Atomic Physics

P. Ewart

Contents

1 Introduction

The structure of atoms and how light interacts with them is responsible for the appearance of the visible world. The small scale of atoms and the properties of nuclei and electrons required a new kind of mechanics to describe their behaviour. Quantum Mechanics was developed in order to explain such phenomena as the spectra of light emitted or absorbed by atoms. So far you have studied the physics of hydrogen and helium as illustrations of how to apply Quantum Theory. There was a time, a few seconds after the Big Bang, when the Universe consisted only of hydrogen and helium nucleii. It took another 300,000 years for atoms, as such, to form. Things, however, have moved on and the universe is now a much more interesting place with heavier and more complicated atoms. Our aim is now to understand *Atomic Physics*, not just to illustrate the mathematics of Quantum Mechanics. This is both interesting and important, for Atomic Physics is the foundation for a wide range of basic science and practical technology. The structure and properties of atoms are the basis of Chemistry, and hence of Biology. Atomic Physics underlies the study of Astrophysics and Solid State Physics. It has led to important applications in medicine, communications, lasers etc, as well as still providing a testing ground for Quantum Theory and its derivatives, Quantum Electrodynamics. We have learned most about atoms from the light absorbed or emitted when they change their internal state. So that is a good place to begin.

2 Radiation and Atoms

We will make extensive use of "models" in this course to help us get a feel for the physics. A favourite model for theorists is the "two-level atom" i.e. one with only two eigenstates ψ_1, ψ_2 with energy eigenvalues E_1, E_2 respectively ($E_2 > E_1$). The wave functions have, in general, a time dependence.

$$\psi_1 = \phi_1(x)e^{iE_1t/\hbar} \tag{1}$$

$$\psi_2 = \phi_2(x)e^{iE_2t/\hbar} \tag{2}$$

When the atom is perturbed it may be described by a wave function that is a linear combination of ψ_1 and ψ_2: $\psi = a\psi_1 + b\psi_2$ giving the probability amplitude. We observe, however, a probability density: the modulus squared; $|\psi|^2$. This will have a term

$$ab\psi_1\psi_2^* e^{i(E_1-E_2)t/\hbar} \tag{3}$$

This is a time oscillating electron density with a frequency ω_{12}:

$$e^{i(E_1-E_2)t/\hbar} = e^{-i\omega_{12}t} \tag{4}$$

So

$$E_2 - E_1 = \hbar\omega_{12} \tag{5}$$

This is illustrated schematically in figure 1.

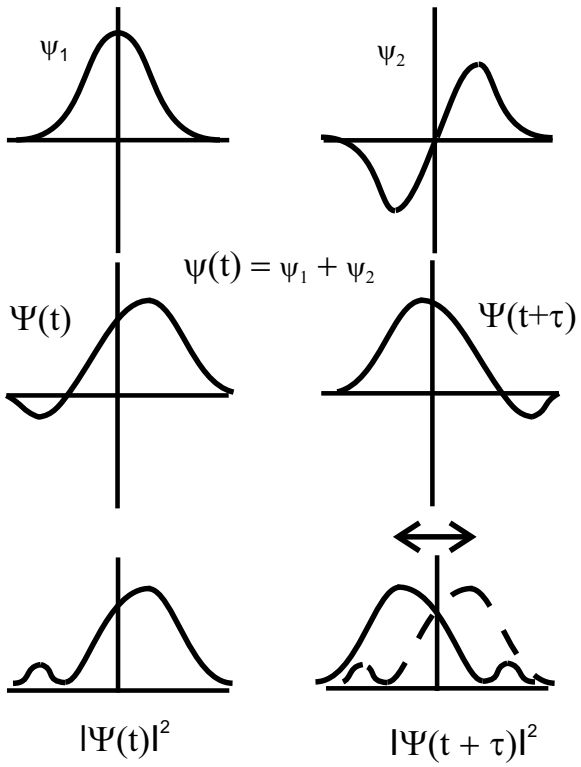


Figure 1: Evolution of the wavefunction of a system with time.

So the perturbation produces a charge cloud that oscillates in space – an oscillating dipole. This radiates dipole radiation. Whether or not we get a charge displacement or dipole will depend on the symmetry properties of the two states ψ_1, and ψ_2. The rules that tell us if a dipole with be set up are called "selection rules", a topic to which we will return later in the course.

2.1 Width and Shape of Spectral Lines

The radiation emitted (or absorbed) by our oscillating atomic dipole is not exactly monochromatic, i.e. there will be a range of frequency values for ω_{12}. The spectral line observed is broadened by one, or more, processes. A process that affects all the atoms in the same way is called "Homogeneous Broadening". A process that affects different individual atoms differently is "Inhomogeneous Broadening".

Examples of homogeneous broadening are lifetime (or natural) broadening or collision (or pressure) broadening. Examples of inhomogeneous broadening are Doppler broadening and crystal field broadening.

2.1.1 Lifetime Broadening

This effect may be viewed as a consequence of the uncertainty principle

$$\Delta E \Delta t \sim \hbar \tag{6}$$

Since $E = \hbar\omega$, $\Delta E = \hbar\Delta\omega$ and if the time uncertainty Δt is the natural lifetime of the excited atomic state, τ, we get a spread in frequency of the emitted radiation $\Delta\omega$

$$\Delta\omega\ \tau \sim 1 \quad \text{or} \quad \Delta\omega \sim \frac{1}{\tau} \tag{7}$$

The lifetime, τ, is a statistical parameter related to the time taken for the population of the excited state to decay to $1/e$ of its initial value. This exponential decay is reflected in the experimental decay of the amplitude $E(t)$ of the light wave emitted. The frequency (or power) spectrum of an exponentially decaying amplitude is a Lorentzian shape for the intensity as a function of frequency

$$I(\omega) \sim \frac{1}{(\omega - \omega_0)^2 + (1/\tau)^2} \tag{8}$$

The full width at half-maximum, FWHM, is then

$$\Delta\omega \sim 2\left(\frac{1}{\tau}\right) \tag{9}$$

A typical lifetime $\tau \sim 10^{-8}$ sec.

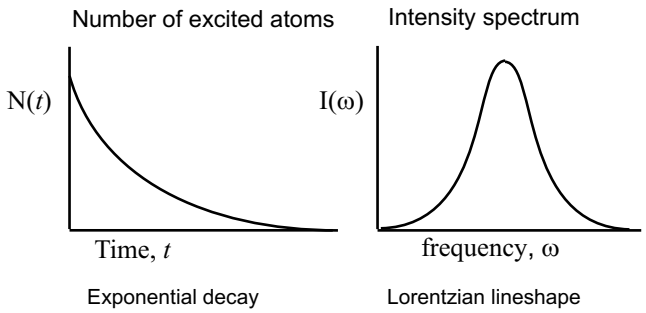


Figure 2: Decay of excited state population $N(t)$ leads to similar exponential decay of radiation amplitude, giving a Lorentzian spectrum.

2.1.2 Collision or Pressure Broadening

A collision with another atom while the atom is radiating (oscillating) disrupts the phase of the wave. The wave is therefore composed of various lengths of uninterrupted waves. The number of uninterrupted waves decays exponentially with a $1/e$ time τ_c, which is the mean time between collisions. At atmospheric pressure this is typically $\sim 10^{-10}$ sec. The exponential decay of the coherent oscillations again leads to a Lorentzian lineshape.

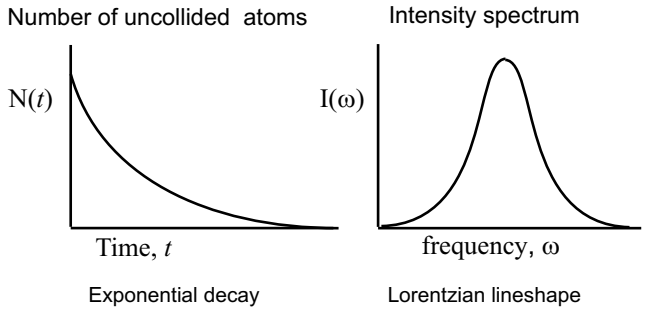


Figure 3: Decay in number of undisturbed atoms radiating leads to decay in amplitude of wave undisturbed by a phase changing collision. The associated frequency spectrum is again Lorentzian.

2.1.3 Doppler Broadening

Atoms in a gas have a spread of speeds given by the Maxwell-Boltzmann distribution. The Doppler shift of the light emitted is therefore different for the atoms moving at different speeds. There is then

a spread of Doppler shifted frequencies leading to a broadening of the spectral line. Since different atoms are affected differently this Doppler Broadening is "Inhomogeneous broadening". The Doppler shift is given by:

$$\omega = \omega_0 \left(1 \pm \frac{v}{c}\right) \tag{10}$$

The spread, or width of the line, is therefore

$$\Delta\omega_D \sim \omega_0 \frac{v}{c} \tag{11}$$

Since $\omega_0 \sim 10^{15}$ rad s^{-1} and $v \sim 10^2$ ms^{-1} we find

$$\Delta\omega_D \sim 2\pi 10^9 \text{ rad s}^{-1} \tag{12}$$

$$\Delta\nu \sim 10^9 \text{ Hz} \tag{13}$$

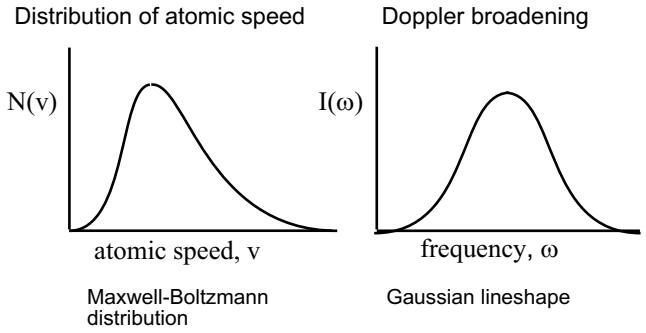


Figure 4: Maxwell-Boltzmann distribution of speeds and the associated Doppler broadening giving a Gaussian lineshape.

2.2 Atomic Orders of Magnitude

We have already started getting a feel for typical values of parameters associated with spectral lines, so now is a good time to do the same for some other aspects of atoms. It is important to have some idea of the orders of magnitude associated with atoms and their structure. We will be looking at atoms more complex than Hydrogen and Helium for which we can't solve the Schrodinger Equation exactly. We have to do the best we can with approximate solutions using Perturbation Theory. So we need to know when perturbation theory will give a reasonable answer. We will be particularly interested in the energy level structure of atoms so let's start there.

First a word about units. Atoms are small $\sim 10^{-10}$ m and their internal energies are small; $\sim 10^{-19}$ Joule. So a Joule is an inconvenient unit. We will use electron Volts, eV, which is $2 \sim 10^{-19}$ J. We have seen that atoms emit or absorb light in the UV or visible range – say ~ 500 nm. This must then be of the order of the energy changes or binding energy of the outer, most loosely bound electron.

Now a wavelength $\lambda \sim 500$ nm corresponds to a frequency, $\nu \sim 6 \times 10^{14}$ Hz. So from $E = h\nu$ we find $E \sim 40 \times 10^{-34} 10^{14} \sim 4 \times 10^{-19}$ J $\sim$ 2eV.

We could also check this using the size of an atom ($\sim 10^{-10}$ m) using the Uncertainty Principle.

$$\Delta p \, \Delta x \sim \hbar \tag{14}$$

$$\Delta p \sim p \sim \frac{\hbar}{\Delta x} \tag{15}$$

$$E = \frac{p^2}{2m} \sim \left(\frac{\hbar}{\Delta x}\right)^2 /(2m) \sim \text{ a few eV} \tag{16}$$

We can compare ~ 2 eV with thermal energy or $\sim kT$ i.e. the mean Kinetic energy available from heat, $\frac{1}{40}$ eV. This is not enough to excite atoms by collisions so atoms will mostly be in their ground state.

2.2.1 Other important Atomic quantities

Atomic size: Bohr radius

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} = 0.53 \times 10^{-10} \text{ m} \tag{17}$$

Ionisation Energy of Hydrogen

$$E_H = \frac{me^4}{(4\pi\epsilon_0)^2 2\hbar^2} = 13.6 \text{ eV} \tag{18}$$

Rydberg constant

$$R = \frac{E_H}{hc} = 1.097 \times 10^7 \text{ m}^{-1} \tag{19}$$

R is useful is relating wavelengths λ to energies of transitions since wavenumber $\bar{\nu} = \frac{1}{\lambda}$ in units of m^{-1}.

Fine structure constant

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \sim \frac{1}{137} \tag{20}$$

This is a dimensionless constant that gives a measure of the relative strength of the electromagnetic force. It is actually also the ratio of the speed v_e of the electron in the ground state of H to c, the speed of light. $\alpha = v_e/c$ and so it is a measure of when relativistic effects become important.

Bohr magneton

$$\mu_B = \frac{e\hbar}{2m} \sim 9.27 \times 10^{-24} \text{ JT}^{-1} \tag{21}$$

This is the basic unit of magnetic moment corresponding to an electron in a circular orbit with angular momentum $\hbar$, or one quantum of angular momentum.

As well as having orbital angular momentum the electron also has intrinsic spin and spin magnetic moment $\mu_S = 2\mu_B$. A proton also has spin but because its mass is ~ 2000 larger than an electron its magnetic moment is ~ 2000 times smaller. Nuclear moments are in general ~ 2000 times smaller than electron moments.

2.3 The Central Field Approximation

We can solve the problem of two bodies interacting with each other via some force e.g. a star and one planet with gravitational attraction, or a proton and one electron – the hydrogen atom. If we add an extra planet then things get difficult. If we add any more we have a many body problem which is impossible to solve exactly. Similarly, for a many electron atom we are in serious difficulty – we will need to make some approximation to simplify the problem. We know how to do Hydrogen; we solve the Schrödinger equation:

$$-\frac{\hbar^2}{2m}\nabla^2\psi - \frac{Ze^2}{4\pi\epsilon_0 r}\psi = E\psi \tag{22}$$

We can find zero order solutions – wave functions ψ that we can use to calculate smaller perturbations e.g. spin-orbit interaction.

The Hamiltonian for a many electron atom however, is much more complicated.

$$\hat{H} = \sum_{i=1}^{N} \left(-\frac{\hbar^2}{2m}\nabla_i^2 - \frac{Ze^2}{4\pi\epsilon_0 r_i} \right) + \sum_{i>j} \frac{e^2}{4\pi\epsilon_0 r_{ij}} \tag{23}$$

We ignore, for now, other interactions like spin-orbit. We have enough on our plate! The second term on the r.h.s. is the mutual electrostatic repulsion of the N electrons, and this prevents us from separating the equation into a set of N individual equations. It is also too large to treat as a small perturbation.

We recall that the hydrogen problem was solved using the symmetry of the central Coulomb field – the $1/r$ potential. This allowed us to separate the radial and angular solutions. In the many electron case, for most of the time, a major part of the repulsion between one electron and the others acts towards the centre. So we replace the $1/r$, hydrogen-like, potential with an effective potential due to the nucleus and the centrally acting part of the $1/r_{ij}$ repulsion term. We call this the Central Field $U(r)$. Note it will not be a $1/r$ potential. We now write the Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}_1 \tag{24}$$

$$\text{where } \hat{H}_0 = \sum_i \left\{ -\frac{\hbar^2}{2m}\nabla_i^2 + U(r_i) \right\} \tag{25}$$

$$\text{and } \hat{H}_1 = \sum_{i>j} \frac{e^2}{4\pi\epsilon_0 r_{ij}} - \sum_i \left\{ \frac{Ze^2}{4\pi\epsilon_0 r_i} + U(r_i) \right\} \tag{26}$$

$\hat{H}_1$ is the residual electrostatic interaction. Our approximation is now to assume $\hat{H}_1 << \hat{H}_0$ and then we can use perturbation theory.

The procedure is to start with just $\hat{H}_0$. Since this is a central potential the equations are separable. Solutions for the individual electrons will have the form:

$$\psi(n, l, m_l, m_s) = R'_{n,l}(r) Y_l^m(\theta, \phi) \chi(m_s) \tag{27}$$

So far, in this approximation the angular functions $Y_l^m(\theta, \phi)$ and spin functions $\chi(m_s)$ will be the same as for hydrogen. The radial functions will be different but they will have some features of the hydrogenic radial functions. The wave function for the whole atom will consist of antisymmetric products of the individual electron wave functions. The point is that we can use these zero order wavefunctions as a basis set to evaluate the perturbation due to the residual electrostatic interactions $\hat{H}_1$. We can then find new wavefunctions for the perturbed system to evaluate other, presumed smaller, perturbations such as spin-orbit interaction. When it comes to the test we will have to decide in any particular atom, which is the larger of the two perturbations – but more of that later.

2.4 The form of the Central Field

Calculating the form of $U(r_i)$ is a difficult problem. We can, however, get a feel for the answer in two limiting situations. Firstly, imagine one electron is taken far away from the nucleus, and the other electrons. What form of potential will it see? Clearly, there are then Z protons surrounded apparently by $Z - 1$ electrons in a roughly spherically symmetric cloud. Our electron then sees, at large r, a Hydrogen-like $1/r$ potential. Secondly, what happens when our electron goes "inside" the cloud of other electrons? Here, at small r, it sees Z protons and feels a Z/r potential.

The potential then looks like it has the form

$$U(r) = Z_{\text{eff}}(r) \frac{e^2}{4\pi\epsilon_0 r} \tag{28}$$

Where $Z_{\text{eff}}(r)$ varies from Z at small r to 1 at large r.

Note, for most of the important space $U(r)$ is not a $1/r$ potential.

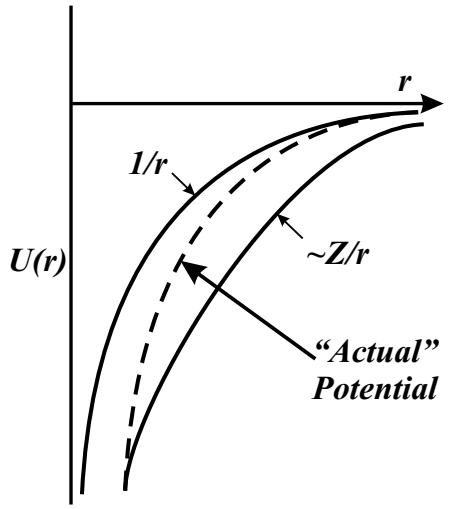


Figure 5: "Actual" potential experienced by an electron behaves like $1/r$ at large r and Z/r for small r. The intermediate r case is somewhere between the two.

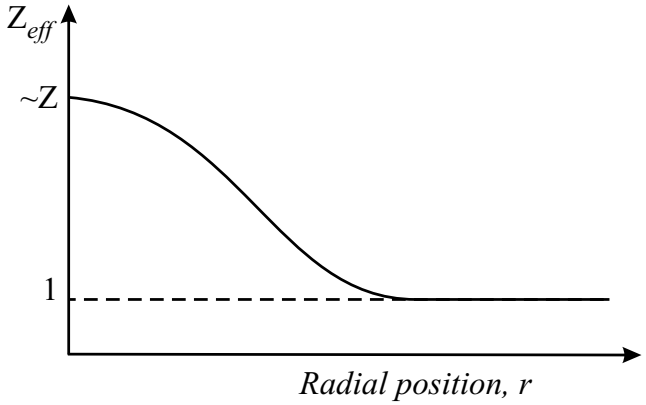


Figure 6: The variation of the effective Z that the electron feels as a function of r.

2.5 Finding the Central Field

The method of finding a good approximation to the Central Field is based on a "Self consistent field". This was first done by Hartree and his method was further developed by Fock so that we now call it the Hartree-Fock method. Hartree's basic idea was to find the best form of the potential by a series of iterations based on some initial guess. It works as follows.

(1) Guess a "reasonable" form for $U(r)$

(2) Put this guess for $U(r)$ into the Schrodinger equation and solve to find approximate wavefunctions

(3) Use these wave functions to calculate the charge distribution of the electrons

(4) Find the potential set up by this charge distribution and see if it matches the original guess for $U(r)$.

(5) Iterate this procedure until a self-consistent solution is found.

This procedure is ideally suited to numerical solution by a computer. The solutions for the wavefunctions will therefore be numerical i.e. they can't be expressed by nice analytical formulae. Hartree's orginal method used simple product wavefunctions which were not correctly antisymmetric. Each of the electrons in this model move in a potential set up by the other electrons and, as a result, the potentials are not the same for all the electrons and so the individual electron

states are not automatically orthogonal. The Hartree-Fock method uses an anti-symmetric basis set of wavefunctions. These are constructed using determinants where columns represent quantum states of individual electrons. This means that interchanging any column automatically changes the sign and makes the states correctly antisymmetric. The product states for each electron contain both space and spin functions. The potential is assumed to be the same for all the electrons. The potential is varied so as to produce the minimum energy for the system. This is the Variational Principle and has the same effect as finding a self-consistent field. The Hartree-Fock method is now the most commonly used way of finding wave functions and energy levels for many electron atoms. The wavefunctions produced are again numerical rather than analytic.

3 The Central Field Approximation

To recap, we have lumped together the Coulomb attraction of the Z protons in the nucleus with the centrally acting part of the mutual electrostatic repulsion of the electrons into $U(r)$. This Central Field goes like $1/r$ at large r and as Z/r at small r.

At in-between values of r things are more complicated – but more interesting! The important point is that we can, in many cases, treat the residual electrostatic interaction as a perturbation, so the Hamiltonian for the Schrödinger equation will be

$$\hat{H} = \hat{H}_0 + \hat{H}_1 + \hat{H}_2 + ... \tag{29}$$

$\hat{H}_1$ will be the perturbation due to residual electrostatic interactions, $\hat{H}_2$ that due to spin-orbit interactions. We will mostly deal with the cases where $\hat{H}_1 > \hat{H}_2$ but this won't be true for all elements. We can also add smaller perturbations, $\hat{H}_3$ etc due to, for example, interactions with external fields, or effects of the nucleus (other than its Coulomb attraction). The Central Field Approximation allows us to find solutions of the Schrödinger equation in terms of wave functions of the individual electrons:

$$\psi(n, l, m_l, m_s) \tag{30}$$

The zero-order Hamiltonian $\hat{H}_0$ due to the Central Field will determine the gross structure of the energy levels specified by n, l. The perturbation $\hat{H}_1$, residual electrostatic, will split the energy levels into different *terms*. The spin orbit interaction, $\hat{H}_2$ further splits the terms, leading to *fine structure* of the *energy levels*. Nuclear effects lead to *hyperfine structures* of the levels.

Within the approximation we have made, so far, the quantum numbers m_l, m_s do not affect the energy.The energy levels are therefore *degenerate* with respect to m_l, m_s. The values of m_l, m_s or any similar magnetic quantum number, specify the *state* of the atom.

There are $2l + 1$ values of m_l i.e. $2l + 1$ states and the energy level is said to be $(2l + 1)$-fold degenerate. The only difference between states of different m_l (or m_s) is that the axis of their angular momentum points in a different direction in space. We arbitrarily chose some z-axis so that the projection on this axis of the orbital angular momentum $\underline{l}$ would have integer number of units (quanta) of $\hbar$. (m_s does the same for the projection of the spin angular momentum $\underline{s}$ on the z-axis).

(Atomic physicists are often a bit casual in their use of language and sometimes use the words energy level, and state (e.g. Excited state or ground state) interchangeably. This practice is regrettable but usually no harm is done and it is unlikely to change anytime soon.)

3.1 The Physics of the Wave Functions

You will know (or you should know!) how to find the form of the wave functions $\psi(n, l, m_l, m_s)$ in the case of atomic hydrogen. In this course we want to understand the physics – the maths is done in the text books. (You may like to remind yourself of the maths after looking at the physics presented here!) Before we look at many electron atoms, we remind ourselves of the results for hydrogen.

3.1.1 Energy

The energy eigenvalues, giving the quantized energy levels are given by:

$$E_n = \left\langle \psi_{n,l,m_l} \left| \hat{H} \right| \psi_{n,l,m_l} \right\rangle \tag{31}$$

$$= -\frac{Z^2 m e^4}{(4\pi\epsilon_0)^2 \, 2\hbar^2 n^2} \tag{32}$$

Note that the energy depends only on n, the Principal quantum number. The energy does not depend on l – this is true ONLY FOR HYDROGEN! The energy levels are degenerate in l. We represent the energy level structure by a (Grotrian) diagram.

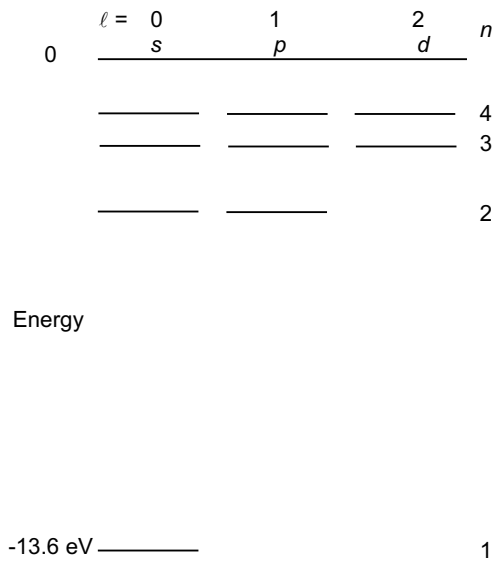


Figure 7: Energy level structure of hydrogen, illustrating how the bound state energy depends on n but not l.

For historical reasons, the states with $l = 0, 1, 2$ are labelled s, p, d. For $l = 3$ and above the labels are alphabetical f, g, h etc.

3.1.2 Angular Momentum

The wavefunction has radial and angular dependence. Since these vary independently we can write:

$$\psi_{n,l,m_l}(r,\theta,\phi) = R_{n,l}(r)Y_l^{m_l}(\theta,\phi) \tag{33}$$

$R_{n,l}(r)$ and $Y_l^{m_l}$ are radial and angular functions normalized as follows:

$$\int R_{n,l}^2(r)r^2dr = 1 \qquad \int \left|Y_l^{m_l}(\theta,\phi)\right|^2 d\Omega = 1 \tag{34}$$

The angular functions are eigenfunctions of the two operators $\hat{l}^2$ and $\hat{l}_z$:

$$\hat{l}^2 Y_l^{m_l}(\theta,\phi) = l(l+1)\hbar^2 Y_l^{m_l}(\theta,\phi) \tag{35}$$

$$\hat{l}_z Y_l^{m_l}(\theta,\phi) = m_l\hbar Y_l^{m_l}(\theta,\phi) \tag{36}$$

Where $l(l+1)$ and m_l are the eigenvalues of $\hat{l}^2$ and $\hat{l}_z$ respectively, such that

$$l = 0, 1, 2...(n-1) \qquad -l \leq m_l \leq l \tag{37}$$

The spin states of the electron can be included in the wave function by multiplying by a spin function χ_s which is an eigenfunction both of $\hat{s}^2$ and $\hat{s}_z$ with eigenvalues $s(s+1)$ and m_s respectively. ($s = 1/2$ and $m_s = \pm 1/2$)

Roughly speaking, the angular functions specify the shape of the electron distribution and the radial functions specify the size of the orbits.

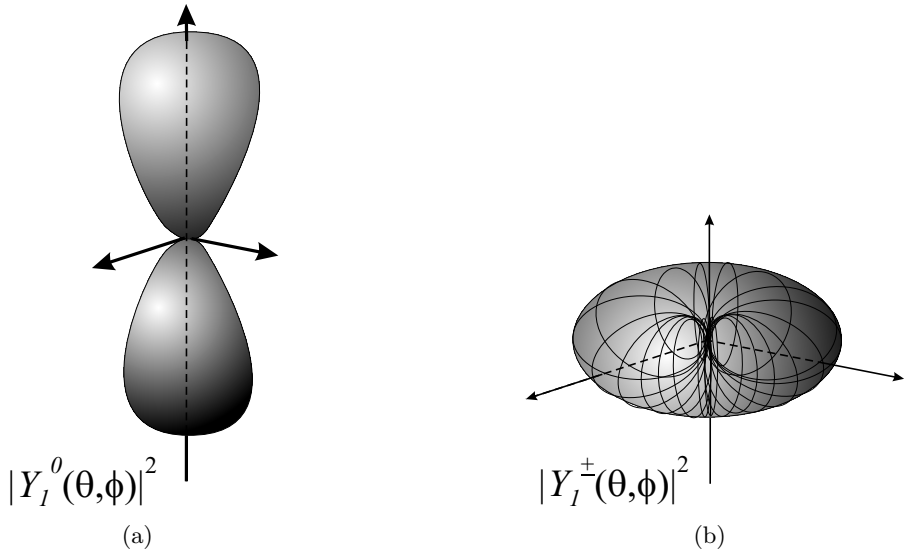
$l = 0$, s-states, are non-zero at the origin ($r = 0$) and are spherically symmetric. Classically they correspond to a highly elliptical orbit – the electron motion is almost entirely radial. Quantum mechanically we can visualise a spherical cloud expanding and contracting – breathing, as the electron moves in space.

$l \geq 1$. As l increases, the orbit becomes less and less elliptical until for the highest $l = (n-1)$ the orbit is circular. An important case (i.e. worth remembering) is $l = 1$ ($m_l = \pm 1, 0$)

$$Y_1^1 = -\left\{\frac{3}{8\pi}\right\}^{1/2} \sin\theta e^{i\phi} \tag{38}$$

$$Y_1^{-1} = +\left\{\frac{3}{8\pi}\right\}^{1/2} \sin\theta e^{-i\phi} \tag{39}$$

$$Y_1^0 = \left\{\frac{3}{4\pi}\right\}^{1/2} \cos\theta \tag{40}$$



(a) (b)

Figure 8: The angular functions, spherical harmonics, giving the angular distribution of the electron probability density.

If there was an electron in each of these three states the actual shapes change only slightly to give a spherically symmetric cloud. Actually we could fit two electrons in each l state provided they had opposite spins. The six electrons then fill the sub-shell ($l = 1$) In general, filled sub-shells are spherically symmetric and set up, to a good approximation, a central field.

As noted above the radial functions determine the size i.e. where the electron probability is maximum.

3.1.3 Radial wavefunctions

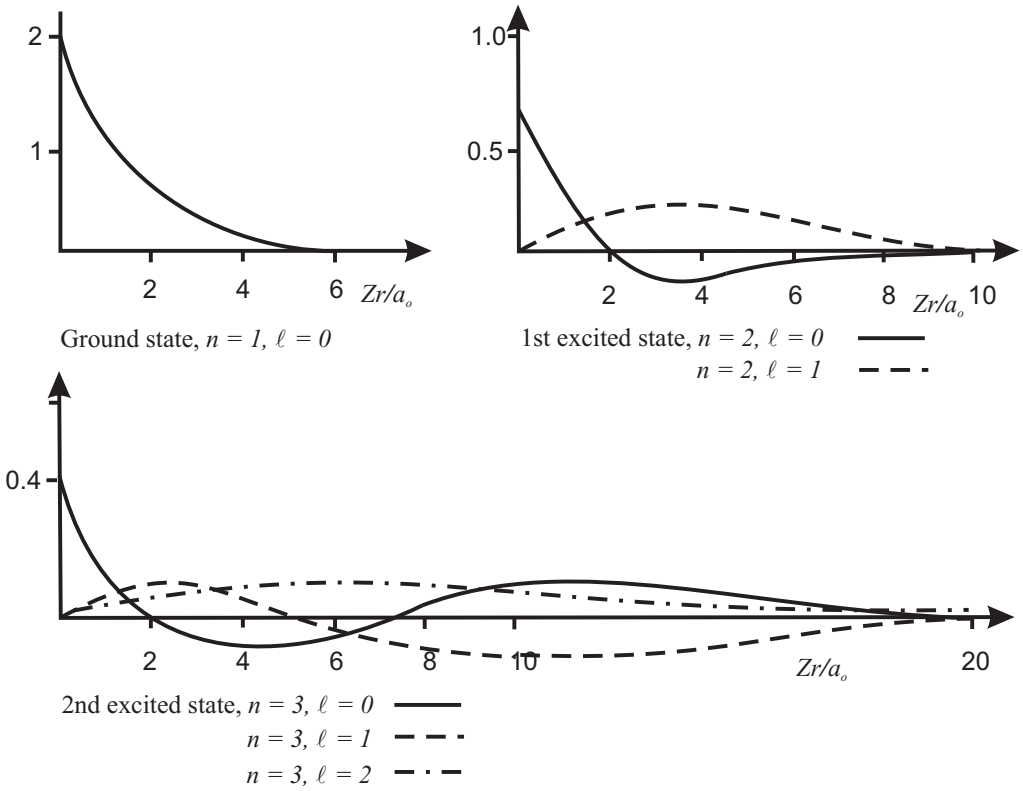


Figure 9: Radial functions giving the radial distribution of the probability amplitude.

NB: Main features to remember!

- $l = 0$, s-states do not vanish at $r = 0$.
- $l \neq 0$, states vanish at $r = 0$ and have their maximum probability amplitude further out with increasing l.
- The size, position of peak probability, scales with $\sim n^2$.
- The $l = 0$ function crosses the axis $(n - 1)$ times ie. has $(n - 1)$ nodes.
- $l = 1$ has $(n - 2)$ nodes and so on.
- Maximum $l = n - 1$ has no nodes (except at $r = 0$)

3.1.4 Parity

The parity operator is related to the symmetry of the wave function. The parity operator takes $r \rightarrow -r$. It is like mirror reflection through the origin.

$$\theta \rightarrow \pi - \theta \quad \text{and} \quad \phi \rightarrow \pi + \phi \tag{41}$$

If this operation leaves the sign of ψ unchanged the parity is *even*. If ψ changes sign the parity is *odd*. Some states are not eigenstates of the parity operator i.e. they do not have a definite parity. The parity of two states is an important factor in determining whether or not a transition between them is allowed by emitting or absorbing a photon. For a dipole transition the parity must change.

In central fields the parity is given by $(-1)^l$. So $l = 0, 2$ even states, s, d etc. are even parity and $l = 1, 3$, odd states are odd parity. Hence dipole transitions are allowed for s↔p or p↔d, but not allowed for s↔s or s↔d.

3.2 Multi-electron atoms

3.2.1 Electron Configurations

Before considering details about multi-electron atoms we make two observations on the Central Field Approximation based on the exact solutions for hydrogen.

Firstly, for a central field the angular wave functions will be essentially the same as for hydrogen. Secondly, the radial functions will have similar properties to hydrogen functions; specifically they will have the same number of nodes.

The energy level-structure for a multi-electron atom is governed by two important principles:
a) Pauli's Exclusion Principle and
b) Principle of Least Energy

Pauli forbids all the electrons going into the lowest or ground energy level. The Least Energy principle requires that the lowest energy levels will be filled first. The $n = 1, l = 0$ level can take a maximum of two electrons. The least energy principle then means the next electron must go into the next lowest energy level i.e. $n = 2$. Given that l can take values up to $n = 1$ and each l-value can have two values of spin, s, there are $2n^2$ vacancies for electrons for each value of n. Our hydrogen solutions show that higher n values lead to electrons having a most probable position at larger values of r. These considerations lead to the concept of shells, each labelled by their value of n.

Now look at the distribution of the radial probability for the hydrogenic wave functions $R_{n,l}(r)$. The low angular momentum states, especially s-states have the electrons spending some time inside the orbits of lower shell electrons. At these close distances they are no longer screened from the nucleus' Z protons and so they will be more tightly bound than their equivalent state in hydrogen. The energy level will depend on the degree to which the electron penetrates the core of inner electrons, and this depends on l. Therefore the energy levels are no longer degenerate in l. To this approximation m_l and m_s do not affect the energy. We can therefore specify the energy of the whole system by the energy of the electrons determined by quantum numbers n and l. Specifying n, l for each electron gives the electron configurations: nl^x where x indicates the number of electrons with a given nl.

3.2.2 The Periodic Table

As atomic science developed the elements were grouped together according to perceived similarities in their physical or chemical properties. Several different "periodic tables" were proposed. The one we all learned in our chemistry classes has been adopted since it has a basis in the fundamental structure of the atoms of each element. The physical and chemical properties, e.g. gas, solid, liquid, reactivity etc are seen to be consequences of atomic structure. In particular they arise largely from the behaviour of the outer shell – or most loosely bound – electrons.

Using the Pauli Principle and the Least energy Principle we can construct the configurations of the elements in their ground states:-

H:	$1s$		
He:	$1s^2$		
Li:	$1s^2$	$2s$	
Be:	$1s^2$	$2s^2$	
B:	$1s^2$	$2s^2 2p$	
C:	$1s^2$	$2s^2 2p^2$	
...			
Ne:	$1s^2$	$2s^2 2p^6$	
Na:	$1s^2$	$2s^2 2p^6$	$3s$

Everything proceeds according to this pattern up to Argon: $1s^2 2s^2 2p^6 3s^2 3p^6$

At this point things get a little more complicated. We expect the next electron to go into the 3d sub-shell. As we have seen, however, a 3d electron is very like a 3d electron in hydrogen: it spends most of its time in a circular orbit outside the inner shell electrons. An electron in a 4s state however goes relatively close to the nucleus, inside the core, and so ends up more tightly bound - i.e. lower energy, than the 3d electron. So the next element, Potassium, has the configuration

K:	$1s^2 2s^2 2p^6$	$3s^2 3p^6$	$4s$
Ca:			$4s^2$

The 3d shell now begins to fill

Sc: $1s^2 2s^2 2p^6$ $3s^2 3p^6 3d$ $4s^2$

The 3d and 4s energies are now very similar and at Chromium a 3d electron takes precedence over a 4s electron

Va:	$3s^2 3p^6$	$3d^3 4s^2$
Cr:		$3d^5 4s$
Mn:		$3d^5 4s^2$

As the 3d shell fills up, the successive elements, the transition elements, have interesting properties as a result of the partially filled outer shell – but this isn't on the syllabus!

There are, however, two features of the periodic table that are worth noting as consequences of the Central Field Approximation.

Rare gases These elements are chemically inert and have high ionization potentials (the energy needed to pull off a single electron). This is not because, as is sometimes (often) stated, that they have "closed shells". They don't all have closed shells i.e. all states for each n are full. They all do have filled s and p sub-shells

He:	$1s^2$	Kr:	$(\dots)4s^2 4p^6$
Ne:	$1s^2 2s^2 2p^6$	Xe:	$(\dots)5s^2 5p^6$
A:	$1s^2 2s^2 2p^6 3s^2 3p^6$	Rn:	$(\dots)6s^2 6p^6$

As we noted earlier this leads to a spherically symmetric charge distribution. Since electrons are indistinguishable all the electrons take on a common wave function. The point is that this results in a higher binding energy for each one of the electrons. So it is harder for them to lose an electron in a chemical bond – they are chemically inert and have high ionization energies.

Alkalis These are the next elements to the rare gases and have one electron outside the full (sp) sub-shells. This outer, or valence, electron therefore moves in a hydrogen like central potential. The electron is generally well-screened by $(Z-1)$ inner electrons from the nucleus and is easily lost to a chemical bond (ionic or co-valent). They are chemically reactive and have low ionization energies which don't change much from one alkali to another.

3.3 Gross Energy Level Structure of the Alkalis: Quantum Defect

As noted already, the single outer electron in an alkali moves in a potential that is "central" to an excellent approximation. We are ignoring, at this stage, any other perturbations such as spin-orbit interaction. As an example we consider sodium. The ground stage (lowest energy level) has the configuration:

(closed shells) 3s.

We know the 3s penetrates the core (inner shells) and is therefore more tightly bound – lower energy – than a 3s electron would be in Hydrogen. When the electron is excited, say to 3p, it penetrates the core much less and in a 3d state its orbit is virtually circular and very close the $n = 3$ level of Hydrogen.

The higher excited states $n > 3$ will follow a similar pattern. The thing to notice – and this turns out to be experimentally useful – is that the degree of core penetration depends on l and very little on n. As a result the deviation from the hydrogenic energy level is almost constant for a given l as n increases.

The hydrogen energy levels can be expressed as

$$E_n = -\frac{R}{n^2} \tag{42}$$

For alkalis, and to some degree for other atoms too, the excited state energies may be expressed as

$$E_n = -\frac{R}{(n^*)^2} \tag{43}$$

Where n^* is an effective quantum number. n^* differs from the equivalent n-value in hydrogen by $\delta(l)$ i.e. $n^* = n - \delta(l)$.

$\delta(l)$ is the quantum defect and depends largely on l only. It is found empirically, and it can be shown theoretically, to be independent of n. Thus all s-states will have the same quantum defect $\delta(s)$; all p-states will have the same $\delta(p)$, etc and $\delta(s) > \delta(p) > \delta(d)$.

	$\delta(s)$	≈ 1.34
For Na:	$\delta(p)$	≈ 0.88
	$\delta(d)$	≈ 0

For heavier alkalis, the $\delta(l)$ generally increases as the core is less and less hydrogen-like. The ionization potential however is almost constant as noted previously. Although the increase in Z leads to stronger binding than in hydrogen the ground state electron starts in a higher n, and this almost exactly off-sets the increased attraction of the heavier nucleus.

Everyone knows that the discrete wavelengths of light emitted, or absorbed, by an atom are discrete because they arise from transitions between the discrete energy levels. But how do we know which energy levels? This is where the quantum effect is very useful.

We also need to remember that a transition involves a change in l of $\pm$ 1 ($\Delta l = \pm 1$). A transition from a lower state will therefore take the atom to an energy level with angular momentum differing

by ± 1 from the initial level. For simplicity, so that the general idea becomes clear, we consider transitions from the ground state. If the ground state (level) has $l = 0$ then there will then be only a single set of transitions to levels with angular momentum one unit larger i.e. $l = 1$. The spectrum of absorption lines will consist of a series of lines of increasing wavenumber (energy) corresponding to transitions to excited states with increasing principal quantum number n. Since these excited levels all have the same value of l they should have the same value of quantum defect. So if we can identify lines with the same quantum defect they will belong to the same series i.e. have the same angular momentum quantum number l. The procedure is then as follows.

(1) Measure the wavelength of each absorption line in the series, λ_n.

(2) Calculate the corresponding wavenumber $\overline{\nu}_n = \lambda_n^{-1}$

(3) Estimate the value of the series limit $\overline{\nu}_\infty$. This corresponds to the ionization potential or Term value for the ground state, T_{n_o}.

(4) Find the Term value, T_n, or binding energy of an excited level from the difference between $\overline{\nu}_\infty$ and $\overline{\nu}_n$:

$$T_{n_i} = T_{n_o} - \overline{\nu}_n$$

(5) The Term value is given by the hydrogenic formula:

$$T_{n_i} = \frac{R_\infty}{(n_i^*)^2}$$

where n_i^* is the effective quantum number given by:

$$n_i^* = n - \delta(l)$$

and $\delta(l)$ is the quantum defect.

(6) Finally we plot $\delta(l)$ against T_{n_i}. This should be a straight line indicating a constant value of $\delta(l)$. Deviation from a straight line indicates that we have not estimated $\overline{\nu}_\infty$ (i.e. T_{n_o}) correctly.

(7) The procedure is then to try a different value of T_{n_o} until we obtain a straight line. This process is ideally suited to computation by a program that minimizes the deviation from a straight line for various values of T_{n_o}.

Once we have found the most accurate value of T_{n_o} from the spectrum we can put the energy levels on an absolute energy scale. We have then determined the energy level structure of the atom for this set of angular momentum states. The procedure can be carried out using data from emission spectra. In this case a range of different series will be be provided from the spectrum. It remains to find the quantum defects for each series of lines and assign values of l to each series of levels. From the discussion above the quantum defects will decrease for increasing l.

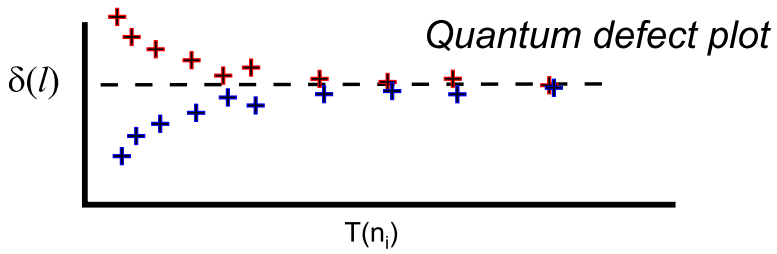


Figure 10: Plot of quantum defect $\delta(l)$ against Term value, showing effect of choosing T_{n_o} either too large or too small. The correct value of T_{n_o} yields a horizontal plot.

4 Corrections to the Central Field: Spin-Orbit interaction

The Central Field Approximation gives us a zero-order Hamiltonian $\hat{H}_0$ that allows us to solve the Schrödinger equation and thus find a set of zero-order wavefunctions ψ_i. The hope is that we can treat the residual electrostatic interaction (i.e. the non-central bit of the electron-electron repulsion) as a small perturbation, $\hat{H}_1$. The change to the energy would be found using the functions ψ_i.

The residual electrostatic interaction however isn't the only perturbation around. Magnetic interactions arise when there are moving charges. Specifically we need to consider the magnetic interaction between the magnetic moment due to the electron spin and the magnetic field arising from the electron orbit. This field is due to the motion of the electron in the electric field of the nucleus and the other electrons. This spin-orbit interaction has an energy described by the perturbation $\hat{H}_2$. The question is: which is the greater perturbation, $\hat{H}_1$ or $\hat{H}_2$?

We may be tempted to assume $\hat{H}_1 > \hat{H}_2$ since electrostatic forces are usually much stronger than magnetic ones. However by setting up a Central Field we have already dealt with the major part of the electrostatic interaction. The remaining bit may not be larger than the magnetic spin-orbit interaction. In many atoms the residual electrostatic interaction, $\hat{H}_1$, does indeed dominate the spin-orbit. There is, however, a set of atoms where the residual electrostatic repulsion is effectively zero; the alkali atoms. In the alkalis we have only one electron orbiting outside a spherically symmetric core. The central field is, in this case, an excellent approximation. The spin-orbit interaction, $\hat{H}_2$ will be the largest perturbation – provided there are no external fields present. So we will take the alkalis e.g. Sodium, as a suitable case for treatment of spin-orbit effects in atoms. You have already met the spin-orbit effect in atomic hydrogen, so you will be familiar with the quantum mechanics for calculating the splitting of the energy levels. There are, however, some important differences in the case of more complex atoms. In any case, we are interested in understanding the physics, not just doing the maths of simple systems. In what follows we shall first outline the physics of the electron's spin magnetic moment $\underline{\mu}$ interacting with the magnetic field, $\underline{B}$, due to its motion in the central field (nucleus plus inner shell electrons). The interaction energy is found to be $\underline{\mu} \cdot \underline{B}$ so the perturbation to the energy, ΔE, will be the expectation value of the corresponding operator $\langle \hat{\underline{\mu}} \cdot \underline{B} \rangle$.

We then use perturbation theory to find ΔE. We will not, however, be able simply to use our zero order wavefunctions $\psi_0(n, l, m_l, m_s)$ derived from our Central Field Approximation, since they are degenerate in m_l, m_s. We then have to use degenerate perturbation theory, DPT, to solve the problem. We won't have to actually do any complicated maths because it turns out that we can use a helpful model – the Vector Model, that guides us to the solution, and gives some insight into the physics of what DPT is doing.

4.1 The Physics of Spin-Orbit Interaction

What happens to a magnetic dipole in a magnetic field? A negatively charged object having a moment of inertia I, rotating with angular velocity $\underline{\omega}$, has angular momentum, $\mathrm{I}\underline{\omega} = \underline{\lambda}\hbar$. The energy is then $E = \frac{1}{2}\mathrm{I}\omega^2$

We suppose the angular momentum vector $\underline{\lambda}\hbar$ is at an angle θ to the z-axis. The rotating charge has a magnetic moment

$$\underline{\mu} = -\gamma\underline{\lambda}\hbar \tag{44}$$

The sign is negative as we have a negative charge. γ is known as the gyromagnetic ratio. (Classically $\gamma = 1$ for an orbiting charge, and $\gamma = 2$ for a spinning charge).

If a constant magnetic field $\underline{B}$ is applied along the z-axis the moving charge experiences a force – a torque acts on the body producing an extra rotational motion around the z-axis. The axis of rotation of $\underline{\lambda}\hbar$ precesses around the direction of $\underline{B}$ with angular velocity ω'. The angular motion of our rotating change is changed by this additional precession from ω to $\omega + \omega'\cos(\theta)$. If the angular momentum $\underline{\lambda}\hbar$ was in the opposite direction then the new angular velocity would be $\omega - \omega'\cos(\theta)$.

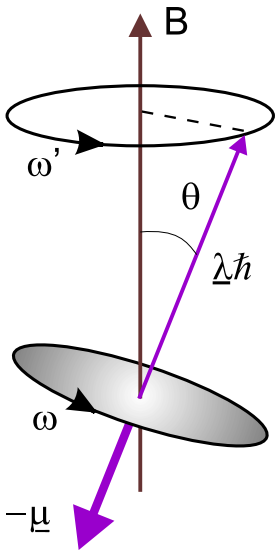


Figure 11: Illustration of the precession (Larmor precession) caused by the torque on the magnetic moment $\underline{\mu}$ by a magnetic field $\underline{B}$.

The new energy is, then

$$E' = \frac{1}{2}\mathrm{I}(\omega \pm \omega' \cos\theta)^2 \tag{45}$$

$$= \frac{1}{2}\mathrm{I}\omega^2 + \frac{1}{2}\mathrm{I}(\omega' \cos\theta)^2 \pm \mathrm{I}\omega\omega' \cos\theta \tag{46}$$

We now assume the precessional motion ω' to be slow compared to the original angular velocity ω. $\omega' << \omega$ so $(\omega' \cos\theta)^2 << \omega^2$ and we neglect the second term on the r.h.s. The energy change $\Delta E = E' - E$ is then

$$\Delta E = \mathrm{I}\omega\omega' \cos\theta \tag{47}$$

$$= \lambda\hbar\omega' \cos\theta \tag{48}$$

Now the precessional rate ω' is given by Larmor's Theorem

$$\omega' = -\gamma B \tag{49}$$

So

$$\Delta E = -\gamma\lambda\hbar B \cos\theta \tag{50}$$

Hence

$$\Delta E = -\underline{\mu} \cdot \underline{B} \tag{51}$$

So $-\underline{\mu} \cdot \underline{B}$ is just the energy of the precessional motion of $\underline{\mu}$ in the $\underline{B}$-field.

(Note that $\lambda \cos\theta\hbar$ is the projection of the angular momentum on the z-axis ($\underline{B}$-direction)

Quantum mechanically this must be quantized in integer units of $\hbar$ i.e. $\lambda \cos\theta = m$, the magnetic quantum number. So the angular momentum vector $\underline{\lambda}\hbar$ can take up only certain discrete directions relative to the $\underline{B}$-field. This is the space or orientation quantization behind the Stern-Gerlach experiment.)

4.2 Finding the Spin-Orbit Correction to the Energy

Now that we know the perturbation Hamiltonian $\hat{H}_2 = -\hat{\underline{\mu}} \cdot \underline{B}$ we will use perturbation theory to find the expectation value of the change in energy; $\Delta E = -\left\langle \hat{\underline{\mu}} \cdot \underline{B} \right\rangle$ The calculation proceeds in the following steps;

- Find the $\underline{B}$-field due to the electron motion
- Set up the operator for $-\hat{\underline{\mu}} \cdot \underline{B}$
  This will have a dependence on radial position, r, and the spin and orbital angular momenta $\underline{s}$ and $\underline{l}$
- Different relative orientations of $\underline{s}$ and $\underline{l}$ will have different energies of interaction and correspond to different total angular momentum $\underline{j} = \underline{l} + \underline{s}$
- Different values of $\underline{j}$ will thus have different energy resulting in a splitting as $j = l \pm 1/2$

4.2.1 The B-Field due to Orbital Motion

The relativistic effect of an electron moving with velocity $\underline{v}$ in an electric field $\underline{E}$ is to set up a $\underline{B}$ field given by:

$$\underline{B} = -\frac{\underline{v} \times \underline{E}}{c^2} \tag{52}$$

We re-express this using the momentum $\underline{p} = m\underline{v}$

$$\underline{B} = -\frac{1}{mc^2}\underline{p} \times \underline{E} \tag{53}$$

In our central field $\underline{E}$ is purely radial: Writing $\underline{E} = |\underline{E}|\frac{\underline{r}}{|\underline{r}|}$ then we have

$$\underline{B} = -\frac{1}{mc^2}\underline{p} \times \underline{r}\frac{|\underline{E}|}{|\underline{r}|} \tag{54}$$

Now $\underline{r} \times \underline{p} = \underline{l}$, angular momentum, so $\underline{p} \times \underline{r} = -\underline{l}$

$$\underline{B} = \frac{1}{mc^2}\frac{|\underline{E}|}{|\underline{r}|}\hat{\underline{l}} \tag{55}$$

(We shall put operator signs only on angular momentum operators.)

We can express $|\underline{E}|$ in terms of the electrical potential $\phi(r)$

$$|\underline{E}| = \frac{\partial \phi(r)}{\partial r} \tag{56}$$

and the Central Field potential energy would be

$$e\phi(r) = U(r) \tag{57}$$

Then

$$|\underline{E}| = -\frac{1}{e}\frac{\partial U(r)}{\partial r} \tag{58}$$

So finally we have

$$\underline{B} = \frac{1}{emc^2}\frac{1}{|\underline{r}|}\frac{\partial U(r)}{\partial r}\hat{\underline{l}} \tag{59}$$

NB: $U(r)$ = Potential energy of the electron in the Central Field.

In the rest frame of the electron, this is the field it experiences as it apparently "sees" the nucleus orbiting it at radius r.

4.2.2 The Energy Operator

The intrinsic magnetic moment of the electron, due to its spin is

$$\underline{\mu}_s = -g_s \frac{\mu_B}{\hbar} \underline{\hat{s}} \tag{60}$$

g_s is the gyro-magnetic ratio for a spinning charge. Relativistic quantum theory (Dirac) gives $g_s = 2$, but QED corrections give a value slightly larger than 2. The energy operator for the spin-orbit interaction is

$$-\underline{\mu}_s \cdot \underline{B} = g_s \frac{\mu_B}{\hbar} \frac{1}{emc^2} \frac{1}{r} \frac{\partial U(r)}{\partial r} \underline{\hat{s}} \cdot \underline{\hat{l}} \tag{61}$$

For many-electron atoms we don't know exactly what $U(r)$ is. The important thing to note, however, is that

$$-\underline{\mu}_s \cdot \underline{B} \propto \frac{1}{r} \frac{\partial U(r)}{\partial r} \underline{\hat{s}} \cdot \underline{\hat{l}} \tag{62}$$

This is a positive quantity with a set of values determined by the relative orientation of $\underline{\hat{s}}$ and $\underline{\hat{l}}$. For a hydrogen-like atom we can put $U(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}$ and so:

$$-\underline{\mu}_s \cdot \underline{B} = \frac{\mu_0}{4\pi} 2Z g_s \mu_B^2 \frac{1}{r^3} \frac{\underline{\hat{s}} \cdot \underline{\hat{l}}}{\hbar^2} \tag{63}$$

This result still isn't quite right. We have not taken account properly of the relativistic transformation between the rest frames of the electron and the nucleus. We actually observe the electron orbiting in the rest frame of the nucleus. This is the Thomas Precession (its derivation is not in the syllabus) and has the effect of dividing the result by a factor 2.

Hence

$$-\underline{\mu}_s \cdot \underline{B} = \frac{\mu_0}{4\pi} Z g_s \mu_B^2 \frac{1}{r^3} \frac{\underline{\hat{s}} \cdot \underline{\hat{l}}}{\hbar^2} \tag{64}$$

Our task now is to find the expectation values of the operators $1/r^3$ and $\underline{\hat{s}} \cdot \underline{\hat{l}}$.

4.2.3 The Radial Integral

The expectation value of $1/r^3$ is given by the integral: $\langle \psi_i | \, 1/r^3 \, | \psi_i \rangle$. The position of the electron doesn't have a definite value. The wave functions ψ_i are not eigenfunctions of the operator $1/r^3$.

The position r is not a constant of the motion. So we have to evaluate the integral

$$\left\langle \frac{1}{r^3} \right\rangle = \int_0^\infty R_{n,l}^2(r) \frac{1}{r^3} r^2 dr \tag{65}$$

Using hydrogen radial functions we find:

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{Z^3}{n^3 a_0^3 l(l+1/2)(l+1)} \tag{66}$$

Notice that this gets smaller rapidly with increasing l. So it is wrong to conclude from the $< \underline{\hat{s}} \cdot \underline{\hat{l}} >$ bit of the operator that the spin-orbit effect increases with l.

The physics here is that lower l electrons move closer to the inner regions of the central field. Here they experience stronger electric fields. Add to this that they get faster as they approach the near point to the nucleus in their elliptical orbit, then the magnetic interaction increases for lower l. For $l = 0$, however, the $\underline{\hat{s}} \cdot \underline{\hat{l}}$ term is zero and there is no spin-orbit effect at all! So s-terms are not split. Our result shows that the spin-orbit splitting is larger for high Z, low n and low l (but zero for $l = 0$!)

4.2.4 The Angular Integral: Degenerate Perturbation Theory

We now need to evaluate $< \hat{\underline{s}} \cdot \hat{\underline{l}} >$. Can we use our zero-order wave functions $\psi_i(n, l, m_l, m_s)$ to do this? The answer is, sadly, no. The problem is that these eigenstates are degenerate in m_l, and m_s. Why is that a problem? Recall first the result of non-degenerate perturbation theory. The first order energy shift induced by a perturbation $\hat{H}'$ is the expectation value:

$$\Delta E = \langle \psi_i | \hat{H}' | \psi_i \rangle \tag{67}$$

If, however, there are degenerate eigenstates, then we can use this result only if

$$\langle \psi_i | \hat{H}' | \psi_j \rangle = 0 \quad \text{for all } j \neq i \tag{68}$$

i.e. the perturbation matrix is diagonal. To see why this is necessary we need to recall that Perturbation Theory works so long as the state of the atom is not changed too much by the perturbation. In this case we can use the unperturbed wavefunctions ψ to find the expectation value of the change in energy and the small change in the wavefunction. Now recall the change in the wavefunction $\Delta\psi$ is given by:

$$\Delta\psi_i = \sum_{j \neq i} \frac{\langle \psi_i | \hat{H}' | \psi_j \rangle}{E_i - E_j} \psi_j \tag{69}$$

All the terms in the sum must be small or else the wavefunctions will be changed too much and Perturbation Theory will not be valid.

Now if we have degenerate states e.g. for some state ψ_j, $E_i = E_j$ then we have a serious problem! We have a zero on the bottom line. We could rescue the situation if we could ensure that the numerator also went to zero i.e

$$\langle \psi_i | \hat{H}' | \psi_j \rangle = 0 \tag{70}$$

What we need is a new basis set of eigenfunctions ϕ_i such that $\langle \phi_i | \hat{H}' | \phi_j \rangle = 0$ for all $i \neq j$.

We can make a new basis set using our original zero order solutions of the Schrödinger equation by forming linear combinations of these functions. However, we need to be careful.

Suppose we have two energy eigenstates ψ_i and ψ_j. A linear combination of these is in general *not* also an eigenstate). Consider

$$\phi = a\psi_i + b\psi_j \tag{71}$$

Where a, b are arbitary constants. Then

$$\hat{H}_0 \phi = aE_i\psi_i + bE_j\psi_j \tag{72}$$

So

$$\hat{H}_0 \phi \neq \text{constant} \times \phi \tag{73}$$

ϕ will, however, be an eigenstate of $\hat{H}_0$, if, and only if, $E_i = E_j$ i.e. the states ψ_i and ψ_j are degenerate. This is good news for it means we can form a new basis set of eigenfunctions ϕ by taking linear combinations of our degenerate set $\psi_i(n, l, m_l, m_s)$. Is this a lucky coincidence – or is it a result of some choice we made in finding the zero order solutions? Well, it isn't luck! We made the arbitrary choice of the z-axis as our axis of reference. If we had chosen another axis we would have found different functions, These different functions would have been linear combinations of the set referred to the z-axis. The lesson from all this is that we need to choose a set of eigenfunctions as our basis set so that the perturbation will be diagonal. This is why we choose the z-axis and apply our perturbing fields along this axis.

One final point; the perturbation matrix needs to be diagonal only within the degenerate set. There may be other off-diagonal elements but they have an energy denominator $(E_i - E_j)$ and provided this is large enough compared to $< \psi_j | \hat{H}' | \psi_i >$ these elements won't contribute much to the perturbation.

We can summarise this by looking at an example of two degenerate states ψ_1 and ψ_2.
We can write the perturbation energy as

$$\Delta E_1 = \langle\psi_1|\,\hat{H}'\,|\psi_1\rangle \tag{74}$$
$$\text{only if} \quad \langle\psi_1|\,\hat{H}'\,|\psi_2\rangle = 0 \tag{75}$$
$$\text{However, if} \quad \langle\psi_1|\,\hat{H}'\,|\psi_2\rangle \neq 0 \tag{76}$$

then we make new zero order functions ϕ_1 and ϕ_2 using

$$\phi_1 = a\psi_1 + b\psi_2 \tag{77}$$
$$\phi_2 = b^*\psi_1 - a^*\psi_2 \tag{78}$$

And find the values of a, b that make

$$\langle\phi_1|\,\hat{H}'\,|\phi_2\rangle = 0 \tag{79}$$

Then we can write:

$$\Delta E_1 = \langle\phi_1|\,\hat{H}'\,|\phi_1\rangle\,,\ \Delta E_2 = \langle\phi_2|\,\hat{H}'\,|\phi_2\rangle \tag{80}$$

4.2.5 Degenerate Perturbation theory and the Vector Model

Finding the right coefficients in our linear combination of zero-order eigenfunctions results in a diagonal matrix. There are mathematical methods for diagonalizing a matrix that can be used.

This procedure, a unitary transformation, is formally equivalent to rotating the axes of our frame of reference to a new set of axes where there are no cross-terms i.e. the matrix is diagonal in the new basis. If we want to avoid off-diagonal elements, when we have degenerate states, we need to choose the right basis or reference frame. We can then find the expectation values of an operator are given by its eigenvalues. Degeneracy usually is associated with the angular momentum and we know the eigenvalues of angular momentum operators. So we can use the symmetry of the situation to guide us to the proper basis set. These new basis eigenfunctons will be labelled by quantum numbers associated with an operator corresponding to a *constant of the motion*.

This leads us to a helpful model; the Vector Model, enabling us to visualise the physics in DPT.

Consider the classical behaviour of a rotating charge in a magnetic field along the z-axis. The angular momentum vector $\underline{\lambda}$ is at angle θ to the z-axis and the Larmor precession causes the vector $\underline{\lambda}$ to trace out a cone at constant angle to the z-axis. $\underline{\lambda}$ and its projection on z is a constant of the motion.

The projection on the x or y axis, however is not constant. It therefore makes sense to refer our motion to the z-axis. The wavefunction can be labelled quantum mechanically by a quantum number m i.e. the projection of $\underline{\lambda}$ on the z-axis is $m\hbar$. This illustrates the key basis of the Vector Model: Quantum Mechanics represents a classical constant of the motion by an operator which commutes with the Hamiltonian. So eigenstates of the Hamiltonian will also be eigenstates of this operator. Then the expectation values, representing the time average of the classical vector, are given by the eigenvalues of the operator. The quantum numbers defining the eigenfunctions are then good quantum numbers. In the example case just described, the vector $\underline{\lambda}$ could represent the total angular momentum $\underline{j}$. Quantum mechanically we can know both $\hat{\underline{j}}^2$ and $\hat{\underline{j}}_z$, with eigenvalues $j(j+1)$ and m_j respectively, and j and m_j are good quantum numbers.

4.2.6 Evaluation of $\left\langle \hat{\underline{s}} \cdot \hat{\underline{l}} \right\rangle$ using DPT and the Vector Model

We now return to finding the expectation value of the operator $\left\langle \hat{\underline{s}} \cdot \hat{\underline{l}} \right\rangle$. We start with the zero order solutions to the Schrödinger equation, $\psi(n, l, m_l)$. Including the spin, quantum numbers s, m_s, our basis states are:

$$|n, l, m_l, s, m_s\rangle \tag{81}$$

We find that

$$\langle n, l, m_l, s, m_s| \, \hat{\underline{s}} \cdot \hat{\underline{l}} \, |n, l, m_l, s, m_s\rangle \neq 0 \quad \text{for all} \;\; m_l \neq m_l', m_s \neq m_s' \tag{82}$$

This means that these states are not suitable for evaluating the perturbation.

As explained above, we need to find linear combinations of the degenerate functions which will give a diagonal matrix. We choose linear combinations that produce eigenstates of the total angular momentum operators $\hat{\underline{j}}^2$ and $\hat{\underline{j}}_z$ where $\hat{\underline{j}} = \hat{\underline{l}} + \hat{\underline{s}}$. These will also be eigenstates of $\hat{\underline{l}}^2$ and $\hat{\underline{s}}^2$

$$\hat{\underline{j}}^2 = \left(\hat{\underline{l}} + \hat{\underline{s}}\right)^2 \quad ; \quad \hat{\underline{j}}_z = \hat{\underline{l}}_z + \hat{\underline{s}}_z \tag{83}$$

Then

$$\hat{\underline{s}} \cdot \hat{\underline{l}} = \frac{1}{2}\left(\hat{\underline{j}}^2 - \hat{\underline{l}}^2 - \hat{\underline{s}}^2\right) \tag{84}$$

And the eigenstates are $|n, l, s, j, m_j\rangle$.

We then find that the off-diagonal elements

$$\langle n, l, s, j, m_j| \, \hat{\underline{s}} \cdot \hat{\underline{l}} \, |n, l, s, j, m_j\rangle = 0 \quad \text{unless} \;\; j = j' \text{ and } m_j = m_j' \tag{85}$$

The diagonal elements (eigenvalues) are then

$$\langle n, l, s, j, m_j| \, \hat{\underline{s}} \cdot \hat{\underline{l}} \, |n, l, s, j, m_j\rangle = \frac{1}{2}\left\{j(j+1) - l(l+1) - s(s+1)\right\}\hbar^2 \tag{86}$$

This is the result we need; the new basis set using quantum numbers j, m_j provides a perturbation matrix that is diagonal. We have dealt with the problem of the degenerate states. Remember that the problem arose from our choice of the z-axis. In the presence of spin-orbit interaction m_l, m_s were not good quantum numbers. In other words the projections of $\hat{\underline{l}}$ and $\hat{\underline{s}}$ on our arbitrary z-axis were not constants of the motion. We can visualise this using our vector model.

We represent the angular momentum by vectors of magnitude given by $|\hat{\underline{l}}|$, $|\hat{\underline{s}}|$, $|\hat{\underline{j}}|$.

When there is no spin-orbit interaction $\hat{\underline{l}}$ and $\hat{\underline{s}}$ take up fixed orientations in space. Their projections on the z-axis are $m_l\hbar$ and $m_s\hbar$ which are constant. When spin-orbit interaction takes place $\hat{\underline{l}}$ and $\hat{\underline{s}}$ precess around their mutual resultant $\hat{\underline{j}}$. Their projections on the z-axis, are no longer constant, m_l, m_s are undefined and no longer good quantum numbers.

Now since there is no external torque the total angular momentum $\hat{\underline{j}}$ is a constant of the motion. The projection of $\hat{\underline{j}}$ on the z-axis, $\hat{\underline{j}}_z$ is also a constant, and quantized in multiples of $\hbar$ as $m_j\hbar$. the operators corresponding to these vectors (angular momenta) are:

$$\hat{\underline{j}} = \hat{\underline{l}} + \hat{\underline{s}} \tag{87}$$

j takes values from $|l - s|$ to $|l + s|$.

$$\text{Also} \quad \hat{\underline{j}}^2 = \left(\hat{\underline{l}} + \hat{\underline{s}}\right)^2 \tag{88}$$

with eigenvalues $j(j+1)\hbar$ and

$$\hat{\underline{j}}_z = \hat{\underline{l}}_z + \hat{\underline{s}}_z \tag{89}$$

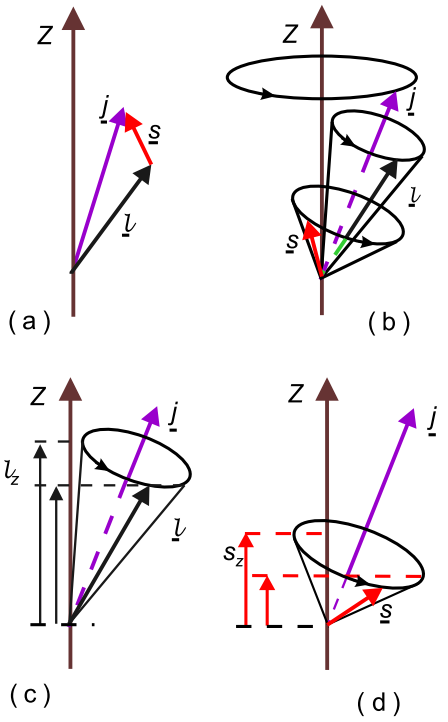


Figure 12: (a) Orbital $\underline{l}$ and spin $\underline{s}$ angular momenta add (couple) to give a resultant $\underline{j}$. With no magnetic interaction the vectors remain fixed in space. (b) Spin-orbit coupling results in precessional motion of $\underline{l}$ and $\underline{s}$ around their resultant $\underline{j}$. (c) and (d) Projections of $\underline{l}$ and $\underline{s}$ on z-axis, l_z and s_z respectively, are no longer constants of the motion and hence m_l and m_s are not good quantum numbers in the presence of spin-orbit coupling

with eigenvalue $m_j\hbar$.

For a single electron $s = \pm 1/2$ so we have two values of $j = l \pm 1/2$.

The key is to choose the basis to suit the perturbation. For example, if there was no spin-orbit interaction and we apply an external B-field, then $|n, l, s, m_l, m_s\rangle$ is a suitable basis. The vectors $\hat{\underline{l}}$ and $\hat{\underline{s}}$ will precess independently around the B-field direction (z-axis). The projections $m_l\hbar$ and $m_s\hbar$ respectively are constants of the motion so m_l, m_s are good quantum numbers.

If we have both spin-orbit interaction and an external field and the interactions are of comparable strength, then the vector model doesn't help. In this case $\hat{\underline{l}}$ and $\hat{\underline{s}}$ will precess around B_{ext} at about the same angular speed as they precess around their mutual resultant $\hat{\underline{j}}$. The motion becomes very complicated and neither m_l, m_s nor j and m_j are good quantum numbers.

The vector model works so long as one perturbation is much stronger than the other.

If the external field is weak, $\hat{\underline{l}}$ and $\hat{\underline{s}}$ couple together to form $\hat{\underline{j}}$ i.e they precess rapidly around $\hat{\underline{j}}$. The interaction with B_{ext} causes $\hat{\underline{j}}$ to precess relatively slowly around B_{ext} with a constant projection on the field axis given by m_j.

If the external field is very strong then $\hat{\underline{l}}$ and $\hat{\underline{s}}$ precess independently around B_{ext} with projections m_l and m_s. Because of their rapid precession around B_{ext}, $\hat{\underline{l}}$ and $\hat{\underline{s}}$ do not combine to form a constant $\hat{\underline{j}}$. The spin-orbit interaction energy, in this case, can be found using the states $|n, l, s, m_l, m_s\rangle$ because the stronger interaction with B_{ext} has removed (raised) the degeneracy.

In conclusion, we found the perturbation to the energy had two terms; a radial integral and a product of the spin and orbit vectors. The result for the radial integral was equation (66) and the result for the spin-orbit term was equation (86). So using these results we could calculate the expectation value of the spin-orbit perturbation.

4.3 Spin Orbit Interaction: Summary

The perturbation,

$$\hat{H}_2 \propto \frac{1}{r}\frac{\partial U(r)}{\partial r}\underline{\hat{s}} \cdot \underline{\hat{l}} \tag{90}$$

The change to the energy of the states is then:

$$\Delta E_{\mathrm{SO}} = \langle\psi|\, \hat{H}_2\, |\psi\rangle \tag{91}$$

The radial integral, gives a factor $\sim \beta_{nl}$

The angular momentum integral $= \left\langle \underline{\hat{j}}^2 - \underline{\hat{l}}^2 - \underline{\hat{s}}^2 \right\rangle$

We choose a basis set of eigenfunctions to diagonalize this perturbation: $|n, l, s, j, m_j\rangle$

Hence

$$\Delta E_{\mathrm{SO}} = \beta_{nl}\frac{1}{2}\left\{j(j+1) - l(l+1) - s(s+1)\right\} \tag{92}$$

Where for hydrogen

$$\beta_{n,l} = \frac{\mu_0 Z^4 g_s \mu_b^2}{4\pi}\frac{1}{n^3 a_0^3 l(l+1/2)(l+1)} \tag{93}$$

For a single-electron atom, e.g. alkalis, j has two values. $j = l \pm 1/2$. Thus the energy levels are split, each value of j has a different ΔE_{SO}.

We now look at this effect in alkali atoms as an example.

4.4 Spin-Orbit Splitting: Alkali Atoms

We consider the case of sodium, an atom with one electron outside a Central Field. The ground state has $n = 3$, $l = 0$. There is therefore no spin-orbit splitting of this level, since the electron has no orbital motion. For the first excited state, $n = 3$, $l = 1$. If there was no spin-orbit interaction we could use eigenfunctions

$$|n, l, s, m_l, m_s\rangle = |3, 1, 1/2, m_l, m_s\rangle \tag{94}$$

where $m_l = +1, 0, -1$ and $m_s = +1/2, -1/2$ i.e. we have 6 degenerate states. When we include the spin-orbit interaction we use the basis $|n, l, s, j, m_j\rangle$, where $j = 1/2$ or $3/2$ so the energy level is split into two.

For $j = 1/2$, $m_j = +1/2, -1/2$, i.e. two degenerate states and
for $j = 3/2$, $m_j = 3/2, 1/2, -1/2, -3/2$ i.e. four degenerate states.

Each level is shifted by

$$\Delta E_{\mathrm{SO}}(3p) = \beta_{3p}(-1),\ j = 1/2 \tag{95}$$

$$\Delta E_{\mathrm{SO}}(3p) = \beta_{3p}(+1/2),\ j = 3/2 \tag{96}$$

Note that the "centre of gravity" of the level remains the same. The product of the degeneracy $(2j+1)$ i.e. number of states times the shift is the same for each of the shifted levels. This result is a consequence of the fact that magnetic interactions have no overall effect on the energy of the atom, since the magnetic forces do no work. For $n = 3$, $l = 2$, a 3d state we have

$$\Delta E_{\mathrm{SO}}(3d) = \beta_{3d}(1),\ j = 5/2 \tag{97}$$

$$\Delta E_{\mathrm{SO}}(3d) = \beta_{3d}(-3/2),\ j = 3/2 \tag{98}$$

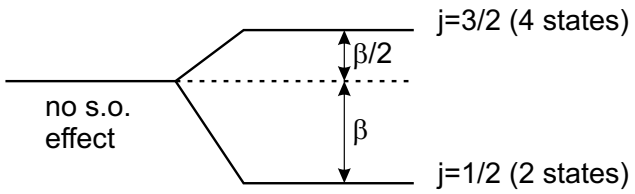


Figure 13: Diagram showing energy splitting of 3p level in an alkali.

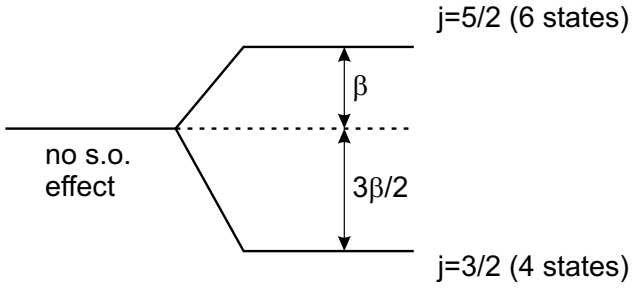


Figure 14: Diagram showing energy splitting of 3d level in an alkali.

Note that $\beta_{3p} > \beta_{3d}$ since the 3p electron penetrates the inner core of electrons far more than the 3d electron.

The actual size of the splitting is set by the value of the radial integral $\beta(n,l)$. For hydrogen we noted this was

$$\beta_{n,l} = \frac{\mu_0 Z^4 g_s \mu_B^2}{4\pi} \frac{1}{n^3 a_0^3 l(l+1/2)(l+1)} \tag{99}$$

This applies to hydrogen-like atoms i.e. ions with all electrons except one stripped off. For a neutral atom, however, the electrostatic force of the central field will be less and so the dependence on Z is reduced from the Z^4 in a hydrogenic ion. We can get an approximate idea of the Z-dependence by considering the Z_{eff} in our Central Field to be a "step function" i.e. to change abruptly from its effective value Z_{o} ("o" for outer) outside the core to Z_{in} ("in" for inner) inside the core. This is a very crude approximation, and we can make it even more crude by supposing $Z_{\text{o}} = 1$ (total screening by $(Z-1)$ electrons) and $Z_{\text{in}} = Z$ (no screening at all).

The total energy of the spin-orbit interaction will be made up of two parts, each with the form:

$$\Delta E_{\text{SO}} \propto Z \left\langle \frac{1}{r^3} \right\rangle \tag{100}$$

The outer part, $Z = 1$, is very small compared to the contribution from the inner part, $Z >> 1$, so we neglect the contribution from the Z_{o} bit. The inner part is hydrogenic so:

$$\left\langle \frac{Z}{r^3} \right\rangle_{\text{inner}} = \frac{Z^4}{n^3 a_0^3 l(l+1/2)(l+1)} \tag{101}$$

The contribution will be weighted by the fraction of the time the electron spends inside the core. This fraction is indicated by the quantum defect or by the effective quantum number n^*. Using the Bohr theory we can show that the fraction inside the core is

$$\left(\frac{n^3}{Z_{\text{in}}^2}\right) / \left(\frac{n^{*3}}{Z_{\text{o}}^2}\right) \tag{102}$$

So the contribution from the inner part becomes:

$$\left\langle \frac{Z_{\text{in}}}{r^3} \right\rangle_{\text{inner}} \times \frac{n^3 Z_{\text{o}}^2}{n^{*3} Z_{\text{in}}^2} \tag{103}$$

Hence

$$\beta_{n,l} \sim \frac{\mu_0}{4\pi} g_s \mu_B^2 \frac{Z_{\text{in}}^2 Z_{\text{o}}^2}{n^{*3} a_0^3 l(l+1/2)(l+1)} \tag{104}$$

So approximately, the size of the spin-orbit splitting scales with $Z_{\text{in}}^2 \approx Z^2$.

We now need to consider atoms other than alkalis, where residual electrostatic interaction, can't be ignored. Before we do that we define a way of indicating the total angular momentum of atoms as a whole, and not just individual electrons with l, s, j, m_l, m_s, m_j, etc.

4.5 Spectroscopic Notation

So far we have seen how the physical interactions affecting the energy of an atom can be ordered according to their strength. We use perturbation theory, or the Vector Model, by starting with the largest interaction – the Coulomb force of the Central Field. This led to the idea of the electron configuration; $n_1 l_1 n_2 l_2 n_3 l_3$ etc. with each individual electron's angular momentum labelled by a lower case letter; s, p, d, f, etc. For "single-electron" atoms, like alkalis, we defined a total angular momentum $\hat{\underline{j}} = \hat{\underline{l}} + \hat{\underline{s}}$. The spin-orbit interaction led to energy levels labelled by their value of j, i.e. $(l + 1/2)$ or $(l - 1/2)$. What do we do when there is more than one electron to worry about? We anticipate the result of the next section where we consider the residual electrostatic interaction between electrons.

If we were to ignore any spin-orbit interaction then the energy levels are completely determined by electrostatic forces only. In this case the space and spin systems within the atom must separately conserve angular momentum. This suggests that we can define a total orbital angular momentum by the vector sum of the individual orbital momenta. We will label such total angular momenta by capital letters; thus the total orbital angular momentum is:

$$\underline{L} = \sum_i \underline{l}_i \tag{105}$$

For simplicity we will consider two-electron atoms i.e. atoms with two valence electrons outside closed inner shells. So $\underline{L} = \underline{l}_1 + \underline{l}_2$. In the same way we define the total spin of the atom to be

$$\underline{S} = \sum_i \underline{s}_i = \underline{s}_1 + \underline{s}_2 \tag{106}$$

These can be visualized using our vector model:

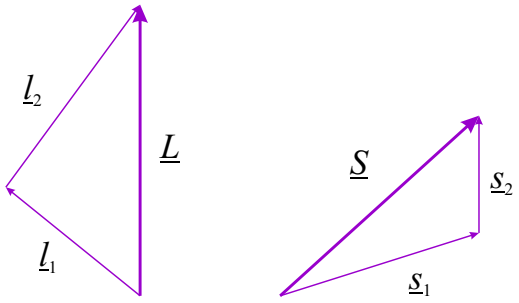


Figure 15: Diagram showing the coupling of all $\underline{l}_i$ to $\underline{L}$ and all $\underline{s}_i$ to $\underline{S}$.

Quantum mechanically these angular momenta are represented by operators $\hat{\underline{L}}^2$ and $\hat{\underline{L}}_z$

$$\underline{L}^2 = \left(\hat{\underline{l}}_1 + \hat{\underline{l}}_2\right)^2 \tag{107}$$

With eigenvalues $L(L+1)\hbar^2$ where $|l_1 - l_2| \leq L \leq l_1 + l_2$, and

$$\underline{L}_z = \hat{\underline{l}}_{z1} + \hat{\underline{l}}_{z2} \tag{108}$$

With eigenvalues $M_L\hbar$ where $-L \leq M_L \leq L$.

Similarly we have total spin operators.

$$\underline{S}^2 = (\hat{\underline{s}}_1 + \hat{\underline{s}}_2)^2 \tag{109}$$

With eigenvalues $S(S+1)\hbar^2$ where $|s_1 - s_2| \leq S \leq s_1 + s_2$, and

$$\underline{S}_z = \hat{\underline{s}}_{z1} + \hat{\underline{s}}_{z2} \tag{110}$$

With eigenvalues $M_S\hbar$ where $-S \leq M_S \leq S$.

M_L and M_S are the quantized projections of $\underline{L}$ and $\underline{S}$ respectively on the z-axis (the axis of quantization). The total angular momentum is found by adding (or coupling) the total orbital and spin momenta: to give $\underline{J}$:

$$\hat{\underline{J}}^2 = \left(\hat{\underline{L}} + \hat{\underline{S}}\right)^2 \tag{111}$$

With eigenvalues $J(J+1)\hbar^2$ and

$$|L - S| \leq J \leq L + S \tag{112}$$

Also:

$$\hat{\underline{J}}_z = \hat{\underline{L}}_z + \hat{\underline{S}}_z \tag{113}$$

With eigenvalues $M_J\hbar$ where $-J \leq M_J \leq J$ i.e. M_J is the projection of $\hat{\underline{J}}$ on the quantization axis.

By convention $L = 0, 1, 2, 3$ is labelled S, P, D, F etc as with the s, p, d, f for single electrons. The values of L and S specify a *Term*. The spin-orbit interaction between $\hat{\underline{L}}$ and $\hat{\underline{S}}$ will split each term into levels labelled by J. For a given $\hat{\underline{L}}$ there will be $(2S+1)$ values of J and so $(2S+1)$ is known as the multiplicity

Spin-orbit interaction will split each term into levels according to its value of J. Thus the multiplicity tells us the number of separate energy levels (i.e. number of J values) except, however when $L = 0$, when there is no spin-orbit splitting and there is only a single level.

Each level is degenerate in J; there are $(2J+1)$ values of M_J; the projection of $\hat{\underline{J}}$ on an external axis. Each value of M_J between $-J$ and $+J$ constitutes a state, and the degeneracy is raised by applying an external magnetic field in which the different orientations of $\hat{\underline{J}}$ relative to the field direction will have different energy.

In general then we have the following notation for an energy level:

$$n_1 l_1 n_2 l_2 ... \quad {}^{2S+1}L_J \tag{114}$$

This reflects the hierarchy of interactions:

Central Field configuration, $n_1 l_1 n_2 l_2 \ldots$
Residual Electrostatic $\rightarrow$ Terms, $L = S, P, D \ldots$
Spin-Orbit $\rightarrow$ Level, $J = |L - S| \rightarrow L + S$
External Field $\rightarrow$ State, $M_J = -J \rightarrow +J$

The ground level of sodium can therefore be written as $1s^2 2s^2 2p^6 3s\ {}^2S_{1/2}$.

It is common practice to drop the configuration of the inner shells and give only that for the outer, valence, electrons. Thus the first two excited levels are 3p ${}^2P_{1/2,3/2}$ and 3d ${}^2D_{3/2,5/2}$. these are sometimes referred to as excited "states", but this is an example of the loose terminology used by Atomic Physicists!

We can now draw a full Energy Level diagram for sodium with all the levels labelled by appropriate quantum numbers.

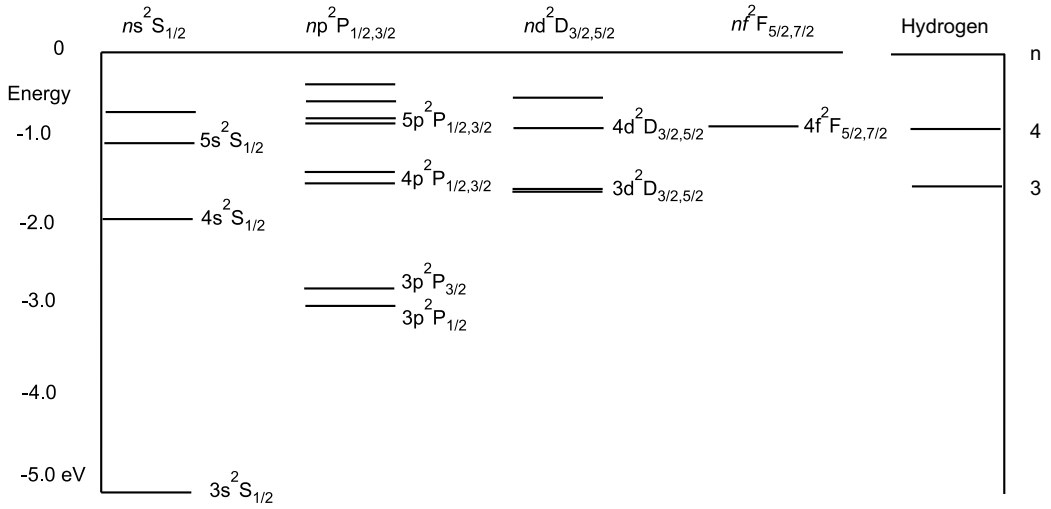


Figure 16: Na energy level diagram showing fine structure (spin-orbit splitting) greatly exaggerated. Note that the difference between energy levels the same n but of increasing l become smaller as they approach the hydrogenic energy level for the corresponding n. Note also that the difference between energy levels of the same n but different l becomes smaller with increasing n.

5 Two-electron Atoms: Residual Electrostatic Effects and LS-Coupling

So far we have made the Central Field Approximation which allows us to estimate the gross structure of the energy levels. This introduced quantum numbers n and l allowing us to define the electron configuration. We noted two perturbations that need to be considered in going beyond the Central Field Approximation. These were, $\hat{H}_1$, the residual electrostatic interaction and $\hat{H}_2$, the spin-orbit interaction. Perturbation theory requires that we apply the larger perturbation first. In the specific case of a one-electron atom in a Central Field we could ignore $\hat{H}_1$. The perturbation due to spin-orbit effects then led us to wave functions labelled by the quantum numbers j, m_j. In the case of an atom with two electrons in a Central Field we have to consider first $\hat{H}_1$, and this will lead us to different quantum numbers. We have anticipated this result by introducing these quantum numbers L, S and J. In the following sections we will look more closely at the physics of how electrostatic interactions lead to LS-coupling. It needs to be stressed at the outset that LS-coupling is the most appropriate description when electrostatic effects are the dominant perturbation. We can then apply the smaller spin-orbit perturbation $\hat{H}_2$ to these LSJ-labelled states.

All this is conveniently visualized using the Vector Model, where the strongest interactions lead to the fastest precessional motions. The residual electrostatic perturbation $\hat{H}_1$ causes the individual orbital angular momenta $\underline{l}_1$ and $\underline{l}_2$ to couple to form $\underline{L} = \underline{l}_1 + \underline{l}_2$. Similarly the spin angular momenta $\underline{s}_1$ and $\underline{s}_2$ couple to give a total spin $\underline{S} = \underline{s}_1 + \underline{s}_2$. So $\underline{l}_1$ and $\underline{l}_2$ precess rapidly around $\underline{L}$ and $\underline{s}_1$ and $\underline{s}_2$ precess rapidly around $\underline{S}$. $\underline{L}$ and $\underline{S}$ are then well-defined constants of the motion.

The weaker, spin-orbit interaction $\hat{H}_2$ then leads to a slower precession of $\underline{L}$ and $\underline{S}$ around their resultant $\underline{J}$. As in the single electron case, the precession rate – and hence the energy shift – depends on the relative orientation of $\underline{L}$ and $\underline{S}$. The result is a fine-structure splitting of the terms defined by $\underline{L}$ and $\underline{S}$ according to the value of $\underline{J}$.

We have been assuming that $\hat{H}_1 >> \hat{H}_2$ but this may not always be the case. We have seen already that $\hat{H}_2$ increases with Z^2. So our assumption that electrostatic dominates over magnetic, spin-orbit, coupling will become less valid for heavy elements. When this happens L and S cease to be good quantum numbers. The total angular momentum J will continue to be a good quantum number, but it will be a result of adding the angular momenta of the individual electrons in a different way. In the limit that $\hat{H}_2 >> \hat{H}_1$, individual electron momenta $\underline{l}_i$ and $\underline{s}_i$ couple to give $\underline{j}_i = \underline{l}_i + \underline{s}_i$. The total $\underline{J}$ is then the vector sum of $\underline{j}_1$ and $\underline{j}_2$. This jj-coupling is, however, not on the syllabus.

In the following we will take Magnesium as our example of a two-electron atom. We will look first at the gross energy level structure arising from the Central Field Approximation. Secondly, we will consider the residual electrostatic effect (perturbation $\hat{H}_1$) which splits the energy levels into terms labelled by L and S. We will think about why these terms have different energies. Thirdly, we apply the smaller $\hat{H}_2$, spin-orbit, perturbation and examine the resulting fine-structure splitting.

5.1 Magnesium: Gross Structure

We choose Mg as our typical two-electron atom rather than Helium because, in the Helium case, the spin-spin interaction is of comparable magnitude to the other perturbations. In heavier elements this is less so. We can, however, use the basic results found for Helium to help us understand the behaviour of our two valence electrons in Magnesium. Magnesium is basically sodium with an extra proton in the nucleus and an extra electron outside. The configuration of the ground energy level is $1s^2 2s^2 2p^6 3s^2$. If we restrict ourselves to exciting only one of the 3s electrons, the excited configurations will be $(\ )3snl$. The electron remaining in the 3s orbit acts as a "spectator" electron and so the excited electron moves in a Central Field not much different from that of sodium. The basic energy level structure will be roughly the same as sodium, but the energy levels will be more negative owing to the increased Coulomb attraction of the nucleus.

What will be the effect of the perturbations $\hat{H}_1$ and $\hat{H}_2$?

5.2 The Electrostatic Perturbation

We will eliminate $\hat{H}_2$ for the moment by considering the configuration 3s4s; allowing us to concentrate on the effects arising from there being two electrons with electrostatic repulsion between them. It may seem, at first sight, that we could deal with this configuration entirely within the Central Field Approximation. After all, the spectator, 3s, electron is a spherically symmetric charge and so also is the 4s electron. Could their mutual interaction not be contained within the Central Field? The answer is "no" because the fields the two electrons see is not the same.

So we will have a "left-over" perturbation $\hat{H}_1$ and it doesn't have to be non-central.

In the Central Field Approximation we can find wave functions for each electron in our 3s4s configuration; $\psi(3s)$ and $\psi(4s)$. What will be the wavefunction for the system as a whole? We could form a product wavefunction $\psi_1(3s)\psi_2(4s)$. Such a wavefunction implies we can identify individual electrons by the subscripts 1 and 2. The particles, are for the moment, distinguishable. Now, can we use this wave function to find our perturbation energy due to $\hat{H}_1$, denoted by ΔE_1?

In other words, is $\Delta E_1 = \langle\psi_1(3s)\psi_2(4s)|\,\hat{H}_1\,|\psi_1(3s)\psi_2(4s)\rangle$?

It is fairly obvious that if we swapped the labels 1 and 2 we would get the same energy. So the states: $|\psi_1(3s)\psi_2(4s)\rangle$ and $|\psi_2(3s)\psi_1(4s)\rangle$ are degenerate. We therefore have to use Degenerate Perturbation Theory. This means we have to find correct linear combinations of our zero-order wavefunctions. The correct linear combination wave function will diagonalize the perturbation matrix. So we can use the form of the perturbation to guide us to the correct combination

$$\hat{H}_1 = -\sum_i \frac{Ze^2}{4\pi\epsilon_0 r_i} + \sum_{i>j} \frac{e^2}{4\pi\epsilon_0 r_{ij}} - \sum_i U(r_i) \tag{115}$$

Since our zero order wavefunctions are eigenfunctions of $U(r_i)$, the third term will not give us any trouble. The first term is also a single-electron operator so this will not cause a problem either. We are left with the $1/r_{ij}$ operator and need to find a wavefunction that will give a diagonal matrix.

We can form two orthogonal functions as follows:-

$$\phi_1 = a\psi_1(3s)\psi_2(4s) + b\psi_1(4s)\psi_2(3s) \tag{116}$$

$$\phi_2 = b^*\psi_1(3s)\psi_2(4s) - a^*\psi_1(4s)\psi_2(3s) \tag{117}$$

It remains to find a and b to make the diagonal matrix elements vanish i.e.

$$\langle\phi_1|\,V\,|\phi_2\rangle = 0 \tag{118}$$

Where V is the two-electron electrostatic repulsion. This perturbation does not distinguish electron 1 from electron 2, so $|a| = |b|$ and for normalization $|a|, |b| = 1/\sqrt{2}$. We now have

$$\phi_1 = \frac{1}{\sqrt{2}}\left(\psi_1(3s)\psi_2(4s) + \psi_1(4s)\psi_2(3s)\right) \tag{119}$$

$$\phi_2 = \frac{1}{\sqrt{2}}\left(\psi_1(3s)\psi_2(4s) - \psi_1(4s)\psi_2(3s)\right) \tag{120}$$

We note now that interchanging the labels does not change ϕ_1, and simply changes the sign of ϕ_2. These functions have a definite exchange symmetry. We shall return to this later.

The off-diagonal matrix element is now:

$$\frac{1}{2}\langle\psi_1(3s)\psi_2(4s) + \psi_1(4s)\psi_2(3s)|\,V\,|\psi_1(3s)\psi_2(4s) - \psi_1(4s)\psi_2(3s)\rangle \tag{121}$$

$1\uparrow$ $2\uparrow$ $3\uparrow$ $4\uparrow$

Multiplying out the terms in the integral we find:

$$1 \times 3 = \langle\psi_1(3s)\psi_2(4s)|\, V\, |\psi_1(3s)\psi_2(4s)\rangle = \mathrm{J} \quad (122)$$
$$2 \times 4 = -\langle\psi_1(4s)\psi_2(3s)|\, V\, |\psi_1(4s)\psi_2(3s)\rangle = -\mathrm{J} \quad (123)$$
$$2 \times 3 = \langle\psi_1(4s)\psi_2(3s)|\, V\, |\psi_1(3s)\psi_2(4s)\rangle = \mathrm{K} \quad (124)$$
$$1 \times 4 = -\langle\psi_1(3s)\psi_2(4s)|\, V\, |\psi_1(4s)\psi_2(3s)\rangle = -\mathrm{K} \quad (125)$$

So the total element is zero as required!

The diagonal elements now give us our energy eigenvalues for the change in energy:-

$$\langle\phi_1|\, V\, |\phi_1\rangle = \Delta E_1 = \mathrm{J} + \mathrm{K} \quad (126)$$
$$\langle\phi_2|\, V\, |\phi_2\rangle = \Delta E_2 = \mathrm{J} - \mathrm{K} \quad (127)$$

J is known as the direct integral and K is the exchange integral. (The wavefunctions in K differ by exchange of the electron label).

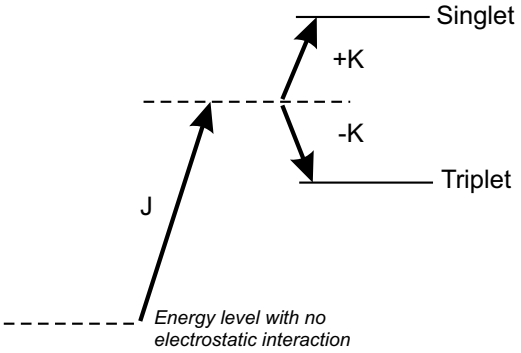


Figure 17: The effect of residual electrostatic interaction is to shift the mean energy by the direct integral J and to split the singlet and triplet terms by the exchange integral 2K

So the energy of any particular configuration will be split by an amount 2K. This corresponds to the two energies associated with the normal modes of coupled oscillators. For example, two pendula, or two masses on springs, when coupled together will settle into correlated or anti-correlated motions with different energies. Note that for our 2-electrons the effect arises purely because of the mutual electrostatic interaction. Note also that this has nothing to do with the distinguishability or indistinguishability of the electrons!

Our result shows that the term of the configuration described by ϕ_1, the symmetric state, has a higher (less negative) energy than the anti-symmetric state ϕ_2. To understand why we need to think a bit more about symmetry and the physics of the interactions between the electrons.

5.3 Symmetry

Degenerate Perturbation Theory forced us to choose spatial wavefunctions that have a define symmetry; either symmetric or antisymmetric. Why? It was because interchanging the labels on the operator could make no physical difference to the result, $\frac{1}{r_{12}} = \frac{1}{r_{21}}$. This must always be the case when we are dealing with identical particles. They must always be in a sate of definite symmetry; either symmetric, (+), or antisymmetric, (-). All electrons in our universe are antisymmetric, so we need to combine our spatial wavefunctions with an appropriate spin wavefunction to ensure our overall antisymmetric state.

Each electron spin may be up ↑ or down ↓. We can therefore construct for two electrons combinations that are either symmetric or antisymmetric:

$$\uparrow_1\uparrow_2, \qquad \downarrow_1\downarrow_2, \qquad \frac{1}{\sqrt{2}}\{\uparrow_1\downarrow_2 + \downarrow_1\uparrow_2\}: \text{ Symmetric} \tag{128}$$

$$\frac{1}{\sqrt{2}}\{\uparrow_1\downarrow_2 - \downarrow_1\uparrow_2\}: \text{ Antisymmetric} \tag{129}$$

The three symmetric spin states form a *triplet* and combine with the antisymmetric space states, ϕ_2. This triplet term energy was found to be shifted down i.e. more strongly bound. Physically, this is because the antisymmetric space function describes a state where the electrons avoid the same region of space. The mutual repulsion from the e^2/r_{12} potential is therefore reduced and the electrons are held more strongly. The antisymmetric state, on the other hand, is a *singlet* state. Since the electrons have opposite spins, equivalent to different spin quantum numbers, they can occupy the same space without violating the Pauli Principle. Increased overlap in space increases the effect of the $1/r_{12}$ repulsion, throwing the electrons further out in the Central Field. This results in a lower binding energy. The singlet terms thus lie above the triplet term for a given configuration.

5.4 Orbital effects on electrostatic interaction in LS-coupling

We have spent a long time considering the effect of the spin on the energy levels of the LS-coupled states. There is, however, another aspect of the electrostatic interaction that we have overlooked; the way the energy depends on L. The Central Field Approximation takes account of the spherically averaged charge distribution. The value of L for a given configuration depends on the relative orientation of the individual angular momenta $\underline{l}_1$ and $\underline{l}_2$. If either of these is zero e.g. 3s then the spatial overlap with the other electron in, say, a 3p orbit will be the same no matter where the axis of the 3p orbit points. Consider, however, what happens if both electrons are in p-orbits say 3p4p. Now $\underline{L} = \underline{l}_1 + \underline{l}_2 = 2, 1$ or 0 i.e. D, P or S Terms. The Vector Model shows that the axes of the individual orbital angular momenta are fixed in one of the three spatial orientations to give quantized values of L. The electron wavefunctions, and probability densities, are like doughnut shapes. The different relative orientations of their axes will lead to different spatial overlaps of the electrons. Consequently, the amount of mutual repulsion will be different in each case. This helps explain why the electrostatic interaction leads to terms of different energy and are labelled by L and S.

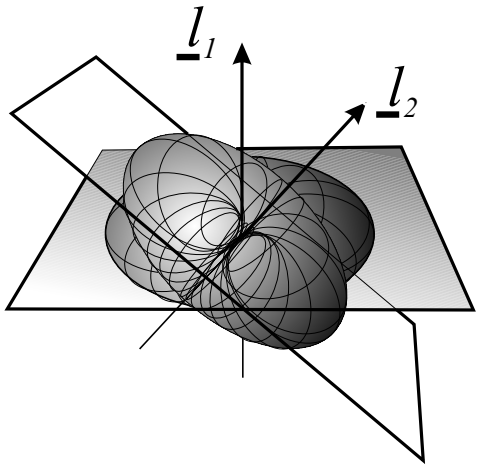


Figure 18: Electron orbitals $\underline{l}_1, \underline{l}_2$ have an overlap depending on their relative orientation and so different vector sums have different electrostatic energy.

5.5 Spin-Orbit Effects in 2-electron Atoms

So far we have applied only the electrostatic perturbation $\hat{H}_1$ to the configurations set by the Central Field. This electrostatic interaction led to singlet and triplet terms labelled by L and S. The 3s4s configuration in Mg will have two terms; ^{1}S and ^{3}S, with ^{1}S lying above ^{3}S in the energy level diagram. Since there is no orbital motion there is no spin-orbit interaction. The total angular momentum $\underline{J} = \underline{L} + \underline{S} = 0$ or 1. The complete designations are therefore
$3s4s\,^1S_0$ and $3s4s\,^3S_1$
Although the multiplicity, $2S+1$, is 3 in the triplet state, there is only one energy level since J has only the value 1, in this case.

Let's now consider the configuration 3s3p. Again $S = 0$ or 1 and we have singlet and triplet terms. $L = 1$, so the terms are ^{1}P and ^{3}P. J, however, in the triplet term has values $2, 1, 0$. These three values arise from each of three relative orientations of $\underline{L}$ and $\underline{S}$. This is clear in our Vector Model:

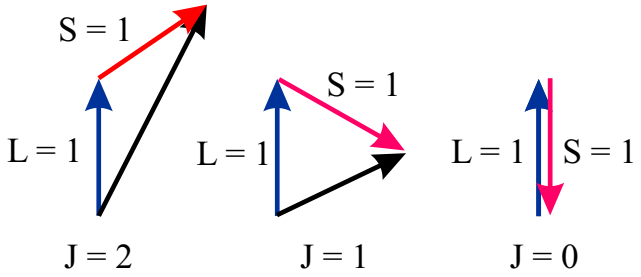


Figure 19: Vector addition of $\underline{L}$ and $\underline{S}$ to give total angular momentum $\underline{J}$.

The magnetic interactions between the total spin moment and the magnetic field established by the total orbital moment, L, are different in each case leading to different energies for the levels of each J value. The triplet, ^{3}P, term is therefore split and acquires fine-structure.

The physical arguments we used for spin-orbit interaction in Na carry over to the Mg case by using the total angular momentum operators $\hat{L}, \hat{S}$ and $\hat{J}$ instead of $\hat{l}, \hat{s}$ and $\hat{j}$. The perturbation to the $|n, L, S, J\rangle$ states is:

$$\Delta E_{\mathrm{SO}} = \left\langle \hat{H}_2 \right\rangle \propto \left\langle \left(\frac{1}{r} \frac{\partial U(r)}{\partial r} \right) \underline{\hat{S}} \cdot \underline{\hat{L}} \right\rangle \tag{130}$$

The radial integral is denoted $\beta_{n,l}$. The angular momentum operator is:

$$\underline{\hat{S}} \cdot \underline{\hat{L}} = \frac{1}{2} \left(\underline{\hat{J}}^2 - \underline{\hat{L}}^2 - \underline{\hat{S}}^2 \right) \tag{131}$$

and, since we have the correct representation in eigenfunctions of the operators on the R.H.S. we find the expectation value is given by the corresponding eigenvalues:

$$\left\langle \underline{\hat{J}}^2 - \underline{\hat{L}}^2 - \underline{\hat{S}}^2 \right\rangle = \frac{1}{2} \left\{ J(J+1) - L(L+1) - S(S+1) \right\} \tag{132}$$

so

$$\Delta E_J = \frac{\beta_{n,L}}{2} \left\{ J(J+1) - L(L+1) - S(S+1) \right\} \tag{133}$$

The separation of the fine-structure levels is found by evaluating ΔE_J for J' and $J'-1$.

$$\Delta E_{J'} - \Delta E_{J'-1} = \beta_{n,l} J' \tag{134}$$

This represents an Interval Rule that is valid so long as two conditions are fulfilled:

(1) LS-coupling is a good description for the terms i.e. when the electrostatic perturbation $\hat{H}_1 >> \hat{H}_2$, the magnetic, spin-orbit interaction.

(2) the perturbation energy is expressible as $\left\langle \hat{\underline{S}} \cdot \hat{\underline{L}} \right\rangle$. This is not the case in Helium where spin-spin interactions are comparable to spin-orbit effects.

The energy levels of the 3s3p terms are therefore

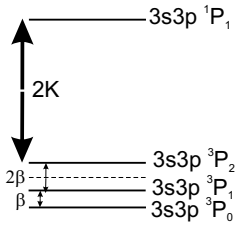


Figure 20: Separation of singlet and triplet terms due to electrostatic interaction and splitting of triplet term by magnetic spin-orbit interaction.

As we go to atoms with higher Z, the magnetic, spin-orbit, interaction increases ($\sim Z^2$) and begins to be comparable to the electrostatic effect. LS-coupling starts to break down and will be shown by two features of the energy levels: (a) The separation of the triplet levels becomes comparable to the separation of singlet and triplet terms of the same configuration. (b) The interval rule is no longer obeyed.

A third indicator of the breakdown of LS-coupling is the occurrence of optical transitions between single and triplet states.

In the LS-coupling scheme the spin states are well-defined and the dipole operator, $e\hat{r}$, does not act on the spin part of the wave function. Thus a dipole transition will link only states with the same value of S. i.e. we have the selection rule $\Delta S = 0$. This rule is well obeyed in Mg and the lighter 2-electron atoms. So we get transitions only between singlet and singlet states, or between triplet and triplet. It is then common to separate out the singlet ad triplet energy levels into separate diagrams.

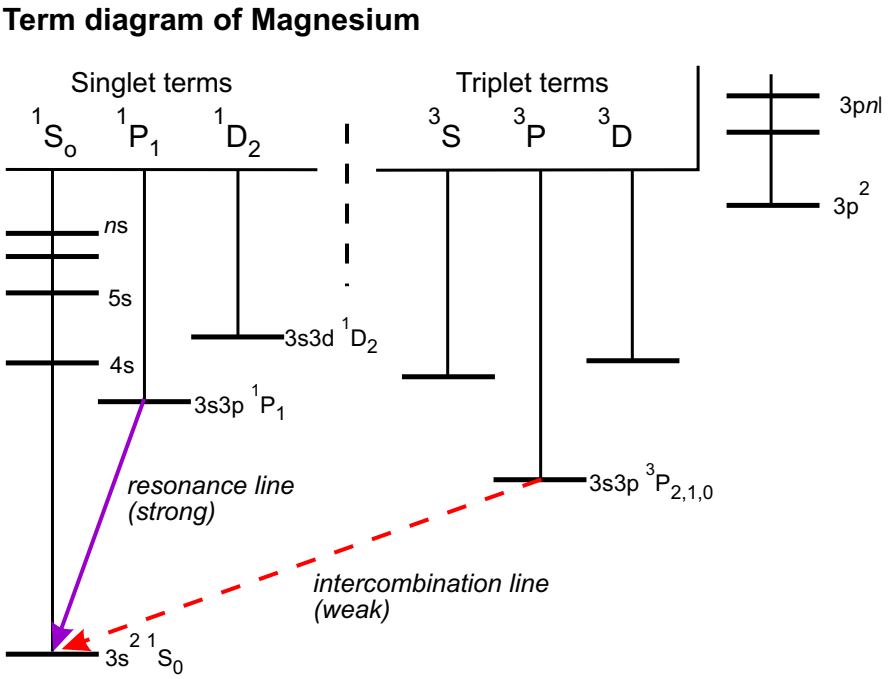


Figure 21: Mg term diagram. The separation of singlet and triplet terms shows schematically that singlet-triplet transitions are forbidden.

The 3s3p $^3P_{2,1,0}$ levels are metastable i.e. they have a long lifetime against radiative decay. The

transition is forbidden by the selection rule on the total spin, $\Delta S = 0$. The 6s6p $^3P_{2,1,0}$ levels in Mercury are also metastable but a transition 6s6p $^3P_1 \rightarrow 6s^2\ ^1S_0$ does occur. Such a transition is known spectroscopically as an intercombination line. The probability is not large, compared to other allowed transitions, but this transition is a strong source of radiation from excited mercury. This is because any atoms ending up in the 6s6p $^3P_1$ level have no other way to decay radiatively. The selection rule $\Delta S = 0$ derives from pure electrostatic interactions that lead to levels labelled by L and S. In mercury the magnetic spin-orbit interaction is larger than in lighter elements and so the energy levels are not based on pure LS-coupled states. This line is the source of light in Hg fluorescent lamps.

6 Nuclear Effects on Atomic Structure

We have secretly made three assumptions about the nucleus so far. We have assumed that the nucleus is stationary, has infinite mass and occupies only a point in space. In fact, the nuclei of all atoms consist of spinning charged objects and so they have a resultant magnetic moment. Secondly, their mass is not infinite, so the nucleus will move with the orbiting electrons around their common centre of gravity. Thirdly the nucleus has a finite size with some shape over which the proton charge is distributed. Each of these effects will make a small change in the energy levels. The magnetic interaction will lead to an effect similar to spin-orbit coupling and give a splitting of the energy levels know as hyperfine structure, hfs. The kinetic energy of the moving nucleus will affect the overall energy of the atomic states. Finally, the size and shape of the charges on the nucleus will affect the Coulomb force on the electrons and hence affect the energy. The term "hyperfine structure" is reserved (at least in Oxford) for magnetic effects of the nuclear spin. The mass and electron field effects will be termed isotope effects, since they depend to some extent on the number of neutrons in a given atom.

6.1 Hyperfine Structure

Both the protons and the neutrons have magnetic moments due to their spin. The total nuclear spin is labelled $I\hbar$. Spinning fermions tend to pair up with another similar particle with a spin in the opposite direction. For this reason nucleii with even numbers of protons and neutrons, the so-called even-even nuclei have $I = 0$. Odd-odd or odd-even nuclei can have integer or half-integer spin: $I = 1/2, 1, 3/2$ etc. These nuclei will have a magnetic dipole moment $\underline{\mu}_I$. Nucleii with integer spins can also have an electric quadruple moment as well. We will be concerned only with the magnetic dipole moment. Nuclear magnetic moments are small compared with electronic moments. Recall the classical relationship between magnetic moment and angular moment $\underline{\lambda}\hbar$:

$$\underline{\mu} = -\gamma\underline{\lambda}\hbar \tag{135}$$

In Quantum Mechanics the spin and orbital moments are:-

$$\underline{\mu}_s = -g_s\mu_B\underline{s} \tag{136}$$

$$\underline{\mu}_l = -g_l\mu_B\underline{l} \tag{137}$$

where $g_s = 2$ and $g_l = 1$ (Note we have used s, l in units of $\hbar$) Now $\mu_B = \frac{e\hbar}{2m_e}$, where m_e = mass of the electron. So nuclear moments are going to be much smaller as the nuclear mass $m_N >> m_e$. So to keep the g-factor of the order of unity we define the nuclear moment by

$$\underline{\mu}_I = g_I\mu_N\underline{I} \tag{138}$$

where g_I is the nuclear g-factor ($\sim$1) and μ_N is the nuclear magneton, related to μ_B by the ratio of electron to proton mass:

$$\mu_N = \mu_B \times \frac{m_e}{m_p} \tag{139}$$

The positive sign on our definition of $\underline{\mu}_I$ is purely a convention; we cannot tell in general whether the magnetic moment will be parallel or antiparallel to $\underline{I}$, it can be either.

The magnetic moments of the electrons and the nucleus will precess around their mutual resultant. The important thing to note is that the smallness of $\underline{\mu}_I$ will mean the interaction with the magnetic field of the electrons will be weak. The precession rate will therefore be slow and angular momenta $\underline{I}$ and $\underline{J}$ will remain well-defined. The interaction can therefore also be treated by perturbation theory and the Vector Model. The perturbation can be expressed:

$$\hat{H}_3 = -\hat{\underline{\mu}}_I \cdot \hat{\underline{B}}_{\rm el} \tag{140}$$

It remains to find the expectation value of this operator, and we will use our vector model to find the answer in terms of the time averaged vector quantities.

6.2 The Magnetic Field of Electrons

The nuclear interaction, being weak, will not follow the motions of individual electron orbits or spins. Their precessions around $\underline{J}$ are too fast for the nuclei to follow. The nuclear moment will therefore effectively see only the time averaged component resolved along $\underline{J}$.

$$\underline{B}_{\text{el}} = (\text{scalar quantity}) \times \underline{J} \tag{141}$$

The electrons in closed shells make no net contribution to the field $\underline{B}_{\text{el}}$. Therefore only the electrons outside the closed shells contribute. Electrons having $l \neq 0$ will have a magnetic field at the nucleus due to both orbital and spin motions. Both of these however depend on $1/r^3$ and so will be small enough to ignore for most cases. For $l = 0$, s-electrons, however, the situation is different. As we noted before, these electrons have $\psi(r) \neq 0$ at $r = 0$. The spin-spin interaction with the nucleus can therefore be strong. This is known as the Fermi contact interaction. This short range interaction dominates the contribution to the energy. In general it is very difficult to calculate but for the case of hydrogen in its ground state an analytical result can be found.

In general the orbital and spin momenta of the electrons provide a field of order $\sim \frac{\mu_0}{4\pi}\mu_B \left\langle \frac{1}{r^3} \right\rangle$. Taking a Bohr radius to give the order of magnitude of r we find:

$$B_{\text{el}} \sim \frac{\mu_0}{4\pi}\frac{\mu_B}{a_0^3} \sim 6\text{T} \tag{142}$$

Putting this with a nuclear momenta of the order of $\mu_N \sim \mu_B/2000$ we find an energy perturbation $\Delta E \sim 50$ MHz. So a transition involving a level with hfs will be split by this order of magnitude in frequency. Since the effect scales with Z it will get up to 100 times larger in heavy atoms i.e. $\sim$ a few GHz. The structure of the splitting will depend on the angular momentum coupling, and to this we now turn.

6.3 Coupling of I and J

Since the nuclear spin magnetic moment is proportional to $\underline{I}$ and the electron magnetic field is proportional to $\underline{J}$ we can write

$$\hat{H}_3 = A_J \underline{I} \cdot \underline{J} \quad \text{and} \quad \Delta E = A_J \left\langle \hat{\underline{I}} \cdot \hat{\underline{J}} \right\rangle \tag{143}$$

The quantity A_J determines the size of the interaction energy. Note that it depends on other factors as well as $\underline{J}$. For example we could have $J = 1/2$ from either a single s-electron or a single p-electron. The former will give a much larger value of A_J. The factor $\underline{I} \cdot \underline{J}$ is dimensionless and will have different values depending on the angle between $\underline{I}$ and $\underline{J}$. We can use our Vector Model to find the result of a DPT calculation as follows:

$\underline{I}$ and $\underline{J}$ couple i.e. add vectorally to give a resultant $\underline{F}$:

$$\underline{F} = \underline{I} + \underline{J} \tag{144}$$

From the vector triangle we find

$$\underline{F}^2 = \underline{I}^2 + \underline{J}^2 + 2\underline{I} \cdot \underline{J} \tag{145}$$

$$\underline{I} \cdot \underline{J} = \frac{1}{2}\left\{\underline{F}^2 - \underline{I}^2 - \underline{J}^2\right\} \tag{146}$$

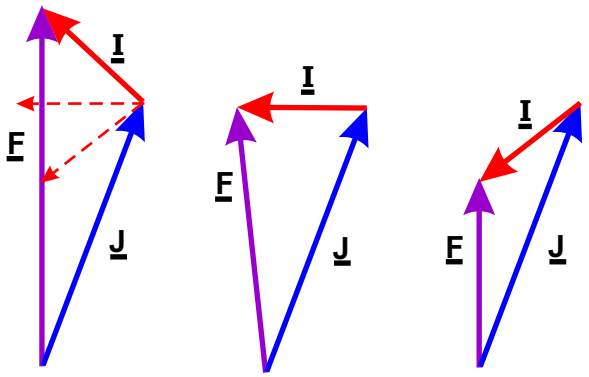


Figure 22: Addition of nuclear spin $\underline{I}$ and total electronic angular momentum $\underline{J}$ to give $\underline{F}$.

The magnitudes of the squared angular momenta are given by their eigenvalues:

$$\Delta E = \frac{A_J}{2}\{F(F+1) - I(I+1) - J(J+1)\} \tag{147}$$

As with spin-orbit coupling this leads to an interval rule:

$$\Delta E_{F'} = \Delta E(F') - \Delta E(F'-1) \sim A_J F' \tag{148}$$

(This provides one way of finding F and so if we know J we can find I, the nuclear spin.)

The interval rule can be messed up if there are additional contributions to the nuclear spin moment such as an electric quadruple moment.

The number of levels into which the hfs interaction splits a level depends on the number of values of F. This, in turn, depends on the coupling of $\underline{I}$ and $\underline{J}$. Our vector model gives us a "triangle rule" whereby the vector addition must yield a quantized value of the total angular momentum $\underline{F}$.

So if $J > I$ then there are $2I+1$ values of F, and if $I > J$ there are $2J+1$ values.

The values of F will be integers in the range: $|I-J| \leq F \leq I+J$. The ordering of the levels i.e. whether $E(F) > E(F-1)$ or vice-versa depends on the sign of A_J. There is, in general, no way to calculate this easily. It depends on nuclear structure, which determines the sign of g_I, and on the direction of $\underline{B}_{\rm el}$ relative to $\underline{J}$. (for single unpaired s-electron $\underline{B}_{\rm el}$ is always antiparallel to $\underline{J}$).

6.4 Finding the Nuclear Spin, I

The nuclear spin I can be found by examining, with high resolution, the structure of a spectral line due to a transition for a level with no, or unresolved, hfs to a hfs-split level. The selection rules of such transitions turn out to be:

$$\Delta F = 0, \pm 1 \text{ but not } 0 \rightarrow 0 \tag{149}$$

There are then three methods that can be used, provided there is sufficient spectral resolution.

Interval rule The level separation, found from the frequency intervals between components of the hfs, is proportional to the F-value of the higher level. This allows us to find F and then, if we know J for the level, we can find I.

Number of spectral lines The number of spectral components will give the number of levels of hfs. The number of levels is $(2I+1)$ for $I < J$ and $(2J+1)$ for $I > J$. (This one works provided we know $J > I$)

Relative Intensities The relative intensity of the spectral components will be proportional to the statistical weight of the hfs levels $(2F + 1)$. So if J is known we can find I.

Usually, in practice, we may need to use evidence from more than one of these methods to be sure. Conventional optical methods are often at or beyond their limit in resolving hfs. There is the additional problem of Doppler broadening which is often larger – or at least comparable to the splitting. Laser techniques, radio frequency methods or a combination of both are used nowadays.

6.5 Isotope Effects

Mass Effects The energy levels determined basically by the Central Field are given by:

$$E_n \sim \frac{Z^2 e^4 m_r}{2\hbar^2 n^2} \tag{150}$$

Where m_r is the reduced mass of the electron – nucleus system. For an infinitely massive nucleus $m_r \to m_e$, the electron mass. For real atoms, however, we need to deal with the motion of the nucleus around the common centre-of-mass. The resulting nuclear kinetic energy is $\frac{p_n^2}{2m_n}$ where p_n, m_n are the momentum and mass respectively of the nucleus. For a two-body system we can treat this by substituting the reduced mass, m_r, for m_e in the Schrödinger equation. The change in energy is then found simply by making the same substitution in the energy eigenvalues.

The result (and you should be able to work this out) is to shift the energy level up by $E(m_r - m_e)/m_e$.

In the case of a two-electron system we will need to know explicitly the electron wavefunctions. The energy change ΔE_{mass} is given by:

$$\Delta E_{\text{mass}} = \frac{p_n^2}{2m_n} = \frac{(\underline{p}_1 + \underline{p}_2)^2}{2m_n} \tag{151}$$

$$= \frac{p_1^2 + p_2^2 + 2\underline{p}_1 \cdot \underline{p}_2}{2m_n} \tag{152}$$

Putting the p_1^2, p_2^2 terms in Schrödinger's equation gives a reduced-mass-type shift. The $\underline{p}_1 \cdot \underline{p}_2$ term, however, needs to be calculated explicitly, usually using Perturbation Theory. The effect of this term is know as the "specific mass shift". The shifts are detected by changes in the frequency of transitions so the effects need to be calculated separately for each level involved. What will be observed is a small difference in the spectral line positions for different isotopes. The change in mass has a decreasing relative effect as the nuclear masses get heavier. Mass effects are therefore more important in light nucleii.

Field Effects The effect of the finite size of the proton in the energy levels in hydrogen can be calculated using perturbation theory. The energy shift is

$$\Delta E = \int_0^\infty \psi^*(r) \Delta V \psi(r) 4\pi r^2 dr \tag{153}$$

ΔV is the change in the potential energy resulting from the difference between a point charge and some spherical distribution, either over the surface or volume of a sphere. This difference is only significant over a very small range of r compared to a Bohr radius a_0. So the wavefunctions are essentially constant over the range of interest and so can come outside the integral.

$$\Delta E \approx |\psi(0)|^2 \int_0^\infty \Delta V(r) 4\pi r^2 dr \tag{154}$$

Only the s-states are significantly affected so using

$$|\psi(0)|^2 = \frac{1}{\pi}\left\{\frac{Z}{na_0}\right\}^3 \tag{155}$$

The result can be derived. For the charge on the surface of the nucleus of radius r_0 we find:

$$\Delta E_{ns} = \frac{Z^4 e^2 r_0^2}{6\pi\epsilon_0 a_0^3 n^3} \tag{156}$$

For the ground state of hydrogen the fractional shift is

$$\frac{\Delta E_{1s}}{E_{1s}} = -\frac{4}{3}\left\{\frac{r_0}{a_0}\right\}^2 \approx -5\times 10^{-10} \tag{157}$$

This is a small effect but important since uncertainty in knowing the size of a proton affects precise experiments to study QED effects.

7 Selection Rules

Now that we have a better picture of atomic energy levels and have begun to see how transitions between them are so important for our study of them, it is time to revisit the subject of selection rules. These are the rules that tell us whether or not an atom can change from one particular state to another. We will consider only electric dipole transitions; the rules for magnetic dipole, electric quadruple etc are different and are not on the syllabus (thankfully). One way to approach transitions is to use time-dependent Perturbation Theory and its result in Fermi's Golden Rule. That also isn't on our syllabus but we want to use some of the underlying physics.

In the first lecture we looked at the wavefunctions for two states ψ_1 and ψ_2 involved in a transition. Physically we had to generate an oscillating dipole – an accelerated charge – in order to create electromagnetic radiation. The solutions of the time-dependent Schrödinger equation, TDSE, have the form

$$\psi_i = \phi(r, \theta, \phi) e^{iE_i t/\hbar} \tag{158}$$

So the perturbation matrix elements are:

$$\langle \phi_i | - e\underline{r} | \phi_j \rangle \, e^{i(E_i - E_j)t/\hbar} \tag{159}$$

The diagonal elements $\phi_i = \phi_j$ represent stationary states and have no oscillation term, so atoms in stationary states do not absorb or emit light. As we saw in Lecture 1 the off-diagonal elements do give an oscillation of frequency $\omega_{12} = |E_1 - E_2|/\hbar$.

If the spatial part of the matrix element is zero, however, we will not get any radiation i.e. the transition is forbidden. The spatial integral will be determined by the quantum, numbers in the two states $\phi_1(n_1, l_1, \text{etc})$ and $\phi_2(n_1, l_1, \text{etc})$. It is the changes in these quantum numbers that give us our selection rules. We need

$$\langle n_1, l_1, n_2, l_2 ... L, S, J, M_J | \sum_i -er_i \left| n_1', l_1', n_2', l_2' ... L', S', J', M_J' \right\rangle \neq 0 \tag{160}$$

7.1 Parity

The off-diagonal matrix elements represent the atom in a super-position of two eigenstates ψ_1 and ψ_2. Any linear combination of solutions of the TDSE will also be an eigenstate [NB: this is not true for the time-independent Schrödinger equation]. Now the spatial wavefuctions all have a definite parity, either even, (+) or odd, (-). The dipole operator $-e\underline{r}$ has odd parity i.e. changing $\underline{r}$ to $-\underline{r}$ changes the sign.

If we consider, for the moment, a single electron making the jump from one state to another then the contribution to the integral from one particular location (x, y, z) is:

$$-e \cdot \phi_{nl}^*(x, y, z) \left[\hat{\underline{i}} x + \hat{\underline{j}} y + \hat{\underline{k}} z \right] \phi_{n'l'}(x, y, z) \tag{161}$$

If the parity of the two states is the same, then the product will be the same at the "opposite" point $(-x, -y, -z)$, The operator, however, will have the opposite sign and so when we integrate over all space the result is zero. Therefore, to get a non-zero dipole matrix element, the parity must change. The parity, as we noted before, is given by $(-1)^l$, where l is the electron's orbital angular momentum quantum number. So we have our first selection rule:

$$\Delta l = \pm 1 \tag{162}$$

This makes physical sense because a photon has angular momentum of one unit of $\hbar$. So conservation of angular momentum demands $\Delta l = \pm 1$, for one electron to jump.

7.2 Configuration

Could more than one electron jump? Consider an atom with two electrons at r_1 and r_2 with initial and final configurations 1s2p ($n_1 l_1 n_2 l_2$) and 3p3d ($n_3 l_3 n_4 l_4$). the matrix element is then

$$\langle \psi_1(1s)\psi_2(2p)| \underline{r}_1 + \underline{r}_2 |\psi_1(3p)\psi_2(3d)\rangle \tag{163}$$

$$= \langle \psi_1(1s)| \underline{r}_1 |\psi_1(3p)\rangle \times \langle \psi_2(2p)|\psi_2(3d)\rangle + \langle \psi_2(2p)| \underline{r}_2 |\psi_2(3d)\rangle \times \langle \psi_1(1s)|\psi_1(3p)\rangle \tag{164}$$

$$= 0 \tag{165}$$

owing to the orthogonality of the eigenfunctions i.e. $\langle \psi_i(nl)|\psi_i(n'l')\rangle = 0$

So this means the configuration can change by only a single electron jump.

The angular momentum does not depend at all on n, so n can change by anything.

Our basic rules are therefore:

$$\Delta n = \text{anything} \tag{166}$$

$$\Delta l = \pm 1 \tag{167}$$

And the photon carries one unit of angular momentum.

7.3 Angular Momentum Rules

The basic rules apply to the total angular momentum J. So provided the changes in J in a transition allow for one unit of $\hbar$ to be taken by the photon we can make a vector triangle rule to give

$$\Delta J = 0, \pm 1 \quad \text{but not} \quad 0 \to 0 \tag{168}$$

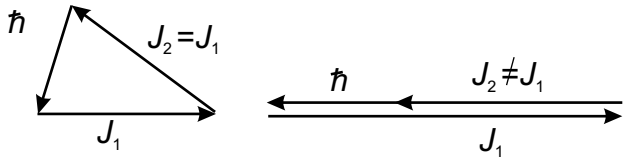


Figure 23: Selection rules reflect conservation of angular momentum including $\hbar$ for the emitted/absorbed photon.

Similar arguments apply to changes in L i.e.

$$\Delta L = 0, \pm 1 \quad \text{but not} \quad 0 \to 0 \tag{169}$$

The rule for S, we have already noted

$$\Delta S = 0 \tag{170}$$

because the dipole operator $-e\underline{r}$ does not operate on the spin part of the wavefunction. Physically this is the sensible notion that the electron spin plays no part in the spatial oscillation.

Finally we noted that changes in M_J are governed by the rule

$$\Delta M_J = 0, \pm 1 \tag{171}$$

There is a peculiar rule, with no obvious physical interpretation, that M_J cannot change from $0 \to 0$ if $\Delta J = 0$. These rules will be apparent only in the presence of an external field that raises the degeneracy in M_J. We consider the effects of external magnetic fields in the next section.

8 Atoms in Magnetic Fields

Atoms have permanent magnetic dipole moments and so will experience a force in an external magnetic field. The effect of magnetic fields on atoms can be observed in astrophysics, eg. in the regions of sunspots, and have some very important applications in basic science and technology. Magnetic fields are used to trap atoms for cooling to within a few nanoKelvin of absolute zero; they play an important role in the operation of atomic clocks and their action on the nuclear spin is the basis of magnetic resonance imaging for medical diagnostics. The basic physics is by now familiar; the magnetic moment of the atom, ($\underline{\mu}_{\text{atom}}$), will precess around the axis of an applied field, $\underline{B}_{\text{ext}}$.

The precessional energy is given by:

$$\hat{H}_{\text{mag}} = -\underline{\mu}_{\text{atom}} \cdot \underline{B}_{\text{ext}} \tag{172}$$

The first question we have to ask is how big is this perturbation energy compared to that of the internal interactions in the atom. We have a hierarchy of interactions in decreasing strength: Central Field, residual electrostatic, spin-orbit, hyperfine. These are represented by the Hamiltonians $\hat{H}_0, \hat{H}_1, \hat{H}_2,$ and $\hat{H}_3$ respectively. The hyperfine interaction is weak compared to that with even the most modest laboratory magnets. So we neglect hfs for the present, but we shall return to it later. At the other end we are unlikely to compete with the Central Field energies of the order of 10eV. (Check this for yourself by estimating $\left\langle \hat{H}_{\text{mag}} \right\rangle$ using μ_B for the atom's magnetic moment and a typical value of B available in a laboratory).

For the same reason $\left\langle \hat{H}_{\text{mag}} \right\rangle$ is usually less than residual electrostatic energies ($\sim$ 1eV) so we then have to decide whether our external magnetic perturbation is bigger or smaller than the internal magnetic, spin-orbit, energy. We can recognise two limiting cases determined by the strength of the external field and the size of the spin-orbit splitting. We define a *weak* field as one for which $\underline{\mu}_{\text{atom}} \cdot \underline{B}_{\text{ext}}$ is less than the fine-structure splitting energy. Conversely, a strong field is one where the external perturbation exceeds the spin-orbit energy.

Just to get the basic physics straight, and to get a feel for what happens, we consider first the simple case where there is no spin-orbit interaction. We select an atom whose magnetic moment is due only to orbital motion and find the effect of B_{ext} on the energy levels. Next we introduce the spin to the problem and the complication of spin-orbit interaction. Then we consider what happens in a strong field and find, pleasingly, that things get simpler again. Finally we look at "weak" and "strong" field effects on hyperfine structure.

8.1 Weak field, no spin

An atom with no spin could be a two-electron atom in a singlet state e.g. Magnesium $3s3p\,^1P_1$. When the field and its perturbation are weak the atomic states will be the usual $|n, l, s, L, S, J, M_J\rangle$ states. In this case, however $J = L$ and we need only the quantum numbers L and M_L . The physics is pictured well by the vector model for an angular momentum $\underline{L}$ and associated magnetic moment $\underline{\mu}_L = -g_l \mu_B \underline{L}$

Recall that $g_L \mu_B$ represents the gyromagnetic ratio for an atom sized circulating charge.

In a magnetic field B_{ext} along the z-axis $\underline{\mu}_L$ (and $\underline{L}$) executes Larmor precession around the field direction with energy:

$$\Delta E_{\text{Z}} = -\underline{\mu}_L \cdot \underline{B}_{\text{ext}} \tag{173}$$

$$= g_L \mu_B \underline{L} \cdot \underline{B}_{\text{ext}} \tag{174}$$

Now $\underline{L} \cdot \underline{B}_{\text{ext}}$ is the projection of $\underline{L}$ onto $\underline{B}_{\text{ext}}$ and in our vector model this is quantized by integer values M_L . Hence

$$\Delta E_{\text{Z}} = g_L \mu_B B_{\text{ext}} M_L \tag{175}$$

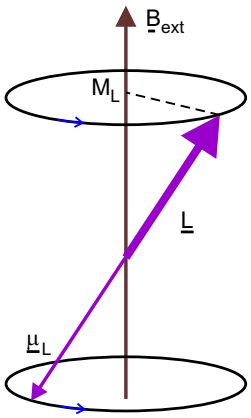


Figure 24: Effect of external magnetic field $\underline{B}_{\text{ext}}$ on an atom with no spin i.e. magnetic dipole due only to orbital motion.

There are $(2L+1)$ values of M_L: $-1 \leq M_L \leq L$ corresponding to the quantized directions of $\underline{L}$ in the field $\underline{B}_{\text{ext}}$. Thus each energy level for a given $\underline{L}$ is split into $(2L+1)$ sub-levels separated by

$$\Delta E'_{\text{Z}} = \mu_B B_{\text{ext}} \tag{176}$$

i.e. independent of L; so all levels are equally split.

The splitting will be observed in transitions between levels. The selection rules on J and M_J will apply to L and M_L in this case $(J = L)$ so

$$\Delta M_L = 0, \pm 1 \tag{177}$$

If we look at our example level $3s3p\,^1P_1$ by a transition to say $3s3d\,^1D_2$ we see 3 states in the $^1P_1$ level and 5 states in the $^1D_2$ level.

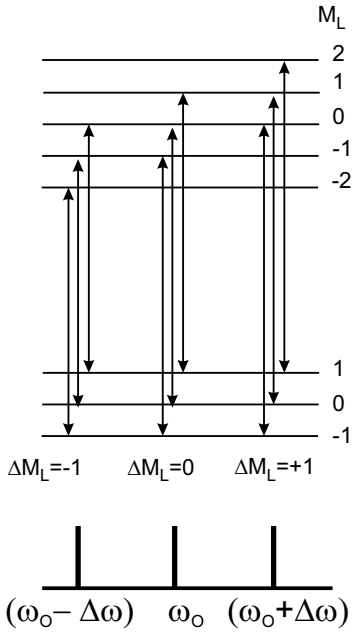


Figure 25: Normal Zeeman effect gives the normal Lorentz triplet due to selection rules $\Delta M_L = 0, \pm 1$ and splitting of all levels into equally spaced sub-levels M_L.

9 transitions are allowed but they form 3 single lines of frequency $\omega_0, \omega_0 \pm \Delta\omega_{\text{Z}}$ where

$$\Delta\omega_{\text{Z}} = \Delta E'_{\text{Z}}/\hbar = \mu_B B_{\text{ext}}/\hbar \tag{178}$$

This is the so-called normal Lorentz triplet. The splitting of the line at ω_0 into 3 components is the Normal Zeeman Effect. (This is why we used Z as the subscript on ΔE_{Z})

It was explained classically, and fully, by Lorentz, long before Quantum Mechanics.

8.2 Weak Field with Spin and Orbit

Since, again, the field is weak there is only a small perturbation so we can use the zero-order wavefunctions defined by the spin-orbit interaction, $|n, L, S, J, M_J\rangle$

The application of an external field will, again, cause the orbital magnetic moment and angular momentum $\underline{\mu}_L$ and $\underline{L}$, to precess. The spin moment and angular momentum will also precess around $\underline{B}_{\mathrm{ext}}$. The perturbation energy operator is then:

$$\hat{H}_4 = g_L \mu_B \underline{L} \cdot \underline{B}_{\mathrm{ext}} + g_S \mu_B \underline{S} \cdot \underline{B}_{\mathrm{ext}} \tag{179}$$

The problem we now have is that $\underline{L}$ and $\underline{S}$ are actually precessing at a faster rate around their resultant $\underline{J}$. The effect of this is that the operators $\underline{L} \cdot \underline{B}_{\mathrm{ext}}$ and $\underline{S} \cdot \underline{B}_{\mathrm{ext}}$ no longer correspond to observables that are constants of the motion.

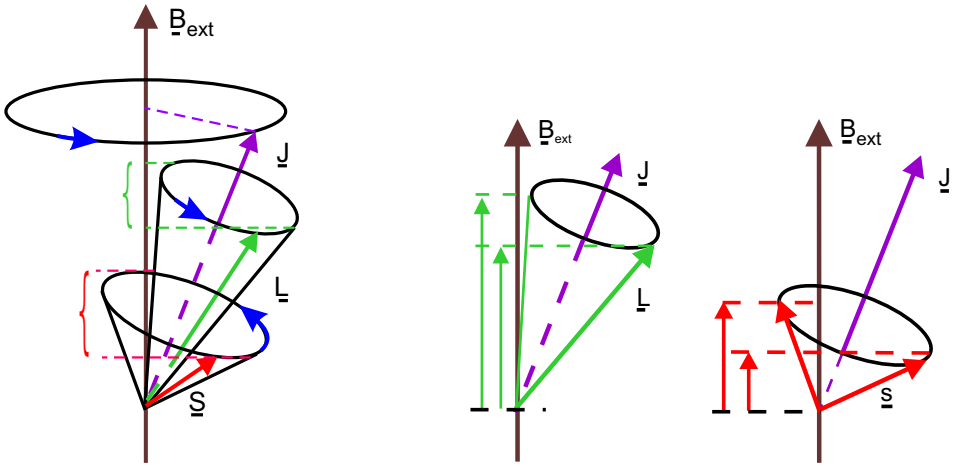


Figure 26: Precession of $\underline{s}$ and $\underline{l}$ around mutual resultant $\underline{J}$ results in projections of $\underline{L}$ and $\underline{S}$ on the field axis (z-axis) not being constants of the motion. So M_L and M_S are not good quantum numbers.

We will use our Vector Model to follow the physics and to help us find the solution. From the diagram we see that the precession of $\underline{S}$ and $\underline{L}$ around the resultant $\underline{J}$ causes the projections of $\underline{S}$ and $\underline{L}$ on the field axis (z-axis) to vary up and down. The diagram suggests that since $\underline{J}$ remains at a constant angle to the z-axis perhaps we could use the total magnetic moment $\underline{\mu}_J$ $(= \underline{\mu}_L + \underline{\mu}_S)$ to calculate the interaction energy i.e.

$$\left\langle -\underline{\mu}_J \cdot \underline{B}_{\mathrm{ext}} \right\rangle \tag{180}$$

There is a snag, however, to this cunning plan. The problem is that $\underline{\mu}_J$ is not parallel or antiparallel to $\underline{J}$. The reason is that the g-factors for orbit and spin are different; $g_L = 1$ and $g_S = 2$. Thus

$$\underline{\mu}_L = \mu_B \underline{L} \quad \text{and} \quad \underline{\mu}_S = 2\mu_B \underline{S} \tag{181}$$

The vector triangle of $\underline{L}$, $\underline{S}$, and $\underline{J}$ is not the same as the triangle of $\underline{\mu}_L$,$\underline{\mu}_S$ and $\underline{\mu}_J$.

All is not lost, however, because the very fact that the spin-orbit precession is so fast means that we can find an effective magnetic moment for the total angular moment that *does* have a constant projection on the field axis, $\underline{B}_{\mathrm{ext}}$. We resolve the vector $\underline{\mu}_J$ into a component along the $\underline{J}$ direction

and a component perpendicular to $\underline{J}$. As $\underline{J}$ precesses around $\underline{B}_{\text{ext}}$ the projection of the perpendicular component of $\underline{\mu}_J$ on z will average to zero. The component along the $\underline{J}$ axis is a constant, which we call the effective magnetic moment, $\underline{\mu}_{\text{eff}}$ i.e.

$$\underline{\mu}_{\text{eff}} = g_J \mu_B \underline{J} \tag{182}$$

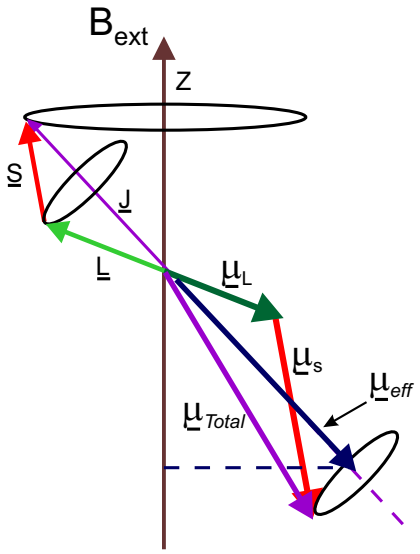


Figure 27: The magnetic moment due to the total angular momentum $\underline{J}$ doesn't lie on the J-axis since the vector sum of spin and orbital magnetic moments $\underline{\mu}_J$ is not parallel to $\underline{J}$ owning to the g-factor for spin being 2 $\times$ g-factor for orbital momentum. The precession of $\underline{s}$, $\underline{l}$ around $\underline{J}$ results in a variation of the projection of $\underline{\mu}_J$ on the z-axis. An effective magnetic moment $\underline{\mu}_{\text{eff}}$ does, however, have a constant projection on the z-axis.

g_J is called the Landé g-factor.

We can now proceed to find the perturbation energy:

$$\Delta E_{\text{AZ}} = g_J \mu_B \left\langle \hat{\underline{J}} \cdot \underline{\hat{B}}_{\text{ext}} \right\rangle \tag{183}$$

$\underline{J} \cdot \underline{B}_{\text{ext}}$ is just the projection of $\underline{J}$ on the z-axis, given by J_z.

$$\Delta E_{\text{AZ}} = g_J \mu_B \underline{B}_{\text{ext}} \left\langle \hat{J}_z \right\rangle \tag{184}$$

and using our $|n, L, S, J, M_J\rangle$ wavefunctions we write the expectation value of $\hat{J}_z$ as its eigenvalue M_J. Hence

$$\Delta E_{\text{AZ}} = g_J \mu_B \underline{B}_{\text{ext}} M_J \tag{185}$$

Landé calculated the form of g_J quantum mechanically, but we can use our Vector Model to get the same result.

Our vector model replaces expectation values by time averages of quantities that are constants of the motion. So we look at our vectors $\underline{L}$, $\underline{S}$ and $\underline{J}$ to find the effective constant projections on the z-axis. We see from the vector diagram that $\underline{L}$ has a constant projection on the $\underline{J}$ axis $\frac{\underline{L} \cdot \underline{J}}{|\underline{J}|}$. This is a vector in the $\underline{J}$ direction so the scalar $\frac{|\underline{L} \cdot \underline{J}|}{|\underline{J}|}$ is multiplied by the unit vector $\frac{\underline{J}}{|\underline{J}|}$, to give $\frac{|\underline{L} \cdot \underline{J}| \underline{J}}{|\underline{J}|^2} = \underline{L}_J$.

Similarly, the projection of $\underline{S}$ on $\underline{J}$ is given by

$$\frac{|\underline{S} \cdot \underline{J}| \underline{J}}{|\underline{J}|^2} = \underline{S}_J \tag{186}$$

The vectors $\underline{L}_J$ and $\underline{S}_J$ have constant time-averaged projections on the axis of $\underline{B}_{\rm ext}$. So the total, time-averaged energy of interaction with the field is:

$$\Delta E_{\rm AZ} = g_L\mu_B\underline{L}_J\cdot\underline{B}_{\rm ext} + g_S\mu_B\underline{S}_J\cdot\underline{B}_{\rm ext} \quad (187)$$

$$= g_L\mu_B\frac{|\underline{L}\cdot\underline{J}|}{|\underline{J}|^2}\underline{J}\cdot\underline{B}_{\rm ext} + g_S\mu_B\frac{|\underline{S}\cdot\underline{J}|}{|\underline{J}|^2}\underline{J}\cdot\underline{B}_{\rm ext} \quad (188)$$

The cosine rule gives the values of $\underline{L}\cdot\underline{J}$ and $\underline{S}\cdot\underline{J}$ in terms of $\underline{J}^2$, $\underline{L}^2$ and $\underline{S}^2$. So with $g_L = 1$ and $g_S = 2$

$$\Delta E_{\rm AZ} = \mu_B\frac{\left[3\underline{J}^2 - \underline{L}^2 + \underline{S}^2\right]}{2|\underline{J}|^2}J_zB_{\rm ext} \quad (189)$$

Now, as per our Vector Model method, we replace the vector (operators) by their magnitudes (eigenvalues);

$$\Delta E_{\rm AZ} = \frac{[3J(J+1) - L(L+1) + S(S+1)]}{2J(J+1)}\mu_BB_{\rm ext}M_J \quad (190)$$

Comparing this with our previous expression we find:

$$g_J = \frac{[3J(J+1) - L(L+1) + S(S+1)]}{2J(J+1)} \quad (191)$$

This is an important result (i.e. remember it – and how to derive it; it is a derivation beloved of Finals examiners.) Its real importance comes from seeing that g_J is the factor that determines the splitting of the energy levels.

$$\Delta E_{\rm AZ} = g_J\mu_BM_JB_{\rm ext} \quad (192)$$

g_J depends on L, S and J and so by measuring the splitting of energy levels we can determine the quantum numbers L, S and J for the level.

We will measure the splitting from the separation of spectral components of a transition in the magnetic field. The energy level splitting will be different in different levels so the pattern observed will not be the simple Lorentz triplet. Consequently, this is known as the Anomalous Zeeman effect. Hence the subscript is AZ on the energy shift $\Delta E_{\rm AZ}$ for spin-orbit coupled atoms in a $\underline{B}$-field.

8.2.1 Anomalous Zeeman Pattern

As an example of the Anomalous Zeeman effect we consider an atom with spin-orbit coupling and sodium is our favourite example. A transition between the $3p^2P_{1/2,3/2}$ levels to the ground level $3s^2S_{1/2}$ illustrates the main features. We look only at the pattern of lines in $^2P_{1/2} \rightarrow ^2S_{1/2}$ and leave the $^2P_{3/2} \rightarrow ^2S_{1/2}$ as an exercise.

First we need to know g_J for each level. Putting in the values of L, S, J for $^2S_{1/2}$ level we find $g_J(3s^2S_{1/2}) = 2$. We could have guessed that! Since there is no orbital motion the magnetic moment is due entirely to spin and the g-factor for spin, $g_S = 2$. Putting in the numbers for $^2P_{1/2}$ we find $g_J(^2P_{1/2}) = 2/3$. We can now draw the energy levels split into two sub-levels labelled by $M_J = +1/2, -1/2$.

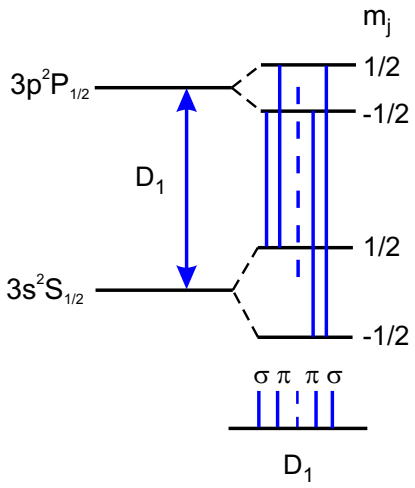


Figure 28: Anomalous Zeeman effect in the $3s3p^2P_{1/2}$ level of Na. Transitions are governed by selection rules $\Delta M_J = 0, \pm 1$.

The transitions are governed by the selection rules: $\Delta M_J = 0, \pm 1$, leading to a pattern of four lines disposed about the zero-field transition frequency, ω_0.

8.2.2 Polarization of the radiation

The magnetic field introduces an axis of symmetry into an otherwise perfectly spherically symmetric system. This has an effect on the polarization of the radiation emitted or absorbed; the oscillating or rotating dipole will have a particular "handedness" in relation to the field, or z-axis. The transition matrix element $\langle \psi_1(r,\theta,\phi)| \, e\underline{r} \, |\psi_2(r,\theta,\phi)\rangle$ represents the change in "position" of the electron and this can be resolved into a motion along the z-axis: $\langle \psi_1| \, ez \, |\psi_2\rangle$ and a circular motion in the x, y plane $\langle \psi_1| \, e(x \pm iy) \, |\psi_2\rangle$. The ϕ-dependence of these integrals can be written as:

$$\langle ez \rangle \sim \int_0^{2\pi} e^{i(m_1 - m_2)\phi} d\phi \tag{193}$$

$$\langle e(x \pm iy) \rangle \sim \int_0^{2\pi} e^{i(m_1 - m_2 \mp 1)\phi} d\phi \tag{194}$$

The z-component is zero unless $|m_1 - m_2| = 0$ i.e. $\Delta m = 0$. Similarly the (x, y) component is zero unless $m_1 - m_2 \mp 1 = 0$ i.e. $\Delta m = \pm 1$

The $\Delta M_J = 0$ transition is thus identified with a linear motion parallel to the $\underline{B}$ field. Viewed at right angles to the field axis this will appear as linearly polarized light, π-polarization.

Viewed along the z-axis, however, no oscillation will be apparent so the π-component is missing in this direction.

Now consider the $(x \pm iy)$ component arising from $\Delta m = \pm 1$ transitions. This appears as a circular motion in the x, y plane around the z-axis. Thus $\Delta m = \pm 1$ transitions give circularly polarized σ^+, σ^- (right/left) light viewed along z, and plane polarized in the x, y plane.

It is possible using appropriate optical devices to determine the "handedness" of the σ^--polarized emissions along the z-axis. The σ^+ and σ^- transitions are either at higher or lower frequency respectively than ω_0 depending on the sign of the oscillating charge. If you can remember the conventional rules of electromagnetism you can work out the sign of the electron. Lorentz (who could remember these rules) was able to use this effect to show that electrons were negatively charged.

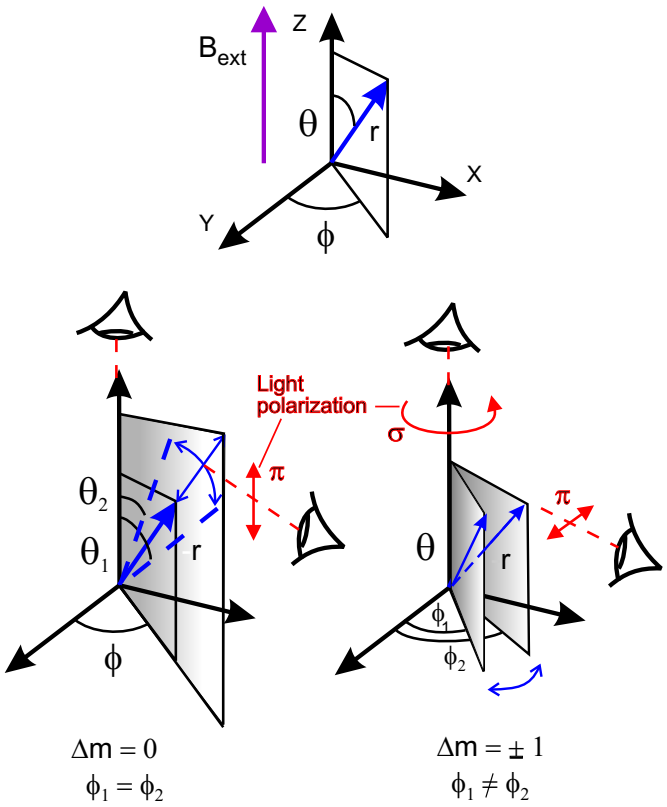


Figure 29: Polarization of Zeeman components viewed along field axis (z-axis) and along a perpendicular axis (x- or y-axis).

8.3 Strong fields, spin and orbit

The word "strong" in this context means that the interaction of the orbit and spin with the external field is stronger than the interaction of the orbit with the spin. The magnetic moments $\underline{\mu}_L$ and $\underline{\mu}_S$ will then precess much more rapidly around $\underline{B}_{\text{ext}}$ than they do around each other. Since $\underline{L}$ and $\underline{S}$ are now essentially moving independently around $\underline{B}$, the vector $\underline{J}$ becomes undefined – it is no longer a constant of the motion. Our spin-orbit coupled wavefunctions $|n, L, S, J, M_J\rangle$ are no longer a good basis for describing the atomic state; J, M_J are not good quantum numbers. Instead we can use LM_L and SM_S to set up a basis set of functions $|n, L, S, M_L, M_S\rangle$. The spin-orbit interaction is now relatively small and we ignore it to this approximation. The energy shifts are now given by the expectation of the operator $\hat{H}_4$:

$$\hat{H}_4 = g_L \mu_B \hat{\underline{L}} \cdot \underline{\hat{B}}_{\text{ext}} + g_S \mu_B \hat{\underline{S}} \cdot \underline{\hat{B}}_{\text{ext}} \tag{195}$$

Now we don't have the complication of coupled spin and orbital motion. Using our $|n, L, S, M_L, M_S\rangle$ basis functions we find

$$\left\langle \hat{H}_4 \right\rangle = \Delta E_{\text{PB}} = (M_L + 2M_S)\mu_B B_{\text{ext}} \tag{196}$$

The energy levels are thus split into $(2L + 1)$ levels labelled by M_L , each of which will be split into levels labelled by M_S.

Consider our $3p^2P_{1/2}$ level of Na.

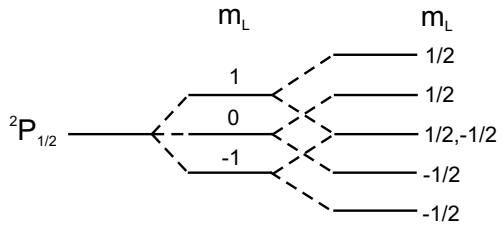


Figure 30: The splitting of a level in a strong field ignoring the effect of spin-orbit interaction.

We can now include the spin-orbit effect as a small perturbation on these states $|n, L, S, M_L, M_S\rangle$. Again, we can use our Vector Model to find the expectation value of the spin-orbit operator $\hat{\underline{S}} \cdot \hat{\underline{L}}$.

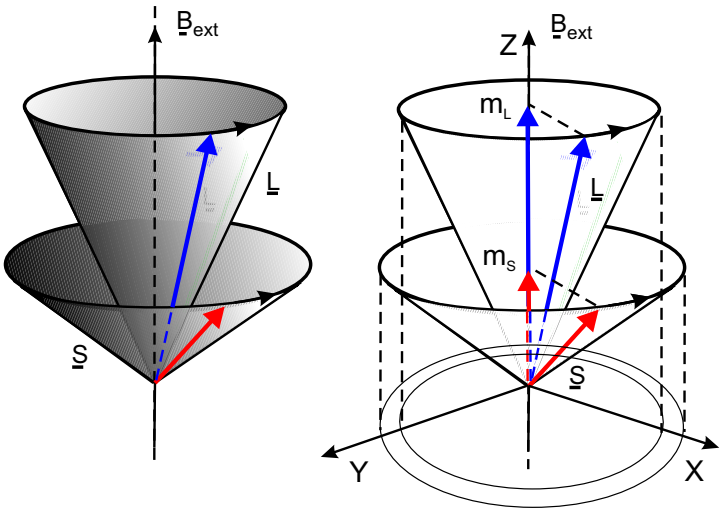


Figure 31: Strong field effect on orbital and spin angular momenta: $\underline{l}$ and $\underline{s}$ precess more rapidly around the external field relative to their mutual precession. The projections of $\underline{l}$ and $\underline{s}$ on the x-y plane average to zero. As a result only the projections on the z-axis remain to define the energy in the field.

Since $\underline{L}$ and $\underline{S}$ are precessing rapidly around the z-axis their components in the x, y-plane will average to zero. We are left only with the z-component so:

$$\xi \left\langle \hat{\underline{S}} \cdot \hat{\underline{L}} \right\rangle = \xi M_S M_L \tag{197}$$

The six states ($M_L = 1, 0, -1, M_S = +1/2, -1/2$) have spin orbit shifts of $(\xi/2, 0, -\xi/2, -\xi/2, 0, \xi/2)$ as shown on the extreme right of the strong field energy level diagram.

The energy level shifts ΔE_{PB}, ignoring the relatively small spin-orbit effect, lead to a set of evenly spaced energy levels. Transitions between these strong field levels are illustrated by the $3p^2P_{1/2} \rightarrow 3s^2S_{1/2}$ transition in Na. The selection rules operating are $\Delta M_L = 0, \pm 1$ and $\Delta S = 0$.

The ΔM_L rule derives from the conservation of angular momentum (including the photon). The $\Delta S = 0$ rule is due, again, to the fact that the dipole operator cannot change the spin. The spin plays no part, so we expect to see a pattern of lines similar to transitions in an atom with no spin i.e. the Normal Zeeman triplet. The allowed transitions do indeed fit this simple pattern. The splitting in the strong field is known as the Paschen-Back effect, hence our ΔE_{PB}.

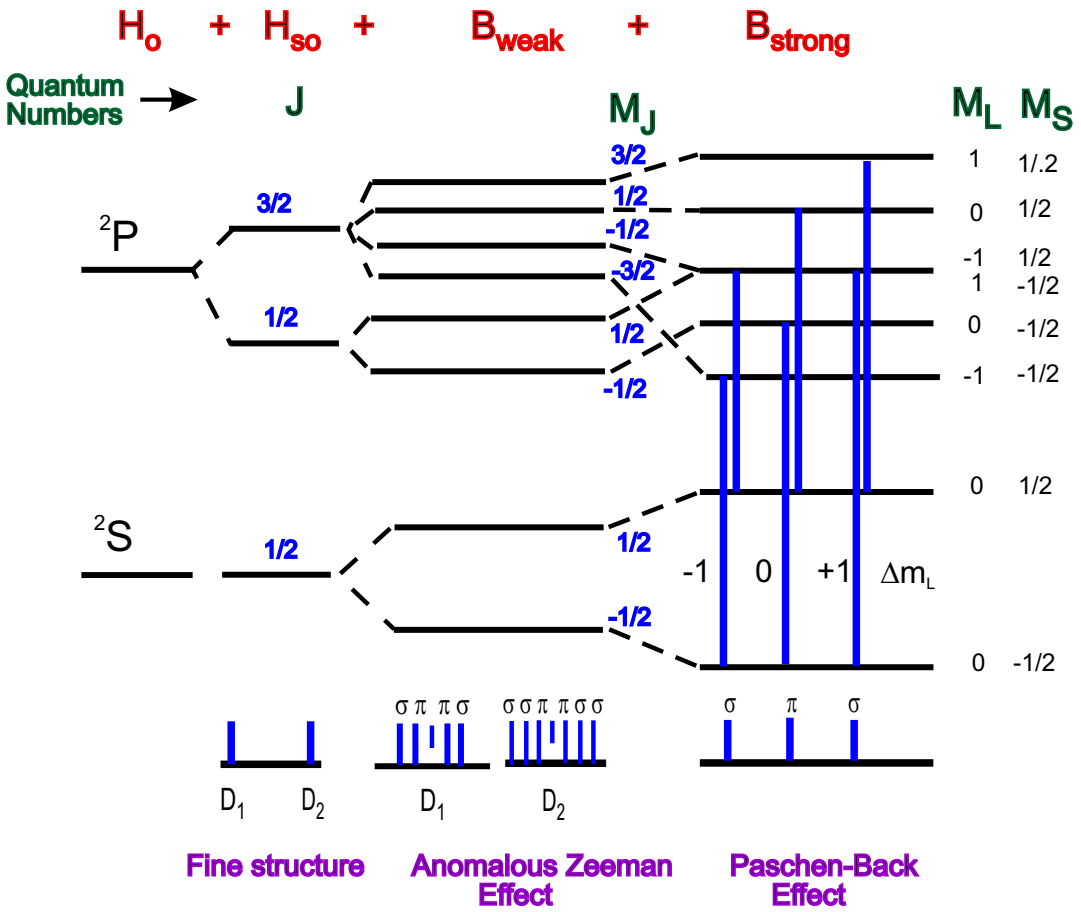


Figure 32: Diagram showing splitting of Na 3p-levels from zero field, through weak to strong field showing change from Anomalous Zeeman to Paschen-Back effect.

8.4 Intermediate fields

When the external field is neither strong nor weak the situation is very complicated. The Vector Model won't work and the result can be calculated quantum mechanically in only the simplest of cases eg. ground state of atomic hydrogen. The energy levels can be found experimentally. The transition from weak to strong fields for the $^2P_{1/2,3/2}$ term is shown in the diagram. Two general rules are obeyed. The state of M_J in a weak field goes over continuously to a state $M_L + M_S = M_J$ in the strong field. This is because the projection of the total angular momentum on the z-axis is a constant of the motion. The second rule is that states of the same M_J never cross. You don't need to know why – in fact you don't need to know anything about "intermediate" fields for this course, but I thought some people might be interested.

8.5 Magnetic field effects on hyperfine structure

We start this section by reminding ourselves of what was said at the beginning of our discussion of magnetic field effects. We consider the interactions in descending order of their perturbation energies. It is the relative magnitude of the interaction *energy* that determines whether an external field is weak or strong relative to the spin-orbit interaction. In the same way an external field will be weak or strong in the context of the hyperfine interaction *with* the electron and nuclear magnetic moments relative to the interaction *between* the electron and nuclear moments. So we compare interaction (i.e. perturbation) energies and not absolute field strengths. Internal fields at the nucleus can be of the order of 100T, so it would be difficult to generate a comparable field using an external magnet.

The nuclear spin moment, however, is very small relative to the electron moments in the ratio of $\sim m_e/m_p$. So the nuclear interaction with the external field is much weaker than that with the electrons. It is possible therefore to generate interaction energies involving the nuclear spin that may be stronger or weaker than the internal energies. The relevant interaction energies are:

$$A_J \underline{I} \cdot \underline{J} + g_J \mu_B \underline{J} \cdot \underline{B}_{\text{ext}} - g_I \mu_N \underline{I} \cdot \underline{B}_{\text{ext}} \tag{198}$$

The first term is the nuclear/electron interaction giving hfs. The second term is the interaction between the magnetic moment $\underline{\mu}_J$ due to total electronic angular momentum $\underline{J}$ and the external field. The third term is the direct interaction of the external field with the nuclear spin $\underline{I}$. Since the nuclear moment $\mu_I << \mu_J$ we can neglect this third term. We are left then only with the hyperfine energy $A_J \underline{I} \cdot \underline{J}$ and the electron interaction energy with the external field. We can now define weak and strong fields in this context:

$$\text{Weak field} \quad A\underline{I} \cdot \underline{J} \quad >> \quad g_J \mu_B \underline{J} \cdot \underline{B}_{\text{ext}} \tag{199}$$

$$\text{Strong field} \quad A\underline{I} \cdot \underline{J} \quad << \quad g_J \mu_B \underline{J} \cdot \underline{B}_{\text{ext}} \tag{200}$$

8.5.1 Weak field

In this case the wave functions defined by the hfs interaction remain a good description of the system, $|(n, L, S, J, I)F, M_F\rangle$. So we evaluate the perturbation energy for the external field interaction in this basis,

$$\langle (n, L, S, J, I)F, M_F | - \underline{\mu}_F \cdot \underline{B} \, | (n, L, S, J, I)F, M_F \rangle \tag{201}$$

$$\Delta E_F = -\underline{\mu}_F \cdot \underline{B}_{\text{ext}} \tag{202}$$

Using our Vector Model, as before, we identify $\underline{F}$, the total angular momentum, and $\underline{F}_z$ its projection on the field axis as constants of the motion. $\underline{F}$ precesses slowly around $\underline{B}$ compared to fast precessions of $\underline{J}$ and $\underline{I}$ around $\underline{F}$.

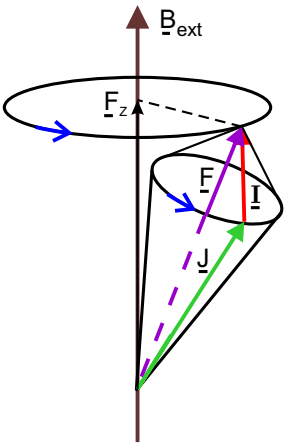


Figure 33: Slow precession of $\underline{F}$ around $\underline{B}_{\text{ext}}$ whilst $\underline{I}$, $\underline{J}$ precess rapidly around $\underline{F}$.

Recall that for $\underline{L}$ and $\underline{S}$ interacting with an external field, we took the projections on $\underline{J}$ first and then projected these components onto the field. We could do exactly the same here with $\underline{J}$ and $\underline{I}$. We have, however, decided to ignore the interaction of $\underline{I}$ with $\underline{B}_{\text{ext}}$. So we take only the projection of $\underline{J}$ onto $\underline{F}$ to define an effective magnetic moment $\underline{\mu}_F$:

$$\underline{\mu}_F = -g_F \mu_B \underline{F} \tag{203}$$

$$= g_J \mu_B \frac{\underline{J} \cdot \underline{F}}{|\underline{F}|^2} \underline{F} \tag{204}$$

[Recall that $g_J \mu_B \underline{J}$ is the effective magnetic moment along the $\underline{J}$ axis.] Using our vector triangle from $\underline{F} = \underline{I} + \underline{J}$ we find

$$\underline{J} \cdot \underline{F} = \frac{1}{2} \left\{ \underline{F}^2 + \underline{J}^2 - \underline{I}^2 \right\} \tag{205}$$

Hence

$$g_F = g_J \frac{F(F+1) + J(J+1) - I(I+1)}{2F(F+1)} \tag{206}$$

The total energy is then, including the hfs term:

$$\Delta E = \frac{A_F}{2} \left\{ F(F+1) - J(J+1) - I(I+1) \right\} + g_F \mu_B \underline{F} \cdot \underline{B}_{\text{ext}} \tag{207}$$

The second term is given by the projection of $\underline{F}$ on $\underline{B}_{\text{ext}}$ i.e. M_F. So each hyperfine level is split by an amount

$$\Delta E_F = +g_F \mu_B B_{\text{ext}} M_F \tag{208}$$

i.e. in the external field each F-level splits into $2F + 1$ levels labelled by $M_F = -F \leq M_F \leq F$

As an example we take ^{39}K in its ground state $4s^2S_{1/2}$. For a $^2S_{1/2}$ state $g_J = 2$, and for ^{39}K, $I = 3/2$, so $F = 2$ or 1. Putting in the numbers we find the g_F-values for each hfs level: $F = 2$; $g_F = 1/2$ and $F = 1$; $g_F = -1/2$. (For $J = 1/2$ levels the values of g_F for each F-value will always be equal and opposite in sign.)

The negative sign on g_F for $F = 1$ arises because the $\underline{J}$-projection on $\underline{F}$ for this level is in the opposite direction to that for $F = 2$. This is easily seen from the vector diagram for $\underline{F}$, $\underline{I}$ and $\underline{J}$. As a result the ordering of the M_F states is inverted for $F = 1$.

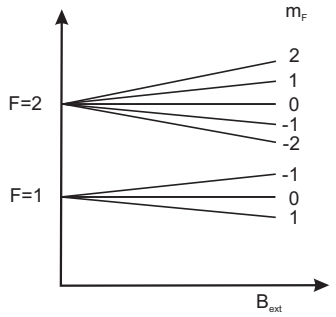


Figure 34: Splitting of hyperfine levels $F = 1$, $F = 2$ in a weak field.

8.5.2 Strong field

In a strong field the electronic angular momentum $\underline{J}$ couples more strongly to $\underline{B}_{\text{ext}}$ than the nuclear spin $\underline{I}$. As a result the total angular momentum $\underline{F}$ is undefined; F, M_F are no longer good quantum numbers. The states are defined now firstly by the angular momentum $\underline{J}$ interacting with $\underline{B}_{\text{ext}}$, and its projection on the field, M_J. As before, this leads to an energy shift:

$$\Delta E_J = g_J \mu_B M_J B_{\text{ext}} \tag{209}$$

What about the nuclear spin $\underline{I}$? We are ignoring its direct interaction with $\underline{B}_{\text{ext}}$. Its interaction with $\underline{J}$, however, needs to be included. The nuclear spin $\underline{I}$, and its moment $\underline{\mu}_I$, wants to precess around $\underline{J}$ and $\underline{\mu}_{\text{eff}}$, its magnetic moment along the $\underline{J}$ axis. Now $\underline{J}$, however, is precessing rapidly around $\underline{B}_{\text{ext}}$. The result is that $\underline{I}$ can follow only the constant component of $\underline{J}$ along the z-axis. The interaction of $\underline{I}$ with the components of $\underline{J}$ in the x, y plane average to zero.

The average energy then is $\propto \langle \underline{I} \cdot \underline{J}_z \rangle$

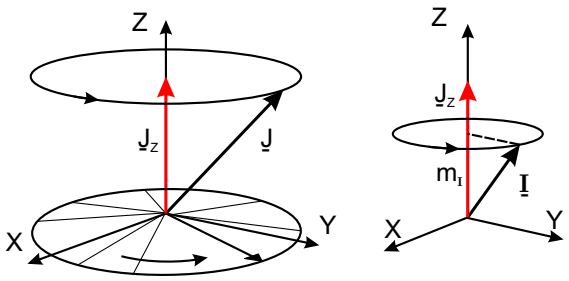


Figure 35: Effect of strong field on hfs: the total electronic angular momentum $\underline{J}$ precesses rapidly around $\underline{B}_{\text{ext}}$ (z-axis). As a result the nuclear spin $\underline{I}$ can follow only the time averaged projection of $\underline{J}$ on the z-axis.

The component $\underline{I}$ along z is just M_I and the J_z component is given by M_J. The nett effect of the $A_J \underline{I} \cdot \underline{J}$ term is to give an energy, $A_J M_I M_J$ so the strong field energy level shifts are:

$$\Delta E_{\text{BG}} = A_J M_I M_J + g_J \mu_B M_J B_{\text{ext}} \tag{210}$$

Although we call this the strong field result, the internal field from the electron's angular momentum $\underline{J}$, is still much stronger than the external field. $\underline{I}$ and $\underline{J}$ are decoupled because $\underline{J}$ precesses much more rapidly than $\underline{I}$ around $\underline{B}_{\text{ext}}$. The $A_J M_I M_J$ term arises from the nuclear spin precession around the time averaged value of $\underline{J}$ which happens to lie along the $\underline{B}_{\text{ext}}$ axis. We stress that we are ignoring the (very slow) precession of $\underline{I}$ around $\underline{B}_{\text{ext}}$ from a weak direct interaction of the $\underline{\mu}_I$ with the external field. The level splitting is known as the Back-Goudsmit effect, hence ΔE_{BG} .

The strong field states are labelled by M_J, with each M_J split into $(2I+1)$ states according to its value of M_I.

The unperturbed hfs levels eg. $F = 1, 2$ in ^{39}K are split by $\sim$ a few hundred MHz. To give a comparable energy an external field needs only to be $\sim 10^{-2}$T. So a strong field in this case is only $\underline{B}_{\text{ext}} \geq 0.1T$. The separation of the M_J levels is linearly proportional to $\underline{B}_{\text{ext}}$ in the strong field regime. The separation of the M_I sub-states remains constant as $\underline{B}_{\text{ext}}$ increases.

The transition between weak and strong field limits can be plotted through intermediate field strengths using similar arguments to that for M_J to M_L, M_S labelled states.

- The total M remains a good quantum number: $M_F \rightarrow (M_I + M_J)$
- States of the same M do not cross.

In addition we note that states with a unique value of M e.g. $M_F = 2$ in the $4s^2S_{1/2}$ (F=2) level of ^{39}K have energies linearly dependent on $\underline{B}_{\text{ext}}$. The "no-crossing" rule for other M-states leads to a "repulsion" of states of same M. The energies eg. $M = 1, 0, -1$ in $F = 2$ bend away from the $M = 1, 0, -1$ states in $F = 1$. This non-linear behaviour is characteristic of states becoming "mixed", i.e. neither M_F nor M_I, M_J are good quantum numbers in the intermediate field regime.

9 X-Rays: transitions involving inner shell electrons

We have been concerned, so far, with the energy levels of the valence electrons i.e. those electrons outside the filled shells (or sub-shells). Transitions between these levels involve photons in the energy range of $\sim$ 10eV down to 10^{-6}eV. These energies correspond to wavelengths of $\sim 10^{-7}$m to 1m. (Ultra Violet, visible, infra-red, microwaves and radio waves). When photon energies are much larger than 10eV i.e. in the keV range then the interaction can disturb the tightly bound, inner-shell electrons. Conversely, transitions of an inner-shell electron from one shell to another will involve emission or absorption of the high energy, short-wavelength photons; X-rays. Usually, however, all the energy levels of the inner shells are occupied, so, in order to allow transitions to occur at all, we first have to create a vacancy in one of the inner shells. The energy required to do this can come from either the absorption of a sufficiently energetic photon, or from the kinetic energy of impact with a high energy electron.

9.1 X-ray Spectra

X-rays are generated when high energy electrons strike a solid target e.g. the metal anode in a "cathode-ray" tube. The spectrum of the X-rays generated in this way consists firstly of a continuous range of wavelengths down to a limiting value corresponding to the maximum energy of the incident electrons. These X-rays are the result of the deceleration of the charged particles and are known as "bremsstrahlung" or "braking radiation".

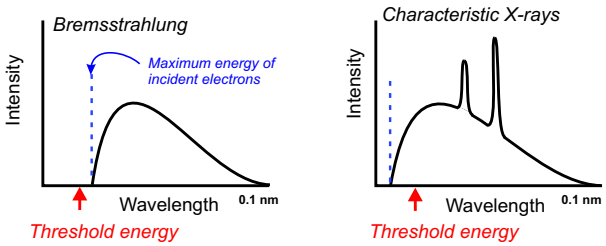


Figure 36: X-ray spectra consist of bremsstrahling and characteristic lines from inner-shell transitions.

When the energy of the electrons is increased above a certain value, for a given target material, sharp peaks i.e. discrete lines, appear superimposed on the continuous "bremsstrahlung". The spectrum of these discrete lines are characteristic of the target element. These characteristic X-rays have the following properties:

- The wavelengths fit a simple series formula.
- All the lines of a particular series appear together once the incident electrons exceed a particular threshold energy.
- The threshold energy for a particular series just exceeds the energy of the shortest wavelength in the series.
- Above a certain energy, no new series appear.

These observations are explained by the following process:

The incident energy (from electron impact) is transferred to an inner-shell electron. If this energy is sufficient the inner-shell electron is raised to a vacant energy level. Now the vacancy energy levels are those lying between the atoms ground state energy and the ionization limit. This is a range of only 10eV or so. If the incident energy is of the order of 10^3eV, or greater, then the most likely result is to ionize the atom i.e. the inner-shell electron escapes with a kinetic energy equal to the

incident impact energy less the binding energy of the inner shell. An electron from a higher inner shell may "fall" into the vacancy. As a result of this transition an X-ray photon is emitted with an energy corresponding to the difference in binding energy of the two shells. For example, creation of a vacancy in the $n = 3$ shell allows electrons from $n = 4$, or higher shells to "fall" into the $n = 3$ vacancy. (Higher energy impacts may eject electrons from deeper shells $n = 2$ and $n = 1$) These transitions are the source of the discrete, characteristic, X-ray lines.

X-ray spectroscopy developed its own nomenclature and it is still (unfortunately) used, so we have to live with a further set of labels. In the context of X-rays the $n = 1, 2, 3$ shells are knows as K, L, M etc. respectively.

9.2 X-ray series

The X-rays emitted in transitions between inner-shell energy levels will have energies corresponding to the difference in binding energy of the electrons in the two shells concerned. The binding energy of an electron in a given shell of quantum number, n, may be expressed using a hydrogenic model:

$$E_n = \frac{R(Z - \sigma_n)^2}{n^2} \tag{211}$$

Where R is Rydberg's constant. σ_n is a screening factor that accounts for the effect of the other electrons.

For the K-shell ($n = 1$) there are 2 electrons. As the ejected electron moves outwards the remaining electron provides a spherically symmetric shell around the nucleus of Z protons and reduces the effective nuclear charge to $(Z - 1)$. The other electrons in higher shells also make a contribution to the screening. The total screening factor, σ_k is then approximately 2. It is difficult to calculate screening factors, although good estimates can be made using atomic structure calculations. Usually we reply on experimental (empirical) values for σ. (The screening factors also depend on the angular momentum of the states involved.)

Transitions from higher shells to a vacancy in the K-shell give rise to a series of lines. The wavenumber ($\overline{\nu} = 1/\lambda$) of these lines will be given by the differences in the binding energies:

K-series:

$$\overline{\nu}_K = R\left\{\frac{(Z - \sigma_K)^2}{1^2} - \frac{(Z - \sigma_i)^2}{n_i^2}\right\} \tag{212}$$

Where $n_i = 2, 3, 4$ etc.

In general:

$$\overline{\nu}_X = R\left\{\frac{(Z - \sigma_i)^2}{n_i^2} - \frac{(Z - \sigma_j)^2}{n_j^2}\right\} \tag{213}$$

With n_i, n_j integers and $n_i < n_j$.

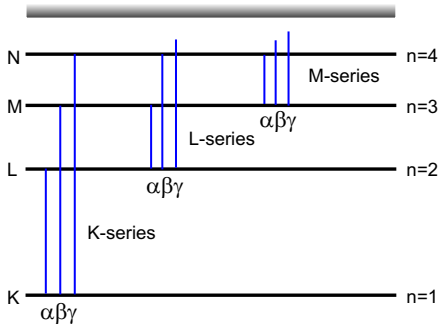


Figure 37: Origin of X-ray series from vacancies created in inner shells.

The longest wavelength series member is labelled α, with successive lines denoted β, γ etc.

9.3 Fine structure of X-ray spectra

A single vacancy in an otherwise full shell has the properties of a single electron in an otherwise empty shell. The X-ray energy levels therefore resemble those of hydrogen or alkali atoms. The energy levels are split into *terms* and the terms are split by spin-orbit interactions giving "fine structure". The energy splitting due to fine structure can be written

$$\Delta E_{\rm fs} = \frac{5.8Z^4}{n^3 l(l+1)} \tag{214}$$

The levels are labelled by quantum numbers (n, l, s, j). The "Z^4" factor results in very large "fine structure" splitting for heavy elements (large Z) eg. 10^7cm^{-1} in Uranium!

This structure was relatively easy to measure and for a long time such measurements gave the most accurate values of α, the fine structure constant. The X-ray lines have a multiplet structure governed by selection rules:

$$\Delta l = \pm 1, \Delta j = 0, \pm 1 \tag{215}$$

e.g. the K_α line becomes a doublet K_{α_1}, K_{α_2}.

9.4 X-ray absorption

Absorption spectra in the visible or UV-range of the spectrum consist of a series of discrete lines whose wavelengths coverage to a series limit; the ionization limit. The strength of the lines decreases also towards the ionization limit. For shorter wavelengths than this limit the absorption is spectrally continuous and continues to decrease in strength as the absorbed wavelengths get shorter.

The discrete absorption spectrum, the series of discrete lines, is the result of transitions of a valence electron to higher, vacanct, energy levels. Beyond the series limit, absorption results in photoionization; the bound electron is excited to an unbound, free state known as the continuum. The probability of the photoionization decreases with increasing energy of the incident photon because it becomes increasingly difficult for the photon/atom system to conserve both energy and momentum in the interaction. The valence electrons are relatively weakly bound to the massive nucleus and so cannot easily transfer momentum to the nucleus.

When the incident photons have X-ray energies they may raise an inner shell electron to an empty shell (valence shell) or eject it from the atom (photoionization). On the scale of X-ray energies the valence shell energies are negligible so photoionization is the most likely result. Once the threshold energy for photoionizing an inner-shell electron is exceeded there is a sudden increase in the absorption probability. The sharp increase in absorption coefficient associated with such a threshold is called an *absorption edge*.

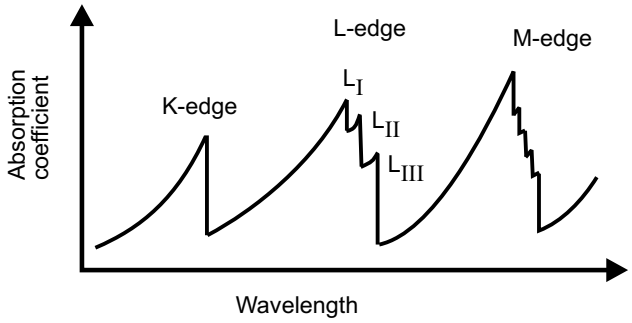


Figure 38: Absorption spectra are characterised by sharp edges with fine-structure. The absorption coefficient falls off for X-ray energies above each edge.

These absorption edges are labelled according to the quantum number of the shell; K edge ($n = 1$) L edge ($n = 2$) M, N etc. Each edge exhibits the fine structure of the shells. Thus the M edge ($n = 3$) has 5 subsidiary edges associated with spin-orbit splitting of the angular momentum states, $^2S_{1/2}$, $^2P_{1/2,3/2}$, $^2D_{3/2,5/2}$. The L-edge has 3 steps and the K-edge is single.

When an absorption edge is examined with high spectral resolution it may be found to consist of a few broad peaks that merge into the continuum. These features cover a range of typically 10eV and correspond to transitions from the inner shell to one of the vacancy valence electron shells. These absorption features are broadened by "lifetime broadening" since the Einstein A-coefficient scales with ν^3, the cube of the transition frequency. (Recall that A=$1/\tau$, where τ is the lifetime of the upper state against radiative decay.) As a result of this broadening only a few of these transitions can be resolved.

Above an absorption edge, the absorption coefficients drops off until the photon energy exceeds the next inner shell binding energy and a new edge is observed. At the K-edge, for example, the photon energy is capable of ejecting an L or M shell electron. It is more likely, however to eject the K-shell electron which is more strongly bound to the nucleus than any of the higher shell electrons. As a result the excess momentum is more effectively transferred to the nucleus. So those electrons held most strongly to the nucleus are most effective in absorbing X-ray photons. This effect partly explains the rapid change in absorption at the edge and the fall-off in absorption for energies above each edge.

9.5 Auger Effect

The creation of a vacancy in, say, the K-shell is followed by one of two processes.

- Emission of characteristic X-rays as already described.
- The ejection of a second electron and emission of longer wavelength X-rays. This is the Auger effect.

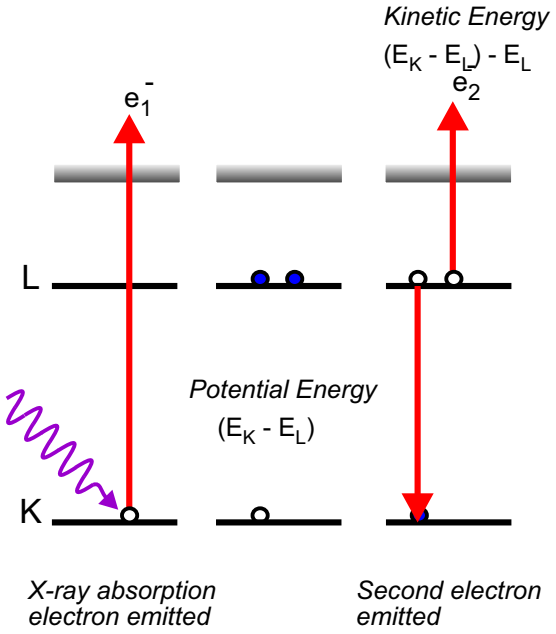


Figure 39: Ejection of a second electron following ejection from an inner shell, resulting from excess energy given to an electron when one electron in the same shell falls to a lower vacancy.

The Auger effect arises because the vacancy in the lower shell (in this case the K-shell creates potential energy that is shared by all the L-shell (and higher shell) electrons. When one L-electron

falls in the K-shell vacancy it can give up its energy either as an X-ray (K_α line) or as kinetic energy to another L-shell electron. The donated energy is $E_K - E_L$ and may exceed the L-shell binding energy. If $(E_K - E_L) > E_L$ then this L-shell electron has enough kinetic energy to escape. The resulting ejected electron has kinetic energy $(E_K - E_L) - E_L = E_K - 2E_L$.

This Auger effect is analogous to autoionization from doubly excited states in two electron atoms.

There are now two vacancies in the L-shell that can be filled by electrons "falling" from higher shells. This leads to emission of longer wavelength X-rays than the K-series or further Auger processes may occur.

10 High Resolution Laser Spectroscopy

10.1 Absorption Spectroscopy

Conventional absorption spectroscopy uses a spectrograph or a monochromator to record the spectrum of "white" light that has passed through an absorbing gas ("White" light is a broad, continuous spectrum from a black body.) A spectrograph records the whole spectrum at once on a photographic film or CCD camera. In a monochromator the film or CCD camera is replaced by a slit (exit slit) similar to the entrance slit of the spectrograph. The spectrum is recorded by scanning the instrument e.g. by rotating the grating, so that the frequency of the light exiting the slit is swept across a range of frequencies. Absorption lines then show up as dips in the recorded intensity as a function of frequency (or wavelength). The resolution is determined by the narrow range of frequencies that are passed by the exit slit, the bandwidth.

The resolving power of spectrographs is limited to about $10^4 - 10^5$. Fabry-Perot interferometers can achieve resolving powers of $\sim 10^6 - 10^7$. Taking 10^5 as a typical value for resolving power and 10^{14}Hz as an order of magnitude for optical frequencies we find conventional absorption spectroscopy can resolve frequency differences of 10^9 or $\sim$1GHz.

1GHz is also the order of magnitude of Doppler broadening of atomic spectral lines. So if we want to improve the resolution of spectral measurements we need to do two things; increase the resolving power and eliminate the Doppler broadening. Laser spectroscopy allows us to do both. (Note that pressure broadening is also typically $\sim$1GHz at atmospheric pressure but this can be reduced by reducing the pressure in the absorbing gas to a few mbar.)

10.2 Laser Spectroscopy

10.3 Spectral resolution

Instead of using a wide continuous spectrum and scanning a monochromatic filter across it, an alternative method is to use a monochromatic light source and to scan its frequency across the spectrum. Certain kinds of laser provide just such a frequency tunable, monochromatic light source. [How these lasers work is not part of the Atomic Physics syllabus. Briefly, however, they work as follows. A laser is essentially a large Fabry-Perot interferometer, consisting of a cavity between two highly reflecting mirrors. Light bouncing between the mirrors forms standing waves of wavelength λ. The cavity length is an integer number of half-wavelengths $L = m\lambda/2$. The wavelengths $\lambda_m = 2L/m$ are called cavity modes with frequencies $\nu_m = m(\frac{c}{2L})$. The cavity modes form a comb of frequencies separated by the Free Spectral Range, $c/(2L)$, that are essentially the Fabry-Perot transmission peaks. The sharpness of each peak is set partly by the Finesse of the cavity i.e. by the reflectivity of the mirrors. The spectral width of the modes of a laser however will be much sharper than those of the simple Fabry-Perot. This is because there is an amplifier inside the cavity that works by stimulated emission. More intense parts of the mode spectrum are amplified at the expense of weaker parts and so the peaks grow much faster than the wings of the modes. When additional filters are placed inside the cavity a single mode may come to dominate the laser action. All of the power goes into this single mode and we have an essentially monochromatic source. By adjusting the filters inside the cavity and also the cavity length, the frequency of the mode can be tuned continuously over a small range.] The spectral width of a laboratory single-mode laser is typically $\sim$1MHz i.e. 10^3 less than the Doppler width, giving a potential resolving power of $10^{14}/10^6 \sim 10^8$. It now remains to reduce or eliminate the Doppler width.

10.4 "Doppler Free" spectroscopy

Doppler broadening comes about because atoms in a gas have a Maxwell-Boltzmann distribution of speeds. The radiation emitted (or absorbed) by atoms moving with speed v is shifted from its stationary value ν_0 to ν:

$$\nu = \nu_0 \left(1 \pm \frac{v}{c}\right) \tag{216}$$

The number of atoms radiating this frequency is given by the number having speed between v and $v + dv$

$$dN = N_0 e^{-\frac{Mv^2}{2kT}} dv \tag{217}$$

Where N_0 is a constant, M is the atomic mass, k is Boltzmann's constant and T is the absolute temperature. The intensity radiated (or absorbed) as a function of frequency is therefore

$$I(\nu) = I(\nu_0) \exp\left[-\frac{Mc^2}{2kT}\left(\frac{\nu - \nu_0}{\nu_0}\right)^2\right] \tag{218}$$

This is a Gaussian curve with half-width (FWHM):

$$\Delta\nu_D = \frac{2\nu}{c}\left[\frac{2kT}{M}\log_e 2\right]^{1/2} \tag{219}$$

(Note this is roughly $\Delta\nu_D \sim \nu(v/c)$, where v is the approximate speed of the atoms $\sim$ speed of sound.) This expression for the "Doppler width" suggests some ways to reduce its value.

Firstly we could cool the atoms i.e. reduce T. Practically we can do this only by cooling the atoms using liquid nitrogen (77°K) in a discharge tube. Since $\Delta\nu_D \propto \sqrt{T}$ this gains us only a factor of ~ 2.

Secondly we could use low frequency transitions since $\Delta\nu_D \propto \nu$. Radiofrequency transitions can be detected between hfs levels and this reduces the Doppler width to negligible proportions. Often, however, we want to study optical transitions (UV, visible) so we now turn to methods using lasers to reduce or eliminate the Doppler width.

10.4.1 Crossed beam spectroscopy

The simplest way to reduce the Doppler effect of moving atoms is to observe them at right angles to their velocity. (This eliminates the first-order Doppler shift, but not the , much smaller, relativistic shift.) Atoms are heated in an oven, allowed to escape through a slit and are collimated by a second slit. A tunable laser beam crosses the resulting atom beam at 90° and so "sees" atoms with essentially zero velocity. The laser light is absorbed only when it is exactly resonant with the atoms' transition frequency. The absorption is detected using the light emitted by the atoms as they decay from the excited state. This is more sensitive that measuring the very small change in the transmitted light intensity. Atomic beams have very low density ($\sim 10^{13}\text{cm}^{-3}$) and are usually only a few millimetres wide, so the percentage absorption is negligible. Furthermore a single atom can absorb and radiate many times as it passes through the laser beam and photons can be detected with good efficiency.

10.4.2 Saturation Spectroscopy

When all the atoms in a gas are in the ground state the gas absorbs strongly light at a transition frequency. If some atoms are excited to the upper state the absorption will be reduced.

Atoms decay from an excited state at a rate A_{21}. (A_{21} is the Einstein coefficient for spontaneous emission of a photon in decay from state 2 to state 1.) Stimulated absorption occurs at a rate that depends on the flux of photons or the energy density of the resonant radiation $\rho(\nu)$. The absorption rate is then $B_{12}\rho(\nu)$ where B_{12} is the Einstein coefficient for stimulated absorption. When the

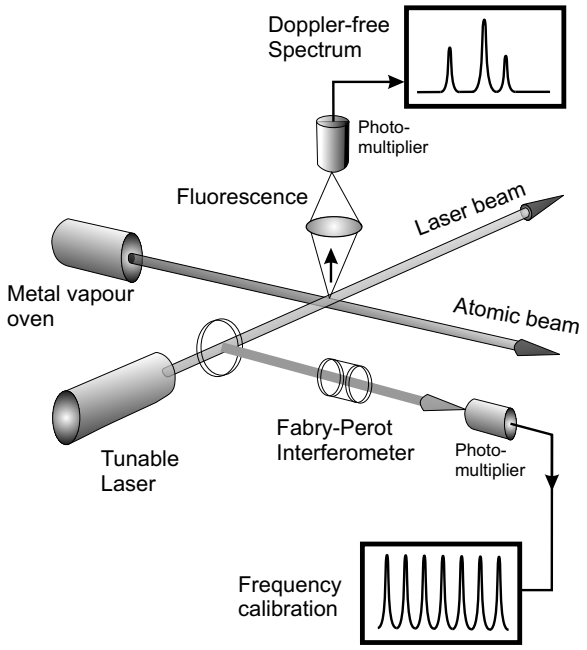


Figure 40: Atomic beam spectroscopy. The light from a tunable laser crosses the atomic beam at right angles and so eliminates the first order Doppler effect. When the frequency of the light is resonant with an atomic transition the resulting absorption is detected by the fluorescence emitted by the atoms.

absorption and decay rates are equal the atoms are said to be "saturated". The absorption coefficient α therefore depends on the intensity of the light, I:

$$\alpha = \frac{\alpha_0}{1 + I/I_{\rm sat}} \tag{220}$$

α_0 is the unsaturated absorption coefficient and $I_{\rm sat}$ is the intensity at which the stimulated absorption rate equals the spontaneous decay rate; $\mathrm{B}_{12}\rho(\nu) = \mathrm{A}_{21}$.

Saturation spectroscopy uses an essentially monochromatic laser beam split into a weak "probe" and a strong "pump" beam. The pump and probe beams are arranged to propagate in opposite directions (usually at a small angle) through the absorbing atoms in a cell or oven. The intensity of the probe beam is monitored as the laser frequency is tuned across the absorption line at frequency ν_0.

When the pump beam is blocked the spectrum of the probe beam will yield a Doppler broadened, Gaussian, profile. At each frequency ν the detuning from resonance is $\Delta\nu$, and the probe is absorbed by the sub-set of atoms that are moving to give a Doppler shift equal to the detuning: $\Delta\nu = \pm\nu(v/c)$

If $\Delta\nu$ is positive i.e. $(\nu - \nu_0) > 0$ then only atoms moving towards the incoming light are shifted into resonance and absorb the light. (Similarly if ν is tuned to a lower frequency than ν_0 only atoms moving away from the light come into resonance.)

When the strong pump beam is allowed into the cell and the laser has detuning $\Delta\nu = +\nu(v/c)$ the probe is resonant with atoms moving towards it. The pump however is resonant with atoms moving away from it and so interacts with a different sub-set (velocity class) of atoms from those absorbing the probe.

Only when $\Delta\nu = 0$ i.e. $\nu = \nu_0$ do pump and probe beams interact with the same atoms; those with essentially zero velocity along the laser beams. Now the pump intensity $I_{\rm pump} \sim I_{\rm sat}$ so the absorption is reduced by saturation. The transmitted intensity of the probe is then increased at exact resonance and a narrow "spike" is observed on the probe spectrum. This is the "Doppler-free"

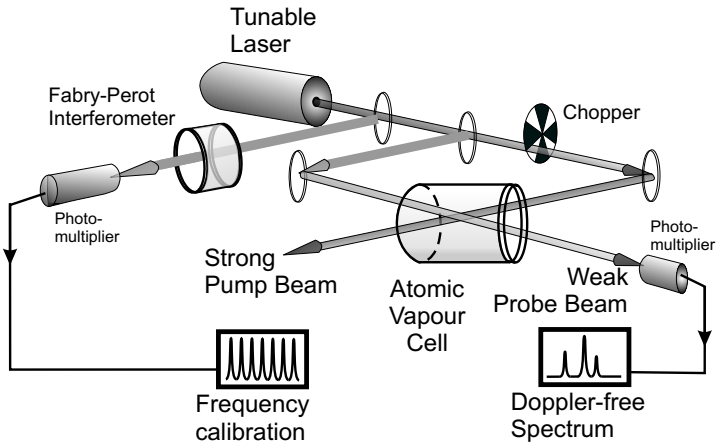


Figure 41: Saturated absorption spectroscopy. A beam from a tunable laser is split into a strong (pump) beam and a weak (probe) beam which intersect at a small angle in the absorbing gas. When the frequency of the light is resonant with an atomic transition the absorption of the weak probe beam is reduced owing to saturation by the pump. The chopper modulates the pump beam so that the saturation effect is detected above the noise on the detector of the probe beam.

spectrum. Usually the pump beam is "chopped" on-and-off and the difference signal is recorded to remove the Doppler broadened "background" to the probe signal.

10.4.3 Two-photon-spectroscopy

It is possible for an atom to absorb two photons simultaneously provided their combined energies correspond to the energy gap between the ground and some excited state E_{12} (Note that this requires two photons to arrive within the volume of an atom within the lifetime of the excited state. The photon-flux, or intensity, required can be provided only by lasers.) The energy required is: $E_{12} = 2h\nu$

The laser beam can be arranged to give two beams in opposite directions through the atoms. An atom moving with velocity $+v$ relative to one beam is Doppler shifted by $+(v/c)\nu$ but by $-(v/c)\nu$ relative to the other beam. So if the atom absorbs one photon from each beam the energy absorbed is

$$E = h\nu(1 + v/c) + h\nu(1 - v/c) = 2h\nu \tag{221}$$

i.e. independent of the velocity, and the Doppler effect is eliminated.

The two-photon absorption is detected usually by detecting photons emitted as the excited atom decays to a lower excited level. The selection rules for a two-photon transition correspond to two, sequential, single photon transitions via a "virtual" state. Thus we apply the usual dipole selection rules twice to find the allowed transitions. So two-photon transitions link states of the same (not opposite) parity. $\Delta L = 0$ or ± 2. The atom will usually decay to the ground state by two, single-photon, transitions, either or both of which may be used to detect the original two-photon absorption.

Two photons may, of course, be absorbed from just one of the beams and such absorption will be Doppler broadened, The area under this Gaussian curve will be roughly equal to that under the narrow, Doppler-free, peak so its level is very low relative to the Doppler-free peak. The broad, low, Doppler broadened background may be eliminated by ensuring absorption of one photon from each, oppositely-going, beam. This can be done using the selection rules for $\Delta M_J = 0, \pm 2$ and circularly polarized beams. Absorption from a single beam of two photons with the same circular polarization leads to $\Delta M_J = 2$ (to conserve angular momentum.) Thus a $\Delta M_J = 0$ transition needs one photon from each beam corresponding to $(\sigma^+ + \sigma^-)$. This restricts the pairs of levels that can be used to

those where $J = 0$ or $1/2$ in both level e.g. $3s^2S_{1/2} \rightarrow 4s^2S_{1/2}$ has no $\Delta M_J = \pm 2$ transitions and only $\Delta M_J = 0$ is allowed.

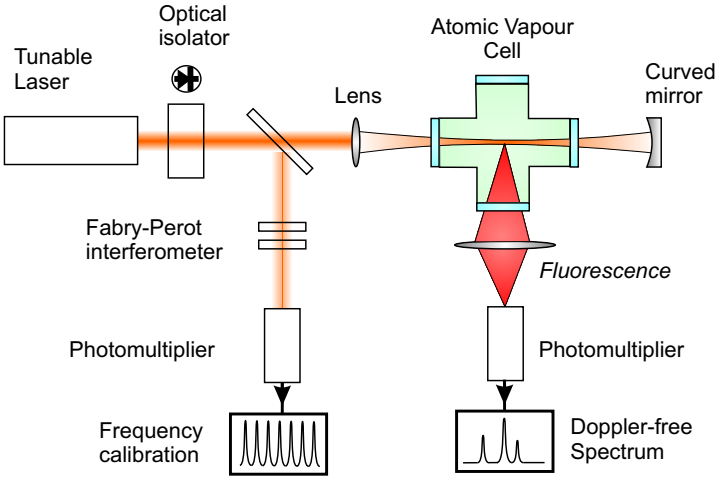


Figure 42: Two-photon spectroscopy. A tunable laser is tuned to the frequency corresponding to half the energy required for a transition. The transition is detected following absorption of two photons by the subsequently emitted fluorescence. The optical isolator (diode) prevents light feeding back into the laser that would cause instability in the laser frequency.

Two experimental details are worth noting. First, it is usually necessary to use an "optical isolator" to prevent laser light being sent back with the laser itself. If a retro-reflecting mirror is used to generate two oppositely going beams then the feedback will upset the laser operation causing instability in the laser frequency.

(An optical isolator is a device that passes light in only one direction. They usually use a polarizer and a Faraday rotator – a piece of special glass, inside a strong magnet, which rotates the plane of polarization.)

Secondly, the lasers usually need to be focussed to a small spot to generate the required intensity.

10.5 Calibration of Doppler-free Spectra

High resolution, Doppler-free, laser spectroscopy is usually concerned with measuring small frequency differences. The change in the laser frequency, or wavelength, therefore needs to be monitored during the scan. This is usually done using a stabilized Fabry-Perot interferometer. Part of the laser beam is split off and passed though this device. Transmission peaks occur when the frequency matches the condition $\nu = mc/(2d)$, where m is an integer and d is the separation of the F.P. plates. These peaks provide frequency markers every free-spectral range $c/(2d)$ and are recorded simultaneously with the Doppler-free spectrum obtained using any of the above methods.

Calibrated wavelength measuring devices based on Michelson interferometers, Fizeau interferometers or Fabry-Perots can also provide absolute wavelength values if needed. Alternative means of absolute wavelength measurement use simultaneously recorded absorption spectra of Iodine or Tellurium, whose wavelengths are know to high accuracy and give many lines across a wide spectrum.

10.6 Comparison of "Doppler-free" Methods

We have described three methods of "Doppler-free" laser spectroscopy. There are several other methods in addition to these three. How do we select an appropriate method? We need to consider advantages and disadvantages of each method, so we need to consider factors such as the following.

(1) Physical properties of atoms If an element has a low vapour pressure it may be easier to form an atomic beam, than to generate a vapour of sufficient density for saturated or two-photon absorption.

(2) Availability of laser source Tunable lasers are available in most of the visible and near infra-red. Various methods can be used to generate UV or mid-infra-red light but the power available may not be sufficient for two-photon or saturation methods.

(3) Experimental complications Saturation spectroscopy creates complex spectra owing to the possibility of "cross-over" resonances. These are extra signals generated when the laser is tuned exactly half-way between two adjacent absorption lines. The atoms with the correct velocity to be in resonance with the pump on one transition are also in resonance with the probe on the other transition.

(4) Advantages of two-photon methods Two-photon allows transitions to states not normally accessible to normal absorption from the ground state. This is because such transitions are forbidden by single photon selection rules. Thus, also, very highly excited states can be reached near the ionization limit. Allowed transitions of similar energies would need UV lasers which are more difficult to make.