Second Year Atomic Physics 2019/20
Part I - Hydrogen

Richard Thompson

December 20, 2019

1   Introduction

In this lecture we will explain the background to this course, how it will be taught, and discuss the importance of atomic physics in the development of our understanding of the world around us.

1.1   Why is atomic physics important?

Everything around us is made of atoms! If we want to understand how the world works we need to understand how atoms work: what they are made of; what holds them together; what states they can exist in; and how they interact with light. These are the topics covered in this course. In the third year "*Light and Matter*" course you will learn more about these topics and also find out about how atoms interact with each other to create molecules.

Historically, atomic physics has led to many advances in our understanding of fundamental aspects of physics, such as

- The interpretation of optical spectra
- The development of quantum mechanics
- The origin and properties of spin
- Quantum electrodynamics (QED)
- The study of quantum mechanical entanglement
- Tests of subtle aspects of quantum optics
- Creation and study of Bose-Einstein condensates (BEC)

Why has atomic physics had this important role? Partly because atomic spectra are easily accessible with simple sources (flames, discharges, lamps, the Sun, etc.), optical components (lenses, prisms etc) and detectors (photographic film and of course the eye). More recently, powerful advanced techniques making use of lasers, nonlinear optical devices, photomutiplier tubes, CCD cameras, etc have allowed us to study atomic systems in ever more detail. Insights gained by the study of simple atoms like hydrogen can then be applied to more complex atoms to build our understanding of the world around us.

Atomic physics also has a large number of important applications in science and technology, for example

- Analysis of spectra to determine the chemical composition of samples
- Lasers based on atomic systems
- Atomic clocks
- Accelerometers, gyroscopes and gravimeters

- Quantum cryptography and quantum information processing
- Tests of fundamental theories such as Lorentz invariance, variations of fundamental constants, matter/antimatter asymmetry and even the Standard Model of Particle Physics.

1.2 Motivation for this course

For the above reasons, atomic physics is a core component of all physics degrees. It also illustrates important aspects of how physics works as an approach to understanding the world, because the atom is one of the first contexts in which you will see extensive use of approximation methods to obtain useful results. Atoms are sufficiently complicated that it is not possible to treat them as idealised systems for which exact analytical solutions exist, even for a two-electron system like helium. Very early on in the course we will have to start using quantum mechanical perturbation theory, because otherwise we will not be able to calculate wave functions, energy levels, etc. In this way it is an introduction to the application of physics to "real" systems where the standard equations cannot be solved exactly. We shall see that the *intelligent* application of perturbation theory allows us to understand what is going on and to obtain accurate results, even when we have had to make some approximations. However, it is absolutely vital to do this carefully, taking account of the relative size of different effects, or else we will obtain completely incorrect results.

1.3 Structure of the course

Here is a rough outline of the course:

Lecture 1 Motivation, description of course, importance and nature of atomic physics, spectroscopy, hierarchy of perturbations

Lecture 2 Historical background, Bohr model, successes and limitations, orders of magnitude, atomic units, introduction to the hydrogen spectrum

Lecture 3 Energy levels of the hydrogen atom: Schrödinger equation, separation of variables, angular properties and quantum numbers, effective potential, solutions to the radial Schrödinger equation

Lecture 4 Fine structure of the hydrogen atom: quantum numbers, hydrogen-like systems, spin-orbit interaction, Zeeman effect

Lecture 5 Transitions between electronic states, Einstein A and B coefficients, lifetime of atomic states and natural width, Doppler width

Lecture 6 Radiation of light from atoms: Calculation of transition probabilities, selection rules for hydrogen, metastable states

Lecture 7 Two-electron systems: wave functions, symmetric and antisymmetric states, space and spin parts, quantum numbers

Lecture 8 Helium atom gross structure: approximate solutions, ground state energy, excited states, exchange interaction, selection rules

Lecture 9 Helium atom fine structure: Terms and levels, quantum numbers, Zeeman effect, effect on spectra

Lecture 10 Many-electron systems: shells, periodic table, noble gases and alkalies, conclusion

The main physics undergraduate course book, *University Physics (Young and Freedman, 14th edition)* does not cover a lot of this material but there are short discussions of the Bohr model (sections 39.2 and 39.3) and the quantum mechanics of the hydrogen atom (sections 41.3 to 41.5). Unfortunately, although there are lots of books that cover the whole of atomic physics at a fairly advanced level, they all go into much more detail than is needed for this course. Examples are *Atomic Physics (Foot)*, and *Elementary Atomic Structure (Woodgate)*. Hydrogen is treated at a fairly high mathematical level in *Introduction to Quantum Mechanics (Griffiths)*, which is the recommended text for the "*Quantum Mechanics*" course. The book that best matches this course is *Atomic Spectra (Softley)*. Although aimed at chemists, this is mostly at a suitable level for this course and covers most, though not all, of the material.

Please note that the books by Foot and Woodgate are both available online through the College library.

1.4 Teaching style

There is a large amount of research that shows that lectures are an inefficient way to gain knowledge. You learn much more by actively engaging with the material in private study than you do by passively listening to a lecture. Therefore we will use a different model from the usual "chalk and talk" approach for this course. I will supply you with a handout for each lecture, which will define the content of the course, along with the problem sheets. In order for you to get the most out of the course, you need to read the handouts *before* the next lecture, as I will not go through the material in detail in the lectures. In order to encourage you to do this, there will be a short online quiz on Blackboard for you to work through before each lecture, with a small number of straightforward questions based on the lecture handout. These short quizzes will replace some of the continuous assessment APS questions for this course and the marks for this quiz will therefore count towards the final mark on the module. They will include opportunities for you to flag up any topics that you are struggling with. Each quiz will only take a few minutes to complete.

The lectures will be used for a variety of activities designed to help you to understand better the material you have already read. I will use the lecture time to work through some problems, to discuss some tricky points in the notes, and to work through a small number of derivations. I will also discuss any questions that people have sent to me via the Blackboard quizzes. We will also use *Mentimeter* questions in the lectures. These mentimeter questions will give you feedback on whether you have understood the material in the notes, they will demonstrate whether you can apply your knowledge to new situations, and they will give me feedback on your progress.

The purpose of this sort of approach to the teaching is to help you to actively engage with the material and not just to passively sit in lectures making notes. This type of *active learning* has been shown to be more effective than conventional lectures, but it requires your participation. In order to get the most out of this course, you therefore need to *come to the lectures!* It is not sufficient to watch the lectures on Panopto as you will not then be actively participating. If you join a gym and watch people exercising, that doesn't help you to get fit – you need to put in the effort yourself to get any benefit. In the same way, the amount you get out of lectures depends on how much effort you put in. This course is designed to help you engage more with the material and therefore to get more out of the course.

1.5 Theoretical approach

The approach that I will take to the physics content of the course is to try to help you gain an understanding of real (i.e. not idealised) atomic systems, starting with the most significant effects (the so-called "gross structure", due to the Coulomb interaction) and then dealing with smaller-scale ("fine structure") effects using a series of successive approximations.

We will start from the non-relativistic Schrödinger equation for the system under investigation. This takes into account the Coulomb interaction of the electrons with the nucleus and with each other. We will then have to deal with relativistic effects using perturbation theory. (Alternatively we could start from the Dirac equation, which is inherently relativistic, but this is mathematically much more advanced, and it's still necessary to use approximation methods later, even in this case.) Other physical effects we will have to deal with using perturbation theory include the interaction of atoms with a magnetic field (the Zeeman effect) and the interaction with electromagnetic radiation.

I will not go through long mathematical derivations as I am more interested in giving you an intuitive and physical picture of what is going on rather than writing down lots of equations. You can find more details of the maths in any of the books I have listed (or in any other undergraduate atomic physics textbook).

1.6 Spectroscopy

Spectroscopy is defined as "the study of the interaction between matter and electromagnetic radiation". In practice, this is the way that we obtain information about atoms. Historically, it was the interaction between spectroscopy and theoretical models that led to advances in the theory of quantum mechanics and ultimately to a full understanding of the structure of atoms and their interaction with radiation.

This is still true today, as ever more sensitive and precise spectroscopic techniques are used to test finer aspects of fundamental physics such as the precise predictions of *Quantum Electrodynamics* (QED) for atomic energy levels, the manipulation of quantum information, and physics beyond the Standard Model.

Spectroscopic studies of atoms can be carried out in all regions of the electromagnetic spectrum, from microwaves to x-rays. We will concentrate mainly on optical spectroscopy because most transitions of interest for this course occur in or close to the visible region of the spectrum. Optical spectroscopy takes many forms, including

- Early techniques based on prisms and gratings, used for the study of absorption and emission spectra of pure vapours of atoms or molecules

- High resolution classical optics techniques for analysis of complex spectra like that of the Sun to identify the composition of the source and measure its properties (such as its temperature and pressure)

- Laser spectroscopy for very precise measurements of transitions using various techniques to eliminate the Doppler effect, such as laser cooling, two-photon spectroscopy and saturation spectroscopy.

1.7 Summary

- It is important to understand the structure and interactions of atoms in order to develop a complete theory of matter

- The study of atomic spectra was crucial in the development of the theory of quantum mechanics

- There is still a strong interplay between new developments in fundamental theory and experimental tests on atomic systems

- Atomic physics provides a good example of how the physics of real and complex systems can be studied using perturbation theory and other approximation techniques

- Atomic physics has many applications across science and technology

2 The Bohr model

In this lecture we will introduce the semi-classical Bohr model of the atom and show how it successfully explains some, but not all, of the properties of the hydrogen atom.

2.1 Historical background

In the late 1800s, spectroscopic measurements had shown that each chemical element emits a unique spectrum of distinct visible wavelengths of light. In particular, scientists had identified the spectrum of hydrogen but it appeared to consist of a number of apparently random specific wavelengths without any obvious pattern. The type of spectrometer used consisted of a narrow slit which was illuminated with light from a hydrogen discharge lamp. The light coming through the slit was dispersed with a prism (for example) and therefore spread out to form a spectrum. The specific wavelengths emitted by the source were called "lines" because they are images of the narrow slit and therefore appear as bright lines on the spectrum, which could be observed with the eye (or photographed).

Lines found in the visible region of the spectrum were called the "Balmer series" and included lines at 656 nm, 486 nm, 434 nm,.... and many more, getting closer together and converging to a limit in the ultraviolet (UV) at 364 nm. There is also a series in the deep UV called the "Lyman series" (from 121 nm to 91 nm) and one in the infrared (IR) (the "Paschen series") between 1875 nm and 820 nm. Other series have wavelengths further into the IR. All these series have the same pattern of lines getting closer together and converging to a limit at the shortest wavelength.

In 1890 Rydberg managed to find a pattern to the wavelengths, which was an enormous breakthrough. He found that they all fitted the simple formula

$$\frac{1}{\lambda} = R_\mathrm{H} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \tag{2.1}$$

where n_1 and n_2 are both positive integers and n_1 has the same value for each member of a given series: $n_1 = 1$ for the Lyman series, 2 for the Balmer series and 3 for the Paschen series. n_2 starts at $n_1 + 1$ for the longest wavelength of a series and approaches ∞ at the series limit. The quantity R_H is called the *Rydberg constant*, and for hydrogen it has the approximate value $1.0968 \times 10^7\,\mathrm{m}^{-1}$ (109 680 cm^{-1}).

The purpose of the Bohr model was to try to understand the Rydberg formula in terms of the structure of the hydrogen atom.

2.2 Energy levels in the Bohr model

Hydrogen is the most important atom to understand for many reasons: it is the most abundant atom in the Universe, it is the simplest atomic system, and it has been (and remains) a testing ground for many generations of theories of atomic structure.

The Bohr model (1913) treats hydrogen as a classical system of an electron orbiting a proton in a circular orbit under the influence of the Coulomb interaction between the two. We know now that this is not a physically complete description because this is a system where quantum mechanical effects will be important. However, the reason we will still discuss the Bohr model is because it reproduces the hydrogen spectrum pretty well, it gives us insight into some aspects of atomic structure and it tells us the order of magnitude of several important atomic properties. When we have established what is explained well by the Bohr model and what is not explained, we can then move on to a more appropriate and theoretically rigorous treatment using quantum mechanics.

For the moment we make the approximation that the nucleus has infinite mass and therefore does not move. The electron (of mass m and charge $-e$) therefore experiences an attractive

potential given by

$$V(r) = -e^2/4\pi\epsilon_0 r \tag{2.2}$$

where r is the distance from the origin where the nucleus (of charge $+e$) is located.

In order for the electron to follow a stable circular orbit, the Coulomb force must be balanced by the centrifugal force, i.e.

$$e^2/4\pi\epsilon_0 r^2 = mr\omega^2 \tag{2.3}$$

where ω is the angular frequency of the orbit (equivalently, the velocity v is given by $v = \omega r$). This gives us the following relationship between r and ω:

$$r^3 = e^2/4\pi\epsilon_0 m\omega^2. \tag{2.4}$$

Once we know r, we can calculate the energy of the system, which is given by the sum of the kinetic energy $\frac{1}{2}mv^2 = \frac{1}{2}m\omega^2 r^2$ and the Coulomb energy $-e^2/4\pi\epsilon_0 r$. The total energy is always negative because the system is bound.

Bohr realised that these equations imply that r can take any value and therefore so can the energy of the atom. However, the hydrogen spectrum consists of a set of spectral lines with specific wavelengths, and a photon of a given wavelength (λ) corresponds to a fixed amount of energy (hc/λ). Therefore, this indicates that the atom must have a discrete set of states with particular values of energy. Bohr therefore added a requirement that the angular momentum of the electron (equal to mvr) can only take particular values given by integer multiples of $\hbar$, (i.e. Planck's constant h divided by 2π). This ad hoc requirement has no physical basis at this point but we will see that it reproduces the observed hydrogen wavelengths.

If we therefore write the angular momentum as $mvr = mr^2\omega = n\hbar$ we find that the radius of the orbit for the nth state is given by

$$r_n = \frac{4\pi\epsilon_0 n^2\hbar^2}{me^2} = n^2 a_0 \tag{2.5}$$

where $a_0 = 4\pi\epsilon_0\hbar^2/me^2 \sim 5 \times 10^{-11}$ m and is called the *Bohr radius*.

An alternative way of stating this basic assumption of the Bohr model is to require that the circumference of the electron's orbit is equal to an integer multiple of the de Broglie wavelength of the electron ($\lambda_{\text{dB}} = h/mv$). This gives identical results to requiring the angular momentum to be an integer multiple of $\hbar$.

The energy of the nth state is then found to be

$$E_n = -\frac{me^4}{2(4\pi\epsilon_0)^2 n^2\hbar^2} = -\frac{1}{2} \cdot \frac{e^2}{4\pi\epsilon_0 a_0 n^2}. \tag{2.6}$$

Anticipating the result of this calculation, we can define the quantity R_∞, given by

$$R_\infty = \frac{me^4}{2(4\pi\epsilon_0)^2\hbar^2} \frac{1}{hc} \tag{2.7}$$

and then we can write the energies as $E_n = -R_\infty hc/n^2$. The energy difference between states n_1 and n_2 is then given by

$$\Delta E = E_{n_2} - E_{n_1} = R_\infty hc \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right). \tag{2.8}$$

If we now make the connection between the wavelength of a photon λ and its energy ΔE, i.e. $\Delta E = hc/\lambda$, we find an expression for the wavelength of a photon emitted in a transition between states n_1 and n_2:

$$\frac{1}{\lambda} = R_\infty \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right), \tag{2.9}$$

which has exactly the same form as Rydberg's formula above. The calculated value of R_∞ does not quite match the measured value of the Rydberg constant, R_H, but that is accounted for by taking into account the motion of the nucleus: In the Rydberg formula we need to use the reduced mass $\mu = mM/(m + M)$ (where M is the mass of the proton) instead of the electron mass m. Once this correction is made, the agreement with the measured value of R_H is excellent.

2.3   Limitations of the Bohr model

The Bohr model is a very impressive and simple model that gives surprisingly good results, especially considering that it uses classical physics in a context where it is clearly not appropriate. Its major success is that it explains the wavelengths of all the spectral lines in the hydrogen atom as transitions between a number of specific energy levels, with a clear physical interpretation in terms of stable classical electron orbits. However, it leaves a number of questions unanswered:

- Why is the electron only allowed to be in particular orbits?

- Why is the angular momentum of the electron required to be an integer multiple of $\hbar$?

- Classically, an accelerating charged particle always radiates electromagnetic radiation. The electron is constantly accelerating in its orbit, so why does it not emit radiation and spiral into the nucleus?

- Why does the model not work for any other atoms apart from hydrogen?

- The predicted value of the angular momentum of the electron turns out to have the wrong value – why is this?

2.4   Orders of magnitude

One useful feature of the Bohr model is that it introduces us to the orders of magnitude of various physical quantities in atomic physics:

- The quantity a_0 (the *Bohr radius*), which is equal to roughly $0.053\,\mathrm{nm}$, gives us the scale of the size of atoms. The radius of hydrogen in its ground state ($n = 1$) is equal to a_0. Atoms with more electrons turn out to have a similar size because some of the electrons are in states with $n > 1$, leading to larger radii, but in larger atoms the Coulomb force is greater because the nucleus has more charge, so this brings the size back down again. These effects roughly cancel each other out, leading to a similar size for all neutral atoms.

- The quantity R_∞ (the *Rydberg constant*), which is equivalent to an energy of roughly $13.6\,\mathrm{eV}$, sets the scale of the energy of atomic states, and although some electrons in larger atoms are very tightly bound, the electrons which interact with radiation (the *valence electrons*) tend to have binding energies of this order of magnitude. This means in turn that many atoms have spectra in the optical region of the spectrum, because the energy of a visible photon (400 to 700 nm) is in the range 3.1 to $1.8\,\mathrm{eV}$.

2.5   A note on units

There is a wide range of units in use in books on atomic physics. **You need to be familiar with them all.** Wavelengths are usually expressed in nm or μm but in old books the Ångstrom unit is used ($1\,\text{Å} = 0.1\,\mathrm{nm}$). Energy differences may be expressed in eV ($1\,\mathrm{eV} = 1.6{\times}10^{-19}\,\mathrm{J}$) or in wavenumbers ($\mathrm{cm}^{-1}$). The wavenumber of a transition is defined as $1/$(wavelength in cm), so $1\,\mathrm{cm}^{-1}$ is equivalent to a frequency interval of $30\,\mathrm{GHz}$ or an energy of $1.98{\times}10^{-23}\,\mathrm{J}$. Wavenumbers were used in the early days of spectroscopy for specifying transitions, and this unit has continued to be used in some situations, but I will not use it much. I will usually

use MHz or GHz for small energy differences: $1\,\mathrm{MHz}$ is equivalent to $6.6{\times}10^{-28}$ J and $1\,\mathrm{GHz}$ is equivalent to $6.6{\times}10^{-25}$ J. For quick reference, the relations between these units are:

$$1\,\mathrm{cm}^{-1} = 100\,\mathrm{m}^{-1} \text{ is equivalent to } 30\,\mathrm{GHz} \qquad 1\,\mathrm{eV} \text{ is equivalent to } 2.4{\times}10^{14}\,\mathrm{Hz}\ (8080\,\mathrm{cm}^{-1})$$

Throughout this course I intend to use SI units where possible, but in some places I will use *atomic units* (a.u.). Atomic units are also used in some books. In this set of units we set $\hbar = e = m_e = 4\pi\epsilon_0 = 1$. The atomic unit of energy then comes out to be equal to 27.2 eV (twice the binding energy of the ground state of hydrogen) and the atomic unit of length is the Bohr radius $a_0 = 0.053\,\mathrm{nm}$. The velocity of light turns out to be equal to $1/\alpha = 137.04$ atomic units. Here α is called the fine structure constant and turns up in several contexts in atomic physics. It is a number (with no units) and is given by $\alpha = e^2/4\pi\epsilon_0\hbar c = 1/137.04$.

Probably the most basic unit in atomic physics is the Rydberg constant, R_∞, which is equivalent to the binding energy of the hydrogen atom in its ground state (assuming infinite nuclear mass). It has the value $1.0974{\times}10^7\,\mathrm{m}^{-1}$ or $109\ 740\,\mathrm{cm}^{-1}$, which can also be expressed as 13.6 eV or $3.3{\times}10^{15}$ Hz. The relations below can be used to relate all the units used for energy in atomic physics to each other.

To convert m^{-1} to J, multiply by hc (and for cm^{-1}, multiply by $100hc$).
To convert eV to J, multiply by e.
To convert Hz to J, multiply by h.
To convert Rydbergs to J, multiply by $R_\infty hc = 2.2 \times 10^{-18}$.
To convert Rydbergs to eV, multiply by $R_\infty hc/e = 13.6$.
To convert a.u. to eV, multiply by 27.2.

2.6 Summary

- The hydrogen spectrum consists of a number of series of spectral lines, with each series converging to a limit.

- The Bohr model is a crude semi-classical model for the energy levels of the hydrogen atom and its spectrum. It relies on apparently arbitrary assumptions to make it work.

- However, the Bohr model correctly predicts all the observed wavelengths of the hydrogen spectrum

- The Bohr model also gives other important insights into the structure of hydrogen and correctly predicts a number of atomic properties such as the radius of the hydrogen atom, which is of order 0.05 nm, and the binding energy of the ground state of hydrogen, which is equal to 13.6 eV.

3 The hydrogen atom – gross structure

In this lecture we will solve the Schrödinger equation for hydrogen by splitting it into an angular part (which has a general set of solutions, the *Spherical Harmonics*), and a radial part, which has solutions that are specific to the Coulomb potential.

3.1 The Schrödinger equation for hydrogen

Our starting point is the time-independent Schrödinger equation for an electron in the Coulomb field of an infinitely massive nucleus:

$$\mathcal{H}\psi(\mathbf{r}) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(r) \right]\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{3.1}$$

Here $V(r) = -e^2/4\pi\epsilon_0 r$, where $-e$ is the charge on the electron and the nucleus has charge $+e$.

If we rewrite the Laplacian operator in spherical polar coordinates, we arrive at a separable form of the Schrödinger equation (see the next section):

$$\left[-\frac{\hbar^2}{2m}\nabla_R^2 + \frac{1}{2mr^2}\hat{\boldsymbol{\ell}}^2 + V(r) \right] \psi(r,\theta,\phi) = E\psi(r,\theta,\phi). \tag{3.2}$$

In this equation ∇_R^2 is the radial part of ∇^2 and $\hat{\boldsymbol{\ell}}^2$ is the operator for angular momentum squared. The first and third terms in the square brackets depend only on r, and $\hat{\boldsymbol{\ell}}^2$ in the middle term depends only on θ and ϕ, so the equation separates out into an angular part and a radial part, and the wave function can be written as a product of an angular part $\Phi(\theta,\phi)$ and a radial part $R(r)$.

3.2 Derivation of the separable form of the Schrödinger equation (*non-examinable*)

The Laplacian operator in 3D is given by

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \right] \tag{3.3}$$

but since the angular momentum squared operator $\hat{\boldsymbol{\ell}}^2$ is defined as

$$\hat{\boldsymbol{\ell}}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \right], \tag{3.4}$$

we can re-express ∇^2 as

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\left(-\frac{1}{\hbar^2}\hat{\boldsymbol{\ell}}^2\right) = \nabla_R^2 - \frac{1}{r^2\hbar^2}\hat{\boldsymbol{\ell}}^2 \tag{3.5}$$

where ∇_R^2 is the radial part of ∇^2. This allows us to rewrite the Schrödinger equation as

$$\left[-\frac{\hbar^2}{2m}\nabla_R^2 + \frac{1}{2mr^2}\hat{\boldsymbol{\ell}}^2 + V(r) \right] \psi(r,\theta,\phi) = E\psi(r,\theta,\phi). \tag{3.6}$$

(You have also seen this derivation in more detail in the "*Quantum Mechanics*" lecture 23).

3.3 Angular part of the wave function

We now look at the form of the solutions to the Schrödinger equation, remembering that it can be written as a product of an angular part and a radial part, i.e.

$$\psi(r,\theta,\phi) = \Phi(\theta,\phi)R(r). \tag{3.7}$$

The angular part of the wave function must satisfy the eigenvalue equation

$$\hat{\boldsymbol{\ell}}^2\Phi(\theta,\phi) = K\Phi(\theta,\phi) \tag{3.8}$$

where K is some constant. But since $\hat{\boldsymbol{\ell}}^2$ is the operator for the square of the angular momentum, we can write down the eigenvalue immediately. We find that the eigenvalue is $K = l(l+1)\hbar^2$ (see "*Quantum Mechanics*" lecture 22) and the wave function $\Phi(\theta,\phi)$ is the spherical harmonic $Y_{lm_l}(\theta,\phi)$.

The spherical harmonics are labelled by two quantum numbers: l is the quantum number corresponding to the magnitude of the angular momentum, and m_l is the quantum number

corresponding to the z-component of the angular momentum (sometimes called the *magnetic quantum number*). The eigenvalue of $\hat{\boldsymbol{\ell}}^2$ is $l(l+1)\hbar^2$ and the eigenvalue of $\hat{\boldsymbol{\ell}}_z$ is $m_l\hbar$.

There are restrictions on the values of the quantum numbers: l takes integer values 0, 1, 2, ... and for a given value of l, there are $2l+1$ possible values of m_l, given by the integers $-l \le m_l \le l$.

The restrictions to integer values of the quantum numbers l and m_l can be shown to be a consequence of the fact that there are boundary conditions that have to be satisfied. This can be seen most easily for the ϕ-dependence of the spherical harmonic, which is $e^{im_l\phi}$. The wave function must have the same value if ϕ is replaced by $\phi + 2\pi$, so $e^{im_l 2\pi}$ must equal 1 and this means that m_l must be an integer, since $e^{2\pi i} = 1$.

Note that we have not yet considered the detailed form of the potential $V(r)$. All we have done is to assume that it is a *central* potential, i.e. that it depends on r only and has no angular dependence. Once we make that assumption, we know what the angular form of the wave function will be, and therefore what the angular momentum properties of the solution will be. This is a very powerful statement as it allows us to make many deductions about the angular momentum properties of even complex atoms, so long as we can assume that the potential seen by each electron is central (see Section 7.5).

The expressions for the spherical harmonics involve the associated Legendre polynomials, and they get complicated for all but the smallest values of l and m_l. Here are the expressions for the first few spherical harmonics (*Note: you do not need to learn these expressions*):

$$Y_{00} = \sqrt{\frac{1}{4\pi}} \tag{3.9}$$

$$Y_{10} = \sqrt{\frac{3}{4\pi}} \cos\theta \qquad Y_{1\pm1} = \mp\sqrt{\frac{3}{8\pi}} \sin\theta e^{\pm i\phi} \tag{3.10}$$

$$Y_{20} = \sqrt{\frac{5}{16\pi}}(3\cos^2\theta - 1) \qquad Y_{2\pm1} = \mp\sqrt{\frac{15}{8\pi}} \sin\theta\cos\theta e^{\pm i\phi} \qquad Y_{2\pm2} = \sqrt{\frac{15}{32\pi}} \sin^2\theta e^{\pm 2i\phi} \tag{3.11}$$

Note that all the spherical harmonics are normalised such that

$$\int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi \, |Y_{lm_l}(\theta,\phi)|^2 = 1 \tag{3.12}$$

The simplest case is when $l = m_l = 0$, for which Y_{00} is spherically symmetrical – this state has no angular momentum. Note that this is in contrast to the Bohr model, where the ground state of hydrogen has angular momentum of $\hbar$.

For $l = 1$, there are three possible values of m_l. If we plot $|Y_{10}|^2$ we find that it looks like a dumbbell oriented along the z-axis whereas $|Y_{1\pm1}|^2$ looks like a ring around the z-axis (see demonstrations.wolfram.com/SphericalHarmonics/). It's therefore reasonable that Y_{10} corresponds to a state that has no z-component of angular momentum, because the wave function is largest close to the z-axis. On the other hand, $Y_{1\pm1}$ does have a z-component of angular momentum, because it corresponds to the electron moving in a radial "orbit" in the xy-plane in either a clockwise or anticlockwise direction. The value of $Y_{1\pm1}$ on the z-axis (i.e. at $\theta = 0$) is zero.

Appropriate superpositions of Y_{11} and Y_{1-1} can be used to make wave functions which point along the x- and y-axes like Y_{10} points along the z-axis (note that these superpositions are no longer eigenstates of $\hat{\boldsymbol{\ell}}_z$ though). These types of combinations of spherical harmonics are the basis of the molecular orbitals that are used in chemistry to help in understanding chemical bonds and chemical reactions.

One final point about the spherical harmonics: they all have a definite *parity*. Just as a function $f(x)$ in one dimension is said to have even parity if $f(-x) = f(x)$ and odd parity if $f(-x) = -f(x)$, we can define parity for a function in three dimensions in a similar way. We
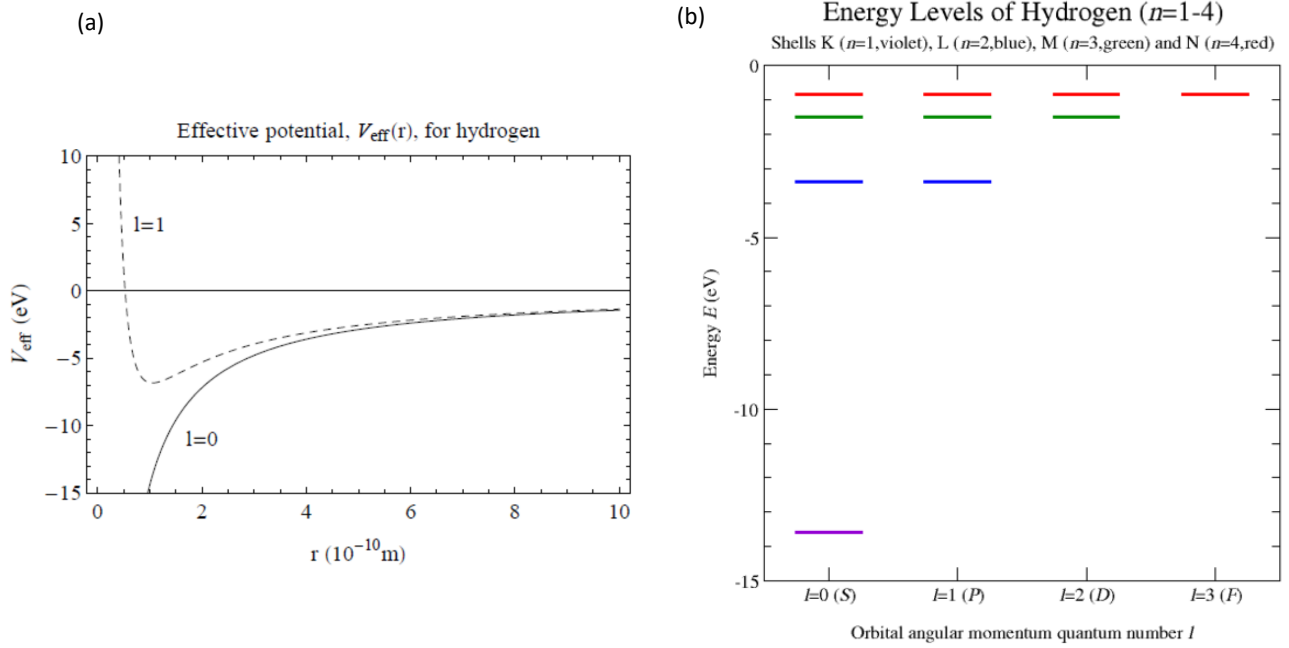
Figure 1: (a) Effective potential and (b) Energy levels of the hydrogen atom

say that a function $f(\mathbf{r})$ has even parity if $f(-\mathbf{r}) = f(\mathbf{r})$ and odd parity if $f(-\mathbf{r}) = -f(\mathbf{r})$. (The substitution of $-\mathbf{r}$ for $\mathbf{r}$ is equivalent to the substitutions $r \to r$; $\theta \to \pi - \theta$; $\phi \to \phi + \pi$ in spherical polar coordinates.) For the spherical harmonics we find that the parity is defined just by the value of l: if l is even then Y_{lm_l} has even parity and if l is odd then Y_{lm_l} has odd parity. This turns out to have enormous significance when we consider transitions between atomic states as it restricts the possible change in the value of l in a transition to $\Delta l = \pm 1$ (this is called a *selection rule* – see Section 6.3).

3.4   Effective radial potential

If we write the wave function as $\psi(r, \theta, \phi) = R(r)Y_{lm_l}(\theta, \phi)$ and substitute it into Equation 3.6, we obtain

$$\left[-\frac{\hbar^2}{2m}\nabla_R^2 + \frac{l(l+1)\hbar^2}{2mr^2} + V(r) \right] R(r) = ER(r). \tag{3.13}$$

where we have inserted the eigenvalue of the angular momentum squared operator $\hat{\boldsymbol{\ell}}^2$. This looks like a one-dimensional Schrödinger equation with a modified potential $V'(r) = V(r) + l(l+1)\hbar^2/2mr^2$ (see Figure 1(a)). This *effective potential* is of exactly the same form as would be obtained in a classical calculation of a particle moving in an orbit in a central potential, when the angular momentum squared has the value $l(l+1)\hbar^2$. The correspondence with the equivalent classical system is very close.

The Coulomb potential in hydrogen is $V(r) = -e^2/4\pi\epsilon_0 r$ but it is best at this stage to allow for a different charge on the nucleus. Any system consisting of a single electron orbiting a nucleus of charge Ze is referred to as a *hydrogen-like system*. Examples include He$^+$ (where $Z = 2$) right up to *highly-charged ions* such as U^{91+} (where $Z = 92$). We therefore write the Coulomb potential more generally as $V(r) = -Ze^2/4\pi\epsilon_0 r$.

3.5   Solutions to the radial wave equation

The solutions to the radial equation can be found analytically. They are written in the form $R_{nl}(r)$, i.e. they depend on r only (as expected) and they are labelled by two quantum numbers.

11

The solutions depend on l because the effective potential $V'(r)$ includes a term that is a function of l. The new quantum number is n, which arises from the boundary conditions on r. The radial part of the wave function $R_{nl}(r)$ is normalised according to the usual formula:

$$\int_0^\infty R_{nl}^*(r)R_{nl}(r)r^2 dr = \int_0^\infty |R_{nl}(r)|^2 r^2 dr = 1 \tag{3.14}$$

where the r^2 comes in because the volume element in spherical polar coordinates is $r^2 \sin\theta dr d\theta d\phi$.

The form of $R_{nl}(r)$ is a polynomial of order $(l-1)$ in Zr/a_0, multiplied by an exponential e^{-Zr/na_0}. This means that all the wave functions decay exponentially at large values of r. The only wave functions that have a finite value at $r=0$ are the ones with $l=0$. For the Coulomb potential ($V(r) = -Ze^2/4\pi\epsilon_0 r$), the solution has the special property that the energy eigenvalue is independent of l. It is given by:

$$E_n = -\frac{1}{(4\pi\epsilon_0)^2}\frac{Z^2 e^4 m}{2\hbar^2 n^2} = -R_\infty hc\frac{Z^2}{n^2} \tag{3.15}$$

and this is just the same quantised energy as found by Bohr with his semi-classical model of the hydrogen atom with $Z=1$. Here R_∞ is the Rydberg constant, given by

$$R_\infty = \frac{me^4}{2\hbar^2(4\pi\epsilon_0)^2}\frac{1}{hc} \tag{3.16}$$

and it has the value $1.09737\times10^7\,\mathrm{m}^{-1}$. The hc is there to convert from energy units into wavenumbers as the Rydberg constant is a spectroscopic unit, to do with the measured wavelengths of spectral lines. An equivalent way to express the Rydberg constant as an energy value is $R_\infty hc \equiv 13.6\,\mathrm{eV}$.

The first three radial eigenfunctions are given by (*Note: you do not need to learn these expressions*):

$$R_{10}(r) = \left(\frac{Z}{a_0}\right)^{3/2} 2e^{-\rho/2} \tag{3.17}$$

$$R_{20}(r) = \left(\frac{Z}{a_0}\right)^{3/2}\frac{1}{2\sqrt{2}}(2-\rho)e^{-\rho/2} \qquad R_{21}(r) = \left(\frac{Z}{a_0}\right)^{3/2}\frac{1}{2\sqrt{6}}\rho e^{-\rho/2} \tag{3.18}$$

$$R_{30}(r) = \left(\frac{Z}{a_0}\right)^{3/2}\frac{1}{9\sqrt{3}}(6-6\rho+\rho^2)e^{-\rho/2} \qquad R_{31}(r) = \left(\frac{Z}{a_0}\right)^{3/2}\frac{1}{9\sqrt{6}}\rho(4-\rho)e^{-\rho/2} \tag{3.19}$$

$$R_{32}(r) = \left(\frac{Z}{a_0}\right)^{3/2}\frac{1}{9\sqrt{30}}\rho^2 e^{-\rho/2} \tag{3.20}$$

where for simplicity we have written these in terms of ρ where $\rho = 2Zr/na_0$ and a_0, the Bohr radius, which is given by $a_0 = 4\pi\epsilon_0\hbar^2/me^2 = 0.053$ nm. Note that some authors use a different definition of ρ so be careful when comparing these equations with other sources. These radial wave functions can be used to calculate the probability $P(r)$ of finding the electron at a particular radius, i.e.,

$$P_{nl}(r)dr = |R_{nl}(r)|^2 r^2 dr \tag{3.21}$$

and they are normalised, so $\int_0^\infty P_{nl}(r)dr = 1$.

The radial wave functions illustrate the general principle that the wave function is zero at the origin except for $l=0$ states; it decays exponentially at large values of r; and the range of r over which it is distributed increases roughly as n^2/Z, i.e. the 'size' of the wavefuncton is roughly $n^2 a_0/Z$. It can be shown that at the origin the electron probability density for the $l=0$ states has the value (*Note: you do not need to learn this expression*):

$$|\psi_{n00}(0)|^2 = \frac{Z^3}{\pi a_0^3 n^3}. \tag{3.22}$$

More examples of these radial functions can be found in quantum mechanics text books.

3.6 Nomenclature and degeneracy

The nomenclature for atomic states was established before the quantum mechanical description was fully understood, so it may appear a bit illogical now. However, the original nomenclature has stuck so we will introduce it now.

Atomic states are labelled by three quantum numbers: n is the *principal quantum number* while l is the *angular momentum quantum number* and m_l is the *magnetic quantum number*, because it is associated with a measurable magnetic moment of the atom, as we will see later. Note that we always use *lower case* letters to refer to the quantum numbers of a single electron.

The principal quantum number n can take any integer value greater than or equal to 1. For each value of n there are n possible integer values of l, which are 0, 1, 2..., $n-1$. For each value of l there are $2l+1$ possible integer values of m_l, i.e. $-l \le m_l \le l$. The state specified by n and l is therefore said to have a *degeneracy* of $2l+1$ because there are $2l+1$ different values that m_l can take. Each m_l value represents a different eigenstate of the system.

In the absence of any external fields, all the states in hydrogen with a given n and l have the same energy. This energy level is labelled "$n(l)$" where (l) is a letter code where "s" stands for $l=0$, "p" stands for $l=1$, "d" stands for $l=2$ and "f" stands for $l=3$.

The lowest energy state of hydrogen (the so-called *ground state*) is therefore written as 1s because it has $n=1$ and $l=0$. Its energy is $E_{1s} = -R_\infty hc$ and its degeneracy is $2l+1 = 1$, i.e. there is only one value of m_l which is $m_l = 0$. The first excited state is 2s ($n=2, l=0$) and has the energy $E_{2s} = -R_\infty hc/4$ with degeneracy $2l+1 = 1$. This is the same energy value as for the state 2p ($n=2, l=1$) which has degeneracy $2l+1 = 3$. For $n=3$ we have 3s, 3p and 3d which all have the energy $-R_\infty hc/9$ and degeneracies 1, 3 and 5 respectively.

Since there are $2l+1$ different values of m_l for each l-state, and there are n different values of l for each n-state, we can find the total degeneracy corresponding to each value of n:

$$\sum_{l=0}^{l=n-1} (2l+1) = n^2. \tag{3.23}$$

This will become important when we start to talk about atoms with more than one electron.

3.7 Summary

We have shown that the solutions to the Schrödinger equation for hydrogen (or for any hydrogen-like atom or ion) can be written in terms of simple product wave functions

$$\psi_{nlm_l}(r, \theta, \phi) = R_{nl}(r) Y_{lm_l}(\theta, \phi) \tag{3.24}$$

where n, l and m_l are all integers given by $n = 1, 2, 3...$; $\ 0 \le l < n$; and $-l \le m_l \le +l$.

These wave functions are simultaneous eigenfunctions of the following three operators for energy ($\mathcal{H}$), total squared angular momentum ($\hat{\boldsymbol{\ell}}^2$) and z-component of angular momentum ($\hat{\boldsymbol{\ell}}_z$):

$$\mathcal{H}\psi = E_n \psi = -R_\infty hc \frac{Z^2}{n^2} \psi \tag{3.25}$$

$$\hat{\boldsymbol{\ell}}^2 \psi = l(l+1)\hbar^2 \psi \tag{3.26}$$

$$\hat{\boldsymbol{\ell}}_z \psi = m_l \hbar \psi. \tag{3.27}$$

All of the results we have derived here will be used later in the course, and that is why it is very important that you understand where all these results come from. In particular, the angular properties are similar for electrons in many different atoms. The dependence of E_{nl} on n only is only true for the Coulomb potential, so in other cases we will find that the energy depends on both n and l, but even so we will see many similarities to hydrogen.

4 The hydrogen atom – Fine structure

In this lecture we will look at various effects that have to be taken into account in order to obtain an exact match between theory and experiment for the hydrogen spectrum, including the finite mass of the nucleus, the spin-orbit interaction, and the Zeeman effect.

4.1 Hydrogen-like systems

We call any system where there is a single charged particle orbiting another charged particle a *hydrogen-like* system. This is because the theoretical treatment of such a system is exactly equivalent to the treatment of hydrogen.

First, consider the question of the nuclear mass. As with the Bohr atom, to get an exact correspondence between the calculated and measured hydrogen energy levels we have to take into account the motion of the nucleus. Let us write the mass of the electron as m and the mass of the nucleus as M. By writing the Hamiltonian in terms of the kinetic energy of both particles, it is possible to show that the Schrödinger equation can be separated into two parts, one of which corresponds to the motion of the centre of mass of the system with the other part describing the relative motion of the particles. The centre of mass motion is not particularly interesting, as it just describes the free motion of a system with total mass $M+m$, but the equation for the relative motion turns out to be identical to the original Schrödinger equation except that m is replaced by μ where μ, the reduced mass of the electron, is equal to $mM/(m+M)$. This means that as M gets bigger, the reduced mass becomes closer to m and the energy levels therefore become closer to those described by R_∞.

Remember that the Rydberg constant R_∞ is proportional to m. In order to get the exact values for the measured wavelengths in hydrogen, we have to replace m in the equation for R_∞ by the electron's reduced mass in the atom, $\mu = mM/(m+M)$, where $M = m_p$ is the mass of the proton. This then gives us a modified Rydberg constant R_H which has the value $1.0968 \times 10^7 \, \text{m}^{-1}$. This is different from R_∞ by about one part in 2000. Other isotopes of hydrogen will have different modified values for the Rydberg constant because their nuclei have different masses (e.g. deuterium has $M = 2m_p$).

An extreme example is *positronium* which consists of an electron and a positron orbiting their common centre of mass. A simple calculation shows that in this case the reduced mass is equal to $m/2$ and so the energy values are all a factor of 2 smaller than those in hydrogen.

Another exotic system is a *muonic atom* where a muon (with a mass approximately 200 times that of the electron) replaces one of the electrons in an atom. This system therefore has binding energies that are in the range of keV rather than eV.

4.2 Theoretical models of atoms

The problem with real atoms is that they are not idealised systems and so if we want to get closer to the exact energy levels, even of a simple system like hydrogen, we have to move from exact solutions of the Schrödinger equation to approximation methods. This is a fundamental feature of the subject of atomic physics, however, it does not stop us making accurate theoretical predictions for the properties of atoms – for instance the best measurements of atomic spectra can be used to determine the Rydberg constant R_∞ to better than one part in 10^9 (1 ppb).

There are a number of approximation methods that can be used but the most important is *perturbation theory*. You have come across this in your *Quantum Mechanics* lectures. We will use time-independent perturbation theory to describe static perturbations in atoms, such as the spin-orbit interaction and the Zeeman effect. We will use time-dependent perturbation theory to describe the interaction of light and matter. In case you need to remind yourself about how perturbation theory works, I have also provided a separate handout about it.

The trick with perturbation theory, when there are several perturbations to consider, is to determine a hierarchy of perturbations and to apply them in order, starting with the largest one. Each perturbation establishes the basis states for the next perturbation. If any two perturbations are of the same order of magnitude, they must be applied at the same time. Often first-order perturbation theory is sufficient, but for accurate results, especially if the energy scales of different perturbations are similar, it may be necessary to use higher orders of perturbation theory.

We start with the atomic energy levels determined by the Schrödinger equation. This is called the *gross structure* of the atom. Typical static perturbations that we need to consider include (for hydrogen) the *spin-orbit interaction* (see Section 4.3), which is the interaction between the spin magnetic moment of the electron and the internal magnetic field generated by the electron's motion in the atom. This determines the *fine structure* of the atom. Then there is the *Zeeman effect* (see Section 4.5), which is the interaction between the magnetic moment of the whole atom and an external magnetic field.

(There is also a small interaction between the resultant total magnetic moment of the electron and the magnetic moment of the nucleus, which results in the *hyperfine structure* of the atom. Finally we may need to consider the *Stark effect*, which is the interaction with an external electric field. However, these topics are not part of this lecture course.)

The main result that we need to use from perturbation theory is that the energy shift ΔE arising from the application of a perturbation $\mathcal{H}'$ in a particular quantum state $\psi(\mathbf{r})$ is just equal to the *expectation value* of that perturbation in that state, i.e.,

$$\Delta E = \langle \mathcal{H}' \rangle = \int \psi^*(\mathbf{r}) \mathcal{H}' \psi(\mathbf{r}) d^3\mathbf{r}. \tag{4.1}$$

The other thing that perturbation theory tells us is that the perturbation will sometimes mix states that are degenerate (i.e., that all have the same energy) in the absence of the perturbation. This happens when these original states are not eigenfunctions of the perturbation: in other words, the quantum numbers that describe these original states are not good quantum numbers in the presence of the perturbation. This means that when the perturbation is applied, new quantum numbers will emerge that describe new states that *are* eigenfunctions of the perturbation. The total number of states cannot be changed by the perturbation, but the quantum numbers that describe the states can change.

4.3 Physical basis of the spin-orbit interaction

This is the most important perturbation for hydrogen. It was discovered by looking at the hydrogen spectrum under high resolution. For example, we already showed that the wavelength of the $n = 3$ to $n = 2$ transition in hydrogen is given by $1/\lambda = R_H(1/4 - 1/9)$, resulting in $\lambda = 656$ nm. However, under high resolution we find that this line is split into two, with a splitting of around 0.016 nm. This is not accounted for by the solutions to the Schrödinger equation.

This splitting arises from the fact that the electron has spin $\mathbf{s}$ which gives rise to a magnetic moment $\boldsymbol{\mu}_s = -(ge/2m)\mathbf{s}$ (note that this is a vector quantity). Here g is the *gyromagnetic ratio* of the electron, which relativistic quantum mechanics (the Dirac equation) predicts has the value 2 exactly.[1] We can therefore write $\boldsymbol{\mu}_s = -(e/m)\mathbf{s}$, or in scalar form $\mu_s = -e\hbar/2m = -\mu_B$ where this equation defines the *Bohr magneton* μ_B.

Now the electron is moving in the electric field of the nucleus. According to relativity, any charged particle moving with a velocity $\mathbf{v}$ in an electric field $\mathbf{E}$ experiences a magnetic field in its moving frame given by $\mathbf{B} = -\frac{1}{c^2}\mathbf{v} \times \mathbf{E}$. The electron's spin magnetic moment $\boldsymbol{\mu}_s$ interacts with

---

[1]In fact the measured value is slightly different from this value because of effects that are described by QED, but it is still very close to 2, so we will use that value from now on.

this magnetic field, resulting in an interaction energy that depends on the relative orientation of $\mathbf{B}$ and $\boldsymbol{\mu}_s$. The perturbation to the energy of the atom is given by $\mathcal{H}' = -\boldsymbol{\mu}_s \cdot \mathbf{B}$ and it can be shown that this can be written as

$$\mathcal{H}' = A_{nl}\hat{\mathbf{s}} \cdot \hat{\boldsymbol{\ell}}/\hbar^2 \tag{4.2}$$

The expectation value of this perturbation gives the *spin-obit interaction energy* of the atom.

4.4 Derivation of A_{nl}

Here we derive an expression for A_{nl} (Equation 4.9), but this derivation is not examinable:

Classically, the magnetic field seen by a particle of charge e moving with a velocity v in an electric field $\mathbf{E}$ is given by

$$\mathbf{B} = -\frac{1}{c^2}\mathbf{v} \times \mathbf{E} = -\frac{1}{ec^2}\mathbf{v} \times \nabla V(r) \tag{4.3}$$

where $V(r) = -Ze^2/4\pi\epsilon_0 r$, since the electron is moving in the electric field of the nucleus of charge Ze. Substituting for $V(r)$ and noting that $\boldsymbol{\ell} = m\mathbf{v} \times \mathbf{r}$, we find that

$$\mathbf{B} = \frac{Ze^2}{4\pi\epsilon_0 r^3 emc^2}\,\boldsymbol{\ell}. \tag{4.4}$$

This magnetic field interacts with the spin magnetic moment of the electron – this is the *spin-orbit interaction.* The energy perturbation associated with this interaction is therefore

$$\mathcal{H}' = -\boldsymbol{\mu}_s \cdot \mathbf{B} = \left(\frac{e}{m}\hat{\mathbf{s}}\right) \cdot \left(\frac{Ze^2}{4\pi\epsilon_0 r^3 emc^2}\hat{\boldsymbol{\ell}}\right) = \frac{Ze^2}{2m^2c^2(4\pi\epsilon_0)}\frac{1}{r^3}\hat{\mathbf{s}} \cdot \hat{\boldsymbol{\ell}}. \tag{4.5}$$

where we have used the definition of the angular momentum $\boldsymbol{\ell} = m\mathbf{v} \times \mathbf{r}$. Note that we have now written the angular momenta as operators $\hat{\mathbf{s}}$ and $\hat{\boldsymbol{\ell}}$. In the last step we have also taken into account an extra factor of $\frac{1}{2}$ that arises from a subtle relativistic effect called the *Thomas precession*, which we won't go into any further. In order to evaluate this expression, we need to calculate the expectation value of $1/r^3$, which for hydrogen can be obtained from the radial wave function. It is possible to derive the general result:

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{1}{l(l + \frac{1}{2})(l + 1)n^3}\left(\frac{Z}{a_0}\right)^3. \tag{4.6}$$

The interaction energy is then given by

$$\Delta E_{nl} = \langle\mathcal{H}'\rangle = -\langle\boldsymbol{\mu}_s \cdot \mathbf{B}\rangle = \frac{Ze^2}{2m^2c^2(4\pi\epsilon_0)}\left\langle \frac{1}{r^3} \right\rangle \langle\hat{\mathbf{s}} \cdot \hat{\boldsymbol{\ell}}\rangle, \tag{4.7}$$

which we write as

$$\Delta E_{nl} = A_{nl}\langle\hat{\mathbf{s}} \cdot \hat{\boldsymbol{\ell}}\rangle/\hbar^2 \tag{4.8}$$

where A_{nl} is given by

$$A_{nl} = \frac{Ze^2}{2m^2c^2(4\pi\epsilon_0)}\frac{\hbar^2}{l(l + \frac{1}{2})(l + 1)n^3}\left(\frac{Z}{a_0}\right)^3. \tag{4.9}$$

Note: You do not need to remember this result.

4.5   Calculation of the spin-orbit energy

In order to evaluate the expectation value of the vector product $\hat{\mathbf{s}} \cdot \hat{\boldsymbol{\ell}}$ we need to think some more about angular momentum, and we will start by thinking about it in a classical way. Since there is no external torque acting on the atom, any coupling between two angular momenta ($\boldsymbol{\ell}$ and $\mathbf{s}$) must result in a constant total angular momentum for the atom as a whole. Call this total angular momentum $\mathbf{j} = \boldsymbol{\ell} + \mathbf{s}$.

The classical interaction energy between and magnetic moment $\boldsymbol{\mu}_s$ and a magnetic field $\mathbf{B}$ is $-\mu_s B \cos\theta$ where θ is the angle between $\boldsymbol{\mu}_s$ and $\mathbf{B}$. The system is isolated, so the energy is constant, hence θ is constant. Then in a classical picture, the angle between the two angular momenta $\boldsymbol{\ell}$ (proportional to $\mathbf{B}$) and $\mathbf{s}$ (proportional to $\boldsymbol{\mu}_s$) must also be constant. However, $\boldsymbol{\ell}$ and $\mathbf{s}$ precess around the total angular momentum $\mathbf{j}$ because the interaction $-\boldsymbol{\mu}_s \cdot \mathbf{B}$ gives rise to an *internal* torque acting on each of them (see Fig 2(a)).

Because of the precession of $\boldsymbol{\ell}$ and $\mathbf{s}$ about their resultant, $\mathbf{j}$, the individual projections of $\boldsymbol{\ell}$ and $\mathbf{s}$ along the z-axis, $\boldsymbol{\ell}_z$ and $\mathbf{s}_z$, are no longer constant. However, since there is no *external* torque acting on $\mathbf{j}$, its projection along the z-axis, $\mathbf{j}_z$, must be a constant of the motion.

Now think about the quantum mechanical description of these angular momenta. Before we apply the spin-orbit interaction, all the states described by the quantum numbers m_l and m_s are degenerate (i.e., the energy of the system does not depend on the orientation of the spin and orbital angular momenta). However, in the presence of the perturbation, the energy of the system now depends on the *relative* orientation of $\boldsymbol{\ell}$ and $\mathbf{s}$, so m_l and m_s are no longer good quantum numbers. Instead, the operator $\hat{\mathbf{j}}$ gives us a new quantum number, j, which is related to the relative orientation of $\boldsymbol{\ell}$ and $\mathbf{s}$ and can take values between $l + s$ and $|l - s|$ in steps of 1. Similarly, the operator $\hat{\mathbf{j}}_z$ gives rise to a new quantum number m_j. In the coupled system, m_l and m_s are no longer good quantum numbers because they do not describe constants of the motion any more. Their operators are not eigenfunctions of the new wave functions that describe the coupled system. The new states, labelled by j and m_j are formed of linear combinations of the m_l and m_s states. The energy now depends on the value of j (but not on the value of m_j).

If we define $\mathbf{j}$ by the equation $\mathbf{j} = \boldsymbol{\ell} + \mathbf{s}$, squaring it yields $\mathbf{j}^2 = \boldsymbol{\ell}^2 + 2\boldsymbol{\ell} \cdot \mathbf{s} + \mathbf{s}^2$. This can be rearranged to give the expression $\boldsymbol{\ell} \cdot \mathbf{s} = \frac{1}{2}(\mathbf{j}^2 - \boldsymbol{\ell}^2 - \mathbf{s}^2)$. We know that when we take expectation values of squared angular momentum operators in quantum mechanics, we get expressions like $\langle \hat{\boldsymbol{\ell}}^2 \rangle = l(l+1)\hbar^2$. We therefore obtain the result

$$\langle \hat{\boldsymbol{\ell}} \cdot \hat{\mathbf{s}} \rangle = \langle \tfrac{1}{2}(\hat{\mathbf{j}}^2 - \hat{\boldsymbol{\ell}}^2 - \hat{\mathbf{s}}^2) \rangle = \tfrac{1}{2}\hbar^2[j(j+1) - l(l+1) - s(s+1)]. \tag{4.10}$$

Note here that because $s = \frac{1}{2}$, j can only take two values: when $\boldsymbol{\ell}$ and $\mathbf{s}$ are parallel, we obtain $j = l + s = l + \frac{1}{2}$ and when they are antiparallel we obtain $j = l - s = l - \frac{1}{2}$ (note that s *always* has the value $\frac{1}{2}$ for electrons).

Combining these results, we can write an expression for the spin-orbit energy:

$$\Delta E = \langle \mathcal{H}' \rangle = -E_n \frac{Z^2 \alpha^2}{n} \frac{[j(j+1) - l(l+1) - s(s+1)]}{2l(l + \frac{1}{2})(l+1)} \tag{4.11}$$

Note: you do not need to learn this expression.

This expression involves the quantity α, which is called the *fine-structure constant*. It is equal to $e^2/4\pi\epsilon_0\hbar c$ and has the approximate value $1/137$. We generally write the spin-orbit energy as $\Delta E = \langle A_{nl}\hat{\boldsymbol{\ell}} \cdot \hat{\mathbf{s}}/\hbar^2 \rangle = \frac{1}{2}A_{nl}[j(j+1) - l(l+1) - s(s+1)]$. For the 2p state of hydrogen (where $l = 1$), this has the value $\Delta E = \frac{1}{2}A_{nl}$ for the $j = \frac{3}{2}$ level and $\Delta E = -A_{nl}$ for the $j = \frac{1}{2}$ level. In general every n, l state in hydrogen splits into two levels due to the spin-orbit interaction, except the s states (with $l = 0$) which do not split. This is because if $l = 0$ there is only one possible value of j, which is $j = s = \frac{1}{2}$.
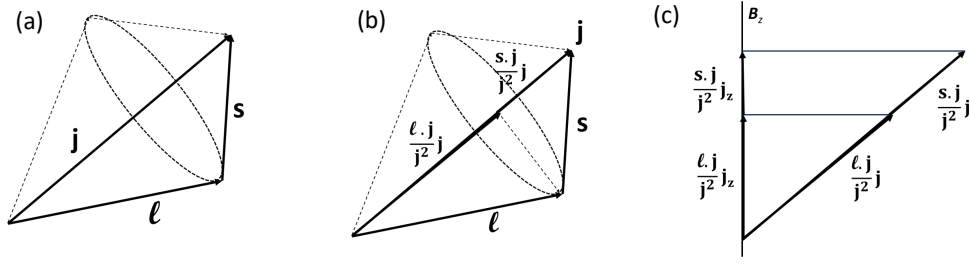
Figure 2: The "vector model" for the spin-orbit interaction and the Zeeman effect.

It is always very important to distinguish between *vector operators* (such as $\hat{\boldsymbol{\ell}}$, $\hat{\mathbf{s}}$ and $\hat{\mathbf{j}}$) and their quantum numbers (l, s and j), which are *scalar quantities*. A vector equation like $\hat{\mathbf{j}} = \hat{\boldsymbol{\ell}} + \hat{\mathbf{s}}$ is always true, but this does not mean that you can write $j = l + s$ because j can take a number of different values. So we always need to be careful about whether we are talking about quantum numbers or vector operators when writing relations between angular momenta.

4.6   Other relativistic effects

The spin-orbit interaction is a relativistic effect. This may not be obvious but in fact anything to do with spin is linked to relativity. Another relativistic effect which is important for hydrogen is an energy shift associated with the relativistic mass correction for the moving electron. This turns out to have the same order of magnitude as the spin-orbit interaction. It can be evaluated by considering the expectation value of the relativistic correction to the kinetic energy of the electron.

Finally, there is a third relativistic contribution, called the *Darwin term*, which affects only s states ($l = 0$). This term has no obvious physical interpretation. Again, this has the same order of magnitude as the spin-orbit interaction.

Combining all three corrections, we obtain an expression (*Note: you do not need to learn this expression*)

$$\Delta E_{nj} = \frac{1}{2}mc^2 \left(\frac{\alpha Z}{n}\right)^4 \left(\frac{3}{4} - \frac{n}{j + \frac{1}{2}}\right). \tag{4.12}$$

Note that this shows that the energy of a state n, j is independent of the value of l.

Although I have treated these relativistic corrections using perturbation theory, we could instead have started from the Dirac equation, rather than the Schrödinger equation. The Dirac equation treats all relativistic effects exactly, without any approximations, and it results in exactly the same expression for the energy levels as that given above. However, it involves difficult mathematics and gives no physical insight into what is going on. And it is still necessary to use perturbation theory for treating other effects such as the Zeeman effect.

Equation 4.12 shows that the energy of a state specified by the quantum numbers n, l, j is independent of l. Experimentally this was found not quite to be the case. The states with $l = 0$ were found to have slightly higher energy than the states with $l > 0$. This energy difference is not explained by the Dirac theory and requires the theory of *Quantum Electrodynamics (QED)* to explain it.

As we said earlier, g, the *gyromagnetic ratio* of the electron, is expected to have the value 2 according to Dirac theory. In practice it is found to be slightly different from this value, by about one part in 10^3. This difference is another consequence of QED. In fact g is known experimentally to about 13 decimal places, and this value is also confirmed by extremely accurate

(and amazingly complex) QED calculations. We will continue to use the value $g = 2$ in the following section as it is generally a very good approximation.

4.7 The Zeeman effect

The Zeeman effect is the interaction of the atom with an external magnetic field $\mathbf{B}$. The interaction arises because the electron's spin is associated with a magnetic moment $\boldsymbol{\mu}_s = -g\mu_B\mathbf{s}/\hbar = -2\mu_B\mathbf{s}/\hbar$ (setting $g = 2$) and also the motion of the electron gives rise to a magnetic moment $\boldsymbol{\mu}_l = -\mu_B\boldsymbol{\ell}/\hbar$. The perturbation results in an interaction $\mathcal{H}' = -(\boldsymbol{\mu}_s + \boldsymbol{\mu}_l) \cdot \mathbf{B}$. Let's assume that the field is in the z-direction, i.e. $\mathbf{B} = B\hat{\mathbf{z}}$. Then the perturbation can be expressed as

$$\mathcal{H}' = -(\boldsymbol{\mu}_s + \boldsymbol{\mu}_l) \cdot \hat{\mathbf{z}}B = (2\hat{\mathbf{s}}_z + \hat{\boldsymbol{\ell}}_z)\mu_B B/\hbar \tag{4.13}$$

We will start by ignoring the spin component of this interaction, so we will just consider the magnetic moment due to the orbital motion of the electron. This is called the *Normal Zeeman Effect* because it was understood on the basis of classical physics before spin was discovered. You have seen a discussion of the Normal Zeeman Effect in the *Quantum Mechanics* course, lecture 24, so we will only discuss it briefly here.

In the case of the Normal Zeeman Effect, the perturbation is just

$$\mathcal{H}' = -\boldsymbol{\mu}_l \cdot \hat{\mathbf{z}}B = \hat{\boldsymbol{\ell}}_z\mu_B B/\hbar \tag{4.14}$$

and since in the absence of spin m_l is a good quantum number, we can simply write

$$\mathcal{H}' = -\boldsymbol{\mu}_l \cdot \hat{\mathbf{z}}B = \hat{\boldsymbol{\ell}}_z\mu_B B/\hbar = \mu_B B m_l \tag{4.15}$$

so the Zeeman effect leads to a shift of each energy level proportional to its value of m_l.

However, in hydrogen the electron spin is always present so it is not possible to ignore it. The Normal Zeeman Effect is not observed in practice. The observed effect was called the *Anomalous Zeeman Effect* because it was not explained until the effects of spin were understood.

The complication arises because m_l and m_s are no longer good quantum numbers due to spin-orbit coupling, so the expectation value of $\hat{\mathbf{s}}_z$ and $\hat{\boldsymbol{\ell}}_z$ will now need to be expressed in terms of j and m_j rather than m_l and m_s. (From now on we will drop the hats on the operators $\hat{\boldsymbol{\ell}}$, $\hat{\mathbf{s}}$ and $\hat{\mathbf{j}}$.)

Assuming that the Zeeman effect is much smaller than the spin-orbit interaction, we need to consider what the expectation value of $\boldsymbol{\ell}_z$ is. Since $\boldsymbol{\ell}$ and $\mathbf{s}$ are coupled together to form $\mathbf{j}$, the component of $\boldsymbol{\ell}$ perpendicular to $\mathbf{j}$ will average to zero, and its component along $\mathbf{j}$ will be constant.

In order to find an expression for this component, we first define the *unit vector* along the direction of $\mathbf{j}$ to be $\mathbf{J} = \mathbf{j}/|\mathbf{j}|$. Then the *magnitude* of the component of $\boldsymbol{\ell}$ along $\mathbf{j}$ is $\boldsymbol{\ell} \cdot \mathbf{J}$ and of course its *direction* is $\mathbf{J}$. The vector expression for the component is therefore $(\boldsymbol{\ell} \cdot \mathbf{J}) \mathbf{J} = (\boldsymbol{\ell} \cdot \mathbf{j}/|\mathbf{j}|^2) \mathbf{j}$ which we can also write as $(\boldsymbol{\ell} \cdot \mathbf{j}/\mathbf{j}^2) \mathbf{j}$ (see Fig 2(b)). The projection of this along the z-axis is therefore $(\boldsymbol{\ell} \cdot \mathbf{j}/\mathbf{j}^2) \mathbf{j_z}$ (see Fig 2(c)). In order to evaluate the expectation value of $\boldsymbol{\ell} \cdot \mathbf{j}$, we again use the relation $\mathbf{j} = \boldsymbol{\ell} + \mathbf{s}$. Writing it as $\mathbf{j} - \boldsymbol{\ell} = \mathbf{s}$, and squaring, we find $\mathbf{j}^2 - 2\boldsymbol{\ell} \cdot \mathbf{j} + \boldsymbol{\ell}^2 = \mathbf{s}^2$, so

$$\langle \boldsymbol{\ell} \cdot \mathbf{j} \rangle = \langle \tfrac{1}{2}(\mathbf{j}^2 + \boldsymbol{\ell}^2 - \mathbf{s}^2) \rangle = \tfrac{1}{2}[j(j+1) + l(l+1) - s(s+1)]. \tag{4.16}$$

We can therefore write

$$\boldsymbol{\ell}_z = \frac{\boldsymbol{\ell} \cdot \mathbf{j}}{\mathbf{j}^2} \cdot \mathbf{j}_z = \frac{[j(j+1) + l(l+1) - s(s+1)]}{2j(j+1)} \cdot \mathbf{j}_z. \tag{4.17}$$

A similar expression can be found for $\mathbf{s}_z$:

$$\mathbf{s}_z = \frac{\mathbf{s} \cdot \mathbf{j}}{\mathbf{j}^2} \cdot \mathbf{j}_z = \frac{[j(j+1) - l(l+1) + s(s+1)]}{2j(j+1)} \cdot \mathbf{j}_z. \tag{4.18}$$

This allows us to evaluate the Anomalous Zeeman Effect as

$$\langle \mathcal{H}' \rangle = \left(2 \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)} + \frac{j(j+1) + l(l+1) - s(s+1)}{2j(j+1)} \right) \mu_B B m_j. \quad (4.19)$$

This is written more compactly as

$$\langle \mathcal{H}' \rangle = \Delta E = g_j \mu_B B m_j \quad (4.20)$$

where

$$g_j = \frac{3j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}. \quad (4.21)$$

So the Zeeman splitting is proportional to B and is different for each value of m_j. The Zeeman interaction lifts the degeneracy of the states. Note that when $l = 0$, j must be equal to s (which always has the value $\frac{1}{2}$) and then $g_j = 2$, which is what we would expect as there is only spin angular momentum present.

As B increases, there will come a point where the approximations start to break down. This is when the magnitude of the Zeeman interaction becomes comparable to the spin-orbit splitting. As we approach this point, we need to include higher orders of perturbation theory in the calculation if we want to have an accurate calculation of the energies of the states. This will result in extra quadratic terms in the expression for the energy shifts due to the Zeeman effect.

4.8 Hyperfine structure (*non-examinable*)

We will just mention hyperfine structure for completeness but we will not give any detailed treatment, and this topic is not examinable. Hyperfine structure is a further splitting of energy levels that arises from the interaction of the atom's internal magnetic field with the magnetic moment of the nucleus (if it has one). In hydrogen the nuclear spin is $I = \frac{1}{2}$ and we always end up with a splitting of each atomic level into two new states labelled by a new quantum number F. If the electronic angular momentum is j, the new quantum number F takes the values $j - \frac{1}{2}$ and $j + \frac{1}{2}$. The splitting between these states is generally very small as the magnetic moment of the nucleus is of order 2000 times less than that of the electron.

4.9 Summary

- Hydrogen-like systems behave in a similar way to hydrogen, but with the appropriate value of the reduced mass and the nuclear charge in the expressions for the energies and wave functions

- The most important perturbation in hydrogen is the spin-orbit interaction between the magnetic moment of the electron and the magnetic field it experiences due to its motion in the nuclear electric field

- The spin-orbit interaction couples $\hat{\boldsymbol{\ell}}$ and $\hat{\mathbf{s}}$ to form a total angular momentum $\hat{\mathbf{j}}$, and states with different values of j now have different energies

- The total relativistic correction to the energy levels is independent of the value of l

- The Zeeman effect arises from the interaction with an external magnetic field B. It splits the energy levels according to the value of m_j, which describes the projection of the total angular momentum on the z-axis. The splitting is proportional to m_j and to B