-2} \sim \cdots \sim \frac{1}{s!} \left(\frac{2}{a}\right)^s. \quad (3.16)$$

The large s coefficients are important when r is large, thus we learn that for large r ,

$$u(r) \sim \sum_s \frac{1}{s!} \left(\frac{2r}{a}\right)^s = e^{2r/a}. \quad (3.17)$$

It seems we have reached an impasse since if $u(r)$ really has this large r behaviour, then from eq. (3.12), we find $R \sim \exp(+r/a)$, which is the solution of eq. (3.10) that we rejected. There is one way out however. If the right-hand side of eq. (3.15) vanishes for some $s_{\max} = n \geq s_0 + 1 = l + 1$, then all coefficients $c_{s \geq n}$ will vanish and thus $u(r)$ will be a polynomial (not an infinite series). This happens when $a = n a_0$. Using eq. (3.11) for a and squaring, we find Bohr's famous formula:

$$E = -\frac{\hbar^2}{2\mu a_0^2 n^2}, \quad (3.18)$$

together with the constraint $l \leq n - 1$, and the solutions are

$$R(r) = R_{n,l}(r) := e^{-r/[na_0]} \sum_{s=l}^{n-1} c_s r^s. \quad (3.19)$$

These solutions need to be multiplied by a normalisation constant $c_{n,l}$, which is determined from the following normalisation condition (see PHYS2003 notes, where the r^2 in the integrand has its origin in the volume element $dV = r^2 \sin \theta dr d\theta d\phi$ in spherical coordinates):

$$\int_0^\infty c_{n,l}^2 R_{n,l}^2 r^2 dr = 1.$$

The explicit mathematical form of some these normalised eigenfunctions ($R_{n,l}$) are given below, in which the first multiplicative factor containing a_0 is the normalisation constant $c_{n,l}$ and the middle term in brackets is the power series:

$$\begin{aligned}
R_{1,0} &= 2(a_0)^{-3/2} \exp(-r/a_0) \\
R_{2,0} &= 2(2a_0)^{-3/2} (1 - r/[2a_0]) \exp(-r/[2a_0]) \\
R_{2,1} &= (2/\sqrt{3})(2a_0)^{-3/2} (r/[2a_0]) \exp(-r/[2a_0]) \\
R_{3,0} &= 2(3a_0)^{-3/2} (1 - 2r/[3a_0] + 2r^2/[27a_0^2]) \exp(-r/[3a_0]) \\
R_{3,1} &= (4\sqrt{2})(3a_0)^{-3/2} [(r/[3a_0])(1 - r/[6a_0])] \exp(-r/[3a_0]) .
\end{aligned} \tag{3.20}$$

Here the first subscript of the function $R(r)$ is the principal quantum number n , the second is the total angular momentum number l , and the Bohr radius is given by:

$$a_0 = \frac{4\pi\hbar^2\varepsilon_0}{\mu e^2} \approx 5.3 \times 10^{-11}\text{m} . \tag{3.21}$$

The radial eigenfunctions are real (i.e. no imaginary part). Radial eigenfunctions for $l = 0$ (i.e. $R_{n,0}$) have $s_0 = l = 0$ and so the power series starts from a constant (i.e. $r^{s_0} = r^0 = 1$). Since $c_{s_0} = 1$ (as mentioned before), which for $s_0 = 0$ means $c_0 = 1$, one calculates c_1 (the term linear in r) from the recurrence relation (eq.3.15) applied to $s = 1$ and $n = 1$. This gives $c_1[1(1+1) - 0(0+1)] = 2[1/(1a_0) - 1/(a_0)]c_0$, and hence $c_1 = 0$ (and consequently all other c_s for $s > 1$ are zero). There is only one term in the power series ($c_0 = 1$) and the normalisation constant $c_{1,0} = 2(a_0)^{-3/2}$. Radial eigenfunctions for $l = 1$ (i.e. $R_{n,1}$) have $s_0 = l = 1$ and so the power series starts from a term linear in r (i.e. $r^{s_0} = r^1 = 1$). For these eigenfunctions $c_{s_0} = c_1 = 1$, and there is no c_0 . One uses the recurrence relation to calculate c_2 from c_1 etc. In Fig.2 the functions $R_{1,0}$, $R_{2,0}$ and $R_{2,1}$ are plotted as a function of r/a_0 . It can be seen that the exponential factor ensures that $R_{n,l}$ decreases as r/a_0 increases. Note that $R_{2,1}$ goes to zero as $r \rightarrow 0$. This is an effect of the term in the Schrödinger equation that depends on $l(l+1)$ (the "repulsive potential" in eq.(3.8)) and is non-zero for $l \neq 0$.

The radial probability density function is given by

$$P_{n,l}(r) = r^2 |R_{n,l}(r)|^2 , \tag{3.22}$$

where $P_{n,l}(r)dr$ is the probability of finding the electron between r and $r + dr$. As mentioned earlier, $R_{n,l}$ are real and the modulus signs are only included for completeness. In Fig.3 the functions $P_{n,l}(r)$ with $n = 1, 2, 3$, each with $l = 0$, are plotted. One can see that the smaller the value of n the more likely the electron is to be found nearer the nucleus. In Fig.4 the functions $P_{n,l}(r)$ with $n = 3$, and $l = 0, 1, 2$ are plotted. The larger the value of l the more likely the electron is to be found nearer the nucleus. The expectation value for r can be found from the integral

$$\langle r \rangle = \int_0^\infty P_{n,l} r dr = \int_0^\infty R_{n,l}^2 r^3 dr . \tag{3.23}$$

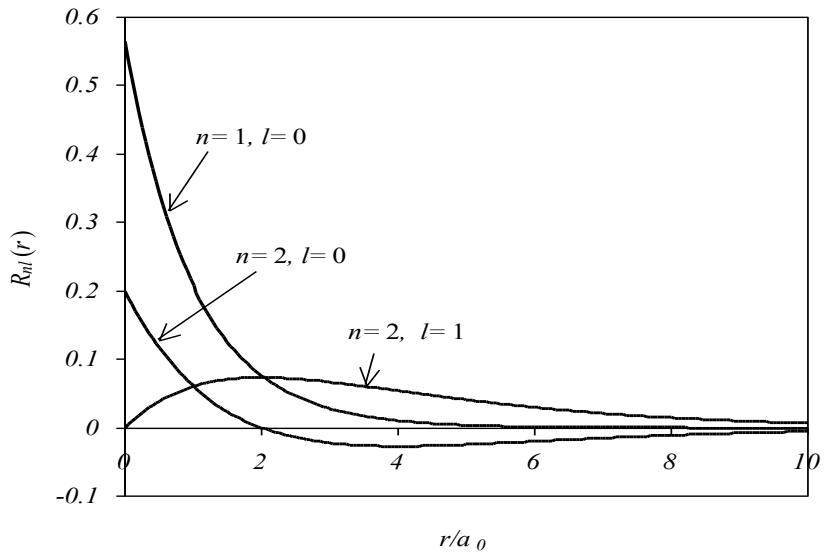


Figure 2: Radial part of the wave function $R_{n,l}(r)$ for $n = 1, l = 0$; $n = 2, l = 0, 1$.

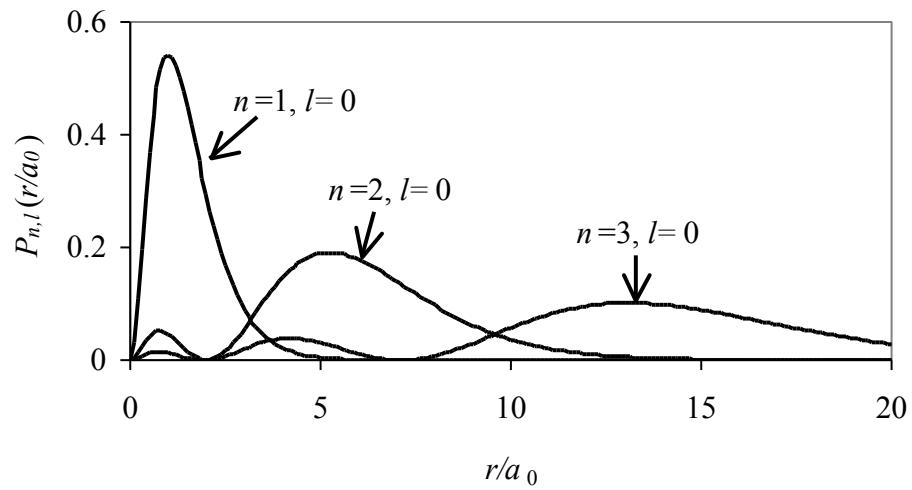


Figure 3: Radial extent of the probability density $P_{n,l}(r) = r^2 |R_{n,l}(r)|^2$ for the hydrogen atom with $n = 1, 2, 3$ and $l = 0$.

The energy associated with these wave functions can also be written as

$$E_n = -\frac{1}{4} \frac{2\mu}{\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{1}{n^2} \approx -\frac{13.6}{n^2} \text{ eV}, \quad (3.24)$$

and is also often presented in the form

$$E_n = -\frac{\alpha^2 \mu c^2}{2n^2},$$

where α is the so-called *fine structure constant* (see Chapter 5):

$$\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c} \approx \frac{1}{137}.$$

The eigenvalues of energy (gross structure) of the hydrogen atom depend solely on the principal quantum number n . The fact that the energies do not depend on l is what is called an accidental degeneracy (all states of l for a given n have the same energy). This is because the potential energy for hydrogen has an exact $1/r$ dependence. In atoms with more than one electron the gross energy levels depend on both n and l because the (more complicated) potential does not have an exact $1/r$ dependence.

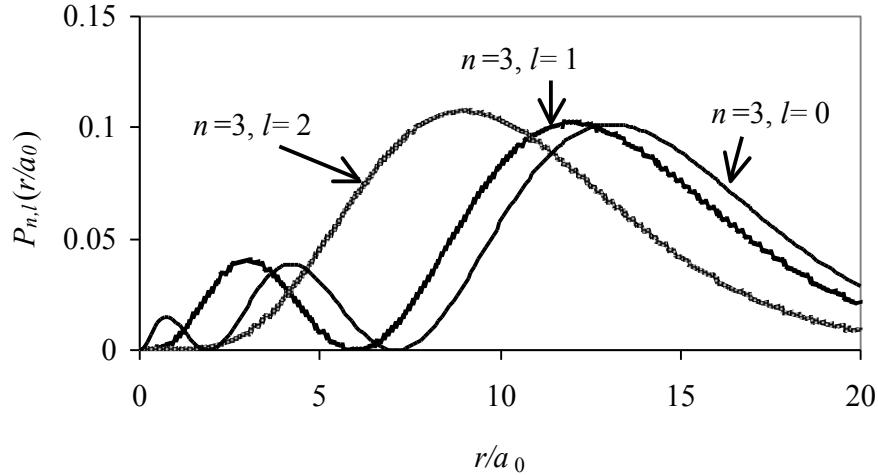


Figure 4: Radial extent of the probability density $P_{n,l}(r) = r^2 |R_{n,l}(r)|^2$ for the hydrogen atom with $n = 3$, and $l = 0, 1, 2$.

3.2 Comparison with the Bohr model (non-examinable)

Non-examinable section starts here

Let us now compare the predictions of quantum mechanics for the hydrogen atom with those of the Bohr model. In the original Bohr model of 1913 the electron in the hydrogen atom is in a circular orbit around the nucleus. There is one quantum number (n), which determines both the energy and the orbital angular momentum. The model was extended to include three quantum numbers and elliptical orbits by Sommerfeld in 1916, although we will just compare quantum mechanics with the original Bohr model.

1. Energy (E_n):

The quantised energies given by eq.(3.24) are the same in both the Bohr model and in quantum mechanics. This is somewhat surprising, given the very different nature of the two models, and the much longer mathematical derivation that was needed in the quantum mechanical calculation.

2. Orbital angular momentum ($|L|$):

A postulate of the Bohr model is that the magnitude of orbital angular momentum $|L|$ is $n\hbar$, with n being an integer. In the quantum mechanical description of the hydrogen atom the quantised values of $|L|$ are not postulated, but instead are a prediction of the theory (following from the postulates of quantum mechanics). Moreover, $|L| = \sqrt{l(l+1)}\hbar$ where l can take any integer value from 0 to $n-1$, instead of being uniquely fixed by the value of n . Hence $|L|$ is (in general) different from its magnitude in the Bohr model in which $|L|$ is equal to an integer multiple of \hbar . If l is large and takes its maximum value of $n-1$ then $|L|$ in quantum mechanics is given by $\sqrt{n(n-1)}\hbar \approx n\hbar$ for large n , in agreement with the Bohr model.

3. Ground state with $|L| = 0$:

The case of $l = 0$ gives rise to (surprisingly) $|L| = 0$. The ground state ($n = 1$) in quantum mechanics necessarily has $l = 0$ (i.e. no angular momentum), while the ground state in the Bohr model has momentum \hbar . This is a significant difference, because quantum mechanics predicts that the ground state of hydrogen has no contribution to the magnetic moment from orbital angular momentum, and this is confirmed experimentally. The state with $l = 0$ does not mean that the electron in the hydrogen atom is at rest (i.e. $|p| = 0$), as might be naïvely expected by applying the classical formula for angular momentum to the case of $|L| = 0$. In quantum mechanics an eigenvector of the angular momentum operator \hat{L}^2 can have a zero eigenvalue, but this state is a superposition of eigenvectors of \hat{p} , and does not correspond to a state with a definite momentum $p = 0$. Thinking classically, the other solution to $|L| = 0$ would be that the electron oscillates in a line through the nucleus, for which the displacement vector and momentum are collinear. However,

quantum mechanics does not describe trajectories, and instead predicts the probability of the electron being found in a spatial region. Of course, classical thinking should not be applied to the interpretation of quantum mechanical results.

4. Radii of orbitals:

In the Bohr model one has circular orbits with a definite value $r_n = \frac{n^2}{Z} \frac{m_e}{\mu} a_0$ for each n , where Z (nuclear charge) is equal to 1 for hydrogen. In quantum mechanics there are orbitals, which are mathematical functions of the coordinates and the quantum numbers n, l, m (i.e. the modulus squared of the wave functions of the eigenvectors of energy), and there is no trajectory for the electron. It so happens that the maximum of $P_{1,0}(r)$ is at the Bohr radius a_0 . The expectation value for the distance of the electron from the nucleus is given by:

$$\langle r \rangle_{n,l} = \frac{n^2}{Z} \frac{m_e}{\mu} a_0 \left(\frac{3}{2} - \frac{l(l+1)}{2n^2} \right). \quad (3.25)$$

For $n = 1, l = 0$ one has $\langle r \rangle_{1,0} = 3a_0/2$, thus showing that the Bohr radius is a good approximation of the size of the hydrogen atom in its ground state. For $l = n - 1$ and large n one has $\langle r \rangle = \frac{n^2}{Z} \frac{m_e}{\mu} a_0$, which is equal to the formula for the radius of the orbits in the Bohr model.

The energies given by eq. (3.24) are the *gross structure* of the atom i.e. the energy levels obtained from only considering the dominant contributions to the Hamiltonian. Including other contributions to the Hamiltonian results in additional energy levels, which is called the *fine structure*, and will be discussed in chapter 5.

Non-examinable section ends here

3.3 Comparison with Experiment

Emission spectra are normally measured using low pressure lamps in which electrons that are emitted from the cathode are accelerated by a high voltage; these electrons collide with the atoms of the gas under study, raising the electrons to higher energy levels. These excited atoms then relax back to their ground state by the electrons emitting electromagnetic energy, giving emission lines at specific wave lengths. Absorption spectra are obtained when light from a source is passed through the gas under study, and the transmitted light will have dark lines (*absorption lines*) at specific wave lengths. The light that was absorbed by the hydrogen is then emitted in random directions to give an emission spectrum. Both spectra are depicted in Figure 5 for visible light.

In the 1880s, before the Bohr model (1913) and quantum mechanics (1925/26), it had already been found that the lines emitted by atomic hydrogen could be

grouped into a number of different series. The first three of these series are called the Lyman, Balmer and Paschen series, with wavelengths given in Table 2. The lines within each series are observed to accumulate towards a limiting value (see Figure 6).

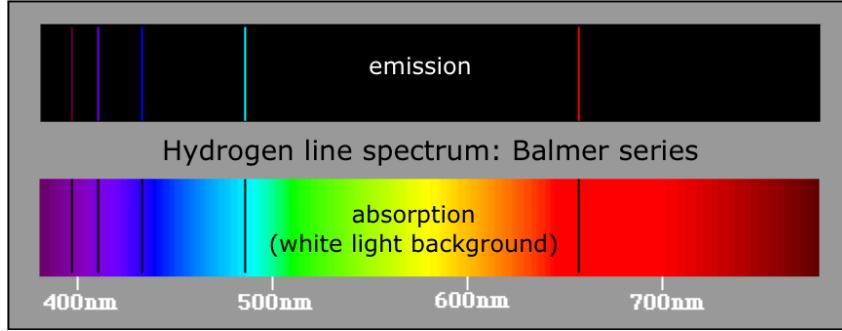


Figure 5: The Balmer series of spectral lines of hydrogen

| Lyman | Balmer | Paschen |
|-------|--------|---------|
| 121.5 | 656.3 | 1874.5 |
| 102.5 | 486.1 | 1281.4 |
| 97.2 | 434.1 | 1093.5 |
| 94.9 | 410.2 | 1004.6 |
| 93.7 | 397.0 | 954.3 |
| 93 | 388.9 | 922.6 |
| 92.6 | | 901.2 |
| 92.3 | | 886.0 |
| 91.9 | | 874.8 |
| | | 866.2 |

Table 2: Wavelengths in nm of some hydrogen electromagnetic emission lines from the Lyman, Balmer and Paschen series.

As discussed in PHYS1011, it was found that the wavelengths of all of these lines could be explained by a formula called the Rydberg formula,

$$\frac{1}{\lambda} = R \left(\frac{1}{q^2} - \frac{1}{p^2} \right), \quad (3.26)$$

where R is called Rydberg's constant ($= 1.097 \times 10^{-2} \text{ nm}^{-1}$), p and q are integers with $p > q$, and λ is the wavelength of the line in the emission/absorption spectrum. The Lyman, Balmer and Paschen series correspond to $q = 1, 2$ and 3 respectively. This formula was derived in the Bohr model with the radical assumption of discrete energy levels in the hydrogen atom, with R being expressed

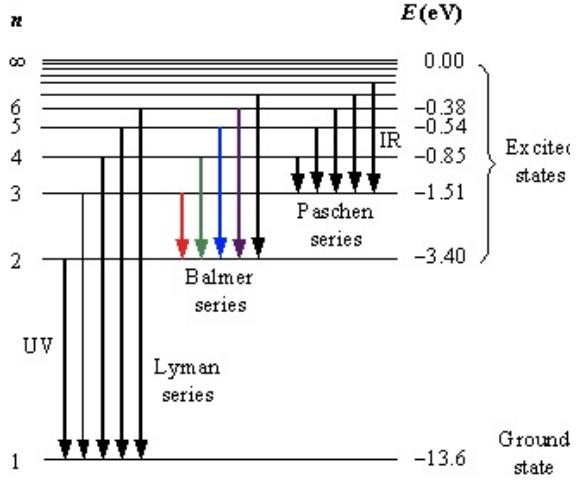


Figure 6: Schematic to explain the origin of the different series observed in the hydrogen emission spectra. The Balmer series has four lines in the visible part of the spectrum.

in terms of fundamental constants. Quantum mechanics obtains the same expression for $1/\lambda$, with the energy levels being eigenvalues of eigenfunctions of energy (obtained by solving the Schrödinger equation). The different series are associated with electrons transferring from all possible higher energy states into a final state which characterises the particular series (Figure 6).

When the emission spectra of hydrogen was measured with higher accuracy (first hints came in 1892, but stronger evidence was provided in 1914 *after* the Bohr model) it was found that some of the lines described by the Rydberg formula were found to consist of multiple lines. This splitting of lines is the fine structure and will be covered in Chapter 5. The fine structure is explained by the presence of additional (sub-dominant) terms in the Hamiltonian, and relativistic effects.

Summary

The solutions to the Schrödinger equation for the bound states of the hydrogen atom are

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r)Y_{lm}(\theta, \phi), \quad (3.27)$$

where $Y_{lm}(\theta, \phi)$ are the spherical harmonic functions and $R_{nl}(r)$ the so-

lutions of the radial equation. Valid quantum numbers are

$$n = 1, 2, 3, \dots \quad (3.28)$$

$$l = 0, 1, \dots, n - 1 \quad (3.29)$$

$$m = -l, -l + 1, \dots, l - 1, l. \quad (3.30)$$

The eigenvalues of the Hamiltonian associated with these eigenfunctions – the energy levels – are

$$E_n = -\frac{\hbar^2}{2\mu a_0^2 n^2}. \quad (3.31)$$

They depend only on the principal quantum number n . As a result, the n^{th} level is n^2 degenerate.

4 Time-Independent Perturbation Theory

Not all quantum mechanical problems can be solved analytically. An equation can always be set up but an analytical solution to the equation is very often not possible for realistic systems. A number of techniques have been developed to address this difficulty. One of these is called the perturbation method. The perturbation method is a so-called approximate analytic technique, meaning that the result is an analytic function which can be manipulated algebraically, but the result is only approximate. This technique is widely used throughout physics and so in this course it will be examinable as a technique not just in the context of atomic physics.

Perturbation theory can be applied to any problem for which a closely related problem can be solved analytically. The idea is that the difference between the two problems can be written down as a small additional term which is added to the Hamiltonian of the analytically solvable problem; this small additional term does not change the solution a great deal, but instead requires only a small correction to the existing eigenvalues and eigenfunctions.

As the mathematics is easier we will start by dealing with a situation where the analytically solvable problem has no degenerate eigenvalues and then return to deal with the degenerate case afterwards.

4.1 Non-degenerate

Let us assume that we have a quantum mechanical system described by the Hamiltonian \hat{H}_o and let the non-degenerate energy eigenvalues be E_j and the eigenfunctions be ψ_j *i.e.*

$$\hat{H}_o\psi_j = E_j\psi_j \quad (4.1)$$

Let there be a small change in the system which does not change the boundary conditions (*i.e.* the perturbation) and let this small change be represented by an additional component to the Hamiltonian \hat{H}_p *i.e.* the perturbed Hamiltonian is given by $\hat{H}_o + \hat{H}_p$. Instead of solving this problem, let us solve a more general problem where the perturbation has a controllable strength *i.e.* let us introduce a parameter, λ , which multiplies the extra part of the Hamiltonian, *i.e.* we are going to solve the problem described by the Hamiltonian $\hat{H}_o + \lambda\hat{H}_p$. At the end of the calculation $\lambda \rightarrow 1$ can be taken to obtain the results for the full strength perturbation \hat{H}_p . Let the new eigenvalues and eigenfunctions be given by E'_j and ψ'_j *i.e.*

$$[\hat{H}_o + \lambda\hat{H}_p] \psi'_j = E'_j \psi'_j \quad (4.2)$$

Now if $\lambda \rightarrow 0$ then the solutions should tend towards that of the original prob-

lem. For *small* perturbations \hat{H}_p (which is multiplied λ) we would expect that the eigenvalues and eigenfunctions only change a little, and so it should be possible to write down the new eigenvalues and eigenfunctions as a power series in λ , *i.e.*

$$E'_j = E_j + \lambda E_j^{(1)} + \lambda^2 E_j^{(2)} + \lambda^3 E_j^{(3)} + \dots \quad (4.3)$$

where the first term, E_j , in the summation, is an eigenvalue of the unperturbed Hamiltonian and $E_j^{(1)}$ etc. are numbers which correspond to the first, second \dots n^{th} order corrections to the energy eigenvalues and

$$\psi'_j = \psi_j + \lambda \psi_j^{(1)} + \lambda^2 \psi_j^{(2)} + \lambda^3 \psi_j^{(3)} + \dots \quad (4.4)$$

where the first term, ψ_j , is an eigenfunction of the unperturbed Hamiltonian and $\psi_j^{(1)}$ etc. are *functions* which correspond to the first, second etc. order corrections to the eigenfunctions.

The two mathematical expressions above, are the mathematical encapsulation of the general idea of perturbation theory.

In order to calculate the corrections, we now exploit the completeness relations which tells us that the each of the corrections to the eigenfunctions can be written as a sum of the original eigenfunctions *i.e.*

$$\psi_j^{(1)} = \sum_i a_{i,j}^{(1)} \psi_i, \quad \psi_j^{(2)} = \sum_i a_{i,j}^{(2)} \psi_i \quad \text{etc.} \quad (4.5)$$

It is important to note all the subscripts and superscripts. The bracketed superscript refers to the order of the correction, *i.e.* what power of λ it depends on, the coefficient has two subscripts which refer to the index of the perturbed, j , and unperturbed, i , eigenfunctions.

The eigenenergy equation (4.2) (the time independent Schrödinger equation) is linear like all eigenvalue/vector equations. This means that we can always multiply ψ'_j by a factor and still have a solution. Normally of course we fix this freedom by insisting that ψ'_j be normalised:

$$\langle \psi'_j | \psi'_j \rangle = 1. \quad (4.6)$$

Here it will prove useful to choose a different normalisation. We can then fix up the normalisation at the end of the calculation by multiplying by the appropriate factor, namely $1/\sqrt{\langle \psi'_j | \psi'_j \rangle}$.

The choice of normalisation that will prove convenient for the following steps is such that in the expansion of ψ'_j over the complete set of original eigenfunctions, the coefficient of the original ψ_j stays as 1. You see from (4.4) that ψ_j appears with unit coefficient. Of course this is so that as $\lambda \rightarrow 0$, we recover the original

eigenfunction. However, when expanded in terms of the complete set of original eigenfunctions, we find ψ_j terms also hiding in $\psi_j^{(1)}$, $\psi_j^{(2)}$ etc., as is clear from (4.5). Requiring that the coefficient of ψ_j is one *for all* λ , therefore requires us to choose $a_{j,j}^{(1)} = 0$, $a_{j,j}^{(2)} = 0$, etc. Being able to set these coefficients to zero will help the calculation we now embark upon, proceed smoothly.

We substitute the sums into the eigenvalue equation to obtain,

$$\begin{aligned} & \left[\hat{H}_o + \lambda \hat{H}_p \right] \left(\psi_j + \lambda \sum_i a_{i,j}^{(1)} \psi_i + \lambda^2 \sum_i a_{i,j}^{(2)} \psi_i + \dots \right) = \\ & \left(E_j + \lambda E_j^{(1)} + \lambda^2 E_j^{(2)} + \lambda^3 E_j^{(3)} + \dots \right) \left(\psi_j + \lambda \sum_i a_{i,j}^{(1)} \psi_i + \lambda^2 \sum_i a_{i,j}^{(2)} \psi_i + \dots \right) \end{aligned} \quad (4.7)$$

This equation is then rearranged so that there are terms which are multiplied by the same power of λ , i.e.

$$\begin{aligned} 0 &= \left[\hat{H}_o - E_j \right] \psi_j \\ &\quad + \lambda \left\{ \hat{H}_o \sum_i a_{i,j}^{(1)} \psi_i + \hat{H}_p \psi_j - E_j^{(1)} \psi_j - E_j \sum_i a_{i,j}^{(1)} \psi_i \right\} \\ &\quad + \lambda^2 \left\{ \hat{H}_o \sum_i a_{i,j}^{(2)} \psi_i + \hat{H}_p \sum_i a_{i,j}^{(1)} \psi_i - E_j^{(2)} \psi_j - E_j^{(1)} \sum_i a_{i,j}^{(1)} \psi_i - E_j \sum_i a_{i,j}^{(2)} \psi_i \right\} \\ &\quad \quad \quad + \lambda^3 \{ \dots \} + \dots \end{aligned} \quad (4.8)$$

In order for this equation to be true for all values of λ , we require that each line above be equal to zero. The first line is equal to zero due to the definition of the original eigenvalues and eigenfunctions. The second line gives the equation:

$$\hat{H}_o \sum_i a_{i,j}^{(1)} \psi_i + \hat{H}_p \psi_j - E_j^{(1)} \psi_j - E_j \sum_i a_{i,j}^{(1)} \psi_i = 0 \quad (4.9)$$

This equation can be simplified by using the orthonormality of the original eigenfunctions. The left-hand side is multiplied by the complex conjugate of each of the original eigenfunctions, ψ_k^* , in turn, and integrated over space, leading to:

$$\sum_i a_{i,j}^{(1)} \int \psi_k^* \hat{H}_o \psi_i dV + \int \psi_k^* \hat{H}_p \psi_j dV - E_j^{(1)} \int \psi_k^* \psi_j dV - E_j \sum_i a_{i,j}^{(1)} \int \psi_k^* \psi_i dV = 0 \quad (4.10)$$

The first term can be simplified to give

$$\sum_i a_{i,j}^{(1)} E_i \int \psi_k^* \psi_i dV + \int \psi_k^* \hat{H}_p \psi_j dV - E_j^{(1)} \int \psi_k^* \psi_j dV - E_j \sum_i a_{i,j}^{(1)} \int \psi_k^* \psi_i dV = 0 \quad (4.11)$$

and then by using orthonormality one can write

$$\sum_i a_{i,j}^{(1)} E_i \delta_{k,i} + \int \psi_k^* \hat{H}_p \psi_j dV - E_j^{(1)} \delta_{k,j} - E_j \sum_i a_{i,j}^{(1)} \delta_{k,i} = 0. \quad (4.12)$$

In the case that $k = j$ the first and final terms sum to zero, leading to

$$E_j^{(1)} = \int \psi_j^* \hat{H}_p \psi_j dV \quad (4.13)$$

i.e. the first order correction to the energy of the j eigenvalue is given by the expectation value of the perturbation \hat{H}_p acting on the original eigenfunction ψ_j . In bra-ket notation this would be written as $\langle \psi_j | \hat{H}_p | \psi_j \rangle$, as in eq. (2.1), and is referred to as a *matrix element*.

The first order correction to the eigenfunction, namely $\psi_j^{(1)}$ can be obtained by computing the coefficients $a_{i,j}^{(1)}$ from (4.12). This works for all the coefficients except $a_{j,j}^{(1)}$. To obtain an expression for $a_{j,j}^{(1)}$ one would have to set $k = j$ in (4.12), but (as we have already seen), the first and last terms cancel (giving eq. (4.13) and no information on $a_{j,j}^{(1)}$). Fortunately by our normalisation condition we already know that $a_{j,j}^{(1)} = 0$. For $k \neq j$ in (4.12) one has:

$$a_{k,j}^{(1)} (E_j - E_k) = \int \psi_k^* \hat{H}_p \psi_j dV, \quad (4.14)$$

and thus the first order correction to the j eigenfunction is

$$\psi_j^{(1)} = \sum_{k \neq j} \frac{\int \psi_k^* \hat{H}_p \psi_j dV}{(E_j - E_k)} \psi_k. \quad (4.15)$$

Note that the sum (which is a sum over k) does not have the $k = j$ term because $a_{j,j}^{(1)} = 0$. The presence of $E_j - E_k$ in the denominator is not problematic because there is no $k = j$ term, and we assumed that none of the eigenvalues were degenerate (i.e. none of the E_k can equal E_j). The case of degeneracy requires further consideration, and is covered in section 4.2.

Having obtained the first order corrections $E_j^{(1)}$ and $\psi_j^{(1)}$, the second order correction to the energy eigenvalues can be obtained (with similar mathematical steps) from the third line in eq.(4.8), leading to:

$$E_j^{(2)} = \sum_{k \neq j} \frac{\left| \int \psi_k^* \hat{H}_p \psi_j dV \right|^2}{(E_j - E_k)}. \quad (4.16)$$

Note that $E_j^{(2)}$ is a sum over k (in contrast to $E_j^{(1)}$ in eq. (4.13) which contained a single matrix element) and involves the modulus squared of the matrix elements. This will (hopefully) ensure that $E_j^{(2)}$ is smaller than $E_j^{(1)}$ for a sufficiently small perturbation \hat{H}_p .

4.2 1D Infinite Potential Well

To illustrate the above, we will do a simple calculation. Our starting system will be a particle in an infinitely deep well of width $2a$ from $x = 0$ to $2a$ (see Quantum Physics Section 2.1). The unperturbed eigenvalues are thus given by

$$E_n = \frac{\hbar^2 \pi^2 n^2}{8ma^2}, \quad (4.17)$$

where n is a positive integer. The associated eigenfunctions are

$$\psi_n(x) = \frac{1}{\sqrt{a}} \sin\left(\frac{n\pi x}{2a}\right). \quad (4.18)$$

The perturbation we will use is to increase the potential from zero in the range $x = a - b$ to $a + b$ to a value V (where of course $b < a$). In this case the first order correction to the energy eigenvalues is given by

$$E_n^{(1)} = \frac{1}{a} \int_{a-b}^{a+b} V \sin^2\left(\frac{n\pi x}{2a}\right) dx = V \left[\frac{b}{a} - (-1)^n \frac{1}{n\pi} \sin\left(\frac{n\pi b}{a}\right) \right] \quad (4.19)$$

The first order correction to the eigenfunctions ψ_j^1 require us to calculate the matrix elements involving the perturbation Hamiltonian and the original eigenfunctions (see eq. 4.15), which are given by

$$\frac{V}{a} \int_{a-b}^{a+b} \sin\left(\frac{n_1\pi x}{2a}\right) \sin\left(\frac{n_2\pi x}{2a}\right) dx = 0 \quad (4.20)$$

unless $n_1 + n_2$ is even, when

$$\begin{aligned} & \frac{V}{a} \int_{a-b}^{a+b} \sin\left(\frac{n_1\pi x}{2a}\right) \sin\left(\frac{n_2\pi x}{2a}\right) dx = \\ & -\frac{2V}{\pi} \left[\frac{(-1)^{\frac{n_1+n_2}{2}}}{n_1+n_2} \sin\left(\frac{(n_1+n_2)\pi b}{2a}\right) - \frac{(-1)^{\frac{n_1-n_2}{2}}}{n_1-n_2} \sin\left(\frac{(n_1-n_2)\pi b}{2a}\right) \right]. \end{aligned} \quad (4.21)$$

To evaluate the integral, split the product of the sine functions into a difference of cosines, and then integrate each cosine separately. Note the following result:

$$\int_{a-b}^{a+b} \cos\left[\frac{(n_1-n_2)\pi x}{2a}\right] dx = \frac{4a}{(n_1-n_2)\pi} \cos\left[\frac{(n_1-n_2)\pi}{2}\right] \sin\left[\frac{(n_1-n_2)\pi b}{2a}\right], \quad (4.22)$$

and note that the result vanishes if $(n_1 - n_2)$ is odd.

If V is sufficiently small (and so the series expansion converges), then these expressions will be close to the true energies and eigenfunctions of the perturbed

1D infinite potential well. Calculating the second order corrections would give a better approximation.

The interesting point to note is the first equation which is zero. This is because the original system has reflection symmetry in $x = a$. This means that the original eigenfunctions are either symmetric or anti-symmetric under the same reflection depending on whether the principal quantum number n is odd or even respectively. In addition, the perturbation has the same symmetry. It is therefore clear that the eigenfunctions of the perturbed system must have the same symmetry as the original system, which means that they cannot be made up of mixtures of the original eigenfunctions containing both symmetric and anti-symmetric eigenfunctions.

4.3 2D Infinite Potential Well

Let us now try and solve a remarkably similar problem to that just discussed but in this case one which has degenerate eigenfunctions for the unperturbed system. As the unperturbed system we will use the 2D square infinite potential well, *i.e.* a system in which the potential is infinite everywhere except in the square region $0 \leq x \leq 2a$, $0 \leq y \leq 2a$. The eigenfunctions for this problem are given by

$$\psi_{n_x, n_y}(x, y) = \frac{1}{a} \sin\left(\frac{n_x \pi x}{2a}\right) \sin\left(\frac{n_y \pi y}{2a}\right), \quad (4.23)$$

which for brevity we refer to as state $|n_x, n_y\rangle$, and the associated eigenvalues are

$$E = \frac{\hbar^2 \pi^2}{8ma^2} (n_x^2 + n_y^2). \quad (4.24)$$

Whilst the ground state of this system is not degenerate, the first excited energy is doubly degenerate.

For the perturbation we consider a small potential spike situated at $(a/2, a/2)$. This will be represented by a Dirac delta function (as in the delta function potential of section 2.5 in PHYS2003) *i.e.*

$$\hat{H}_p = Va^2 \delta(x - a/2) \delta(y - a/2). \quad (4.25)$$

This particular form is chosen to make the integrations below simple.

One can proceed by using the non-degenerate perturbation described up to now and write down the first order perturbations to the energies of the eigenfunctions as

$$E_{1,1}^{(1)} = \frac{V}{4}, E_{1,2}^{(1)} = E_{2,1}^{(1)} = \frac{V}{2}, \dots \quad (4.26)$$

However, we will see shortly that whilst the formula for $E_{1,1}^{(1)}$ (which is not

degenerate) is correct, the formula for the degenerate energy levels $E_{1,2}^{(1)} = E_{2,1}^{(1)}$ is **not correct**.

We could also write down the first order perturbation for the eigenfunctions for the ground state, but when we come to the first order perturbation for the eigenfunction for states $|1, 2\rangle$ and $|2, 1\rangle$ we encounter a problem. The problem is that the matrix elements mixing these two states and the perturbation are non-zero *i.e.*

$$\iint \frac{1}{a} \sin\left(\frac{\pi x}{2a}\right) \sin\left(\frac{\pi y}{a}\right) V a^2 \delta(x-a/2) \delta(y-a/2) \frac{1}{a} \sin\left(\frac{\pi x}{a}\right) \sin\left(\frac{\pi y}{2a}\right) dx dy = \frac{V}{2} \quad (4.27)$$

The coefficient of the term which corresponds to this matrix element in the first order wave function perturbation is

$$\frac{\int \psi_{1,2}^* \hat{H}_p \psi_{2,1} dx dy}{(E_{1,2} - E_{2,1})}, \quad (4.28)$$

which is infinite because $E_{1,2} = E_{2,1}$. This does not just mean that the first order perturbation of the wave functions cannot be calculated, but instead it means that the whole calculation has failed because all higher order terms will also have the same divergent (infinite) terms.

It turns out that we can do the calculation by exploiting the fact that we make an arbitrary choice when we write down the eigenfunctions for a degenerate eigenvalue. This is because any linear combination of the wave functions we write down is also an eigenfunction with the same eigenvalue. What we have to do to fix the perturbation calculation is choose the correct sets of degenerate eigenfunctions so that all of the matrix elements involving the perturbation and two different eigenfunctions from the degenerate set are always zero.

Let us attempt to do this for the two eigenfunctions $|1, 2\rangle$ and $|2, 1\rangle$. We wish to generate two linear combinations of these eigenfunctions, which are orthogonal to each other and give a zero matrix element when combined with the perturbation. We will write these wave functions as $\tilde{\psi}_1 = \alpha \psi_{1,2} + \beta \psi_{2,1}$ and $\tilde{\psi}_2 = \gamma \psi_{1,2} + \delta \psi_{2,1}$. The matrix element involving these two and the perturbation is given by

$$\begin{aligned} & \int \tilde{\psi}_1^* \hat{H}_p \tilde{\psi}_2 dx dy \\ &= \alpha^* \gamma \int \psi_{1,2}^* \hat{H}_p \psi_{1,2} dx dy + \alpha^* \delta \int \psi_{1,2}^* \hat{H}_p \psi_{2,1} dx dy \\ & \quad + \beta^* \gamma \int \psi_{2,1}^* \hat{H}_p \psi_{1,2} dx dy + \beta^* \delta \int \psi_{2,1}^* \hat{H}_p \psi_{2,1} dx dy \\ &= \frac{V}{2} (\alpha^* \gamma + \alpha^* \delta + \beta^* \gamma + \beta^* \delta). \end{aligned} \quad (4.29)$$

Another way of writing this is:

$$\int \tilde{\psi}_1^* \hat{H}_p \tilde{\psi}_2 dx dy = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \begin{pmatrix} V/2 & V/2 \\ V/2 & V/2 \end{pmatrix} \begin{pmatrix} \gamma \\ \delta \end{pmatrix}. \quad (4.30)$$

where the matrix is a representation of the perturbation. It can be seen by inspection that the choice $\alpha = \beta = \gamma = 1/\sqrt{2}$ and $\delta = -1/\sqrt{2}$ causes the matrix element to vanish (linear algebra techniques can be used to derive these values for this particular scenario and for more complicated scenarios with more than two degenerate eigenfunctions). Therefore, the new degenerate eigenfunctions for the unperturbed system we will use for the calculation are $\tilde{\psi}_1 = 1/\sqrt{2} (\psi_{1,2} + \psi_{2,1})$ and $\tilde{\psi}_2 = 1/\sqrt{2} (\psi_{1,2} - \psi_{2,1})$. Using a perturbation with these wave functions gives first order energy corrections

$$E_{1,1}^{(1)} = \frac{V}{4}, \quad E_{(1,2)+(2,1)}^{(1)} = V, \quad E_{(1,2)-(2,1)}^{(1)} = 0, \dots \quad (4.31)$$

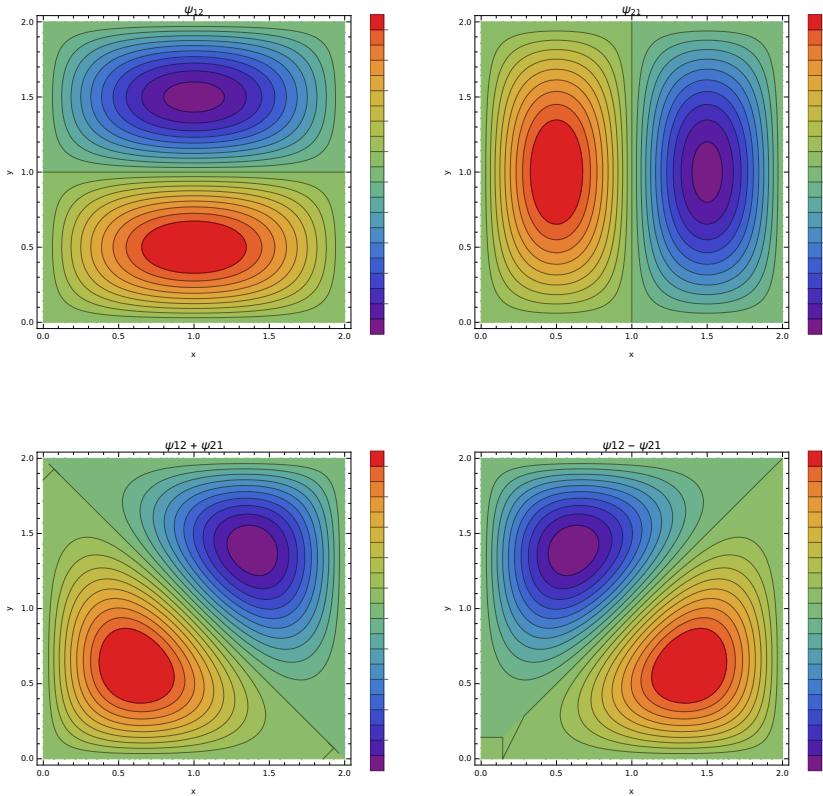


Figure 7: Unperturbed eigenfunctions for $\psi_{1,2}$ and $\psi_{2,1}$ (top row), and the ones modified via the perturbation (bottom row).

We should note that the effect of the perturbation is to break the degeneracy

of the first two excited states. It is this that leads to the breakdown of perturbation theory if we use the wrong unperturbed wave functions. This is because whilst for the unperturbed system all linear combinations of two degenerate wave functions are eigenstates, this is not true for even the smallest perturbation. In this case there is only one combination which are eigenfunctions. If we attempt the solution without first choosing the correct basis wave functions, as soon as the perturbation turns on, the perturbation cause a sudden shift to the correct wave functions which does not agree with the spirit of small changes built into perturbation theory.

Note that the presence of the delta-function spike reduces the symmetry of the system: previously the square could be rotated by 0, 90, 180, and 270 degrees and the system would be the same, but this is no longer true with the presence of the spike. This is reflected in the unperturbed values for $\psi_{1,2}$ and $\psi_{2,1}$. The perturbed wavefunctions will not be symmetric any more if the perturbation is not.

4.4 Stark Effect

Now let us apply the technique of time-independent perturbation theory to an atomic physics problem. What happens if a small electric field is applied to a hydrogen atom? Classically, applying such an electric field causes a decrease in the energy of the atom, since the atom behaves like a dipole in the electric field, and the energy decreases, $\Delta E = -\vec{d} \cdot \vec{E}_{\text{ext}}$, where \vec{d} is the dipole moment.

Quantum mechanically, our unperturbed system is the hydrogen atom model that has already been solved in Chapter 3. The perturbation is an electric field of magnitude E pointing in a specific direction. This direction can be chosen as the z direction without loss of generality. The perturbation term is given by

$$\hat{H}_p = -eEz = -eEr \cos \theta, \quad (4.32)$$

where we have changed from Cartesian to spherical coordinates. Here e is the charge on the electron, and this expression for \hat{H}_p is the classical interaction energy between an electric field and a hydrogen atom (a proton-electron system), with z being the z coordinate of the electron.

Let us consider the effect of this perturbation on the ground state of the hydrogen atom, which is given by the eigenfunction $\psi_{1,0,0} = R_{1,0}Y_{0,0}$ i.e. $\psi_{nlm} = R_{n,l}Y_{l,m}$ is the product of the angular and radial wave functions for $n = 1$, $l = 0$, and $m = 0$ that were derived in Chapters 2 and 3:

$$\psi_{1,0,0} = (\pi a_0^3)^{-1/2} e^{-r/a_0}. \quad (4.33)$$

Since $Y_{0,0} = 1/\sqrt{4\pi}$ there is no angular (θ, ϕ) dependence in $\psi_{1,0,0}$. The first order correction to the energy of this state is given by

$$E_{1,0,0}^{(1)} = \int \psi_{1,0,0}^* \hat{H}_p \psi_{1,0,0} dV = \int dV \frac{e^{-r/a_0}}{(\pi a_0^3)^{1/2}} (-eEr \cos \theta) \frac{e^{-r/a_0}}{(\pi a_0^3)^{1/2}} \quad (4.34)$$

Here dV is the usual volume element in spherical coordinates ($dV = r^2 \sin \theta d\theta d\phi dr$). This matrix element is zero. We can show this to be true because for every point included in the integral there is a twin point which is the opposite direction but the same distance from the origin, *i.e.* the points given by \mathbf{r} and $-\mathbf{r}$ (this corresponds to θ and $\pi - \theta$ respectively. The wave function $\psi_{1,0,0}$ is the same at these two positions, but the perturbation term has the same magnitude but different sign, adding to give zero and thus the whole integral is zero.

The second order correction to the energy of this state is given by

$$E_{1,0,0}^{(2)} = \sum_{k \neq (1,0,0)} \frac{\left| \int \psi_k^* \hat{H}_p \psi_{1,0,0} dV \right|^2}{(E_{1,0,0} - E_k)}. \quad (4.35)$$

It turns out that most of the matrix elements in the sum are zero, but the matrix elements involving ψ_k , with $k = (n, 1, 0)$, are non-zero. Evaluating the sum requires a somewhat lengthy calculation, and gives $E_{1,0,0}^{(2)} = -0.84 \frac{8}{3} a_0^3 E^2$. However, one can see by inspection of eq. (4.35) that the energy is proportional to the magnitude of the electric field squared (because E is a constant in \hat{H}_p and can be taken out of the sum). Moreover, the contribution to the energy is negative (in agreement in sign with the expected shift in classical physics) because the numerators are positive definite and the denominators are negative definite due to $E_{1,0,0} < E_k$.

Let us now consider the effect on the degenerate $n=2$ states. Clearly, we are going to need degenerate perturbation theory, and the first thing we need to do is decide on the correct basis wave functions. The conventional wave functions are written explicitly as follows (where $\psi_{200} = R_{2,0} Y_{0,0}$, $\psi_{210} = R_{2,1} Y_{1,0}$, $\psi_{211} = R_{2,1} Y_{1,1}$, $\psi_{21-1} = R_{2,1} Y_{1,-1}$):

$$\psi_{200}(r, \theta, \phi) = \frac{1}{\sqrt{32\pi a_0^3}} \left(2 - \frac{r}{a_0} \right) e^{-r/2a_0}, \quad (4.36)$$

$$\psi_{210}(r, \theta, \phi) = \frac{1}{\sqrt{32\pi a_0^3}} \frac{r}{a_0} \cos \theta e^{-r/2a_0}, \quad (4.37)$$

$$\psi_{21\pm 1}(r, \theta, \phi) = \mp \frac{1}{\sqrt{64\pi a_0^3}} \frac{r}{a_0} \sin \theta e^{\pm i\phi} e^{-r/2a_0}. \quad (4.38)$$

It turns out that the only matrix elements which are non zero are

$$\int \psi_{2,1,0}^* \hat{H}_p \psi_{2,0,0} dV = \int \psi_{2,0,0}^* \hat{H}_p \psi_{2,1,0} dV = \Delta = 3ea_0 E. \quad (4.39)$$

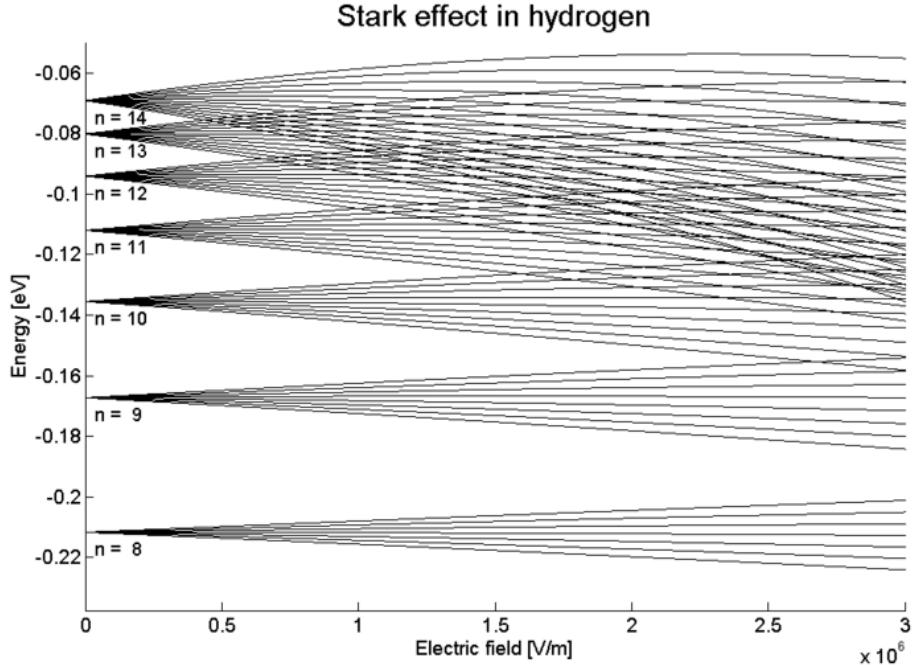


Figure 8: Experimental measurements of the Stark effect in hydrogen.

Thus, the matrix representation of the perturbation for this set of degenerate states is

$$\begin{pmatrix} 0 & \Delta & 0 & 0 \\ \Delta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (4.40)$$

And the set of unperturbed states we require are

$$\frac{1}{\sqrt{2}} (\psi_{2,0,0} + \psi_{2,1,0}), \quad \frac{1}{\sqrt{2}} (\psi_{2,0,0} - \psi_{2,1,0}), \quad \psi_{2,1,1} \quad \text{and} \quad \psi_{2,1,-1}. \quad (4.41)$$

In this case the first order corrections to the energy of the states are

$$3ea_0E, \quad -3ea_0E, \quad 0, \quad 0. \quad (4.42)$$

The shifts in energy calculated above are observed in experiments and are called the quadratic and linear Stark shifts.

Summary

Perturbation theory is a technique to obtain approximate solutions to the Schrödinger equation (for a perturbed system) from a known exact solution (for an unperturbed system). It gives reliable results if the series converges rapidly, and sometimes the first-order term alone is sufficient for a good description of the perturbed system. For degenerate energy levels it is necessary to work in a basis (good states) in which the matrix elements between two degenerate states is zero. Often it is possible to guess what this basis ought to be. Otherwise, it is necessary to use linear algebra.

5 The Fine Structure of the Hydrogen Atom

Whilst the predictions of the quantum model of hydrogen in chapter 3 (the *gross structure*) are a very good approximation to reality, it turns out that in high resolution spectra of hydrogen these predicted lines are in fact split into sets of lines. This is the so-called *fine structure* of hydrogen and means that some terms in Hamiltonian must have been omitted from the model that was studied in chapter 3.

The Schrödinger equation can be thought of as the quantum equivalent to Newton's equation of motion in the sense that it is non-relativistic (i.e. kinetic energy is given by a quantum equivalent of $p^2/2m$ rather than starting from the relativistic expression $E^2 = p^2c^2 + m^2c^4$). Just as with Newton's equations in the classical world, non-relativistic quantum mechanics is a good approximation in the quantum world under many circumstances. However, it is known to fail under other circumstances. The extension of quantum mechanics to make it relativistic was made by Dirac (in 1928) who replaced the Schrödinger equation with a relativistic version called the *Dirac equation*. The Dirac equation is not studied in this course. In the small velocity limit the Dirac equation tends towards the Schrödinger equation and we will treat the difference between the Dirac and Schrödinger equations as a series of perturbations to the Schrödinger equation. For historical reasons the different perturbations have been named and are studied in this chapter.

5.1 Spin-Orbit Coupling

In 1925 Uhlenbeck and Goudsmidt proposed an explanation for the fine structure of hydrogen and the Stern-Gerlach experiment. They proposed that the electron has an intrinsic angular momentum, *i.e.* in classical terms, an electron is spinning around an axis through their centre, but this picture cannot be carried across to the electron. Spin is a physical explanation of the *fourth quantum number* that was introduced by Pauli. Many experiments on atoms could not be explained without the concept of spin. The spin of the electron is $s = \hbar/2$, and a component of the spin of an electron measured in a specific direction can take the values $\pm\hbar/2$. Associated with this spin is a magnetic moment $\boldsymbol{\mu}$ (not to be confused with the earlier symbol μ for reduced mass), which is a *current loop times area enclosed by the current*:

$$\boldsymbol{\mu} = -\frac{ge}{2m}\mathbf{S} \quad (5.1)$$

where $g = 2.002319\dots$ has been measured very precisely, but we will take $g = 2$ in this course. It is known that electric fields and magnetic fields are related (through Lorentz transformations), and what is perceived as an electric field in one frame of reference will be observed in another frame to have a component of a magnetic field. In the rest frame of the nucleus the nucleus generates only

an electric field, but in the rest frame of an electron passing the nucleus the electron also experiences a magnetic field given by

$$\mathbf{B} = \frac{\mathbf{E} \times \mathbf{v}}{c^2}, \quad (5.2)$$

where \mathbf{v} is the velocity of the electron. The electric field due to the nucleus is the standard Coulomb interaction and is given by:

$$\mathbf{E} = \frac{e \mathbf{r}}{4\pi\epsilon_0 r^3}. \quad (5.3)$$

Using this expression for \mathbf{E} the magnetic field in the electron's rest frame is:

$$\mathbf{B} = \frac{e \mathbf{r} \times \mathbf{v}}{4\pi\epsilon_0 c^2 r^3} = \frac{e \mathbf{L}}{4\pi\epsilon_0 m c^2 r^3}. \quad (5.4)$$

Here \mathbf{L} is the orbital angular momentum. It is known that a magnetic moment in the presence of a magnetic field has an energy given by $-\boldsymbol{\mu} \cdot \mathbf{B}$ which is given by

$$-\boldsymbol{\mu} \cdot \mathbf{B} = \frac{ge^2}{8\pi\epsilon_0 m^2 c^2 r^3} \mathbf{L} \cdot \mathbf{S}. \quad (5.5)$$

This is a classical expression for the energy correction and can be converted into a quantum counterpart by replacing \mathbf{L} and \mathbf{S} with operators $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ in the usual way. This extra term in the Hamiltonian is the *spin-orbit* term \hat{H}_{so} can be treated as a quantum mechanical perturbation. It turns out that the above calculation requires multiplication by 1/2 in order to take into account the fact that the orbiting (and hence accelerating) electron is not an inertial frame. The correction for this is called the Thomas precession correction and the corrected energy is:

$$\hat{H}_{so} = \frac{ge^2}{16\pi\epsilon_0 m^2 c^2 r^3} \hat{\mathbf{L}} \cdot \hat{\mathbf{S}}. \quad (5.6)$$

We now need to determine the effect of this perturbation on the hydrogen atom states, and this requires finding the correct basis of eigenfunctions to use in perturbation theory (*good basis*). Note that \hat{H}_{so} commutes with $\hat{\mathbf{L}}^2$, $\hat{\mathbf{J}}^2$, and with J_z . Using the expression for total angular momentum which is given by

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}, \quad (5.7)$$

we can rewrite the perturbation as

$$\hat{H}_p = \frac{ge^2}{32\pi\epsilon_0 m^2 c^2 r^3} (\hat{J}^2 - \hat{L}^2 - \hat{S}^2). \quad (5.8)$$

This perturbation commutes with the operators \hat{J}^2 , \hat{L}^2 , \hat{S}^2 and \hat{J}_z and thus the new basis states are eigenstates of these operators. The first-order correction to the energies of these new states is given by

$$E_{n,j,l,s}^{(1)} = \frac{\hbar^2 g e^2}{32\pi\varepsilon_0 m^2 c^2} \frac{1}{(na_0)^3 l(l+\frac{1}{2})(l+1)} \{j(j+1) - l(l+1) - s(s+1)\}. \quad (5.9)$$

This is often written as

$$E_{n,j,l,s}^{(1)} = -\frac{\alpha^2 g E_n^{(0)}}{4nl(l+\frac{1}{2})(l+1)} \{j(j+1) - l(l+1) - s(s+1)\}, \quad (5.10)$$

where $g \approx 2$, and α is the so-called fine structure constant, which is given by

$$\alpha = \frac{e^2}{4\pi\varepsilon_0 \hbar c} \approx \frac{1}{137}. \quad (5.11)$$

There is no spin-orbit interaction for $l = 0$ and so the above formula for $E_{n,j,l,s}^{(1)}$ only applies to $l > 0$. Note that there is one more quantum number m_j that labels the state, where $\hbar m_j$ is the eigenvalue of \hat{J}_z . However the first-order correction to the energies (and even the exact answer) does not depend on m_j .

5.2 Addition of Angular Momenta

To work out $E_{n,j,l,s}^{(1)}$ for the spin-orbit interaction it is necessary to know the allowed values of the quantum number j . The total angular momentum $\hat{\mathbf{J}}$ has been defined as $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$. Working out the possible values of j and m_j for given l and s is an example of the general problem of addition of angular momenta. This is covered in detail in PHYS6003 (Advanced Quantum Physics) and in this course we only state the result.

In the general case where two different forms of angular momentum are added together, say $\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2$, the rule is that the quantum number j associated with \hat{J}^2 , can take the values from the sum of the quantum numbers associated with \hat{J}_1^2 and \hat{J}_2^2 , which is $j_1 + j_2$, down to the modulus of the difference of these two quantum numbers ($|j_1 - j_2|$) in integer steps, *i.e.* the allowed values of j are

$$|j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2 - 1, j_1 + j_2. \quad (5.12)$$

For each of these possible values of j , the quantum number m_j of the operator \hat{J}_z (z component of angular momentum) can take a value from $-j$ to j in integer steps, *i.e.* the operator $\hat{J}_z = \hat{J}_{1z} + \hat{J}_{2z}$ has quantum number m_j that takes its values from the list:

$$-j, -j + 1, \dots, j - 1, j. \quad (5.13)$$

For the case of $l = 1$ and $s = 1/2$ (*i.e.* $j_1 = 1$ and $j_2 = 1/2$), applying the above rule gives $j = 3/2$ and $1/2$. For a general value of l together with $s = 1/2$ one has $j = l \pm 1/2$. These two values are then used in eq. (5.10) to give the magnitude of $E_{n,j,l,s}^{(1)}$ from the spin-orbit interaction.

5.3 Term Symbol

In general, in order to describe the state of hydrogen involving spin-orbit coupling we need to define n , j , l , s (and m_j). This information (except m_j) is commonly represented as the so-called **Term Symbol**. This details the s , j and l values of the state in the form

$$^{2s+1}l_j, \quad (5.14)$$

where $2s+1$, called the multiplicity, is given as a number. The l value is given by the standard letters associated with the different angular momentum, namely s, p, d etc. The j value is given as a number. This is normally preceded by the value of the principal quantum number, i.e. an electron in a $2p$ orbital could be in the states

$$2^2p_{\frac{3}{2}} \quad \text{or} \quad 2^2p_{\frac{1}{2}}. \quad (5.15)$$

5.4 Relativistic Mass Correction

The second correction comes from the fact that in the Schrödinger equation the kinetic energy is given by $p^2/2m$ (more precisely it is $p^2/2\mu$, but the mass m of the electron is very close in value to the reduced mass μ), whereas in Relativity the kinetic energy is given by

$$\sqrt{p^2c^2 + m^2c^4} - mc^2 = \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \dots \quad (5.16)$$

The second term in the expansion of the kinetic energy is called the relativistic mass correction. It can be shown that the first order energy shifts associated with this perturbation are given by

$$E_{n,j,l,s}^{(1)} = -\frac{\alpha^2}{n^2} \left[\frac{3}{4} - \frac{n}{l + \frac{1}{2}} \right] E_n^{(0)} \quad (5.17)$$

where α is once again the fine structure constant.

5.5 Darwin Correction

The third correction has no classical analogue and is due to the highly non-classical dynamics of electrons. The Coulomb potential is modified by the fact that an electron is not at a definite point. This correction only affects the $l = 0$ states (i.e. when the spin-orbit interaction is zero). To first order it gives a correction to the energy of $l = 0$ states of

$$E_{n,j,0,s}^{(1)} = -\frac{\alpha^2}{n} E_n^{(0)}. \quad (5.18)$$

5.6 Sum of Fine Structure Corrections

Summing together the three fine-structure corrections discussed above (taking $g = 2$), which are all of order $\alpha^2 E_n^{(0)}$, gives an overall first-order energy shift of

$$E_{n,j,l,s}^{(1)} = -\frac{\alpha^2}{n^2} \left[\frac{3}{4} - \frac{n}{j + \frac{1}{2}} \right] E_n^{(0)}. \quad (5.19)$$

5.7 Other effects

It turns out that even the Dirac equation does not fully describe the hydrogen atom, as it treats the electric field as a classical field (which is not correct). When the electric field is treated correctly using quantum electrodynamics (i.e. the field is quantised) it turns out that states with the same j but different l are in fact not degenerate. The energy difference is however extremely small and is called the **Lamb shift**. This shift was first observed in 1947 for the $2p_{1/2}$ and $2s_{1/2}$ states of hydrogen, which both have $j = 1/2$ but $l = 1$ and 0 respectively.

6 The Interaction of Atoms and Electromagnetic Waves

The interaction of atoms with electromagnetic radiation is a key topic in atomic physics. Not only has a great deal been learnt about atoms using spectroscopy but many of the modern applications of atomic physics, such as lasers, are dependent on light atom interactions.

6.1 Time-Dependent Perturbation Theory

Whilst we can do a lot with static perturbations clearly we are also interested in perturbations which change in time, *e.g.* an oscillatory electric field due to light. In this case we need to use the full time dependent, rather than time independent, Schrödinger equation.

We will start by discussing the solution to the unperturbed problem and then extend this discussion to cover the perturbed system. We will assume that the time-independent Schrödinger equation for the system in the absence of the perturbation has been solved to obtain the eigenvalues, E_n , and eigenfunctions, $\psi_n(\mathbf{r})$.

Recall from sec. 2.9 that we can write the general solution of the time dependent Schrödinger equation in the form

$$\psi(\mathbf{r}, t) = \sum_n a_n^{(0)} \psi_n(\mathbf{r}) e^{-i \frac{E_n t}{\hbar}}. \quad (6.1)$$

If we want to calculate the time dependence for a specific wave function, which is known at some time, which we can set as the zero of time, then we only have to determine the amplitude coefficients, a_n , using the orthonormality of the eigenfunctions.

Now that we have solved the unperturbed system, we introduce the time dependent perturbation $\lambda \hat{H}_p(t)$. Once again, we will seek a solution in the form of a series expansion of the eigenfunctions, *i.e.*

$$\psi'_j = \psi_j + \lambda \psi_j^{(1)} + \lambda^2 \psi_j^{(2)} + \lambda^3 \psi_j^{(3)} + \dots. \quad (6.2)$$

which, as each of the wave functions in the sum on the right hand side can be expanded in terms of the original eigenfunctions, can also be written in the form,

$$\psi'(\mathbf{r}, t) = \sum_n \left(a_n^{(0)}(t) + \lambda a_n^{(1)}(t) + \lambda^2 a_n^{(2)}(t) + \dots \right) \psi_n(\mathbf{r}) e^{-i \frac{E_n t}{\hbar}}. \quad (6.3)$$

If we substitute this into the time dependent Schrödinger equation

$$\left[\hat{H} + \lambda \hat{H}_p(t) \right] \psi'(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \psi'(\mathbf{r}, t). \quad (6.4)$$

and collect the terms in different powers of λ (as we did in the time independent perturbation theory), we find that the zeroth order terms give that

$$\frac{da_n^{(0)}}{dt} = 0. \quad (6.5)$$

And the first order terms give

$$\frac{da_n^{(1)}}{dt} = -\frac{i}{\hbar} \sum_j a_j^{(0)} \int \psi_j^*(\mathbf{r}) \hat{H}_p(t) \psi_j(\mathbf{r}) dV \exp\left(-i \frac{(E_j - E_n)t}{\hbar}\right). \quad (6.6)$$

It is obviously possible to obtain the higher order terms but often they are not required.

6.2 Einstein A and B Coefficients

One of the most important contributions to our understanding of the interaction between atoms and electromagnetic waves was made by Einstein in 1916, just three years after the Bohr model was proposed. At the time it was known that atoms could absorb electromagnetic (EM) radiation and that the amount of energy absorbed by a specific transition was proportional to the energy density of the electromagnetic radiation at the frequency of the transition $\rho(\omega)$. It was also known that, once excited, atoms emitted radiation at a rate which is proportional to the number of excited atoms. Einstein's contribution was to apply simple thermodynamic principles to a very revealing thought experiment. He showed that there must be a third way that EM radiation could interact with atoms; this is *stimulated emission* in which incoming EM radiation causes an atom to fall from a higher energy level to a lower energy level. Einstein's derivation is described in what follows.

Consider a group of identical atoms with two energy levels, E_1 and E_2 where $E_1 < E_2$, which is placed inside a black body, *i.e.* a region where the energy density of the electromagnetic spectrum obeys Planck's law.

Planck's law states (where ω is the angular frequency of the radiation):

$$\rho(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3 [\exp(\hbar\omega/k_B T) - 1]}. \quad (6.7)$$

Einstein reasoned that the number N_1 of atoms in the ground state energy level E_1 must be maintained in thermal equilibrium with the population N_2 of atoms in the excited energy level E_2 , through emission and absorption of photons with the appropriate frequency ω :

$$E_2 - E_1 = \Delta E = \hbar\omega. \quad (6.8)$$

This situation is illustrated in Fig.9, where α_{\uparrow} and α_{\downarrow} are the rates at which an atom makes a transition to a higher-energy state and lower-energy state respectively (to be clarified below). When equilibrium is reached the number of atoms in each energy level is constant (and so is the number of photons). Let $N_2(N_1)$ be the number of atoms in the energy level $E_2(E_1)$. One has the following equation for the rate of change of N_2 :

$$\frac{dN_2}{dt} = -N_2 A_{21} - N_2 B_{21}\rho(\omega_0) + N_1 B_{12}\rho(\omega_0). \quad (6.9)$$

The first term on the right-hand side is due to spontaneous emission causing E_2 to E_1 transitions, and is parametrised by A_{21} (Einstein A parameter). It causes a decrease in N_2 and hence the negative sign. The second term on the right-hand side is due to stimulated emission (E_2 to E_1 transitions and so also decreases N_2). It is parametrised by B_{21} (Einstein B parameter) and the energy density of the EM radiation that can cause the transition, $\rho(\omega_0)$, where ω_0 is the *transition frequency* defined by

$$\omega_0 = (E_2 - E_1)/\hbar. \quad (6.10)$$

The third term on the right-hand side is due to absorption and is parametrised by B_{12} (Einstein B parameter) and the energy density of the EM radiation that can cause the transition. This mechanism is causing a transition from E_1 to E_2 and so increases N_2 (hence the positive sign). In equilibrium one has $dN_2/dt = 0$ and so eq. 6.9 becomes:

$$-N_2 A_{21} - N_2 B_{21}\rho(\omega_0) + N_1 B_{12}\rho(\omega_0) = 0. \quad (6.11)$$

This equation can be written in the notation of α_{\uparrow} and α_{\downarrow} :

$$N_1 \alpha_{\uparrow} = N_2 \alpha_{\downarrow}, \quad (6.12)$$

where $\alpha_{\uparrow} = B_{12}\rho(\omega_0)$ and $\alpha_{\downarrow} = A_{21} + B_{21}\rho(\omega_0)$. Eq. 6.9 can be rearranged to give

$$\rho(\omega_0) = \frac{A_{21}}{\frac{N_1 B_{12}}{N_2} - B_{21}}. \quad (6.13)$$

Boltzmann's law from statistical mechanics tells us:

$$\frac{N_1}{N_2} = \frac{\exp(-E_1/k_B T)}{\exp(-E_2/k_B T)} = \exp\left\{\frac{\hbar\omega_0}{k_B T}\right\}. \quad (6.14)$$

Substituting this expression for N_1/N_2 into eq. 6.13 gives

$$\rho(\omega_0) = \frac{A_{21}}{[\exp(\hbar\omega_0/k_B T)B_{12}] - B_{21}}. \quad (6.15)$$

Comparing the expression for $\rho(\omega_0)$ in eq. 6.15 with that in eq. 6.7 (applied to ω_0) gives

$$B_{12} = B_{21}, \quad (6.16)$$

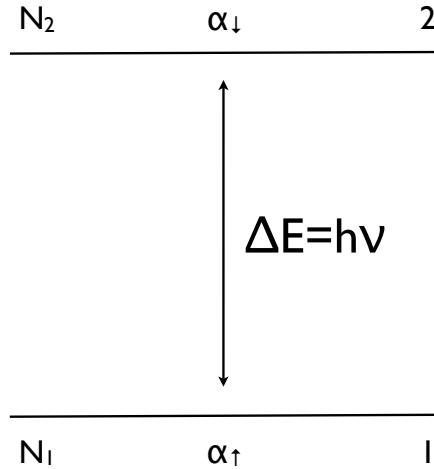


Figure 9: Schematic for populations of a two-level atom with intrinsic rates of emission and absorption.

and

$$\frac{A_{21}}{B_{21}} = \frac{\hbar\omega_0^3}{\pi^2 c^3}. \quad (6.17)$$

Hence the three Einstein coefficients are related and knowledge of one is sufficient to evaluate the other two. Note that the presence of the stimulated emission term B_{21} is necessary in order to reproduce the Planck formula; taking $B_{21} = 0$ would mean that eq. 6.15 is not equivalent to eq. 6.7.

6.3 Fermi's Golden Rule

In order to treat the interaction of light and atoms properly it is necessary to use quantum mechanics for both the atom and the electromagnetic field. However, this is difficult and so it is much more common to use a so-called semi-classical approach, in which the atom is treated with quantum mechanics but the electromagnetic field is treated classically. The disadvantage of this is that whilst absorption and stimulated emission can be modelled this way, we cannot model spontaneous emission. However, at least one of Einstein's coefficients can be predicted and so this can be used to predict the spontaneous emission rate.

A good approximation under most circumstances is that an electromagnetic wave interacts with the atom *via* via its electric field only, and that the electric field is uniform across the extent of the atom. The correction to the energy of an electron in the presence of an electric field is given by $-e\mathbf{r} \cdot \mathbf{E}(t)$. The displacement vector is the displacement between the electron and the proton and

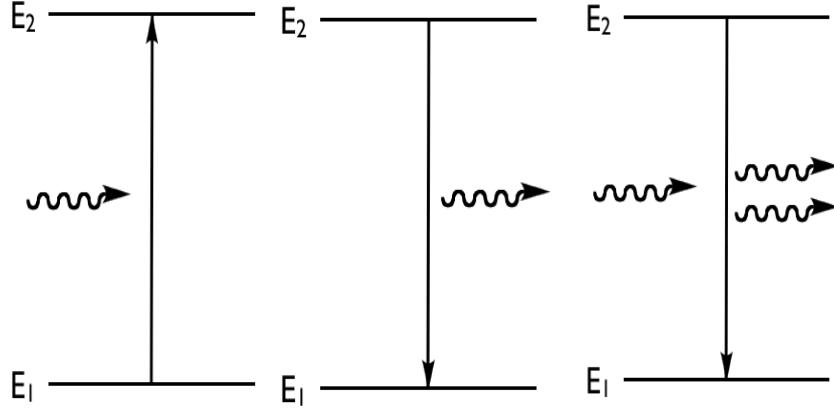


Figure 10: Representation of absorption, spontaneous emission and stimulated emission processes indicating the photons involved and the direction of the electronic transition

so $-e\mathbf{r}$ is the dipole moment of the atom. For this reason, the approximation discussed above is called the *dipole approximation*, often denoted as E1.

The correct time-dependent perturbation within the dipole approximation is thus

$$\hat{H}_p(t) = -e\mathbf{r} \cdot \mathbf{E} \cos \omega t = -e\mathbf{r} \cdot \mathbf{E} \frac{1}{2} (\exp(i\omega t) + \exp(-i\omega t)), \quad (6.18)$$

where the amplitude of \mathbf{E} is constant in time.

We will assume that the atom starts in an eigenstate of the unperturbed system, *i.e.* it is either in the ground state or in an excited state. Let us take this state to be the j^{th} state. In this case all of the zeroth order amplitude coefficients, $a_n^{(0)}$, are zero apart from one, namely $a_j^{(0)} = 1$, and the time dependence of the other coefficients is given to first order in the perturbation by

$$\frac{da_n^{(1)}}{dt} = -\frac{i}{\hbar} \int \psi_n^*(\mathbf{r}) \hat{H}_p(t) \psi_j(\mathbf{r}) dV \exp\left(-i \frac{(E_j - E_n)t}{\hbar}\right). \quad (6.19)$$

Let us define *dipole matrix elements* as

$$\mathbf{D}_{nj} = \int \psi_n^*(\mathbf{r}) [-e\mathbf{r}] \psi_j(\mathbf{r}) dV. \quad (6.20)$$

Then the above expression above can be rewritten as

$$\frac{da_n^{(1)}}{dt} = -\frac{i\mathbf{D}_{nj} \cdot \mathbf{E}}{2\hbar} (\exp(i\omega t) + \exp(-i\omega t)) \exp\left(-i \frac{(E_j - E_n)t}{\hbar}\right). \quad (6.21)$$

We define the transition frequency for the transition to the n^{th} state from the j^{th} state to be $\omega_{jn} = (E_n - E_j) / \hbar$, then integrate over time to obtain

$$a_n^{(1)}(t) = \frac{\mathbf{D}_{nj} \cdot \mathbf{E}}{2\hbar} \left[\frac{1 - \exp i(\omega_{jn} + \omega)t}{\omega_{jn} + \omega} + \frac{1 - \exp i(\omega_{jn} - \omega)t}{\omega_{jn} - \omega} \right]. \quad (6.22)$$

Now let us assume that the n^{th} state is higher in energy than the j^{th} state (*i.e.* we are modeling absorption) and that the radiation is almost resonant with the transition (*i.e.* that $\omega \approx \omega_{jn}$). Hence we the second term in the brackets is much greater than the first and we can neglect the first. This approximation is called the *rotating wave approximation*. In this case the probability of finding the system in the n^{th} state at some time t after $t = 0$ is given by

$$\left| a_n^{(1)}(t) \right|^2 = |\mathbf{D}_{nj} \cdot \mathbf{E}|^2 \frac{\sin^2[(\omega_{jn} - \omega)t/2]}{(\hbar(\omega_{jn} - \omega))^2}. \quad (6.23)$$

In general the electromagnetic excitation has a continuous spectrum of frequencies with an energy density $\rho(\omega)$, which we assume is approximately constant near the transition frequency. We also assume that there is no phase relationship between the different frequencies, which means that to determine the probability of finding the system in the excited state, we have to add/integrate the probabilities for each of the different frequency components rather than adding the amplitudes.

It is known from electromagnetism that the energy density of an electromagnetic wave is given in terms of the electric field by

$$\rho(\omega) d\omega = \frac{1}{2} \varepsilon_0 |\mathbf{E}|^2, \quad (6.24)$$

and so the probability of being in the excited state is given by

$$P_n(t) = \frac{2}{\varepsilon_0} \int_0^\infty |\mathbf{D}_{nj} \cdot \mathbf{e}|^2 \frac{\sin^2[(\omega_{jn} - \omega)t/2]}{(\hbar(\omega_{jn} - \omega))^2} \rho(\omega) d\omega, \quad (6.25)$$

where \mathbf{e} is a unit vector in the direction vector of the electric field. The integral can be shown to give

$$P_n(t) = \frac{\pi |\mathbf{D}_{nj} \cdot \mathbf{e}|^2}{\varepsilon_0 \hbar^2} \rho(\omega_{jn}) t. \quad (6.26)$$

The probability is now seen to be linear in time, which is an artefact of the approximation. In reality, it is only true for short times. Moreover, we are especially interested in the *rate* of absorption, *i.e.* the transition rate $dP_n/dt = W_{nj}$, which is given by

$$W_{nj}(\mathbf{e}) = \frac{\pi |\mathbf{D}_{nj} \cdot \mathbf{e}|^2}{\varepsilon_0 \hbar^2} \rho(\omega_{jn}). \quad (6.27)$$

where we are leaving the electric field direction \mathbf{e} as an argument to remind ourselves that the result is valid for a specific electric field direction. This important equation is a special case of **Fermi's golden rule**. It allows us to calculate emission and absorption rates from first principles.

In many cases, including the very important case of thermal radiation (which was assumed in the derivation of Einstein's A and B coefficients), the direction of the electric field is random. After averaging over polarisation and propagation directions, the absorption rate is

$$W_{nj} = \frac{\pi |\mathbf{D}_{nj}|^2}{3\epsilon_0 \hbar^2} \rho(\omega_{jn}). \quad (6.28)$$

As a result, we find

$$B_{nj} = \frac{\pi |\mathbf{D}_{nj}|^2}{3\epsilon_0 \hbar^2}, \quad (6.29)$$

$$A_{nj} = \frac{|\mathbf{D}_{nj}|^2 \omega_{jn}^3}{3\pi \epsilon_0 \hbar c^3}. \quad (6.30)$$

6.4 Selection Rules

From the calculation discussed above one can see that the rate of absorption is proportional to the vector matrix element \mathbf{D}_{ij} which is given by the integral

$$\mathbf{D}_{ij} = \int \psi_i^*(\mathbf{r}) [-e\mathbf{r}] \psi_j(\mathbf{r}) dV. \quad (6.31)$$

If all the components of this matrix element are zero, then electromagnetic waves cannot cause a transition between the i^{th} and j^{th} state, *i.e.* electromagnetic energy cannot cause an electron to be excited between the two states and energy cannot be emitted as an electron relaxes to the lower energy of the two states from the other.

It turns out that for most pairs of states the matrix element is zero and the transitions are so-called forbidden, and that there are some fairly simple rules which allow us to check which states are allowed and which are forbidden. We will now look at calculating these rules.

To make these calculations simpler we will separate the x , y and z components of the vector matrix element. We choose the z direction to be the direction associated with the magnetic quantum number for angular momentum. These components correspond to replacing the position vector \mathbf{r} in the integral with x , y and z . In spherical polar co-ordinates, $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$ and $z = r \cos \theta$. The integral for the x component of the matrix element when written out in full is given by

$$D_{ij}^x = -e \int_0^\infty \int_0^{2\pi} \int_0^\pi \psi_i^*(\mathbf{r}) (r \sin \theta \cos \phi) \psi_j(\mathbf{r}) r^2 \sin \theta dr d\phi d\theta. \quad (6.32)$$

We know that the wavefunctions for the atomic orbitals of hydrogen can be separated into radial and angular parts, *i.e.* $\psi_{n,l,m_l}(r, \theta, \phi) = R_{n,l}(r)Y_{l,m_l}(\theta, \phi)$. If we put these into the integral above using a n', l' and m' to refer to the i state and n, l and m to refer to the j state, we can then separate the radial and angular integrations and we get,

$$D_{ij}^x = -e \int_0^\infty R_{n',l'}^*(r) R_{n,l}(r) r^3 dr \int_0^\pi \int_0^{2\pi} Y_{l',m'}^*(\theta, \phi) [\sin \theta \cos \phi] Y_{l,m}(\theta, \phi) \sin \theta d\theta d\phi. \quad (6.33)$$

It turns out that the radial integral is never zero. In order to aid calculating the value of the angular integral the following relations can be used:

$$\sin \theta \cos \phi = \sqrt{\frac{2\pi}{3}} (Y_{1,-1} - Y_{1,1}), \quad (6.34)$$

$$\sin \theta \sin \phi = i \sqrt{\frac{2\pi}{3}} (Y_{1,-1} + Y_{1,1}), \quad (6.35)$$

$$\cos \theta = \sqrt{\frac{4\pi}{3}} (Y_{1,0}), \quad (6.36)$$

and

$$Y_{1,m} Y_{l,m'} = A Y_{l+1,m+m'} + B Y_{l-1,m+m'}, \quad (6.37)$$

where A and B are non-zero constants, whose value is not important to this calculation. Using these we can simplify the angular integral above:

$$D_{ij}^x = -e I_R \int_0^{2\pi} \int_0^\pi Y_{l',m'}^*(\theta, \phi) \sqrt{\frac{2\pi}{3}} (Y_{1,-1} - Y_{1,1}) Y_{l,m}(\theta, \phi) \sin \theta d\theta d\phi, \quad (6.38)$$

where I_R denotes the integral of the radial part. Omitting the explicit dependence of the radial function, we obtain

$$D_{ij}^x \propto \int_0^\pi \int_0^{2\pi} Y_{l',m'}^*(AY_{l+1,m-1} + BY_{l-1,m-1} - AY_{l+1,m+1} - BY_{l-1,m+1}) \sin \theta d\theta d\phi \quad (6.39)$$

This integral involves integrating two spherical harmonics multiplied together over the angular co-ordinates. From the orthonormality of the spherical harmonics we know these integrals will be zero unless the angular quantum numbers are the same, *i.e.* $l' = l + 1, m' = m - 1$ or $l' = l - 1, m' = m - 1$ or $l' = l + 1, m' = m + 1$ or $l' = l - 1, m' = m + 1$. The four possibilities can be summarised as $\Delta l = \pm 1, \Delta m_l = \pm 1$ (here we add a subscript l to remind ourselves that we are dealing with z component of orbital angular momentum). If we were to repeat the calculation above for the y component of the matrix

element we would get the same result that the matrix element was zero unless $\Delta l = \pm 1, \Delta m_l = \pm 1$. For the z component the matrix element is zero unless $\Delta l = \pm 1, \Delta m_l = 0$. Thus, the overall selection rules for electromagnetic transitions are $\Delta l = \pm 1$ and $\Delta m_l = 0, \pm 1$.

6.5 Forbidden Transitions

Let us consider some of the transitions which are forbidden and what this means. For instance the $2s$ state ($n= 2, l = 0$) has only one state lower in energy, the $1s$ state ($n = 1, l = 0$). A transition between these two states cannot be caused by electromagnetic waves according to the calculation above, and so if we highly excite atomic hydrogen we would expect that some of the electrons would relax into the $2s$ state and get stuck, according to a naïve interpretation of the selection rules derived above, for ever. Clearly, this is not reasonable. Some other perturbation we have not considered will cause the electrons to relax to the ground state. In fact if we measure the lifetime of electrons in the $2s$ state we find it is approximately 0.125s which compares with a lifetime of the $2p$ state of 10^{-8}s . Clearly, whatever these other perturbations are they are much weaker.

One important simplification we made when calculating the selection rules is that the electric field is uniform in space, the so-called dipole approximation, another is that we could neglect the magnetic field. It turns out that if we treat these two approximations correctly we end up with a hierarchy of perturbations which can be separated into terms of the form dipole, quadrupole, octapole and higher multipoles. Each of these terms has its own selection rules as to which transitions they can cause. However, all of these higher terms are extremely small and so for most purposes they can be ignored.

7 Two-Electron Atoms/Ions

So far we have only dealt with one type of atom: hydrogen. As soon as a second electron is introduced the Schrödinger equation is no longer analytically solvable and thus will require numerical methods or approximations. The step between one and two electrons is a big one. We first deal with two-electron atoms/ions before moving onto Multi-Electron Atoms. Helium is the atom with two electrons.

7.1 Two-Electron Hamiltonian

As with the hydrogen atom one starts by writing out the Schrödinger equation for the problem. In this chapter we just consider the gross structure, although there will be fine-structure corrections (some of which are new effects due to there being more than one electron). Note that we carry on with the approximations that the system is non-relativistic and that the electrostatic interaction is treated classically.

As with hydrogen there will be fine structure and the relativistic mass, spin-orbit and Darwin corrections will all apply. However there are a number of new effects due to there being more than one electron. These include the so-called spin-other-orbit interaction which leads to a perturbation Hamiltonian proportional to the spin on one electron dotted with the orbital momentum of another electron, the spin-spin interaction which comes about because of one spin is in the magnetic field due to the other, and magnetic orbit-orbit interaction in which the magnetic dipole due to one orbit experiences the magnetic field due to the first orbital. Here we will ignore most of these effects and concentrate on the two electrons interacting electrostatically with the nucleus and with each other.

The Schrödinger Hamiltonian for helium is given by

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{2e^2}{4\pi\varepsilon_0 r_1} - \frac{2e^2}{4\pi\varepsilon_0 r_2} + \frac{e^2}{4\pi\varepsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|}, \quad (7.1)$$

where the subscripts on the ∇^2 operators indicate that the derivatives are with respect to the position of one or other of the electrons.

7.2 The Wave function

Firstly we will solve the problem in which we ignore the interactions between the electrons. In this case we can write the Hamiltonian as the sum of two Hamiltonians, one for each electron, which are like that for the hydrogen atom except that the charge of the nucleus is double that of the proton i.e.

$$\hat{H}_o = \sum_i \left[-\frac{\hbar^2}{2m}\nabla_i^2 - \frac{2e^2}{4\pi\varepsilon_0 r_i} \right]. \quad (7.2)$$

The solutions of the two parts of the sum separately are the same orbitals as for the hydrogen atom but scaled due to the different nuclear charge. In order that we can use the solutions later for multi-electron systems let us find these scalings for the general case of an ion of atomic number Z i.e.

$$\hat{H}_o = \left[-\frac{\hbar^2}{2m} \nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r} \right]. \quad (7.3)$$

Let us scale the radial co-ordinate using $r' = Zr$

$$\hat{H}_o = Z^2 \left[-\frac{\hbar^2}{2m} \nabla'^2 - \frac{e^2}{4\pi\epsilon_0 r'} \right]. \quad (7.4)$$

This is just the Hamiltonian for the hydrogen atom with the energies multiplied by Z^2 . Thus the eigenvalues of the new problem are the same as for hydrogen multiplied by Z^2 . The eigenfunctions of the new problem will be the scaled hydrogen eigenfunctions with the radial co-ordinate being $r' = Zr$. In addition they have to be renormalised as they differ from the (normalised) eigenfunctions for hydrogen. Thus the new eigenfunctions are given by

$$\psi_{n,l,m}(\mathbf{r}) = Z^{3/2} \psi_{n,l,m}^H(\mathbf{r}'). \quad (7.5)$$

This is not the final form of the eigenfunction because the wave function must respect the exchange symmetry discussed in section 2.16. Electrons are indistinguishable and are fermions. This means that the overall wave function including the spin part has to be anti-symmetric under exchange of the labels of the two particles, and so the spatial part of the wave function can be either symmetric or anti-symmetric. The symmetric spatial eigenfunction is

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} (\psi_{n_1, l_1, m_1}(\mathbf{r}_1) \psi_{n_2, l_2, m_2}(\mathbf{r}_2) + \psi_{n_1, l_1, m_1}(\mathbf{r}_2) \psi_{n_2, l_2, m_2}(\mathbf{r}_1)), \quad (7.6)$$

and the anti-symmetric spatial eigenfunction is:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} (\psi_{n_1, l_1, m_1}(\mathbf{r}_1) \psi_{n_2, l_2, m_2}(\mathbf{r}_2) - \psi_{n_1, l_1, m_1}(\mathbf{r}_2) \psi_{n_2, l_2, m_2}(\mathbf{r}_1)). \quad (7.7)$$

As shown in Table 1 in Chapter 2, the symmetric spatial eigenfunction must be joined with the anti-symmetric spin state in order to keep the whole wave function anti-symmetric, and vice versa. The \pm sign in the spatial eigenfunction has a physical effect on the energy eigenvalues, which we will calculate in section 7.3. Putting $\mathbf{r}_1 = \mathbf{r}_2$ would mean that the anti-symmetric spatial eigenfunction vanishes when the two electrons are at the same point in space, but the same is not true for the symmetric spatial eigenfunction. Hence the average separation of the anti-symmetric state is greater than that of the symmetric state and this results in different energy eigenvalues.

7.3 Exchange Term

In the limit that the electron-electron interactions are treated as a perturbation we should start with the eigenfunctions of the non-interacting case. In the case that the two electrons are in the same orbital (i.e. same n, l, m quantum numbers) only the symmetric eigenfunction is applicable, because the anti-symmetric eigenfunction would vanish. In the unperturbed case these states are degenerate with energy equal to the sum of the energies of the individual orbitals. It can be shown that above eigenfunctions are the good states for the electron-electron interaction.

If we calculate the first-order perturbation to the energy of these eigenfunctions due to the electron-electron interaction (i.e. using eq. (4.13) as usual) one obtains

$$\Delta E^{(1)} = E_{\text{direct}} \pm E_{\text{exchange}}, \quad (7.8)$$

where the plus sign (minus sign) is for the symmetric (anti-symmetric) spatial eigenfunction and

$$E_{\text{direct}} = \int \psi_{n_1, l_1, m_1}^*(\mathbf{r}_1) \psi_{n_2, l_2, m_2}^*(\mathbf{r}_2) \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} \psi_{n_1, l_1, m_1}(\mathbf{r}_1) \psi_{n_2, l_2, m_2}(\mathbf{r}_2) dV_1 dV_2. \quad (7.9)$$

$$E_{\text{exchange}} = \int \psi_{n_1, l_1, m_1}^*(\mathbf{r}_1) \psi_{n_2, l_2, m_2}^*(\mathbf{r}_2) \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} \psi_{n_1, l_1, m_1}(\mathbf{r}_2) \psi_{n_2, l_2, m_2}(\mathbf{r}_1) dV_1 dV_2. \quad (7.10)$$

In these equations the integral is over both volume elements, *i.e.* each of the three position variables. The second integral differs from the first only by the exchange of the position dependence $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$ in the right-hand product of wave functions.

The first of the two energies is called the direct term and can be rewritten as eq. (7.11). It is what you would write down if you were going to calculate the electron-electron interaction classically, *i.e.* an integral involving two charge densities, where $e|\psi|^2$ is a charge density

$$E_{\text{direct}} = \int \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} |\psi_{n_1, l_1, m_1}(\mathbf{r}_1)|^2 |\psi_{n_2, l_2, m_2}(\mathbf{r}_2)|^2 dV_1 dV_2. \quad (7.11)$$

The second of the two energies is called the exchange term. It has no classical counterpart and arises from requiring that the eigenfunctions respect the exchange symmetry (*i.e.* eq. (7.6) and eq. (7.7)). This term is extremely important because it splits the degeneracy between the symmetric and anti-symmetric wave functions.

A surprising effect of the exchange term is that the energy of the helium atom is thus dependent on its spin state, since the $S = 0$ states (called *parahelium*) are always coupled to a symmetric spatial eigenfunction state and the $S = 1$ states (called *orthohelium*) are always coupled to an anti-symmetric spatial state.

7.4 Comparison with Experiments

Shown below is one method of plotting the energy levels of Helium. What is shown is the energy relative to a singly ionised helium ion in its ground state (taken to be $E = 0$), *i.e.* the energy required to remove an electron from a helium atom while the other electron remains in its ground state. The energies are shown in four columns which correspond to the orbital angular momentum (s, p, d, f only shown) of the excited electron. This is given by the electron configuration which lists the atomic orbitals that the electrons inhabit. Each of these columns has two sets of lines associated with the total spin of the electrons, *i.e.* the singlet $S = 0$ and triplet $S = 1$ states. These columns are indicated by the term symbols at the top of the columns. Also shown for reference are the energy levels of hydrogen with principal quantum number n .

From the diagram it is apparent that the energies are remarkably close to those of hydrogen, apart from the ground state and the lower energy levels. This is perhaps surprising because before taking into account the electron-electron interaction we calculated that the single-particle energies of helium should be a factor of 4 bigger (*i.e.* more negative) than those of hydrogen.

This result is a consequence of electron-electron interaction and the size of the different orbitals. Consider an electron in a highly excited orbital with large n . This electron will be found most of the time considerably further from the nucleus than the electron in the $1s$ orbital. Using Gauss's theorem it is possible to prove that the sum of the nuclear and electron-electron electrostatic potentials for points much further from the nucleus than the $1s$ orbital will be the same as for a point charge of $+e$ at the position of the nucleus. It is said that the $1s$ orbital *screens* one of the nuclear charges, and so the effective nuclear charge (as seen by the excited electron) is less than the actual nuclear charge.

Thus electrons in a highly excited orbital will have energies very close to that of an electron in similar orbitals in hydrogen. Of course, for a limited time the highly excited electron can be found inside the $1s$ orbital and so it experiences a more negative potential *i.e.* it experiences the full (unscreened) nuclear charge. Hence its energy will be slightly lower than the comparable electron in hydrogen. In terms of the direct and exchange integrals the effect of screening is mainly due to the direct integral (the screening effect can also be understood from a classical approach, and the direct term has an analogue in classical physics). The exchange integral will still cause a splitting between the symmetric and anti-symmetric spatial eigenfunctions. However, the exchange integral is maximised when the electrons are close to each other (*i.e.* when $|\mathbf{r}_1 - \mathbf{r}_2|$ is smaller) and thus the exchange integral decreases in magnitude as the principal quantum number of the excited electron becomes larger (and consequently is on average further away from the $1s$ electron).

The $1s1s$ ground state only exists for the $S = 0$ state, and the electron-electron repulsion raises the energy significantly (by about 30 eV) compared to the model without this term. Let us substantiate the statements with some numbers. The

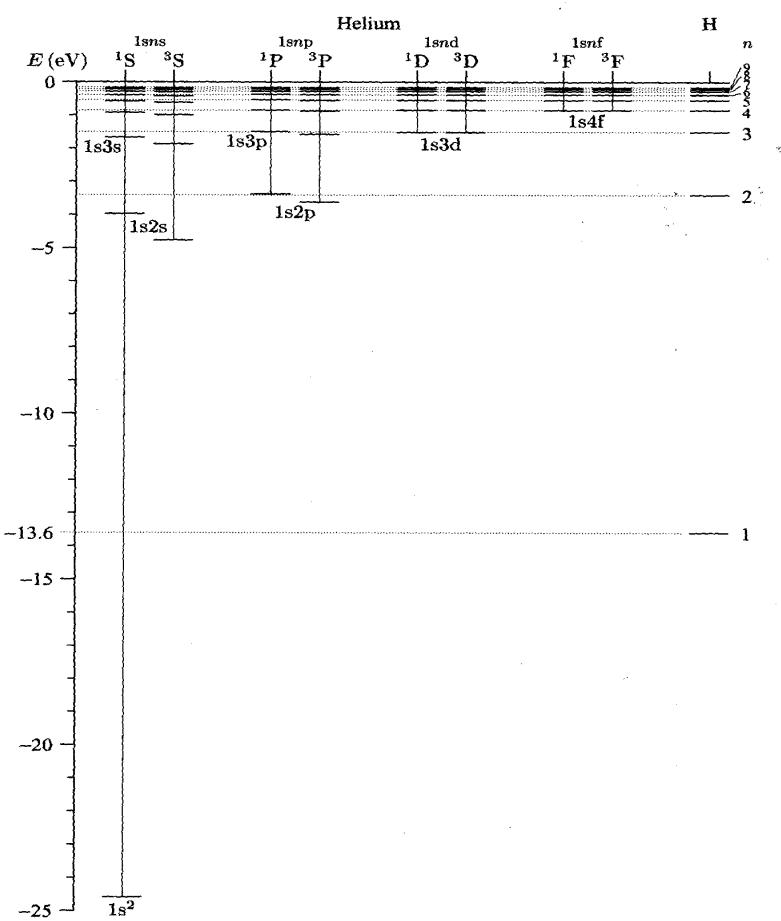


Figure 11: Energy level of Hydrogen and Helium.

ionization energy of Helium is 24.6 eV (the highest among all the atoms in the periodic table), while the energy required to remove the second electron is 54.6 eV. Notice that once the first electron is removed, the system is like a hydrogenic atom by with $Z = 2$, and thus the binding energy is a factor $4 \times -13.6 \text{ eV} = -54.4 \text{ eV}$. However, the inter-electron repulsion is then about $54.4 - 24.6$ or approx 30 eV.

7.5 Optical transitions

In order to understand the optical spectra of helium we have to determine what transitions are optically allowed. Within the dipole (E1), long wavelength (compared to wavelength of atom) approximation the interaction/perturbation Hamiltonian consists of the sum of two dipole operators; one for each electron, i.e.

$$\hat{H}_p = \sum_{i=1,2} -e\mathbf{r}_i \cdot \mathbf{E}. \quad (7.12)$$

The matrix elements calculated with this operator will involve terms of the form,

$$\int \psi_{n1,l1,m1}^*(\mathbf{r}_1) \psi_{n2,l2,m2}^*(\mathbf{r}_2) (-e\mathbf{r}_1) \psi_{n3,l3,m3}(\mathbf{r}_1) \psi_{n4,l4,m4}(\mathbf{r}_2) dV_1 dV_2, \quad (7.13)$$

and a similar integral with \mathbf{r}_2 .

These integrals can always be separated into two factors where only one involves the perturbation i.e.

$$\int \psi_{n1,l1,m1}^*(\mathbf{r}_1) (-e\mathbf{r}_1) \psi_{n3,l3,m3}(\mathbf{r}_1) dV_1 \int \psi_{n2,l2,m2}^*(\mathbf{r}_2) \psi_{n4,l4,m4}(\mathbf{r}_2) dV_2. \quad (7.14)$$

Clearly the integral not involving the dipole operator ($-e\mathbf{r}_1$) is zero unless the initial and final states are the same (due to orthonormality), i.e. only one electron (in this case the electron with position vector (\mathbf{r}_1)) can change state at a time. The integral involving the dipole operator will have the same selection rules as for the hydrogen case. The same is true for the integrals with ($-e\mathbf{r}_2$). Hence there is a new selection rule for atoms with more than one electron: *only one electron can make a transition at a time*.

There is an additional selection rule which involves spin. As the perturbation Hamiltonian does not depend on spin at all this means the spin state cannot change. *This selection rule means that singlet and triplet spin states cannot be linked by optical transitions.* This is why early measurements of the optical spectra of helium were interpreted in terms of two types of helium, referred to as ortho and para helium, as there are effectively two almost hydrogen like series of lines associated with the $S = 0$ and $S = 1$ states of helium respectively.

8 Multi-Electron Atoms

8.1 The Hamiltonian and the Central Field Approximation

The non-relativistic Schrödinger equation for multi-electron atoms has the following Hamiltonian:

$$\hat{H} = \sum_i \frac{-\hbar^2}{2m} \nabla_i^2 + \sum_i \frac{-Ze^2}{4\pi\varepsilon_0 |\mathbf{r}_i|} + \sum_i \sum_{j>i} \frac{e^2}{4\pi\varepsilon_0 |\mathbf{r}_i - \mathbf{r}_j|}. \quad (8.1)$$

In both cases the indices label the electrons (i runs from 1 to atomic number Z , and $j > i$ ensures no double counting of the electron-electron repulsion term).

This is an extremely difficult problem to solve even for a small number of electrons, and analytic solutions are not possible. Numerical techniques and simplifying assumptions are needed in order to progress. One assumption that can be made is that the electron-electron repulsion terms can be split up into spherically symmetric parts associated with each electron (i.e. $|\mathbf{r}_i|$ dependence only, also called a **central potential**), and a component that is non-central. The spherically symmetric part is added to the Coulomb potential for that electron (also $|\mathbf{r}_i|$ dependent only) and the sum of these two contributions is labelled as $U(|\mathbf{r}_i|)$. Hence the Hamiltonian in eq. (8.1) can be rewritten as:

$$\hat{H} = \sum_i \left[\frac{-\hbar^2}{2m} \nabla_i^2 + U(|\mathbf{r}_i|) \right] + \hat{H}_p, \quad (8.2)$$

where

$$\hat{H}_p = - \sum_i U(|\mathbf{r}_i|) + \left(\sum_i \frac{-Ze^2}{4\pi\varepsilon_0 |\mathbf{r}_i|} + \sum_i \sum_{j>i} \frac{e^2}{4\pi\varepsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \right). \quad (8.3)$$

Note that we have added and subtracted the central part of the potential to \hat{H} and \hat{H}_p respectively. In general, the central potentials $U(|\mathbf{r}_i|)$ will be different for each electron, as the electron-electron repulsion for a given electron depends on its own spatial distribution (i.e. its wave function) and the spatial distribution (wave functions) of all of the other electrons. The non-spherical/non-central part of the Hamiltonian denoted by \hat{H}_p in eq. (8.3) (sometimes called the *residual electrostatic interaction*) is the sum of the exact interelectronic Coulomb separations ($\sum_{ij} 1/r_{ij}$) minus the rotational average of the direct minus the exchange potential summed over the electrons. This term can be treated as a perturbation if it is fairly small, and turns out to be the case for atoms due to the fairly symmetric distribution of charges in a given shell, which results in partial cancellation of non-radial forces. Note that the inter-electronic interactions are too strong to be treated as perturbation (apart from Helium), but

with the rotational part subtracted off, the H_p is small enough to warrant a successful perturbative treatment.

If the form of each central potential $U(|\mathbf{r}_i|)$ is known then the energy eigenfunctions of each of the electrons can be solved separately. The central potential is spherically symmetric and so separation of variables is possible. The solutions to the angular equation will be the spherical harmonics. Thus it is only necessary to solve the radial equation, which is given by eq. (3.8) but the Coulomb potential for hydrogen is replaced by $U(|\mathbf{r}_i|)$ in a multi-electron atom. The eigenfunctions for the atom can be formed as the product of these single-particle eigenfunctions. It is important to recall that the radial equation eq. (3.8) has the centrifugal potential term which depends on the orbital angular momentum quantum number l , and this will mean that states with different l will be non-degenerate i.e. for the single-particle eigenfunction the equivalent of the 2s and 2p states of Hydrogen will not be degenerate. The degeneracy in the solutions for hydrogen was a special case, and arose due to the potential being strictly $1/(|\mathbf{r}_i|)$ in eq (3.8). For multi-electron atoms $U(|\mathbf{r}_i|)$ has a more complicated dependence on ($|\mathbf{r}_i|$).

8.2 Self-Consistent Atomic Orbital Calculations

Calculating the central potentials is a challenging task and in practice it is obtained to a desired accuracy by an iterative method, in which new *improved* wave functions are calculated using a central potential that in turn has been calculated with the previously generated wave functions in an iterative process (see eq. (8.4)) below, and then the potential is updated with the improved wave functions i.e. initial wave functions → improved central potential → improved wave functions → further improved central potential etc, in which the "→" means that the Schrödinger equation is solved to obtain the wave functions or eq. (8.4) is used to obtain the improved potential. The starting point is an initial (educated) guess for the wave functions, and the end point will be when the newly generated central potential differs from the previous one by less than a desired precision. This is a so-called **self-consistent solution** and it is possible to show that if this iterative process converges then it converges to the true solution. The formula used to determine the central potential from the previous generation of solutions ψ_i is:

$$U(|\mathbf{r}_j|) = \frac{-Ze^2}{4\pi\epsilon_0 |\mathbf{r}_j|} + \sum_{i \neq j} \int \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_j - \mathbf{r}_i|} |\psi_i|^2 dV_i. \quad (8.4)$$

This formula calculates the effect of the electron-electron repulsion terms and each integral will just give a function of $|\mathbf{r}_j|$ for the j (electron) under consideration ($|\mathbf{r}_j|$ is not integrated over, as $i \neq j$ in the sum). The method just described is called the *Hartree method*. The Hartree method does not take into account particle exchange symmetry. The development of the above method to include the particle exchange symmetry (and therefore exchange energies)

is called the *Hartree-Fock method*. Multi-particle wave functions for electrons must be antisymmetric and will vanish if any two of the electrons have the same quantum numbers. These anti-symmetric multi-particle wave functions in terms of single-particle eigenfunctions can be written down using the Slater determinant, which reproduces the anti-symmetric two-electron eigenfunction for Helium when applied to two particles. In the determinant the subscript 1 on ψ_1 stands for the set of quantum numbers n_1, l_1, m_1 that defines the single-particle eigenfunction:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_1(\mathbf{r}_2) & \dots & \psi_1(\mathbf{r}_N) \\ \psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) & \dots & \\ \vdots & \vdots & \ddots & \\ \psi_N(\mathbf{r}_1) & & & \psi_N(\mathbf{r}_N) \end{vmatrix} \quad (8.5)$$

Before we consider the case of a general atom we will consider the case of what are effectively the simplest set of multi-electron atoms - the alkali metal atoms.

8.3 Alkali Metals

The alkali metals, lithium, sodium, potassium, rubidium, caesium and francium, can be considered as the simplest multi-electron atoms after helium because the highest-energy electron (the valence electron) is in an orbital where the electron is most of its time further from the nucleus than all the other electrons. As we discussed in the case of helium, Gauss's law means that whilst this electron is *outside* the other electrons it experiences a potential which is effectively that of a nucleus with a charge reduced from the real charge by the number of electrons that are closer to the nucleus than the outer electron. This means the outer electron mostly experiences a potential which is that of a single positive charge at the nucleus, *i.e.* its states will be hydrogen like.

Below in Fig.8.3 we show the energy levels of sodium available to the outer (valence) electron assuming that the inner electrons are not excited. The energy levels are measured relative to a singly ionised sodium atom. The dashed lines indicate the sets of levels associated with a particular principal quantum number. The lowest principal quantum number (n) available to the outer electron is 3 as all the lower states are filled by the inner electrons. Also shown on the right are the energy levels of hydrogen. As can be seen, the energies of the d and f (*i.e.* $l = 2$ and 3) subshells of sodium correspond extremely well to those of hydrogen in eq.(3.24). However, the p and s subshells are lower in energy than the same levels in hydrogen. In fact, the energy levels of sodium are well approximated by the formula

$$E_{n,l} = -\frac{Ry}{(n - \delta_l)^2}, \quad (8.6)$$

where $Ry = 13.6$ eV is the Rydberg energy, *i.e.* the ionisation energy of the ground-state of hydrogen, and δ_l is called the *quantum defect* (can be written

as $\delta(l)$ - see below). Eq. (8.6) is just eq. (3.24) for hydrogen but with the term δ_l , which parametrises the difference between the hydrogen energy levels and those of sodium. A positive value of δ_l , which was determined experimentally in the 19th century, gives a lower energy level than the corresponding level for hydrogen, while $\delta_l = 0$ gives the same energy levels as for hydrogen. As the subscript on δ_l suggests, the quantum defect is different for different values of orbital angular momentum. Writing δ_l as $\delta(l)$ one has for sodium $\delta(0) = 1.35$ (s states), $\delta(1) = 0.86$ (p states), $\delta(2) = 0.01$ (d states), and for all higher orbital angular momenta the quantum defect is effectively zero. Moreover, the quantum defect has almost negligible dependence on n .

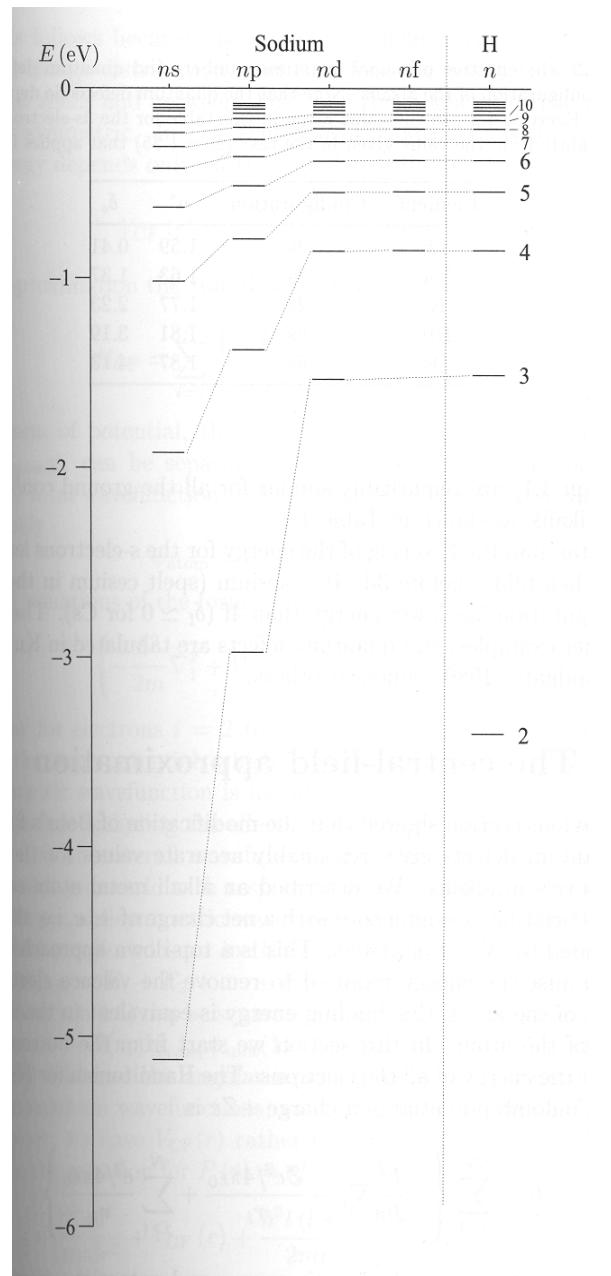
One can understand the origin of the different values of the quantum defect by looking at the form of the radial wave functions of hydrogen for the same n and different l below (which would be expected to be reasonable approximations for the corresponding radial wave functions for the single valence electron in sodium)

This shows that whilst electrons in the $l = 0$ state spend most of their time far from the nucleus they also spend some fraction of their time close to the nucleus. The probability of the electron being found with $r < a_0$ is given by the area under the curve for $r < a_0$, and this area is significantly larger for the $l = 0$ state than for the $l = 1$ and $l = 2$ states. While being very close to the nucleus the electron will be inside the orbits of the other electrons (which are in the $n = 1$ and $n = 2$ shells with average radii for sodium of around $0.1 a_0$ and $0.4 a_0$ respectively) and therefore it will experience a much lower potential (the nucleus has a charge of $+11e$ for sodium) than it would if it were orbiting a hydrogen atom. This will lower its energy compared to the equivalent electron in hydrogen. Another way of saying this is that the screening of the nuclear charge drops off the closer an electron is to the nucleus, and the electron experiences the full nuclear charge part of the time. An electron in the $l = 1$ spends much less of its time close to the nucleus and within the $n = 1$ and $n = 2$ shells than an electron in an $l = 0$ state. This means that the energy of the p is not lowered as much as that of the s state.

The absolute effect of the incomplete screening will drop off as the principal quantum number increases because in higher energy orbitals the distance of the electron from the nucleus increases, and thus the electron experiences the screened nuclear charge most of the time. Hence as n increases the energy levels of sodium are essentially hydrogenic, even for $l = 0$ and $l = 1$ states. These different effects are all captured using the quantum defect in the formula above.

8.4 Periodic Table

From very early on chemists had noticed that certain groups of atoms behave in a similar manner. This led to the development (in the first half of the 19th century) of a range of different systems for arranging the elements into groups. One of the most successful of these was the periodic table by Mendeleev (1869)



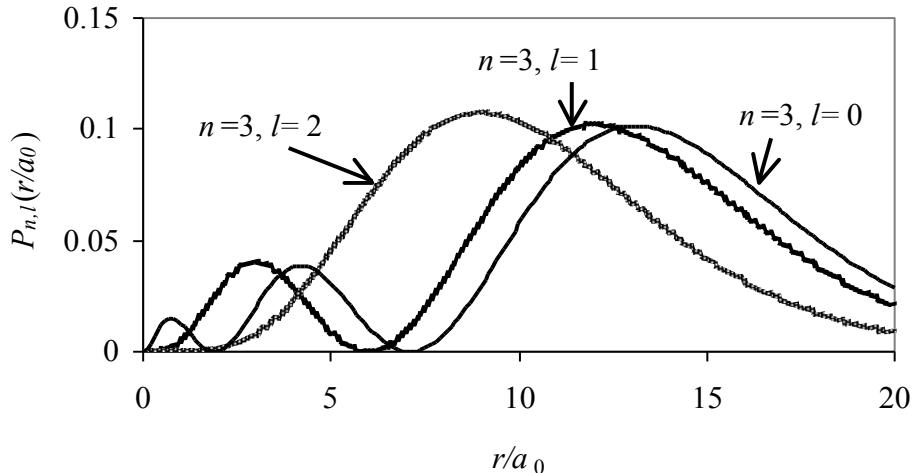


Figure 12: Radial extent of the probability density $P_{n,l}(r) = C_{n,l}r^2 |R_{n,l}(r)|^2$ for the hydrogen atom with $n = 3$, and $l = 0, 1, 2$.

where the atoms were mostly ordered by atomic mass and chemical similarity. A further novel feature of Mendeleev's table was that he assumed there were some elements still to be discovered and he left gaps in his table to take this into account. Some of his predictions (Scandium, Germanium and Gallium) were successful, while others were not. Mendeleev's periodic table was slightly improved by Henry Moseley (1913) who used x-ray spectroscopy to measure the atomic number of atoms and then adjusted Mendeleev's table accordingly to arrange in terms of atomic number. Despite the fact that periodic tables had been written down, a clear theoretical explanation of why certain atoms behaved similarly and why the periodic table should take such a form was not possible before the formulation of quantum mechanics.

Using our understanding of the central potential and the effect of screening of nuclear charge by inner electrons it is now possible to explain the periodic table of elements. As mentioned in section 8.1, the perturbation due to non-central electron-electron interactions is small enough (due the cancelling of non-radial forces caused by the symmetric charge distributions of filled shells) that for the purposes of explaining the periodic table we do not need to take it into account. Thus we can treat each electron in an atom as being associated with a single electron orbital (specified by n, l, m quantum numbers) which can be calculated using the central potential. We can describe the multi-electron energy eigenfunctions by multiplying together the single electron eigenfunctions from which it is made up (and then anti-symmetrising the multi-electron eigenfunction), as was done in the case of Helium. The list of shells (denoted by n) and subshells

(denoted by l) that are filled is called the *electron configuration* (examples given below), and each subshell contains $(2l + 1)$ orbitals, each denoted by m . Each orbital can be occupied by two electrons, one with $m_s = 1/2$ and one with $m_s = -1/2$. In this central field approximation, the energy levels and (spatial) eigenfunctions for these single particle quantum states are labelled by n , l and m and are unique to each multi-electron atom because the central potential is different for each atom due to the different nuclear charge and number of electrons (e.g. the energy of the $3s$ state is different for sodium and rubidium)

We will now consider the electron configuration of the ground state of the different atoms. As the atomic number (Z) of the atom is increased so is the number of electrons (which equals Z) orbiting the nucleus. By definition, in the ground-state of an atom the electrons will be found in the lowest energy orbitals within the constraints of the Pauli exclusion principle. We already know that for helium the two electrons are found in the orbitals which derive from the $1s$ orbitals of hydrogen. When we come to lithium ($Z = 3$) the $1s$ orbital is now filled by the first two electrons and the third electron will have to go into one of the orbitals which derives from the $n = 2$ states of hydrogen. The effect of screening by the inner electrons lowers the energy of the $2s$ state relative to the $2p$ state (as the $2s$ state spends more time closer to the nucleus, as discussed in section 8.3) and so the third electron goes into the $2s$ state. The electron configuration of lithium is given by $1s^2 2s$ or $[\text{He}]2s$. The latter form states where the electron configuration is like that of helium (which has a filled $1s$ orbital) plus an electron in an orbital in the $2s$ subshell. In the case of beryllium ($Z = 4$) the extra electron can also go into the $2s$ orbital, but with opposite spin. In boron ($Z = 5$) the $1s$ and $2s$ subshells are already filled and so the highest-energy electron is in the $2p$ subshell. The $2p$ subshell can take 6 electrons before being filled. The element which fills this subshell is neon. The next element is sodium and as the $n = 1$ and 2 shells are filled its last electron has to go into the $n = 3$ shell, and once again the $3s$ subshell has the lowest energy. In the next element, magnesium, the last electron completes the $3s$ orbital. The next six elements up to argon have their highest energy electrons in the $3p$ subshell. The next element is potassium. One might expect that the $3d$ subshell is the next to be filled. However, it turns out that the $4s$ subshell is lower in energy than the $3d$ subshell due to the fact that the screening is greater for the $3d$ than for the $4s$. Hence potassium and calcium have their highest energy electron in the $4s$ subshell.

The next element is scandium and the $3d$ subshell now starts to be filled (it is lower in energy than the $4p$ subshell). As can be seen from the periodic table below, in general the trend across the ten elements starting from scandium (the first row of the *transition metals*) is for the final electron to be added to the $3d$ subshell. However, in the case of chromium and copper this is not the case. For these two elements, an electron from the $4s$ subshell moves to the $3d$ subshell, and so in comparison with the previous element there is one less electron in the $4s$ but two more electrons in the $3d$. Once the $3d$ subshell has been filled, in the next six atoms the outer electron goes in the $4p$ subshell, ending with krypton.

For the next two atoms the outer electron is in the $5s$ subshell. After this a simple intuitive pattern of the order of filling subshells is not always (although nearly always) correct.

There are a number of trends across the periodic table. One of the most important of these is in the ionisation energy of the elements. Below are plotted the first ionisation energy, i.e. the energy required to make a single positively charged ion from the neutral atom. The first figure shows all of the elements for which the ionisation energy is known and the second a close up of the first figure that shows only the first 40 elements. The ionisation energy of atoms with one electron in an s subshell is clearly the lowest. The ionisation energies show a general trend to increase up to the element just before the next alkali metal. This makes sense because the nuclear charge is increasing, which increases the energy needed to release an electron (because the electrons are more tightly bound). The extra electrons which are being added are at roughly the same average distance from the nucleus as the outer electron of the previous alkali metal, and thus these electrons are relatively ineffective at screening the nuclear charge for each other, and so cannot compensate for the increasing nuclear charge. Note however that there are drops in ionisation energy associated with starting each new shell - this is because the principal quantum number (n) has changed and the average radius of the electron has greatly increased compared to the previous shell (and so is less tightly bound). In addition it should be noted that there are much smaller drops associated with going from half-filled p subshells (3 electrons) to p subshells with 4 electrons.

8.5 Fine Structure of Multi-Electron Atoms

The model of multi-electron atoms discussed above is sufficient for explaining many of the properties of atoms. However, it is possible to refine the model to include fine-structure effects, of which we are going to consider only two of the many possible perturbations. These are i) the non-central component of the electron-electron interaction and ii) the spin-orbit correction. These are given by the following Hamiltonian perturbations:

$$\hat{H}_{\text{non-central}} = - \sum_i U(|\mathbf{r}_i|) + \left(\sum_i \frac{-Ze^2}{4\pi\varepsilon_0 |\mathbf{r}_i|} + \sum_i \sum_{j>i} \frac{e^2}{4\pi\varepsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \right), \quad (8.7)$$

$$\hat{H}_{\text{spin-orbit}} = \sum_i \xi_i (|\mathbf{r}_i|) \hat{\mathbf{l}}_i \cdot \hat{\mathbf{s}}_i. \quad (8.8)$$

Eq. (8.7) is the same as eq. (8.3), but \hat{H}_p has been relabelled as $\hat{H}_{\text{non-central}}$. The expression for $\hat{H}_{\text{spin-orbit}}$ in eq. (8.8) is the generalisation of eq. (5.6) to the case of more than one electron, and the spatial function ξ_i is derived from the central-potential (ξ_i corresponds to the part that multiples $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ in eq. (5.6), in which the spatial dependence $1/r^3$ in that equation originates from the Coulomb

Periodic Table of the Elements

| Atomic Number | Symbol | Name | Electron Configuration |
|---------------|--------|--------------|------------------------|
| 1 | H | Hydrogen | $[He]1s^1$ |
| 2 | He | Helium | $[He]2s^2$ |
| 3 | Li | Lithium | $[He]2s^1$ |
| 4 | Be | Beryllium | $[He]2s^2$ |
| 11 | Na | Sodium | $[Ne]1s^1$ |
| 12 | Mg | Magnesium | $[Ne]2s^2$ |
| 19 | K | Potassium | $[Ar]1s^1$ |
| 20 | Ca | Calcium | $[Ar]2s^2$ |
| 37 | Rb | Rubidium | $[Kr]5s^1$ |
| 38 | Sr | Strontium | $[Kr]5s^2$ |
| 55 | Cs | Cesium | $[Xe]6s^1$ |
| 87 | Fr | Francium | $[Rn]7s^1$ |
| 88 | Ra | Radium | $[Rn]7s^2$ |
| 21 | Sc | Scandium | $[Ar]3d^14s^2$ |
| 22 | Ti | Titanium | $[Ar]3d^24s^2$ |
| 23 | V | Vanadium | $[Ar]3d^34s^2$ |
| 24 | Cr | Chromium | $[Ar]3d^54s^1$ |
| 25 | Mn | Manganese | $[Ar]3d^54s^2$ |
| 26 | Fe | Iron | $[Ar]3d^64s^2$ |
| 27 | Co | Cobalt | $[Ar]3d^74s^2$ |
| 28 | Ni | Nickel | $[Ar]3d^84s^2$ |
| 29 | Cu | Copper | $[Ar]3d^104s^1$ |
| 30 | Zn | Zinc | $[Ar]3d^104s^2$ |
| 31 | Ga | Gallium | $[Ar]3d^104s^24p^1$ |
| 32 | Ge | Germanium | $[Ar]3d^104s^24p^2$ |
| 33 | As | Arsenic | $[Ar]3d^104s^24p^3$ |
| 34 | Se | Selenium | $[Ar]3d^104s^24p^4$ |
| 35 | Br | Bromine | $[Ar]3d^104s^25p^5$ |
| 36 | Kr | Krypton | $[Ar]3d^104s^25p^6$ |
| 37 | Y | Yttrium | $[Kr]4d^55s^2$ |
| 41 | Nb | Niobium | $[Kr]4d^55s^1$ |
| 42 | Mo | Molybdenum | $[Kr]4d^55s^1$ |
| 43 | Tc | Technetium | $[Kr]4d^55s^1$ |
| 44 | Ru | Ruthenium | $[Kr]4d^95s^1$ |
| 45 | Rh | Rhodium | $[Kr]4d^9$ |
| 46 | Pd | Palladium | $[Kr]4d^10$ |
| 47 | Ag | Silver | $[Kr]4d^105s^1$ |
| 48 | Cd | Cadmium | $[Kr]4d^105s^2$ |
| 49 | In | Indium | $[Kr]4d^105s^25p^1$ |
| 50 | Sn | Tin | $[Kr]4d^105s^25p^3$ |
| 51 | Sb | Antimony | $[Kr]4d^105s^25p^5$ |
| 52 | Te | Tellurium | $[Kr]4d^105s^25p^6$ |
| 53 | I | Iodine | $[Kr]4d^105s^25p^6$ |
| 55 | Cs | Cesium | $[Xe]6s^1$ |
| 56 | Ba | Barium | $[Xe]6s^2$ |
| 57 | La | Lanthanum | $[Xe]5d^14s^2$ |
| 58 | Ce | Cerium | $[Xe]4f^15d^16s^2$ |
| 59 | Pr | Praseodymium | $[Xe]4f^25d^16s^2$ |
| 60 | Nd | Neodymium | $[Xe]4f^35d^16s^2$ |
| 61 | Pm | Promethium | $[Xe]4f^45d^16s^2$ |
| 62 | Sm | Samarium | $[Xe]4f^55d^16s^2$ |
| 63 | Eu | Europium | $[Xe]4f^65d^16s^2$ |
| 64 | Gd | Gadolinium | $[Xe]4f^75d^16s^2$ |
| 65 | Tb | Terbium | $[Xe]4f^85d^16s^2$ |
| 66 | Dy | Dysprosium | $[Xe]4f^95d^16s^2$ |
| 67 | Ho | Holmium | $[Xe]4f^105d^16s^2$ |
| 68 | Er | Erbium | $[Xe]4f^115d^16s^2$ |
| 69 | Tm | Thulium | $[Xe]4f^125d^16s^2$ |
| 70 | Yb | Ytterbium | $[Xe]4f^135d^16s^2$ |
| 71 | Lu | Lutetium | $[Xe]4f^145d^16s^2$ |
| 89 | Ac | Actinium | $[Rn]5d^17s^2$ |
| 90 | Th | Thorium | $[Rn]5f^26d^17s^2$ |
| 91 | Pa | Protactinium | $[Rn]5f^66d^17s^2$ |
| 92 | U | Uranium | $[Rn]5f^66d^57s^2$ |
| 93 | Np | Neptunium | $[Rn]5f^66d^57s^2$ |
| 94 | Pu | Plutonium | $[Rn]5f^76d^17s^2$ |
| 95 | Am | Americium | $[Rn]5f^76d^57s^2$ |
| 96 | Cm | Curium | $[Rn]5f^76d^57s^2$ |
| 97 | Bk | Berkelium | $[Rn]5f^96d^17s^2$ |
| 98 | Cf | Flerovium | $[Rn]5f^{10}6d^17s^2$ |
| 99 | Es | Einsteinium | $[Rn]5f^{11}6d^17s^2$ |
| 100 | Fm | Fermium | $[Rn]5f^{11}6d^27s^2$ |
| 101 | Md | Mendelevium | $[Rn]5f^{13}7s^2$ |
| 102 | No | Nobelium | $[Rn]5f^{14}6d^17s^2$ |
| 103 | Lr | Lawrencium | $[Rn]5f^{14}6d^17s^2$ |

Figure 13: The Periodic Table from scienzenotes.org

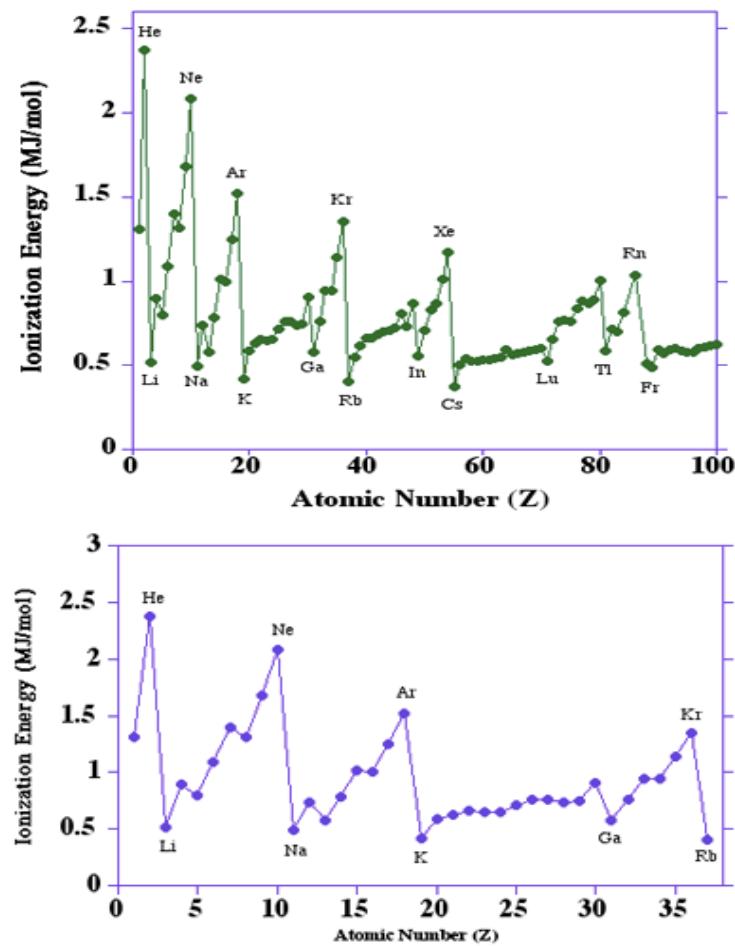


Figure 14: Ionization energy of atoms

potential for a one-electron atom). The exact form of ξ_i is not important for a qualitative discussion and we will not mention it further. Note that we now use the symbols $\hat{\mathbf{l}}_i$ and $\hat{\mathbf{s}}_i$ for the angular momentum operators (i.e. lower case letters), which correspond to the operators for each electron labelled by i . In this chapter the upper case letters $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ (no subscript i) are now used for the total orbital angular momentum and total spin angular momentum operators respectively of *all* the electrons, and the corresponding eigenvalues are given by upper case letters (e.g. L and S). It can be shown that L and S depend only on the angular momentum of unfilled subshells, with zero contribution from filled subshells. For a given electronic configuration (n and l) in the central field approximation (i.e. neglecting all perturbations) the energy of these (n, l) states does not depend on m and m_s i.e. there is degeneracy for a given n and l . The above two perturbations break this degeneracy but in different ways. If one of the perturbations is much stronger than the other then the strong perturbation should be considered first in determining the new eigenstates and eigenenergies. The second (and weaker) perturbation can then be treated by evaluating its effect on the new eigenstates and eigenenergies. It turns out that the $\hat{H}_{\text{non-central}}$ is more important in small Z atoms while $\hat{H}_{\text{spin-orbit}}$ is more important in large Z atoms (for the latter the larger electric field of the nucleus, given by taking $e \rightarrow Ze$ in eq. (5.3) for an atom with atomic number Z), increases in magnitude with increasing Z . The atoms where both perturbations are of comparable magnitude (e.g. Germanium $Z = 32$) are said to be in the *intermediate coupling regime*. We will only consider the case of $\hat{H}_{\text{non-central}}$ being the dominant perturbation, and for non-filled subshells (i.e. the valence electrons) in which the splitting of the energy levels is not negligible. It can be shown that these perturbations have negligible effect on the energy levels in filled subshells due to cancellations of the individual contributions from each electron.

The treatment of the above perturbations is made simpler by the fact that complete orbitals have a zero contribution to the energies due to the perturbations. This is most easily seen in the spin orbit case where for a full orbital for any specific value of l we will have two spin states with opposite spin whose net contribution to the energy will be zero. In the case of the non-central potential the effect is that the sum of the complete set of states associated with a particular orbital angular momentum number, i.e. for $l = 1$ the $m_l = 1, 0$, and -1 states, has a spherical charge distribution and thus does not contribute to the non-central potential. Thus when dealing with the two perturbations above we need only consider electrons in non-filled orbitals.

8.6 LS or Russell-Saunders Coupling

We will treat the case where the non-central potential is the dominant perturbation. This regime is called *LS* (or Russell-Saunders) coupling, and this name comes from the fact that L and S are **good quantum numbers** (see sec-

tion 2.10 for this definition) in this limit. In the central field approximation the Hamiltonian \hat{H} is taken to be eq. (8.2) but without the part \hat{H}_p ($= \hat{H}_{\text{non-central}}$):

$$\hat{H} = \hat{H}_{\text{central}} = \sum_i \left[\frac{-\hbar^2}{2m} \nabla_i^2 + U(|\mathbf{r}_i|) \right]. \quad (8.9)$$

If the non-central potential is the dominant perturbation then one has

$$\hat{H} = \hat{H}_{\text{central}} + \hat{H}_{\text{non-central}}, \quad (8.10)$$

where $\hat{H}_{\text{non-central}}$ is given by eq. (8.7), and \hat{H} is now the (full) gross Hamiltonian. It can be shown that the \hat{H} in eq. (8.10) commutes with the operators $\hat{L}^2, \hat{S}^2, \hat{J}^2 (= \hat{\mathbf{J}} \cdot \hat{\mathbf{J}})$ and \hat{J}_z , where $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$ and \hat{J}_z is the z -component of $\hat{\mathbf{J}}$, with respective quantum numbers J and M_J . Hence the good quantum numbers are L, S, J , and M_J , which together with the energy eigenvalues label the common eigenfunctions of these five commuting operators, denoted by $|E, L, S, J, M_J\rangle$ in ket notation. If the spin-orbit interaction is included one has

$$\hat{H} = \hat{H}_{\text{central}} + \hat{H}_{\text{non-central}} + \hat{H}_{\text{spin-orbit}}. \quad (8.11)$$

In perturbation theory one can then act $\hat{H}_{\text{spin-orbit}}$ (which we are assuming to be much smaller than $\hat{H}_{\text{non-central}}$) on the eigenfunctions of the Hamiltonian in eq. (8.10).

As mentioned earlier, the single particle eigenfunctions that result from solving the Schrödinger equation with the Hamiltonian in the central field approximation in eq. (8.9) are degenerate for specific n and l i.e. no dependence of the energy on m ($-l \leq m < l$) and $m_s = \pm 1/2$. The presence of $\hat{H}_{\text{non-central}}$ causes a splitting of these degenerate energy levels, which have different values of L and S (which are two of the good quantum numbers). This is a similar effect to what was seen in helium with the exchange energy which depended on S . Only the electrons in unfilled shells contribute to the good quantum numbers L and S , which we will now calculate with a specific example.

Taking carbon as an example, with electron configuration $1s^2 2s^2 2p^2$, there are 15 choices of m and m_s for the two electrons in the $2p$ subshell i.e. 15 possible states. This can be seen that there are 6 possible states in the p-orbital (three values of m , each associated with two spin values, and the total ways of distributing 2 electrons in 6 states $= {}^6C_2 = 15$). These 15 states are all degenerate in energy in the central-field approximation, but are non-degenerate when $\hat{H}_{\text{non-central}}$ and $\hat{H}_{\text{spin-orbit}}$ are included. If one of the two electrons were in a different subshell, say the $3p$ (an excited state), this calculation of L and S can be done by the rule for adding two angular momenta in eq. (5.12) applied to the two orbital angular momenta l_1 and l_2 (to obtain L) and to the two spin angular momenta s_1 and s_2 (to obtain S). Pauli's exclusion principle would not be a constraint because the two electrons always have different quantum numbers ($n = 2$ for $2p$ and $n = 3$ for $3p$), and thus all choices of m and m_s are allowed.

However, calculating the values of L and S when both electrons are in the unfilled $2p$ subshell is best done by tabulation due to Pauli's exclusion principle being a constraint. On doing this one has three possibilities for the values of L and S :

- i) $L = 0, S = 0$ (denoted by 1S , where S here means $L = 0$).
- ii) $L = 1, S = 1$ (denoted by 3P , where P means $L = 1$).
- iii) $L = 2, S = 0$ (denoted by 1D , where D means $L = 2$).

The notation 1S , 3P , 1D is the term symbol notation of section 5.3 (without the subscript for the J value), but now with capital letters to denote the value of the orbital angular momentum L , instead of lower case letters which denoted the value of l for a single electron. These three possible combinations of L, S are called **terms**, which together contain the 15 states. If only $\hat{H}_{\text{non-central}}$ were present (with no $\hat{H}_{\text{spin-orbit}}$) then the original 15 states with the same energy in the central field approximation are split into three groups (the above three *terms*) which have different energies. For each term the value of J can be obtained by applying the rule for adding two angular momenta in eq. (5.12) to L and S . This gives the five *levels*

$$^1S_0, ^3P_0, ^3P_1, ^3P_2, \text{ and } ^1D_2. \quad (8.12)$$

The term 3P (with $L = 1, S = 1$) gives values of $J = 0, 1, 2$, while the other terms 1S and 1D have $J = L$ (i.e. just one value of J) because $S = 0$. The spin-orbit interaction causes a splitting in energy of the $^3P_0, ^3P_1, ^3P_2$ levels because the correction to energy from this interaction has a part that depends on J . The terms 1S and 1D receive a change in energy from the spin-orbit interaction, but there is only one value of J for each of these terms and so there is no splitting into different levels (i.e. one term gives rise to only one level, meaning $^1S \rightarrow ^1S_0$ and $^1D \rightarrow ^1D_2$).

The above five levels still contain the original states 15 states, but there is still degeneracy. This degeneracy is lifted if an external magnetic field is applied (section 8.9), and all 15 states will have different energies. The levels 1D_2 and 3P_2 each contain 5 states ($M_J = 2, 1, 0, -1, -2$) because $J = 2$, while 3P_1 contains 3 states ($M_J = 1, 0, -1$) because $J = 1$. The levels 3P_0 and 1S_0 have one state each ($M_J = 0$) because $J = 0$. Hence there are still 15 states in all ($5+5+3+1+1$) which have been reclassified in terms of the good quantum numbers L, S, J, M_J .

8.7 Hund's Rules

Hund's rules allow the determination of the ground-state level of an atom within the LS coupling regime, assuming that the electron configuration (i.e. n, l are specified) is known and that there is only one incomplete subshell. Hund's rules do not apply to atoms with two incomplete subshells e.g. chromium ($3d^5, 4s^1$), copper, silver and gold. Moreover, Hund's rules cannot be applied to determine the order of the energies of excited states of the atom (they only give information

on which level is the ground state). The three rules are to be applied in the following order:

1. Within the constraints of the Pauli exclusion principle the level with maximum multiplicity (i.e. maximum S) is the lowest in energy.
2. Within the constraints discussed in the first rule and constraints from Pauli's exclusion principle the level with the maximum orbital momentum L is the lowest in energy.
3. For less than half-filled subshells the total angular momentum J of the lowest energy state is the minimum possible, i.e. $|L - S|$, and for more than half-filled subshells it is the maximum possible, i.e. $|L + S|$. For exactly half-filled subshells $L = 0$ and so $J = S$.

These rules require that every orbital (i.e. n, l and m are specified) in a subshell is singly occupied with one electron before any one orbital is doubly occupied, and all electrons in singly occupied orbitals have the same z -component of spin of $1/2$.

These rules follow from the discussion in section 8.6, i.e. the effect of spin on the non-central potential is dominant (apply rule 1), the effect of L on the non-central potential is the next most important effect (apply rule 2) and the spin-orbit coupling is the least important (apply rule 3).

Hund's rules mention in several places the maximum and minimum values of angular momenta. We will now look at a method for determining what these maxima and minima are. Hund's rules apply only if there is one unfilled subshell. The possible states associated with this subshell are depicted by a grid in which:

- i) Columns correspond to all the possible m values of the partially filled subshell;
- ii) The two rows correspond to the two possible value of $m_s = \pm 1/2$ for an electron.

For a p subshell the grid is as follows (where the orbitals are the columns):

| | 1 | 0 | -1 |
|----------------|---|---|----|
| $+\frac{1}{2}$ | | | |
| $-\frac{1}{2}$ | | | |

Due to the exclusion principle, at most one electron can go into each entry (box) in the grid. Rules 1 and 2 are realised by filling the boxes (with one electron denoted by a *) from the left of the top row across the row and then from the left of the bottom row to the right. This maximises S (top row with $m_s = 1/2$) and maximises L (fill from left to right). Thus for 4 electrons in a p subshell (e.g. oxygen, sulphur) we have

| | 1 | 0 | -1 |
|----------------|---|---|----|
| $+\frac{1}{2}$ | * | * | * |
| $-\frac{1}{2}$ | * | | |

We then count the total values associated with the rows, *i.e.* the number of marks in the top row minus the number of marks in the bottom row and halve (which is equivalent to summing the z -component of spins for the electrons, $S = \sum m_s$). This is the maximum value the total spin S could take. We now count the total value associated with the columns, *i.e.* total the number of marks in a column times the number at the top of the column (which is equivalent to summing the quantum number m for the electrons, $L = \sum m$). This is the maximum total orbital angular momentum. For the example above we have $S = 1$ $L = 1$. Hund's third rule is then applied to these two numbers, giving $J = 2$ (*i.e.* 1+1, because the subshell is more than half full). The term symbol for this lowest energy level is thus 3P_2 .

8.8 Optical Selection Rules

As has been done throughout the course we will treat optical transitions within the dipole approximation (E1). In this case the perturbation Hamiltonian for light-atom interactions is given by

$$\hat{H}_p = -e \sum_{i=1}^Z \mathbf{r}_i \cdot \mathbf{E} \quad (8.13)$$

This is a generalisation of eq. (7.12) to the case of Z electrons. As with the case of helium this Hamiltonian allows only one electron to change its configuration at a time. The selection rule for the electron undergoing the transition is $\Delta l = \pm 1$ and for all the other electrons $\Delta l = 0$ (no transition). This selection rule has a consequence for the total orbital angular momentum, which is given by the selection rule $\Delta L = 0, \pm 1$ but $L = 0$ cannot go to $L = 0$. The reason why $\Delta L = 0$ is allowed despite the fact that $\Delta l = \pm 1$ for the electron making the transition can be understood by considering the fact that L and l are lengths of vectors. It is thus possible that the vectors corresponding to L before and after the transition can be joined by a unit magnitude vector and still have the same magnitude as long as the initial state is not $L=0$.

The perturbation cannot change the spin (an intrinsic property) and this leads to the selection rule $\Delta S = 0$. The selection rule for J comes from the fact that $\mathbf{J} = \mathbf{L} + \mathbf{S}$. Thus $\Delta J = 0, \pm 1$ but $J = 0$ cannot go to $J = 0$ for the same reason given for L above. It also follows that $\Delta M_J = 0, \pm 1$. The J selection rules can also be derived directly from the fact that a photon is a spin 1 particle. Angular momentum conservation can be applied to the creation or annihilation of a photon leading to the selection rules given above.

8.9 Zeeman Effect

The Zeeman effect is the name given to the effect of an external (uniform) magnetic field on the energy levels of atoms, and is the magnetic analogue of the Stark effect. It was first observed in 1896. The *normal Zeeman effect* (an historical name, which had an explanation in the Bohr model) occurs in atoms with no net spin ($S = 0$) while the *anomalous Zeeman effect* (first observed in 1898 and could not be explained in the Bohr model) occurs in atoms with net spin ($S \neq 0$).

We will be treating the Zeeman effect only for atoms for which LS coupling is a good approximation. First we will discuss the derivation of the Hamiltonian for the perturbation due to an external magnetic field. We will then treat this in the weak field limit, in which the energy associated with the Zeeman effect is much smaller than the energy associated with spin-orbit coupling. The strong field limit (Paschen-Bach effect), in which the energy associated with the Zeeman effect is much bigger than the energy associated with spin-orbit coupling, will only be briefly mentioned.

As was mentioned in section (5.1), using a classical approach it can be shown that an electron orbiting a nucleus with orbital angular momentum $\hat{\mathbf{l}}$ has a magnetic moment $\hat{\boldsymbol{\mu}}_l$ given by

$$\hat{\boldsymbol{\mu}}_l = -\frac{e}{2m}\hat{\mathbf{l}} = \frac{-\mu_B}{\hbar}\hat{\mathbf{l}}, \quad (8.14)$$

when written using the Bohr magneton $\mu_B = e\hbar/2m$.

It was also shown in section (5.1) that a similar equation exists for spin angular momentum but with a factor g_s , which is approximately 2 ($g_s = 2.00232$ being a better approximation):

$$\hat{\boldsymbol{\mu}}_s = \frac{-g_s\mu_B}{\hbar}\hat{\mathbf{s}}. \quad (8.15)$$

Using the result that the interaction energy for a magnetic moment in a magnetic field is given by $-\boldsymbol{\mu} \cdot \mathbf{B}$ we obtain the perturbation Hamiltonian for the Zeeman effect:

$$\hat{H}_{\text{Zeeman}} = \frac{\mu_B}{\hbar} \sum_i (\hat{\mathbf{l}}_i + g_s \hat{\mathbf{s}}_i) \cdot \mathbf{B} = \frac{\mu_B}{\hbar} (\hat{\mathbf{L}} + g_s \hat{\mathbf{S}}) \cdot \mathbf{B}. \quad (8.16)$$

Without loss of generality we can assume that the magnetic field is along the z -axis. This converts the perturbation Hamiltonian to

$$\hat{H}_{\text{Zeeman}} = \frac{\mu_B}{\hbar} (\hat{L}_Z + g_s \hat{S}_Z) B. \quad (8.17)$$

In perturbation theory in the LS coupling regime, with $|L, S, J, M_J\rangle$ as the unperturbed eigenfunctions (not writing the energy eigenvalue E in the ket) the first-order correction $E^{(1)}$ (to be written below as $\Delta E^{(1)}$) is the expectation value:

$$\langle L, S, J, M_J | \hat{H}_{\text{Zeeman}} | L, S, J, M_J \rangle. \quad (8.18)$$

To evaluate this matrix element, which splits up into matrix elements with \hat{S}_Z and \hat{L}_Z , we will use a relationship which applies to both $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ and whose derivation can be found in textbooks:

$$\langle L, S, J, M_J | \hat{S}_Z | L, S, J, M_J \rangle = \frac{\langle L, S, J, M_J | \hat{\mathbf{S}} \cdot \hat{\mathbf{J}} | L, S, J, M_J \rangle}{\hbar J(J+1)} M_J. \quad (8.19)$$

$$\langle L, S, J, M_J | \hat{L}_Z | L, S, J, M_J \rangle = \frac{\langle L, S, J, M_J | \hat{\mathbf{L}} \cdot \hat{\mathbf{J}} | L, S, J, M_J \rangle}{\hbar J(J+1)} M_J. \quad (8.20)$$

In order to calculate the matrix element on the right the following relationships can be used:

$$\hat{S}^2 = \hat{\mathbf{S}} \cdot \hat{\mathbf{S}} = (\hat{\mathbf{J}} - \hat{\mathbf{L}}) \cdot (\hat{\mathbf{J}} - \hat{\mathbf{L}}) = \hat{J}^2 + \hat{L}^2 - 2\hat{\mathbf{J}} \cdot \hat{\mathbf{L}}. \quad (8.21)$$

$$\hat{L}^2 = \hat{\mathbf{L}} \cdot \hat{\mathbf{L}} = (\hat{\mathbf{J}} - \hat{\mathbf{S}}) \cdot (\hat{\mathbf{J}} - \hat{\mathbf{S}}) = \hat{J}^2 + \hat{S}^2 - 2\hat{\mathbf{J}} \cdot \hat{\mathbf{S}}. \quad (8.22)$$

These relations give expressions for $\hat{\mathbf{J}} \cdot \hat{\mathbf{L}}$ and $\hat{\mathbf{J}} \cdot \hat{\mathbf{S}}$, which equal $\hat{\mathbf{L}} \cdot \hat{\mathbf{J}}$ and $\hat{\mathbf{S}} \cdot \hat{\mathbf{J}}$ respectively, because the operators commute. The effect of \hat{J}^2 , \hat{L}^2 , and \hat{S}^2 on $|L, S, J, M_J\rangle$ brings out the eigenvalues. Thus the first-order correction to the energy levels of atomic states due to the Zeeman effect is:

$$\Delta E^{(1)} = g_J \mu_B M_J B. \quad (8.23)$$

where g_J is called Landé g factor, and is given by

$$g_J = \frac{J(J+1) + L(L+1) - S(S+1)}{2J(J+1)} + g_s \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)}. \quad (8.24)$$

This is often approximated by taking $g_s = 2$, which is true to approximately 0.1%. Within this approximation

$$g_J = \frac{3J(J+1) - L(L+1) + S(S+1)}{2J(J+1)}. \quad (8.25)$$

The energy correction depends on M_J (for the *normal Zeeman effect* one has $J = L$). As an example, if we have a multi-electron atom in which one of the levels is described by the term symbol 1D_2 then the five degenerate states associated with the different M_J values will be split into five non-degenerate states with energy shifts (relative to their unperturbed energy) of magnitude:

$$\Delta E^{(1)} = g_J \mu_B M_J B. \quad (8.26)$$

Transitions between these states are possible if the selection rules (including $\Delta M_J = 0, \pm 1$) are obeyed.

For a field of 1 Tesla the energy separation between neighbouring states is 67 meV. The above figure shows the splitting of the states as a function of the magnetic field in the weak field limit. As the magnetic field is increased

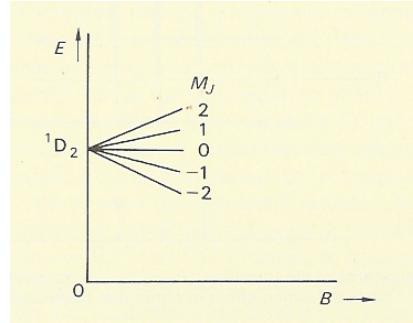


Figure 15: Splitting of the states as a function of the magnetic field in the weak field limit.

(the strong field limit) the perturbation energy due to the Zeeman effect will be greater than the energy level splitting caused by the spin-orbit correction. In this limit there is a different formula for ΔE . The explanation of the Zeeman effect (both *normal* and *anomalous*) was a great triumph for quantum mechanics, and (together with the explanation of the energy levels of helium) it provided further vindication of the theory in 1926 by solving a longstanding experimental puzzle.

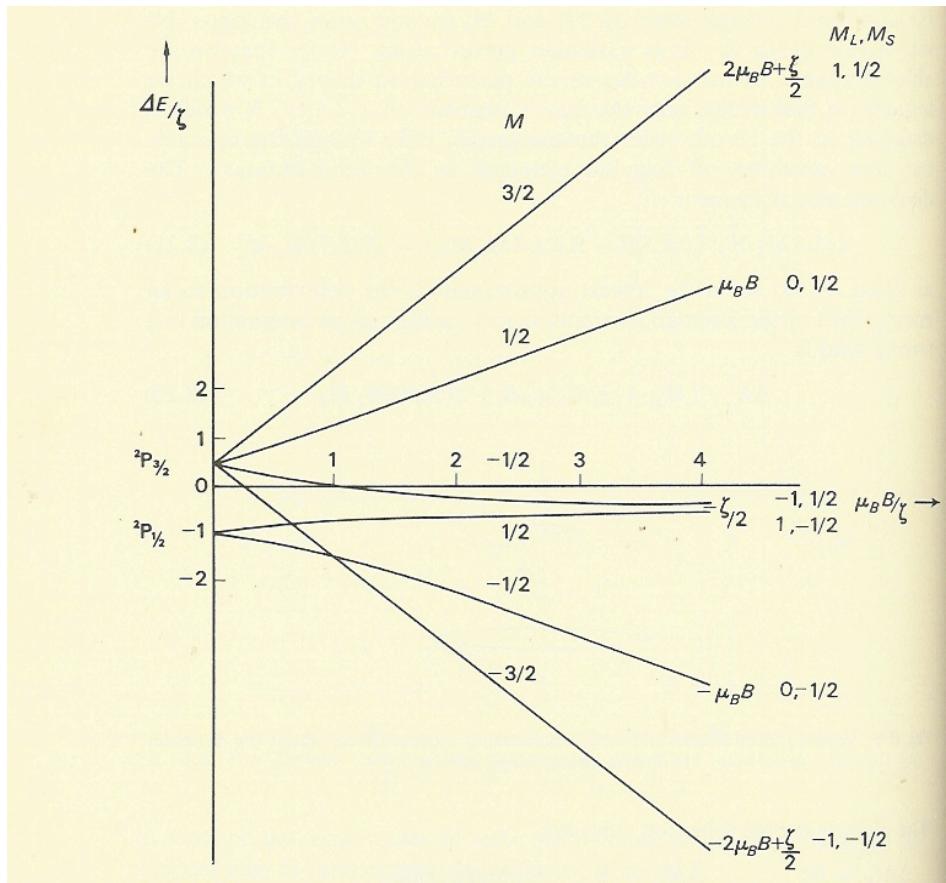


Figure 16: Energy as a function of magnetic field for two initially degenerate states of a multi-electron atom for magnetic fields in the weak, intermediate and strong regimes.