

# Further Quantum Physics

*Concepts in quantum physics and the structure of hydrogen and helium  
atoms*

Prof Andrew Steane

January 18, 2005



# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
1.1	Quantum physics and atoms . . . . .	7
1.1.1	The role of classical and quantum mechanics . . . . .	9
1.2	Atomic physics—some preliminaries . . . . .	9
1.2.1	Textbooks . . . . .	10
<b>2</b>	<b>The 1-dimensional projectile: an example for revision</b>	<b>11</b>
2.1	Classical treatment . . . . .	11
2.2	Quantum treatment . . . . .	13
2.2.1	Main features . . . . .	13
2.2.2	Precise quantum analysis . . . . .	13
<b>3</b>	<b>Hydrogen</b>	<b>17</b>
3.1	Some semi-classical estimates . . . . .	17
3.2	2-body system: reduced mass . . . . .	18
3.2.1	Reduced mass in quantum 2-body problem . . . . .	19
3.3	Solution of Schrödinger equation for hydrogen . . . . .	20
3.3.1	General features of the radial solution . . . . .	21
3.3.2	Precise solution . . . . .	21
3.3.3	Mean radius . . . . .	25
3.3.4	How to remember hydrogen . . . . .	25
3.3.5	Main points . . . . .	25
3.3.6	Appendix on series solution of hydrogen equation, off syllabus . . . . .	26

<b>4</b>	<b>Hydrogen-like systems and spectra</b>	<b>27</b>
4.1	Hydrogen-like systems . . . . .	27
4.2	Spectroscopy . . . . .	29
4.2.1	Main points for use of grating spectrograph . . . . .	29
4.2.2	Resolution . . . . .	30
4.2.3	Usefulness of both emission and absorption methods . . . . .	30
4.3	The spectrum for hydrogen . . . . .	31
<b>5</b>	<b>Introduction to fine structure and spin</b>	<b>33</b>
5.1	Experimental observation of fine structure . . . . .	33
5.2	The Dirac result . . . . .	34
5.3	Schrödinger method to account for fine structure . . . . .	35
5.4	Physical nature of orbital and spin angular momenta . . . . .	36
5.4.1	Dirac notation . . . . .	39
5.4.2	Main points . . . . .	39
<b>6</b>	<b>Perturbation Theory</b>	<b>41</b>
6.1	First order non-degenerate time-independent perturbation theory . . . . .	41
6.1.1	Simple examples . . . . .	42
6.2	2nd and higher orders . . . . .	43
6.2.1	Energy levels ‘repel’ . . . . .	44
6.3	Examples in hydrogen ground state . . . . .	45
6.3.1	Nuclear volume correction . . . . .	45
6.3.2	Relativistic kinetic energy correction . . . . .	47
<b>7</b>	<b>Degenerate perturbation theory</b>	<b>49</b>
7.1	A summary . . . . .	49
7.1.1	Degenerate sub-space . . . . .	50
7.1.2	The general method . . . . .	51
7.2	Constants of the motion . . . . .	51

<i>CONTENTS</i>	5
7.3 Example: charged gyroscopes . . . . .	53
7.3.1 Quantum gyroscopes . . . . .	54
7.4 Appendix: An exact analysis of a simple example . . . . .	55
<b>8 Hydrogen fine structure according to Schrödinger</b>	<b>57</b>
8.1 Relativistic kinetic energy correction . . . . .	57
8.2 Spin-orbit interaction . . . . .	58
8.2.1 Magnetic dipoles . . . . .	58
8.2.2 Magnetic field experienced by the electron . . . . .	59
8.2.3 The Hamiltonian and perturbation calculation . . . . .	60
8.3 Darwin term for $l = 0$ . . . . .	62
8.3.1 Main points . . . . .	63
8.4 Appendix on the correct choice of eigenstates for hydrogen . . . . .	63
8.4.1 Commutation of vector operators . . . . .	64
<b>9 The Zeeman effect</b>	<b>65</b>
9.1 Weak $B$ field: Zeeman effect . . . . .	65
9.1.1 Classical treatment and normal Zeeman effect . . . . .	65
9.1.2 Quantum treatment: Zeeman effect for $s$ states . . . . .	66
9.1.3 Effect on the emission spectrum . . . . .	67
9.2 Strong $B$ field: Paschen-Back effect . . . . .	68
9.3 Uses of the Zeeman effect . . . . .	68
9.3.1 Main points . . . . .	69
<b>10 Helium: gross structure</b>	<b>71</b>
10.1 Main features of the structure, screening . . . . .	71
10.1.1 Ground state . . . . .	73
10.2 The variational method . . . . .	74
10.2.1 Application to helium . . . . .	76
10.2.2 Main points . . . . .	76

<b>11 Exchange symmetry and the Pauli exclusion principle</b>	<b>77</b>
11.1 Identical particles . . . . .	78
11.1.1 Classical situation . . . . .	78
11.1.2 Quantum case: Exchange Principle . . . . .	79
11.1.3 Pauli Exclusion Principle . . . . .	82
11.2 Example to illustrate exchange effects . . . . .	82
11.2.1 Non-identical particles . . . . .	83
11.2.2 Identical Bosons . . . . .	84
11.2.3 Identical Fermions . . . . .	85
11.2.4 Discussion . . . . .	86
11.2.5 More than two particles . . . . .	87
11.2.6 Exchange symmetry of eigenstates of summed operators . . . . .	88
11.2.7 When can identical particles be treated as if they were non-identical? . . . . .	88
11.3 Exchange symmetry in helium . . . . .	90
11.4 Appendix: detailed derivation of states and splitting . . . . .	91
<b>12 Time-dependent perturbation theory</b>	<b>95</b>
12.1 First order time-dependent perturbation theory . . . . .	95
12.1.1 Limit of slow and fast perturbations . . . . .	96
12.2 Periodic perturbation: Fermi Golden Rule . . . . .	98
12.2.1 Forbidden transitions and selection rules . . . . .	101
<b>13 Interaction of atoms with light</b>	<b>103</b>
13.1 Emission of light waves from an atom . . . . .	103
13.1.1 Electric dipole radiation . . . . .	105
13.1.2 Selection rules for electric dipole radiation . . . . .	105
13.1.3 An example of non-Fermi Golden Rule behaviour: magnetic resonance . . . . .	107
13.1.4 Main points . . . . .	107

# Chapter 1

## Introduction

### 1.1 Quantum physics and atoms

The main aim of this lecture course is to take you further into quantum physics. However I want to avoid the subject appearing too abstract, so we shall be studying realistic examples as we go along. Most of these examples will be in the behaviour of the electron(s) in hydrogen and helium atoms. The course therefore also acts as an introduction to some aspects of atomic physics.

The big ideas we will meet are *precise calculations for hydrogen and related systems*, *spin*, *perturbation theory*, the idea of a *degenerate space of states*, *constants of the motion*, *relativistic effects in hydrogen*, *the variational method*, *exchange symmetry*, the treatment of *time-dependent Hamiltonians*, and the basic *interaction between light and matter*.

Some important quantum ideas are not covered, such as a thorough discussion of symmetries and their consequences, a full treatment of angular momentum coupling, the quantisation of fields and quantum entanglement. However, there is enough here to make those further ideas reasonably accessible should the student wish to explore them. The atomic physics covered in these lectures is only a very incomplete introduction, because it is restricted to hydrogen and helium atoms, and even for them several important issues will be ignored, such as the nuclear effects leading to ‘hyperfine structure’, the Zeeman effect in general and the fine structure of helium. The opportunity to get to grips with a tightly bound system of large numbers of electrons (i.e. any one of all the other atoms of the periodic table) is provided in the third year.

Let us consider why we might want to learn about atoms.

If you take an everyday object such as a brick and crumble it up into little pieces, then the smallest piece which you could say is somewhat like a very small piece of brick is an atom. Once you try to go to something smaller, by looking inside the atom, suddenly what you find is nothing at all like a brick or anything from everyday life. It is a springy cloud of electrons surrounding an incredibly dense nucleus; it is an electric system with more resonances than a piano, some of which are the most precisely defined quantities in all of physics, and the classical physics of Newton’s equations completely fails to give any sort of correct description of what is going on.

Atomic physics is the study of those properties of an atom which are associated primarily with the system of electrons, and with how they interact with light. Understanding hydrogen and helium is the beginning of the subject. It is of immense practical importance, from allowing us to probe the outer reaches of space to designing street lamps. It also plays a significant role in physics education by introducing many of the significant concepts of quantum mechanics in a precise way. These can

then be taken up and applied in other areas such as particle physics and condensed matter physics. These concepts include not just quantisation of energy, but also more subtle ideas such as exchange symmetry and many-particle wavefunctions.

Before we try to understand the quantum mechanics of a solid material, it is useful first to understand a single atom. Interestingly, it is also useful to understand an atom first before we try to understand a proton, because the quantum theory of the electron in hydrogen is a lot easier than the quantum theory of the quarks in a proton. In fact, the hydrogen atom is the most precisely understood system in all of physics. For example, the difference in binding energies of the two most strongly bound states of the electron-proton system has been measured<sup>1</sup> to a fractional precision  $1.8 \times 10^{-14}$ . The Schrödinger equation is not precise to this degree because it does not account correctly for special relativity—yet even when we replace the Schrödinger equation by the relativistically correct Dirac equation (outside the syllabus) the theory is still not precise enough. There is a part per million contribution which arises, roughly speaking, from the nature of the vacuum, whose theoretical treatment is provided by the full relativistic quantum theory of the electromagnetic interactions, called Quantum Electrodynamics (QED). With QED, however, we do at last find agreement between experiment and theory to fantastic precision (up to order  $\alpha^8$ , see later for the meaning of this). This agreement is one of the triumphs of modern physics—it allows us to be confident that our understanding of the natural world is built on accurate foundations. If the human race knows anything at all, it knows what happens when an electron orbits a proton.

Atomic physics allows us to explore our fundamental understanding of Nature in several other ways. Continuing the theme of precision, we can for example test for a variation in time of fundamental ‘constants’ such as the charge on the electron; such variation is certainly small since if a variation by one part in  $10^{14}$  had taken place during the year 2000 then the atomic clocks in Paris would have noticed because comparisons of frequencies of different atoms are sufficiently precise<sup>2</sup>! The mass of the  $Z$  boson, normally associated with vast particle accelerator experiments, was measured in the basement of the Clarendon laboratory using two optical tables, associated lasers, polarizers, and a tube of bismuth gas in 1996 to a precision<sup>3</sup> of 5%, by detecting a transition in bismuth atoms which violates the principle that left-handed and right-handed fundamental processes must occur equally often. Such ‘parity violating’ processes are only caused by the so-called *weak* interaction which is mediated by the  $Z$  particle; this is the link between the observed optical effect (a polarization rotation with a special frequency-dependence) and the underlying fundamental physics.

The interaction of light and matter is central to atomic physics because the primary way in which light is produced or absorbed in the natural world is by interactions between photons and electrons in atoms or molecules. This is in contrast to radio waves, for example, which interact with electrons moving in larger objects such as a metal wire. Almost all the light we see streaming to and fro in the world, bringing to us a sense of the leaves on the trees and the waves on the sea shore and the stars around us, was emitted by atoms or molecules in the leaf or the water or the star, rather than by something smaller or larger. In this sense, we may truly say that if you want to see what atoms look like, *look around you*.

Properties of atoms are basic tools of several other branches of physics—for example the speed of stars is deduced from the Doppler shift of the light they emit, and the reason it is possible to know the light wavelength emitted by the star is because we can recognise the signatures of the atoms in the outer layer of the star, which only emit at a set of characteristic wavelengths. The set is recognisable, but is found uniformly shifted in frequency. Observation of many other properties of stars also relies on arguments from atomic physics, for example we can deduce magnetic fields from the Zeeman effect. Experiments involving plasmas, such as fusion power research, similarly require detailed knowledge of

<sup>1</sup>Niering *et. al*, Phys. Rev. Lett. **84**, 5496 (2000), also Holzworth *et. al* in Atomic Physics 17: Proceedings of the XVII International Conference on Atomic Physics (ICAP 2000) (American Institute of Physics, New York 2001). The result:  $f_{1s-2s} = 2466\ 061\ 413\ 187\ 103 \pm 46$  Hz.

<sup>2</sup>C. Salomon *at al.* in “Cold Atom Clocks,” Atomic Physics 17: Proceedings of the XVII International Conference on Atomic Physics (ICAP 2000), pp 23-40 (American Institute of Physics, New York 2001).

<sup>3</sup>The experiment was precise to 2%; the theoretical extraction of the  $Z$  mass relied on highly complex atomic structure calculations which were precise to 5%.



the properties of free atoms and ions. Finally, to return to the humble street light; this is based on some simple properties of sodium atoms.

### 1.1.1 The role of classical and quantum mechanics

Quantum mechanics is the only way to understand physics properly. However, because classical mechanics is so much more familiar, it a very useful tool for providing physical insight and estimates of various quantities such as energies, distances, etc. One should develop a feel for both.

## 1.2 Atomic physics—some preliminaries

The atomic physicist counts “one, two, lots.” That is to say, hydrogen and helium are both special cases, and the rest are understood in terms of concepts borrowed from them. Hydrogen is special because it can be solved exactly (ignoring quantum electrodynamics) and because the various relativistic effects are all of a similar order of magnitude, while in other atoms one of these tends to dominate (the spin-orbit interaction, more of this later). Helium is special because the many-electron averaging which we use in atoms with many electrons is not so successful, and also various spin effects are all of the same order of magnitude (see chapter 5).

The following table gives the values of constants which will be much used and should be committed to memory to 3 significant figures. Their meaning will become clear as we go along.

Fine structure constant	$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c}$	1/137.03599976
Rydberg energy	$E_R = \frac{1}{2}m\alpha^2c^2$	13.6056981 eV
Bohr radius	$a_0 = \hbar/m\alpha c$	0.529177249 Å
Bohr magneton	$\mu_B = e\hbar/2m$	$h \times 13.9962418$ GHz/T
electron $g$ factor	$g = 2(1 + \frac{\alpha}{2\pi} - 0.328\frac{\alpha^2}{\pi^2} + \dots)$	$2 \times 1.0011596$

The diagram in the lecture shows how to think about an atom. It is useful to pick hydrogen as a starting point, and then go on to a ‘typical’ atom having atomic number of order 10. Here are the order-of-magnitude sizes of some physical features which come up:

	feature	hydrogen ( $Z = 1$ )	range for $Z = 1$ to 100	units
	mean radius	0.79	0.5—2.7	Å
	ionization potential	13.6	4—24	V
	r.m.s. speed of an electron	0.01	0.001—1	$c$ (speed of light)
	electric field at mean radius	$10^{11}$	$10^{11}$ — $10^{12}$	V/m
	electric field at edge of nucleus	$10^{21}$	$10^{21}$ — $2 \times 10^{22}$	V/m
	mean magnetic field experienced by outer electron	3	3—1000	T
	magnetic field experienced by the nucleus	100	10—1000	T
	electric dipole moment of the atom	0	0	Cm
	magnetic dipole moment of the atom	$10^{-23}$	0— $4 \times 10^{-23}$	J/T

The aim at this stage is to note the order of magnitude and the fact that for some properties a wide range is possible, while others do not vary much. In particular, notice that the electric and magnetic fields inside any atom are huge. The large electric field is simply because the region inside the atom is at a very small distance (one angstrom down to one fermi) from the charge-carrying particles (electrons, protons). The large magnetic field is because the electrons are moving quickly through this large electric field, at speeds approaching the speed of light. These statements will be further clarified and justified as we go on.

### 1.2.1 Textbooks

There are many excellent textbooks on quantum mechanics. My own favourite is R. Shankar, “Principles of Quantum Mechanics”, Plenum Press. Shankar starts with the mathematics of state vectors and operators, and only later introduces wavefunctions. This will suit strong students but possibly not the majority. I also like the book by Gasiorowicz. Merzbacher is good as a further reference once you have got some familiarity with the subject. The treatment by A. Peres is more at a graduate level but it is very good on basic interpretational issues and may be an attractive additional resource for philosophers for that reason.

The atomic physics we are dealing with at this stage is only a small part of the complete subject, but in any case you will be learning more in the third year. The following are useful in different ways:

C. Foot, “Atomic Physics”, Oxford University Press, highly recommended. It covers the right material for 2nd and 3rd year, including some of the B2 as well as B1 material, and good on experimental methods as well as the basic theory.

Woodgate, “Elementary Atomic Structure,” excellent and thorough, though aimed at a level which may test the average student.

Eisberg and Resnick, good summary, some very nice pictures, I recommend you read it. We will see that some aspects of the discussion of exchange symmetry are less clear than they might be.

Leighton, “Principles of Modern Physics,” chapters 5,6,7,8. Still a good clear introduction, though is slightly muddled when dealing with exchange symmetry, so not recommended for that.

Haken and Wolf is accessible and fills in experimental details which are missing elsewhere.

Kuhn, “Atomic Spectra” an old book, is a good source of information on general atoms (spectra, term diagrams) and is thorough on X ray atomic spectra. We won’t need it at this stage but I wanted you to be aware of it.

T. P. Softley, “Atomic Spectra”, summary of physics of atoms from a chemist’s point of view. It is succinct and down-to-earth, which is helpful, but some subtler points are not treated thoroughly.

## Chapter 2

# The 1-dimensional projectile: an example for revision

A cannon shoots a rubber ball straight up into the air: what happens?

### 2.1 Classical treatment

We shall not analyze the gunpowder explosion. Instead, we take the muzzle velocity  $v_m$  (the velocity of the rubber ball as it leaves the cannon) as a ‘given’. Hence, at time  $t_1$  we have:

$$\text{initial conditions: position } x = L, \text{ momentum } p = mv_m \quad (2.1)$$

Notice that two variables (position and momentum) are sufficient and necessary to give complete information about the state of motion of the rubber ball.

Equations of motion: Newton’s 2nd law ( $F = ma$ ) tells us the rate of change of one variable ( $p$ ) while the rate of change of the other variable ( $x$ ) is given by the definition of momentum:

$$\begin{aligned} F &= \frac{dp}{dt} \\ \frac{p}{m} &= \frac{dx}{dt} \end{aligned} \quad (2.2)$$

Note that both these equations are first-order differential equations. Together they are a pair of coupled first-order differential equations. The standard way to solve them is to substitute the 2nd into the first and thus get a single second-order differential equation.

The force is the familiar gravitational force  $F = -mg$ . In order to make the comparison with the quantum treatment more obvious, we shall write this as  $F = -dV/dx$  where  $V(x)$  is the potential energy function for a uniform gravitational field:

$$V(x) = mgx \quad \text{for } x > 0 \quad (2.3)$$

(we shall assume the ground is at  $x = -R$  where  $R$  is the radius of the rubber ball, so when  $x = 0$  the ball is just touching the ground).

The solution of these classical equations of motion is simple and well-known. The result is

$$p(t) = mv_m - mgt, \quad (2.4)$$

$$x(t) = L + v_m t - \frac{1}{2}gt^2. \quad (2.5)$$

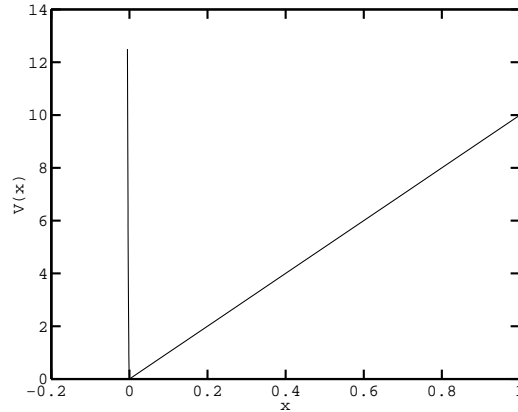


Figure 2.1: Potential energy function for the bouncing ball. The quadratic behaviour for  $x < 0$  is there but is hard to see on the graph because the elastic force is large compared to gravity.

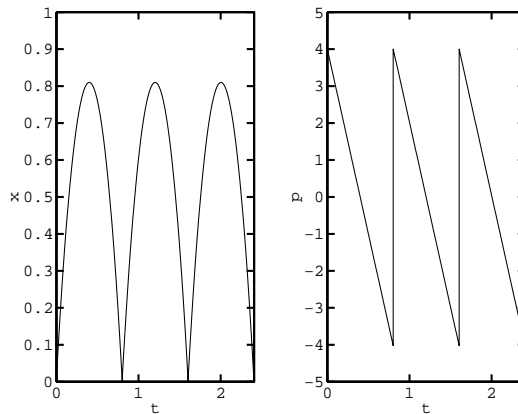


Figure 2.2: Classical solution for the bouncing ball. The left hand graph shows the position as a function of time, the right hand graph shows the momentum as a function of time.

This solution is accurate until  $x$  reaches zero, at which point the ball hits the ground.

To analyse the bounce on the ground, the standard method is to claim that energy is conserved, and that the state of motion of the ground (i.e. of planet earth) is unaffected. However, in order to keep the treatment uniform, we shall instead again use the method of solving the equations of motion (2.2). To do this, we need to know the potential energy of the ball as it impacts on the ground. As the ground and the ball are squeezed, the potential energy rises sharply. There is still a contribution from gravitational potential energy, but now most of the energy is elastic potential energy so we shall ignore the small gravitational part. We shall assume Hooke's law (force proportional to displacement) for the elastic force, so:

$$V(x) = \frac{1}{2}\kappa x^2 \quad \text{for } x \leq 0 \quad (2.6)$$

where  $\kappa$  is the spring constant which is large enough so that only a small displacement of the ball is enough to reduce its velocity to zero and then throw it back upwards again.

The equations of motion (2.2) are once again easy to solve: they give simple harmonic motion for the period when  $x \leq 0$ , and then the ball flies upwards again.

## 2.2 Quantum treatment

### 2.2.1 Main features

State at time just before 0: the wavefunction  $\psi$  of the ball is a smooth function, with no wiggles, peaked near  $x = 0$ . The expectation value of momentum  $\langle \psi | \hat{p} | \psi \rangle = 0$ .

Result of impulse produced by the gunpowder (time  $t = 0$ ): the wavefunction is now centred near  $x = L$ , and now has many wiggles (i.e. the real part of  $\psi$  oscillates with wavelength  $\lambda = h/(mv_m)$  where  $h$  is Planck's constant and  $v_m$  is the muzzle velocity). These wiggles imply that the second derivative  $d^2\psi/dx^2$  is large, and hence the ball has a large kinetic energy. Note that this illustrates the fact that the wavefunction contains *all* the information about the state of motion of the ball. For example, the wiggles result in an expectation value of momentum  $\langle \psi | \hat{p} | \psi \rangle = mv_m$ .

Further motion: the main peak of the wavefunction rises upwards until it reaches a height  $H$ . The wavefunction also spreads out (diffraction). Meanwhile the wiggles become more spread out (i.e. a larger wavelength), indicating that the momentum is getting smaller, until at  $x = H$  there are once again no wiggles and the mean momentum is zero. Subsequently, the peak of the wavefunction moves back down towards  $x = 0$  and the wiggles again appear.

As the front-most part of the wavefunction reaches the ground at  $x = -R$ , it is reflected, and for a short time a standing wave is set up as the reflected part of the wavefunction interferes with the part still moving towards the ground. Then the peak in the wavefunction moves upwards and the main features of the motion are repeated. The main thing which does not exactly repeat is the overall spreading out of the wavefunction: this continues until eventually the wavefunction is spread over the full range of heights from the ground to  $H$  and slightly higher.

### 2.2.2 Precise quantum analysis

As in the classical treatment, we shall not analyse the gunpowder explosion. Instead, we take the wavefunction  $\psi(x, 0)$  of the rubber ball as it leaves the cannon as a given. To be precise,  $t = 0$  is taken to be the time when the mean position of the ball is  $\langle \psi | \hat{x} | \psi \rangle = L$ , and we assume that by this time the gunpowder explosion has ceased to provide any further energy, and the mean momentum reaches the value  $\langle \psi | \hat{p} | \psi \rangle = mv_m$ .

$$\text{initial conditions: } \psi(x, 0) = \frac{1}{(2\pi\sigma^2)^{1/4}} e^{ikx} e^{-(x-L)^2/4\sigma^2} \quad (2.7)$$

where  $\hbar k = mv_m$  and  $\sigma$  is the spread in position. We assume  $\sigma \ll L$  so that this wavefunction does not reach the ground to any significant degree, and therefore we can neglect the influence of the ground at this point. Notice that specifying the initial wavefunction is sufficient and necessary to give complete information about the initial state of motion of the rubber ball.

To analyse the subsequent motion, we need to solve Schrödinger's equation:

$$\hat{H}\psi(x, t) = i\hbar \frac{\partial \psi(x, t)}{\partial t} \quad (2.8)$$

where  $\hat{H}$  is the Hamiltonian, which is the operator for total energy, which can be written

$$\hat{H} \equiv \frac{\hat{p}^2}{2m} + V(x) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x). \quad (2.9)$$

Hence Schrödinger's equation is

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x)\psi(x, t) = i\hbar \frac{\partial \psi(x, t)}{\partial t}. \quad (2.10)$$

This is a single partial differential equation, in contrast to the classical case where we had a pair of ordinary differential equations. Both the quantum and the classical case have in common that the equation is a first-order differential equation in time. The potential energy function  $V(x)$  is exactly the same in the quantum and classical treatments, equations (2.3) and (2.6) and figure 2.1.

Since we have the initial conditions, i.e. complete wavefunction at a given time, (2.7), and the equation of motion is first order in time, it follows that the equation of motion completely specifies the subsequent evolution of the wavefunction. In other words there is a unique solution  $\psi(x, t)$  for all subsequent times and positions.

The standard way to find this solution is as follows. First, we find the solutions of the time-independent Schrödinger equation  $\hat{H}\phi = E\phi$ , i.e.

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \phi(x)}{\partial x^2} + V(x)\phi(x) = E\phi(x). \quad (2.11)$$

This is itself a difficult mathematical problem. If we assume it can be solved, then we have a set of functions  $\phi_i(x)$ , each of which is an energy eigenstate for the problem we are dealing with, i.e. that of motion in a uniform gravitational field with a quadratic potential barrier (the ground) at the bottom. The energy eigenstates have energies  $E_i$ .

Now, since the Hamiltonian is hermitian, and since the eigenfunctions of any hermitian operator form a complete set, it follows that the initial state  $\psi(x, 0)$  can be written as a superposition of energy eigenfunctions:

$$\psi(x, 0) = \sum_i c_i \phi_i(x) \quad (2.12)$$

Note that this is a function purely of  $x$ , because it is the wavefunction only at one particular instant of time,  $t = 0$ . To find the coefficients, multiply (2.12) by  $\phi_j^*$  and integrate, and use the fact that the eigenfunctions are mutually orthogonal:

$$\int \phi_j^* \psi(x, t_1) dx = \sum_i c_i \int \phi_j^* \phi_i(x) dx = c_j \quad (2.13)$$

This mathematical procedure should by now be very familiar to you. I will underline its importance by re-writing the final result (2.13):

$$c_j = \langle \phi_j | \psi \rangle \quad (2.14)$$

and hence

$$\psi = \sum_i \langle \phi_i | \psi \rangle \phi_i(x) \quad (2.15)$$

$$= \sum_i \phi_i(x) \langle \phi_i | \psi \rangle. \quad (2.16)$$

In Dirac notation this takes the memorable form

$$|\psi\rangle = \sum_i |\phi_i\rangle \langle \phi_i | \psi \rangle. \quad (2.17)$$

The precise values of the coefficients  $c_i$  are given by (2.13). However, we should also use our physical insight to tell us some general features. Note that the total energy of the rubber ball is of order  $(1/2)mv_m^2 + mgL$ . Therefore we should expect the largest coefficients  $c_i$  for those energy levels which have energies close to this. Also, if  $(1/2)mv_m^2$  is large compared to the ground-state energy  $E_1$ , then we should expect the coefficient  $c_1$  to be small, etc.

So far we have done two things: (1) found the energy eigenfunctions, and (2) expressed the initial state as a superposition of energy eigenfunctions. It is now very simple to find the wavefunction at all

subsequent times. We know that each energy eigenfunction will evolve in time simply as  $\exp(-iE_i t/\hbar)$ , so the complete wavefunction at any later time is

$$\psi(x, t) = \sum_i c_i \phi_i(x) e^{-iE_i(t-t_1)/\hbar} = \sum_i \langle \phi_i | \psi(x, 0) \rangle \phi_i(x) e^{-iE_i(t-t_1)/\hbar}. \quad (2.18)$$

Note that this solution includes all the bouncing motion and everything. This completes the mathematical treatment of the evolution of the quantum system (assuming that we have been able to discover the energy eigenfunctions  $\phi_i$  and their energies  $E_i$ ). It remains to provide a *physical* statement to say what this mathematical result tells us about the rubber ball. For this we quote one of the basic postulates of the theory: the probability of finding the rubber ball between  $x$  and  $x + dx$  at time  $t$  is  $|\psi(x, t)|^2 dx$ .

Now we shall use the general machinery of quantum mechanics to add some further useful information.

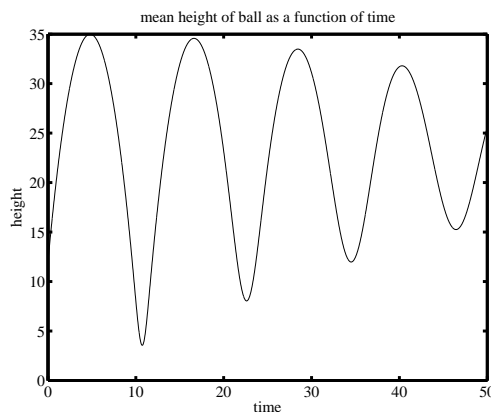


Figure 2.3: Quantum solution for the bouncing ball. The graph shows the expectation value of the height of the ball as a function time.

First, we know that quantum mechanical expectation values obey the laws of motion of classical mechanics. This implies, for example, that should the mean position of our rubber ball return to  $\langle \psi | \hat{x} | \psi \rangle = L$  then we know the mean momentum must at that time reach the value *precisely*  $\langle \psi | \hat{p} | \psi \rangle = mv_m$ . More generally, the classical results shown in figure 2 also apply quite closely to the quantum expectation values, see figure 3, the difference arising from the fact that the expectation value of the potential function  $\langle V \rangle$  is not quite the same as the function  $V$  itself, especially near  $x = 0$ .

Secondly, we can give further very useful physical statements. If, when the ball has wavefunction  $\psi(x, t)$ , a physical property  $Q$  (such as energy, or momentum, etc.) is measured, then the measurement will yield one of the eigenvalues of the operator  $\hat{Q}$ , and the probability that eigenvalue  $q_i$  is obtained is given by

$$P(\text{obtain } q_i) = |\langle u_i | \psi \rangle|^2 \quad (2.19)$$

where  $u_i(x)$  is the eigenfunction of  $\hat{Q}$  associated with eigenvalue  $q_i$ .

Combining this general result with equation (2.18) we can see that the measurement of total energy has a special property: the probability to obtain any particular energy  $E_i$  does not change in time. This is to be compared with the classical result that the *total* energy does not change at all as the ball goes through its motion. On the other hand, we could instead ask about measurement of just the kinetic energy (i.e. the operator  $-(\hbar^2/2m)\partial^2/\partial x^2$ ). For this observable, we would find that the probability of getting a high result (large kinetic energy) is large at first, when the ball leaves the cannon, and then falls as the ball slows down near the top of its trajectory, and then increases again as the ball falls to the ground, etc.





## Chapter 3

# Hydrogen

### 3.1 Some semi-classical estimates

When we wish to understand some aspect of the physical world, whether in this area of physics or any other, it is good practice to begin with what we call a “back of the envelope” calculation. This is any sort of very rough calculation which gives a useful initial impression of what might be going on, and some estimated order of magnitude.

To this end, let us consider a semi-classical model of hydrogen, as follows. Consider a classical particle of charge  $-e$  moving at speed  $v$  in a circular orbit of radius  $r$  about a fixed charge  $+e$ . The force is  $e^2/4\pi\epsilon_0 r^2$  so Newton’s law “ $F = ma$ ” gives

$$\frac{e^2}{4\pi\epsilon_0 r^2} = m \frac{v^2}{r} \quad \implies \quad \frac{e^2}{4\pi\epsilon_0 r} = mv^2. \quad (3.1)$$

This shows that the potential energy is equal to minus twice the kinetic energy, and it is also a relationship between  $v$  and  $r$ . Next (this is the quantum part), we will use Heisenberg’s Uncertainty Relation  $\Delta x \Delta p \geq \hbar/2$ . This is a precise statement about standard deviations, but in the present context we use it in a rough way. It is reasonable to guess that the radius  $r$  of the classical orbit gives the right order of magnitude for  $\Delta x$ , and similarly the value of  $v_x$  varies between  $\pm v$  so  $mv$  gives an order of magnitude estimate for  $\Delta p$ . We argue that in general it is found that the wavefunction for the ground state in a single potential well gives spreads  $\Delta x$  and  $\Delta p$  such that  $\Delta x \Delta p \simeq \hbar$ . Thus we obtain

$$rmv \simeq \hbar. \quad (3.2)$$

Combining this with (3.1) we have

$$r \simeq \frac{4\pi\epsilon_0 \hbar^2}{e^2 m} \simeq 0.5 \times 10^{-10} \text{ m}$$

and

$$v = e^2/4\pi\epsilon_0 \hbar \simeq c/137 \simeq 2000 \text{ km/s.}$$

Hence the total energy is (P.E. + K.E.) =  $(-K.E.) = -(1/2)mv^2 \simeq 13.6 \text{ eV}$ .

The radius and speed given by this semi-classical argument are good guides to what we will find in the full quantum treatment. We will find the radius  $r$  is the Bohr radius  $a_0$  which is close to the mean radius of the ground state of hydrogen. The speed  $v$  is  $\alpha c$  where  $\alpha$  is the fine structure constant, to be discussed, and this equals the r.m.s. speed of the electron in the ground state of hydrogen. Finally, it so happens that the energy came out exactly right: the ground state energy of hydrogen is indeed  $-13.6 \text{ eV}$ .

Notice that we could have obtained the estimate (3.2) by claiming that the angular momentum  $rmv$  should be of order 1 unit of  $\hbar$ , instead of by using the Uncertainty Principle. That would be a fair thing to do, but it turns out that the true quantum ground state doesn't have any orbital angular momentum, so the 'classical orbit' picture isn't completely successful.

In the limit of large energy and angular momentum quantum numbers, certain superpositions of quantum energy eigenstates give probability densities which do look like a lump of high probability (a wavepacket) going around the nucleus in an elliptical orbit, and thus the quantum and classical descriptions merge. There is one possible area for confusion, however: the classical orbit with the largest angular momentum for given energy is the circular one, while the quantum wavefunction with the *least* angular momentum for a given energy is the spherically symmetric one.

The elliptical orbits of the classical solution, when combined with some statements about quantisation of angular momentum, are called Bohr orbits. Bohr in fact arrived at the correct energy levels of hydrogen by an argument like the one above, but made precise rather than 'rough and ready', in that his criteria were to make specific assumptions about quantisation of angular momentum, and to ensure that the limit of high quantum numbers agreed with classical physics (the *correspondence principle*). I will not discuss the Bohr theory here, however. Although it has historical importance, I feel it may confuse more than it illuminates at this stage. What we need to do next to understand hydrogen is write down and solve the Schrödinger equation for the system.

## 3.2 2-body system: reduced mass

The concept of *reduced mass* for a 2-body system arises in classical mechanics, and also in quantum mechanics.

I won't go through the full textbook treatment for the classical case (since you can look it up in a textbook), but I want to comment on it in order to convey physical insight.

Consider for example an electron and a proton initially at rest and a distance of  $r = 1$  metre apart, with the proton at the origin and the electron on the positive  $x$  axis. The force on the electron owing to the Coulomb attraction by the proton is  $F = e^2/(4\pi\epsilon_0 r^2) = 2.3071 \times 10^{-28}$  N in the negative  $x$  direction. The force on the proton owing to the Coulomb attraction by the electron is  $2.3071 \times 10^{-28}$  N in the positive  $x$  direction. The electron therefore accelerates in the negative  $x$  direction with acceleration  $a_e$  given by  $|a_e| = F/m_e = 253.26 \text{ ms}^{-2}$ , and the proton accelerates in the positive  $x$  direction with acceleration  $a_p = F/m_p = 0.14 \text{ ms}^{-2}$ .

Now consider the relative coordinate  $x = x_e - x_p$ , i.e. the position of the electron relative to the proton. The acceleration  $d^2x/dt^2$  of the relative coordinate is

$$\frac{d^2x}{dt^2} = \frac{-F}{m_e} - \frac{F}{m_p} \quad (3.3)$$

$$= \frac{-F}{\mu} \quad (3.4)$$

$$\text{where } \mu = \left( \frac{1}{m_e} + \frac{1}{m_p} \right)^{-1} = \frac{m_e m_p}{m_e + m_p}. \quad (3.5)$$

This example gives the essential physical insight: when 2 bodies mutually attract (or repel), then the acceleration of the relative coordinate is the sum of the two accelerations of the particles, and the resulting equation of motion for the relative coordinate is *exactly the same as the equation of motion for a single particle with mass  $\mu$*  given by equation (3.5). This mass is called the *reduced mass*, because it is smaller than either of the masses of the two particles involved.

In the example just given, the two accelerations were 253.26 and  $0.14 \text{ ms}^{-2}$ , so they add to make 253.40

$\text{ms}^{-2}$ . Using the concept of reduced mass, we get the same result by first calculating  $\mu = 9.1044 \times 10^{-31}$  kg ( $= 0.9995 m_e$ ), with force  $F = 2.3071 \times 10^{-28}$  N as before, leading to acceleration  $F/\mu = 253.40 \text{ ms}^{-2}$ .

The only thing I have omitted here is the generalization to 3 dimensions: this is left as an exercise for the reader, or you may like to look it up.

Notice that the example using an electron and a proton shows that the reduced mass is a small correction to the lighter mass when the masses of the two bodies are very different. The fact that we are interested in getting this calculation exactly right is a first example of the fact that atomic physics is an especially precise part of physics.

### 3.2.1 Reduced mass in quantum 2-body problem

To show how the reduced mass concept carries over to quantum mechanics, we first write the complete energy operator of a two-body system such as hydrogen. To do this, we must add the kinetic energies of the proton and the electron, and the potential energy from their Coulomb interaction:

$$\hat{H}_{\text{full}} = \frac{\hat{p}_p^2}{2m_p} + \frac{\hat{p}_e^2}{2m_e} - \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_e - \mathbf{r}_p|} \quad (3.6)$$

Next, write down the position of the centre of mass:

$$\mathbf{R} = \frac{m_p\mathbf{r}_p + m_e\mathbf{r}_e}{m_p + m_e} \quad (3.7)$$

and also define the relative coordinate

$$\mathbf{r} = \mathbf{r}_e - \mathbf{r}_p. \quad (3.8)$$

The operator for the momentum of the system as a whole is  $\hat{P} = -i\hbar\nabla_R$  and the operator for the momentum associated with changes in the relative coordinate is  $\hat{p} = -i\hbar\nabla_r$ , where

$$\nabla_R \equiv \left( \frac{\partial}{\partial X}, \frac{\partial}{\partial Y}, \frac{\partial}{\partial Z} \right), \quad \nabla_r \equiv \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

Finally, prove (exercise for the reader) that the total kinetic energy of the system can be written

$$\frac{\hat{p}_p^2}{2m_p} + \frac{\hat{p}_e^2}{2m_e} = \frac{\hat{P}^2}{2(m_e + m_p)} + \frac{\hat{p}^2}{2\mu} \quad (3.9)$$

i.e. it is now expressed as a sum of kinetic energy of translation of the system as a whole, plus kinetic energy associated with the motion of the relative coordinate, where the latter term involves the reduced mass.

By expressing the kinetic energy this way, the Hamiltonian becomes a sum of two parts, where the first part depends only on  $\mathbf{R}$  and the second part depends only on  $\mathbf{r}$ :

$$\hat{H}_{\text{full}} = \left( \frac{\hat{P}^2}{2(m_e + m_p)} \right) + \left( \frac{\hat{p}^2}{2\mu} - \frac{e^2}{4\pi\epsilon_0|\mathbf{r}|} \right) \quad (3.10)$$

This means the complete Schrödinger equation for the two-body system is *separable*. Writing the time-independent Schrödinger equation for convenience, it separates into one equation for the motion of  $\mathbf{R}$ :

$$\frac{-\hbar^2}{2(m_e + m_p)} \nabla_R^2 \Psi(\mathbf{R}) = E_{\text{cm}} \Psi(\mathbf{R}), \quad (3.11)$$

and a separate equation for the motion of  $\mathbf{r}$ :

$$\frac{-\hbar^2}{2\mu} \nabla_r^2 \psi(\mathbf{r}) - \frac{e^2}{4\pi\epsilon_0 r} \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (3.12)$$

The energy  $E_{\text{cm}}$  is kinetic energy of translational motion of the centre of mass of the system, the energy  $E$  is the sum of kinetic and potential energy in the centre of mass frame. The total energy of the pair of bodies is  $E_{\text{cm}} + E$ . We are usually interested in a stationary atom, where  $E_{\text{cm}} = 0$ .

### 3.3 Solution of Schrödinger equation for hydrogen

Hydrogen consists of an electron orbiting a proton. By far the greatest contribution to the potential energy of the system is the simple Coulomb attraction. Therefore to an excellent approximation the Schrödinger equation for the system is the one we have just derived, equation (3.12). Recall that the first term  $-\hbar^2 \nabla^2 / 2\mu$  is the kinetic energy in the centre of mass frame, and the second term is the potential energy which is negative because the particles attract one another. The reason this equation is approximate rather than exact is that it ignores relativistic corrections which are small but non-negligible owing to the high velocities involved, and it treats the proton as a point-like particle whereas in fact it is composed of quarks and gluons and has a non-zero volume.

To solve (3.12), first, since the potential is spherically symmetric (called a *central* potential), treat the problem in spherical polar coordinates. The Laplacian operator  $\nabla^2$ , multiplied by  $-\hbar^2/2\mu$  to express kinetic energy, is in spherical polar coordinates

$$\frac{-\hbar^2}{2\mu} \nabla^2 = \frac{-\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{\hat{\mathbf{L}}^2}{2\mu r^2} \quad (3.13)$$

where  $\hat{\mathbf{L}}$  is the operator for angular momentum (defined by  $\hat{\mathbf{L}} = \mathbf{r} \wedge \hat{\mathbf{p}}$ ). The first term is the contribution to kinetic energy coming from radial motion, i.e. motion in and out, while the second term is the kinetic energy of angular motion, i.e. rotation. (You might recognize the form  $\hat{\mathbf{L}}^2/2\mu r^2$  as  $\hat{\mathbf{L}}^2/2I$  where  $I = \mu r^2$  is the moment of inertia.)

Multiplying the whole equation by  $r^2$ , and remembering that  $\hat{\mathbf{L}}^2$  is a function of angular variables only, not radial ones, it is possible to see that the Schrödinger equation can be solved by the method of separation of variables. There exists a complete set of solutions having the form  $\psi(r, \theta, \phi) = R(r)Y(\theta, \phi)$ . The  $Y$  part you should recall from previous lectures, it is an eigenstate of the operator  $\hat{\mathbf{L}}^2$ , with eigenvalue  $l(l+1)\hbar^2$  where  $l = 0, 1, 2, \dots$  is a non-negative integer. It is also an eigenstate of the operator for the  $z$ -component of angular momentum  $\hat{L}_z$ , having eigenvalue  $m\hbar$  where  $m$  is an integer in the range  $-l \leq m \leq l$ . Therefore we will label the  $Y$  by the subscripts  $l, m$  to remind us of this:

$$\hat{\mathbf{L}}^2 Y_{l,m}(\theta, \phi) = l(l+1)\hbar^2 Y_{l,m}(\theta, \phi) \quad (3.14)$$

$$\hat{L}_z Y_{l,m}(\theta, \phi) = m\hbar Y_{l,m}(\theta, \phi). \quad (3.15)$$

The  $Y$  functions are called the *spherical harmonics*.

The radial part remains to be solved. It is a solution of the equation

$$\left( \frac{-\hbar^2}{2\mu} \frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d}{dr} \right) + \frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right) R_E(r) = ER_E(r). \quad (3.16)$$

We have written  $V(r)$  rather than the explicit form  $-e^2/4\pi\epsilon_0 r$  as a reminder that the results so far are valid for any central potential.

### 3.3.1 General features of the radial solution

Before solving this equation precisely, it is very useful at this stage to use our general knowledge to estimate the approximate form of the solutions. The term  $\hbar^2 l(l+1)/2mr^2$  is rotational kinetic energy, but it plays a role in the radial equation mathematically equivalent to a contribution to radial potential energy. It is called the ‘centrifugal barrier’ since it prevents the system reaching  $r = 0$  when it has non-zero angular momentum. Combining the centrifugal barrier with the Coulomb attraction, the net ‘potential energy’ is shown in figures 3.1 and 3.2 for  $l = 0, 1, 2, 3$ . For each value of  $l$ , we see that an infinite number of solutions is possible, with increasing energy and increasing numbers of nodes in the wavefunction, see figure 3.3. The energy curves in figure 3.1 imply that the wavefunctions with larger  $l$  will have more energy, and will be associated with a larger mean value of  $r$ , i.e. the electron is on average further away from the proton. These general features will be true no matter what the detailed form of  $V(r)$ , (as long as it roughly resembles a  $1/r$  curve), but for the case where  $V(r)$  varies exactly as  $1/r$  there is a set of coincidences: the  $k$ 'th member of the set of energy levels for a given value of  $l$  (counting up from the one of least energy) has exactly the same energy as the  $k + 1$ 'th member of the set associated with the next smaller angular momentum,  $l - 1$ . This fact is illustrated by the horizontal lines in figure 3.2 which show the energy level values for hydrogen.

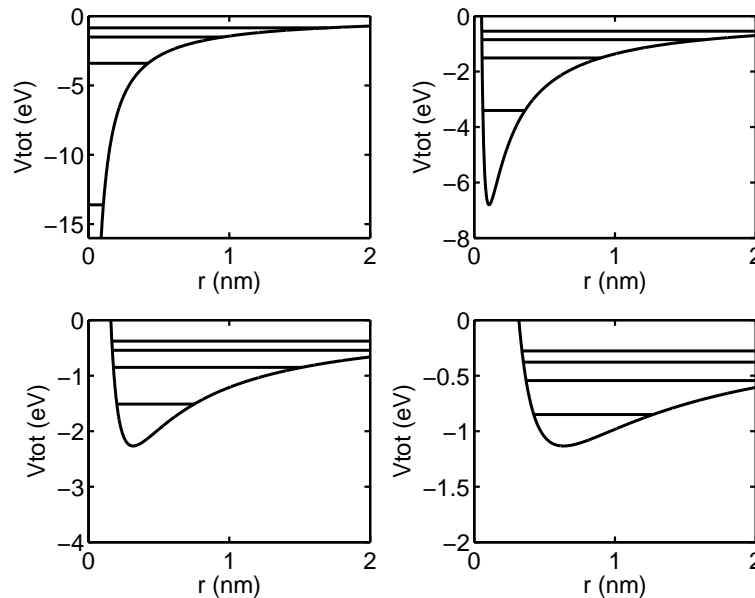


Figure 3.1: The full curves show the sum of the Coulomb energy and the centrifugal barrier appearing in the radial Schrödinger equation for hydrogen, for  $l = 0, 1, 2, 3$  (top left, top right, bottom left, bottom right respectively). The horizontal lines show the lowest four energy levels in each case. Note that the vertical scales are different.

### 3.3.2 Precise solution

To gain some more precise information, we next use a standard trick, which is to write  $R_E(r) = P_E(r)/r$ . Note that  $R_E(r)$  is the radial part of the wavefunction, while  $P_E(r)$  is simply a related

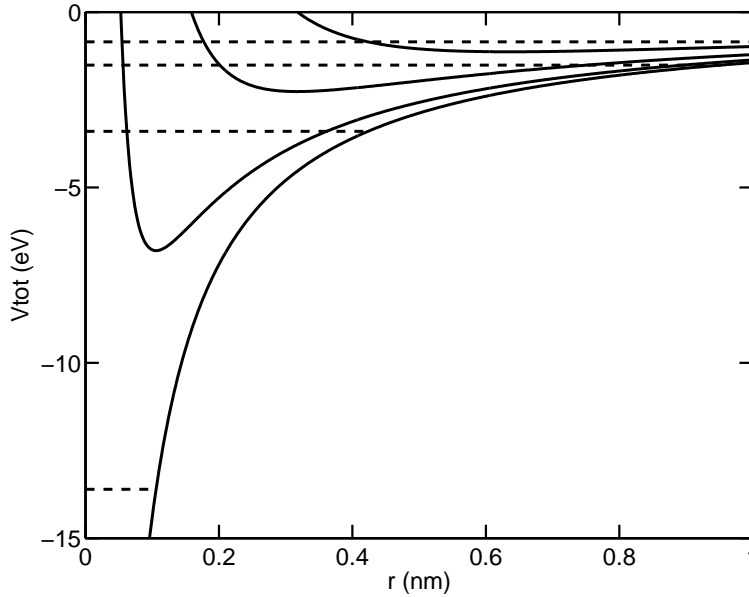


Figure 3.2: All the curves of figure 3.1 plotted on a single graph. For hydrogen ( $1/r$  potential) it is found that energy levels of different  $l$  exactly equal one another, so for the  $n$ 'th level (starting from  $n = 1$  as the lowest) there are  $n$  different radial wavefunctions all of the same energy.

function which we have introduced for convenience<sup>1</sup>. Substituting it in, we find

$$\frac{-\hbar^2}{2\mu} \frac{d^2 P_E}{dr^2} + \left( V(r) + \frac{l(l+1)\hbar^2}{2\mu r^2} \right) P_E = E P_E. \quad (3.17)$$

Now let's guess the solutions. Since  $V(r)$  is only linear in  $1/r$ , the so-called "centrifugal barrier" term  $l(l+1)/r^2$  will dominate when  $r \rightarrow 0$ . Trying  $P \sim r^\beta$  we obtain  $\beta(\beta-1) = l(l+1)$  which has solutions  $\beta = l+1$  or  $(-l)$ . Rejecting the latter as unphysical (tends to infinity at origin), we find

$$R_E(r) \propto r^l \text{ as } r \rightarrow 0. \quad (3.18)$$

This is already a useful observation: we see that the states of least angular momentum,  $l = 0$ , can be finite at the origin (the wavefunction is proportional to  $r^0 = 1$ ) while states of higher and higher angular momentum are flatter and flatter at the origin. This is important because *the origin is where the nucleus is*: we see that states of low angular momentum are more sensitive to the details of the nucleus, such as the fact that it is not really a point particle.

As  $r \rightarrow \infty$  the potential and the centrifugal term fall off, so the kinetic energy dominates, and the solution is one which falls off exponentially as  $\exp(-\kappa r)$  where  $\kappa^2 \hbar^2 / 2\mu = -E$  (this is a positive quantity since  $E$  is negative for bound states). The equation can be solved by a further change of function, followed by the method of series solution (see appendix). The energy eigenvalues are found to be

$$E = -\frac{\mu e^4}{(4\pi\epsilon_0)^2 2\hbar^2} \frac{1}{(k+l+1)^2} \quad (3.19)$$

and the associated wavefunctions are also obtained. They should be looked up in the textbooks. The quantum number  $k$  is an integer in the range  $k = 0, 1, 2, \dots$ , which counts through the set of energy

<sup>1</sup>Well,  $P$  doesn't have a direct physical interpretation, but  $|P|^2$  does. It should be obvious to you, after a little thought, that  $|P|^2 = r^2 |R|^2$  is interesting because it is (when correctly normalised) the probability density for finding the electron at  $r$ , i.e. anywhere in the spherical shell of radius  $r$ , thickness  $dr$ .

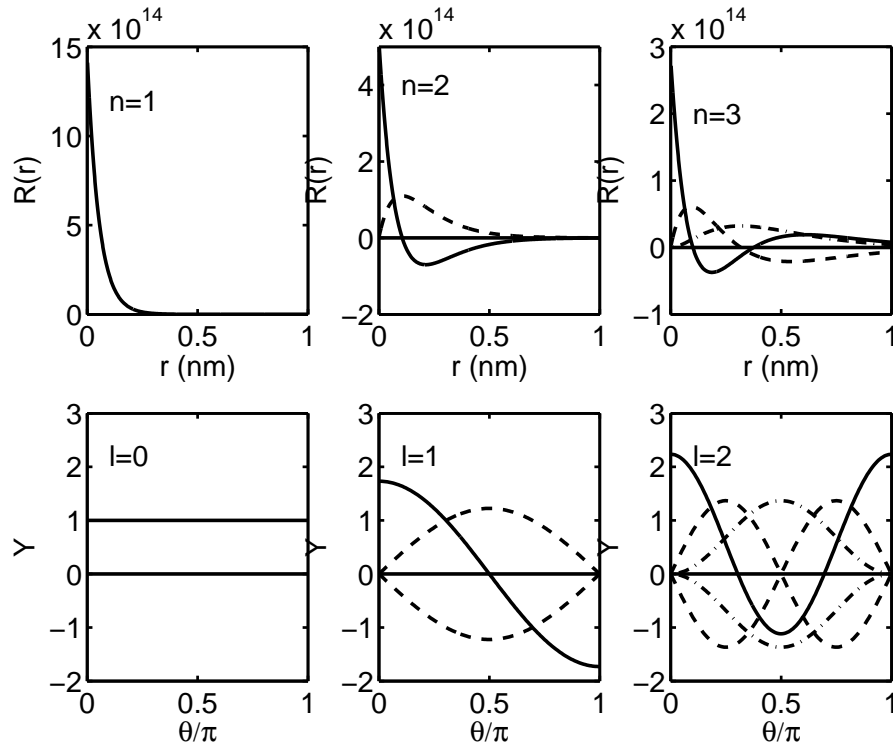


Figure 3.3: Wavefunctions for hydrogen. The top three graphs show the radial part  $R(r)$  for  $n = 1, 2, 3$  and all values of  $l$  (i.e.  $l = 0 \dots n - 1$ ). The line type indicates the value of  $l$ :  $l = 0, 1, 2$  for full curve, dashed, dash-dot, respectively. The bottom three graphs show the  $\theta$  dependence of  $Y_{l,m}(\theta, \phi)$  for  $l = 0, 1, 2$  and all values of  $m$  (i.e.  $m = -l \dots l$ ). The line type indicates the value of  $m$ :  $m = 0, \pm 1, \pm 2$  for full curve, dashed, dash-dot, respectively. The normalization is such that  $|Y|^2$  integrates to  $4\pi$  over all solid angles.

levels of given  $l$ . We met it above when discussing the approximate expected form of the wavefunctions.  $k$  is equal to the number of nodes (zero crossings) in the radial part of the wavefunction.

The wavefunction of the ground state is worth remembering, it is:

$$\psi(r, \theta, \phi) = Ae^{-r/a_0} \quad (3.20)$$

where  $A$  is a normalisation constant which need not be memorized since it can be easily calculated.  $a_0$  is the Bohr radius. Note that the left hand side of the expression emphasizes that the complete wavefunction over all space has been written down, in spherical polar coordinates. It so happens that this ground state wavefunction does not have any  $\theta$  and  $\phi$  dependence: it is spherically symmetric, but other states are not necessarily spherically symmetric (those with  $l > 0$  are not).

There are two important simplifications of equation (3.19). First we notice that the energies are all given by a constant divided by an integer squared. The integer  $k+l+1$  is called the *principal quantum number* and is indicated by the symbol  $n$ . We can now drop  $k$  since it is given by  $n$  and  $l$ , and we note that  $n$  and  $l$  have the ranges:

$$n = 1, 2, 3, \dots \infty \quad (3.21)$$

$$l = 0, 1, 2 \dots n - 1 \quad (3.22)$$

Both  $n$  and  $l$  arise for any central potential, not just the  $1/r$  potential, and  $n - l - 1$  is always the number of nodes in the radial wavefunction, but for potentials with some other dependence than  $1/r$  the energy depends on both  $n$  and  $l$ . This case arises for atoms other than hydrogen.

The next simplification of equation (3.19) is to express the constant in a more physically meaningful way. A useful insight is gained by considering the *fine structure constant*  $\alpha$  defined by

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \quad (3.23)$$

where  $c$  is the speed of light. Consider the dimensions of this quantity:  $e^2/4\pi\epsilon_0$  has the dimensions of an energy multiplied by distance, so  $e^2/4\pi\epsilon_0 c$  has dimensions of energy multiplied by time (Joule-seconds in SI units), therefore  $\alpha$  is dimensionless. The fine structure constant is absolutely central to the quantum physics of electromagnetic interactions: it is a combination of fundamental constants which compares electromagnetic quantities with Planck's constant. The value of  $\alpha$  is  $7.29735 \times 10^{-3}$ , and it is easy to remember this as  $\alpha \simeq 1/137$  which is accurate to one part in 4000.

You can show (see problems) and it is easy to remember that the mean speed (strictly, the r.m.s. speed) of the electron in the ground state of hydrogen is equal to  $\alpha c$ . Furthermore, the potential energy is equal to minus twice the kinetic energy (this is a consequence of the virial theorem) and therefore the total energy equals minus the kinetic energy:

$$E = -\frac{1}{2}\mu\alpha^2 c^2 \quad (3.24)$$

for the ground state. You should confirm that this agrees with equation (3.19), that is, it is just a convenient way of expressing the answer obtained by solving the Schrödinger equation. We will see later (chapter 5) that the expression in terms of  $\alpha c$  arises naturally out of a relativistically correct treatment (i.e. one valid for speeds of order  $c$ ).

Finally, then, all the energy levels of hydrogen are given by

$$E_n = -\frac{1}{2}\mu\alpha^2 c^2 \frac{1}{n^2}. \quad (3.25)$$

The quantity

$$R_\infty = \frac{1}{2} \frac{m_e \alpha^2 c^2}{\hbar c} = 10\,973\,731.568\,549(83) \text{ m}^{-1} \quad (3.26)$$

is called the Rydberg, where  $m_e$  is the mass of a free electron ( $\sim 9.10939 \times 10^{-31}$  kg). Note that by convention the Rydberg is defined in terms of the full electron mass, not the reduced mass. The dimensions of the Rydberg are inverse distance, so it is directly related to wavenumber (inverse of wavelength) measurements. The ionisation energy of the hydrogen ground state is not exactly 1 Rydberg multiplied by  $\hbar c$  but slightly less than this (even before relativistic corrections are taken into account) owing to the reduced mass.

The ground state wavefunction is found to be

$$\psi_1(r, \theta, \phi) = \frac{1}{(\pi a_0^3)^{1/2}} e^{-r/a_0} \quad (3.27)$$

where the constant  $a_0 = 4\pi\epsilon_0\hbar^2/\mu e^2$  is called the Bohr radius. The Bohr radius is important because it gives the size of the hydrogen atom—that is, we know the wavefunction does not have a sharp cut-off, but  $a_0$  gives the distance scale over which the wavefunction is falling off as  $r$  increases. It is useful to remember  $a_0$  in terms of  $\alpha$ , which can be done by noting that an electron moving at speed  $\alpha c$  at a distance of  $a_0$  from some point has angular momentum  $\hbar$  about that point, hence  $a_0\mu\alpha c = \hbar$ , from which

$$a_0 = \frac{\hbar}{\mu\alpha c} \simeq 0.5 \text{ \AA} \quad (3.28)$$

### Comment

What is the meaning of  $n$ ? Out of the smoothness of the Schrödinger equation an integer, a discrete thing, has emerged. Why? Because out of all the continuous infinity of candidate solutions to the



$$\begin{aligned}
\langle r^k \rangle &= \int_0^\infty \int_0^\pi \int_0^{2\pi} \psi(r, \theta, \phi)^* r^k \psi(r, \theta, \phi) r^2 \sin \theta dr d\theta d\phi = \int_0^\infty R(r)^* r^k R(r) 4\pi r^2 dr \\
\langle r^2 \rangle &= (5n^2 + 1 - 3l(l+1)) n^2 a_0^2 / 2 \\
\langle r \rangle &= (3n^2 - l(l+1)) a_0 / 2 \\
\langle r^{-1} \rangle &= 1 / (n^2 a_0) \\
\langle r^{-2} \rangle &= ((l+1/2)n^3 a_0^2)^{-1} \\
\langle r^{-3} \rangle &= (l(l+1/2)(l+1)n^3 a_0^3)^{-1}
\end{aligned}$$

Table 3.1: Expectation values of some powers of  $r$  for hydrogen energy eigenstates.

Schrödinger equation, only a discrete set behave correctly as  $r \rightarrow 0$  and  $r \rightarrow \infty$ . The wavefunction is like a standing wave which has to have a specific form at these two ‘boundaries’, and in between it can oscillate up and down an integer number of times. To be specific, it passes through zero  $n - (l + 1)$  times. The reason we choose to call  $n$  the ‘principal quantum number’ rather than  $n - (l + 1)$  is to draw attention to the fact that the energy eigenvalue depends primarily on  $n$ . Indeed, for the  $1/r$  potential, by a mere chance (actually by a hidden symmetry) the energy depends only on  $n$  and not on  $l$  at all. However,  $n$  is primarily associated with the eigenfunctions, rather than the eigenvalues, and an equivalent quantum number will appear whenever we have a central potential, of whatever dependence on  $r$ . The eigenvalues will then still be related to  $n$  and  $l$ , but not by a simple formula.

### 3.3.3 Mean radius

It will be valuable later to have some information about the mean value of the radial coordinate, and of some of its powers. This is given in table 3.1. The expressions for the ground state are easy to derive and are treated in the problems. Notice the result  $\langle 1/r \rangle = 1/a_0$  for the ground state, and notice that the main features of the variation with  $n$  and  $l$  are as expected from the shape of the energy curves in figure 3.2 and the wavefunctions in figure 3.3.

### 3.3.4 How to remember hydrogen

The ground state Bohr orbit has	$v/c = \alpha$
The ground state energy is	$E_1 = -T_1$
	$= -\frac{1}{2}m\alpha^2 c^2$
an electron orbiting at speed $\alpha c$ at the Bohr radius	
has one unit of angular momentum, so	$mva_0 = \hbar$
from which	$a_0 = \hbar/(m\alpha c)$
and more generally	$\langle r \rangle \propto n^2$

### 3.3.5 Main points

The main points of this chapter have been:

reduced mass; Schrödinger equation for a central potential;  $RY_{l,m}$ ;  $n, l, m$  quantum numbers and their meaning;  $E_R = -\frac{1}{2}m_e\alpha^2 c^2$ ; Bohr radius; the form of the wavefunctions; hydrogen-like systems and scaling with  $Z$ ; the semi-classical picture.

### 3.3.6 Appendix on series solution of hydrogen equation, off syllabus

This calculation is “off syllabus” in the sense that it would not normally be set as an atomic physics question in finals, but as a part of mathematical methods in physics it is a standard example of the series solution of a differential equation, which you should know how to handle.

The radial Schrödinger equation for a  $1/r$  potential will be solved, using previous knowledge of the angular part of the solution. This is a long calculation, and the detailed working between steps will involve further algebra which is not set down here. In other words, there is no *quick* derivation of the familiar  $E \sim 1/n^2$  answer. The derivation will nevertheless illuminate some important features of the solution wavefunctions, and the meaning of the quantum number  $n$ .

With the clue that the solution falls off exponentially at large  $r$ , adopt the dimensionless variable

$$\begin{aligned} \rho &= \sqrt{-2mE} r / \hbar \quad [\text{N.B. } E < 0 \text{ so the sqrt is real}] \\ \text{and let } P_E &= e^{-\rho} u \end{aligned} \tag{3.29}$$

$$\Rightarrow 0 = \frac{d^2 u}{d\rho^2} - 2 \frac{du}{d\rho} + \left[ \left( \frac{2m\alpha^2 c^2}{-E} \right)^{1/2} \frac{1}{\rho} - \frac{l(l+1)}{\rho^2} \right] u \quad [\text{using } V(r) = e^2/4\pi\epsilon_0 r]$$

$$\text{Try } u = \rho^{l+1} \sum_{k=0}^{\infty} C_k \rho^k \quad [\text{remember } u \sim r^{l+1} \text{ at } r \rightarrow 0]$$

$$\Rightarrow \frac{C_{k+1}}{C_k} = \frac{-(-2m\alpha^2 c^2/E)^{1/2} + 2(k+l+1)}{(k+l+2)(k+l+1) - l(l+1)} \quad [\text{after some work!}]$$

$$\begin{aligned} \Rightarrow E &= -\frac{\frac{1}{2}m\alpha^2 c^2}{(k+l+1)^2} \quad [\text{for the series to terminate}] \\ &= -\frac{E_1}{n^2} \quad [\text{Hydrogen energy levels}] \end{aligned}$$

$$\text{where } n = k+l+1, \quad k = 0, 1, 2, \dots; \quad l = 0, 1, 2, \dots$$

(3.30)

## Chapter 4

# Hydrogen-like systems and spectra

### ‘Spectroscopic notation’

A spherically symmetric quantum system always has energy eigenstates of definite total orbital angular momentum, so the quantum number  $l$  arises. For historical reasons, this quantum number  $l$  is associated with a letter, according to the following table:

$l =$	0	1	2	3	4	5	...
	s	p	d	f	g	h	...

(the letters arose from observations on spectra: ‘s’ for sharp, ‘p’ for principal, ‘d’ for diffuse, ‘f’ for fundamental (for some reason!), then g,h,...)

Using this notation the radial part of the wavefunctions of hydrogen can be indicated by giving the values of  $n$  and  $l$  in the form shown by the table:

$n$	$l$	notation	radial wavefunction
1	0	1s	$(\pi a_0^3)^{-1/2} \exp(-r/a_0)$
2	0	2s	$(8\pi a_0^3)^{-1/2} (1 - r/2a_0) \exp(-r/2a_0)$
2	1	2p	$(8\pi a_0^3)^{-1/2} (r/2\sqrt{3}a_0) \exp(-r/2a_0)$
3	0	3s	etc.
3	1	3p	... look them up in Leighton/Woodgate/...!
3	2	3d	

etc.

In this table, the radial part of the wavefunction has been normalised such that  $\int_0^\infty R^* R 4\pi r^2 dr = 1$ , but note that other ways of normalising are possible, depending on how the normalization of the full wavefunction  $R(r)Y(\theta, \phi)$  is apportioned between the radial and angular parts—a matter of convention and convenience.

## 4.1 Hydrogen-like systems

Our understanding of hydrogen can be applied almost unchanged to any system involving a pair of oppositely charged particles bound to each other. We simply adjust for the charge and mass of the particles. Any atom can be converted into a hydrogen-like ion by removing all but the last electron. Then if the atomic number is  $Z$  we have a nucleus of charge  $Z$  and the reduced mass of the electron is slightly changed since the nucleus is now heavier. We can immediately solve the problem without

significant further work by looking carefully at the Schrödinger equation

$$\frac{-\hbar^2}{2\mu}\nabla^2\psi_E - \frac{Ze^2}{4\pi\epsilon_0 r}\psi_E = E\psi_E, \quad (4.1)$$

and reasoning as follows: the only change compared to hydrogen is the  $Z$  which multiplies  $e^2/4\pi\epsilon_0$  in the equation (the fact that the reduced mass will not be the same as for hydrogen is already implicitly allowed for since  $\mu$  in the equation is the reduced mass). When we solve for the behaviour, everywhere in the maths where previously  $e^2/4\pi\epsilon_0$  occurred, there will now instead be  $Ze^2/4\pi\epsilon_0$ . It follows that we can just write down the answer, using the rule that wherever you see  $e^2$ , now put  $Ze^2$ . When we write the solution in terms of the fine structure constant  $\alpha$  for convenience, it means that wherever previously we had  $\alpha$  we now have  $Z\alpha$ . Here is what we find:

$$\begin{array}{ll} \text{The ground state Bohr orbit has} & v/c = Z\alpha \\ \text{The ground state energy is} & E_1 = -T_1 \\ & = -\frac{1}{2}\mu Z^2\alpha^2 c^2 \\ \text{from the expression for } a_0, \text{ we deduce} & \langle r \rangle \propto n^2/Z\mu \end{array}$$

Notice that these results make physical sense: a highly charged nucleus pulls the electron in closer to it, and it can move at higher speeds while remaining bound.

The case of deuterium has  $Z = 1$  but the mass of the nucleus is approximately twice that of hydrogen [Problem: how many times precisely?: Look it up in a book of physical constants], therefore the reduced mass of the electron changes, and all the energy levels are slightly shifted. Such a difference between the energy levels of different isotopes of the same element is called an *isotope shift*<sup>1</sup>.

Two other interesting cases are positronium, and muonic atoms. Positronium is the system of an electron and a positron orbiting one another. The two particles will annihilate each other after about a microsecond (or 10 nanoseconds, depending on the state, see Leighton p. 624) but up until that time they form a hydrogen-like system with the important property that both particles are truly point-like: there is no internal structure, in contrast to the quarks inside the proton in hydrogen. This means the properties can be calculated even more precisely than for hydrogen, and one can test in particular the theory of the Coulomb interaction down to very small distances, small compared to the size of a proton, by very precise measurements of the spectrum.

If atoms are placed in a beam of muons from a particle accelerator, occasionally an atomic electron will be knocked out and a muon captured. The muon is 206.8 times heavier than an electron, so its wavefunctions have a much smaller Bohr radius. It soon decays to the ground state and then owing to the small Bohr radius it sits inside the majority of the charge distribution provided by the electrons. Therefore it experiences almost the whole nuclear charge and its energies are approximated by the hydrogen formula with  $Z\alpha$  where  $Z$  is almost the full charge of the nucleus, and  $\mu$  is now the reduced muon mass. The only thing is, the muon orbit can be so small that it spends most of its time right inside the nucleus! Therefore the whole calculation needs to be reworked, we can no-longer assume a point-like nucleus and measurements on muonic atoms yield much information on nuclear properties.

Finally, another important hydrogen-like system is antihydrogen (!): a positron bound to an anti-proton. Accurate comparisons between the spectra of hydrogen and antihydrogen are just becoming feasible in laboratory experiments, and needless to say they give direct evidence of the nature of the fundamental symmetries of particles and their anti-particles. Antihydrogen, being the anti-atom which is easiest to fabricate, will also provide the first experimental test to ensure that gravity acts the same on particles and anti-particles. [Problem: why can't we test that using anti-protons or positrons which are more readily available from particle accelerators?]

<sup>1</sup>Isotopes of a given element have the same number of electrons and protons, but differ in the number of neutrons in the nucleus.

## 4.2 Spectroscopy

The main experimental techniques to investigate the structure of free atoms are emission and absorption spectroscopy. In absorption spectroscopy, light from a source with a continuous spectrum over a wide enough range is directed through a vapour cell containing a vapour of the atoms to be investigated. The spectrum of the transmitted light is measured, and found to contain a continuous background interrupted by narrow absorption lines at those wavelengths which the atoms can absorb. In emission spectroscopy atoms in a vapour are excited, typically by means of electron bombardment in a discharge, and the spectrum of the emitted light is measured. It is found to consist of a set of narrow lines.

The workhorse instrument for examining the gross structure of the spectrum is the diffraction grating. This is especially so for emission spectroscopy, since many of the wavelengths for neutral atoms are in the range near-ultra-violet to near-infra-red, which can be investigated using gratings. In absorption spectroscopy, wavelengths are generally shorter (for a simple reason we will examine in a moment) and diffraction gratings with sufficient reflectivity may not be available. A quartz prism spectrometer can be used in the ultra-violet region.

The details of emission and absorption spectroscopy are well described in two experiments on the practical course: I recommend obtaining the manuscripts and doing the experiments. The main points to consider in any form of spectroscopy are resolution, i.e. the smallest separation between spectral features which can be resolved by the instrument, and signal to noise ratio, in this case determined by the overall transmission from source to detector.

I will not describe the grating spectrometer in detail here. However, it is poor practice to try to learn physics without some idea of how the experimental methods work, so I will outline some of the considerations, to show a selection of the issues which arise. I will then quote without proof a fundamental result concerning the resolution of any optical spectroscopic instrument.

### 4.2.1 Main points for use of grating spectrograph

A diagram was given in the lecture. The main considerations for correct use of the instrument are

- Is the grating of high quality? It needs to be flat to less than a wavelength (e.g.  $\lambda/10$ ) of the light being examined, and have good reflectivity at the wavelengths of interest.
- Is full use made of the light available from the source? A condenser lens is used to focus the light onto the entrance slit. Note, however, that too narrow a focus will not help since the beam after passing through the slit expands again, and much of it will miss the collimation lens altogether if the focus was too tight.
- Is full use of the width of the grating made? The whole width of the grating must be illuminated to obtain the best resolution. The condenser lens should be chosen appropriately to bring this about.
- Is the light properly collimated at the grating? The collimation lens must be carefully adjusted to one focal length from the slit; this can be done by observing the transmitted light through a telescope which is known to be accurately focussed at infinity.
- Is a sharp image of the slit formed in the detection plane? The imaging lens is adjusted while the slit image is observed.
- Is the slit as narrow as it usefully can be? The minimum useful slit size is determined by the imperfect sharpness of the image owing to diffraction effects associated with the finite size of the diffraction grating and collimation/imaging optics. Below this limit, further reduction of slit size merely causes the image to be less bright.

- Is stray light reaching the detector reduced to a minimum?

### 4.2.2 Resolution

The resolution of all optical spectroscopy instruments can be understood by one simple result:

*The minimum resolvable wavenumber difference is equal to the inverse of the longest path difference over which interference occurs in the instrument.*

(Wavenumber is one over wavelength,  $\tilde{\nu} = 1/\lambda$ ).

For example, for a diffraction grating of width  $w$  used at normal incidence, if one of the diffracted orders for some wavelength  $\lambda$  is found to be at angle  $\theta$ , then the longest path difference for interference is  $w \sin \theta$ , therefore the minimum resolvable wavenumber difference is

$$\delta\tilde{\nu} = \frac{1}{w \sin \theta} \quad (4.2)$$

Combining this with the fact that the  $n$ 'th order diffraction occurs at angle  $\theta$  given by  $d \sin \theta = n\lambda$  for a grating of spacing  $d$  we have

$$\delta\tilde{\nu} = \frac{d}{wn\lambda} \quad (4.3)$$

$$\Rightarrow \frac{\delta\tilde{\nu}}{\tilde{\nu}} = \frac{1}{nN} \quad (4.4)$$

where  $N = w/d$  is the number of lines in the grating.

High-quality diffraction gratings are readily available with widths of a few centimetres, and really good ones go up to around 20 cm. Taking  $\sin \theta$  of order 0.5, equation (4.2) gives  $\delta\tilde{\nu} \simeq 0.1 \text{ cm}^{-1}$  for the best gratings. Note that it is common to discuss atomic spectra in wavenumber units, rather than wavelength or frequency. This is because wavenumber is proportional to energy, so useful for getting insight into the energy level structure, and because it is not always good practice to convert it into energy or frequency, because the value of Planck's constant may not be as accurately known as the measured wavenumbers, therefore such a conversion would involve a large loss in precision, or would require different texts to make clear what value for Planck's constant was being used.

### 4.2.3 Usefulness of both emission and absorption methods

Emission spectroscopy can be used to examine all the transitions of any atom (within technical limits), so it might be asked whether absorption spectroscopy adds anything. Both are useful because the spectra of atoms (and of anything else more complicated than an atom) are extremely complex and hard to interpret. One is faced with a plethora of lines and a lot of ingenuity is required to make sense of the data. The great boon of the absorption method is that a simpler spectrum is obtained: if the atoms are at typical temperatures (below a few thousand degrees C) then they will almost all be in the ground state, therefore only transitions from the ground state are observed. This enables the ground state transitions to be identified in the emission spectrum, which is the first step towards

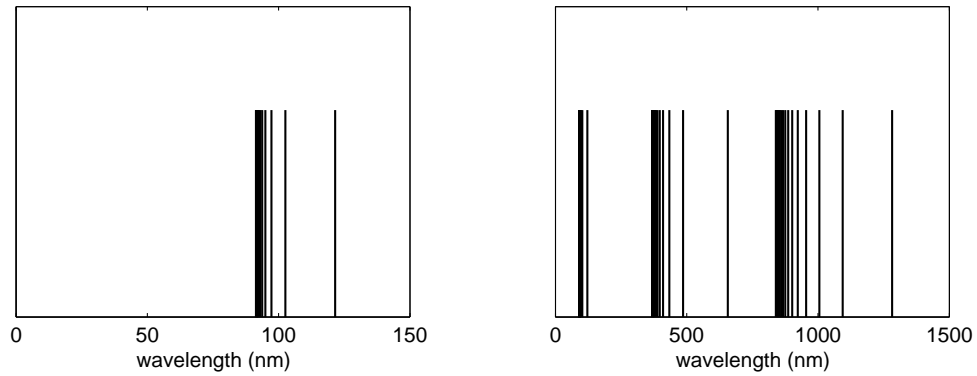


Figure 4.1: **Left:** The complete absorption spectrum of hydrogen. Only the positions of the spectral lines are shown, not their relative intensities. There are many lines close together near the series limit at 91 nm, but no further lines of wavelength shorter than this or longer than 122 nm. **Right:** Partial emission spectrum of hydrogen. Only the positions of the spectral lines are shown, not their relative intensities. The group near 100 nm are at the same wavelengths as the absorption spectrum. There are further lines at longer wavelengths, not shown.

interpreting it.

### 4.3 The spectrum for hydrogen

For hydrogen the energy level formula  $E_n = -E_1/n^2$  enables us to calculate the transition frequencies: for a transition from  $n = j$  to  $n = k$  the frequency is

$$\nu = \frac{|E_k - E_j|}{h} = \frac{\mu}{m_e} R_\infty c \left| \frac{1}{k^2} - \frac{1}{j^2} \right|.$$

The absorption spectrum is shown in figure 4.1: it consists of a sequence of lines of increasing frequency (decreasing wavelength), the longest wavelength is approximately 122 nm (make sure you confirm this by calculation, to be sure you know where it comes from), the lines get closer and closer at shorter wavelengths, up to a ‘series limit’ at approximately 91 nm (confirm this yourself too). [Problem:  $91/122 \simeq 0.75$ . Is the ratio between the series limit and the longest wavelength exactly  $3/4$ ? If not then with what degree of precision do you expect it to agree approximately with  $3/4$ ? (Hint: consider whether the Schrödinger equation we have been using is exact).]

The emission spectrum has lines at the same wavelengths as the absorption spectrum, and also further lines at the frequencies of all the other transitions. These occur as a set of identifiable groups or series, each being the set of transitions to a given lower level. The first few series are named after atomic physicists who first studied them, Lyman ( $n = 1$ ), Balmer ( $n = 2$ ), Paschen ( $n = 3$ ), Brackett ( $n = 4$ ) and Pfund ( $n = 5$ ). Only the Balmer and Paschen series were first detected experimentally. After Bohr introduced his model, he predicted the presence of the others; their subsequent experimental observation was an important success of the model. The series limit of each group enables the energy of the common lower level to be deduced immediately. Similar series occur in more complicated atoms, where this observation is a useful pointer to aid the interpretation process.

The transition  $n = 3$  to  $n = 2$  is the beautiful red “Balmer-alpha” line of wavelength 656 nm.

[Problem. Light from a distant star has a sequence of absorption lines at 134.3 nm, 113.3 nm, 107.5 nm, 104.9 nm, and further lines getting closer and closer as they approach a limit near 101 nm. It is argued that these lines are due to absorption by hydrogen gas in the atmosphere of the star. Is this explanation consistent with the spectrum of hydrogen? If so, how fast is the star moving relative to us?]

### What is going on when atoms emit and absorb light

We will discuss the coupling between atoms and light more thoroughly towards the end of the course. For now, I would simply like to give some guides to your physical intuition.

When an atom (or any other physical system) is in an energy eigenstate, it is in a certain important sense a *static* system. It is correct to picture the wavefunction  $\psi$  as oscillating in time, and the electron de Broglie waves as moving from place to place, but in an energy eigenstate all this wave motion always has the overall result that the probability distribution  $|\psi|^2$  is completely static. Therefore when you think about an atom in one of its energy eigenstates, if you want to contemplate the net charge distribution within the atom, you should imagine something completely fixed, as unmoving as if it were set in concrete.

When an atom emits or absorbs light, on the other hand, it is temporarily in a superposition of energy eigenstates and the charge distribution oscillates. The oscillation continues for as long as it takes the atom to complete the absorption or emission process. The oscillation frequency is the frequency of the emitted or absorbed light.

Another interesting point is that when it comes to absorption and emission, a given atom is a completely different ‘beast’, depending on what state it is in. For example, hydrogen in the ground state has resonant frequencies at 2466 THz, 2923 THz, 3083 THz, etc., and can only absorb, while hydrogen in the 4s state has resonant frequencies at 3083 THz, 617 THz, 160 THz at which it can emit, and 74 THz, 114 THz, etc. at which it can absorb.

### Main points

The main points of this chapter have been: spectroscopic notation  $s, p, d$  etc.; hydrogen-like systems and  $Z^2$  scaling; isotope shift from reduced mass; basic features of experimental method for emission and absorption spectroscopy; use of and resolution of the diffraction grating spectrometer; the hydrogen spectrum; physical picture of absorption and emission process.



## Chapter 5

# Introduction to fine structure and spin

### 5.1 Experimental observation of fine structure

When the emission spectrum of an element such as sodium or potassium is observed, using a standard grating spectrometer, then careful examination of the emission lines reveals that most of them are not a single line at a single wavelength, but a pair of wavelengths close together, a so-called *doublet*. Other elements produce groups containing more lines close together, triplets, quadruplets, etc. When the energy separation within each group is small compared to the energy separation between one group and another, then this phenomenon is called *fine structure*.

It is found that the hydrogen emission spectrum has complicated fine structure, the emission lines appearing as groups of several lines together. In hydrogen the fine structure is especially small and almost undetectable by grating spectrographs, higher resolution instruments such as Fabry-Perot etalons have to be used. Many of the groups of lines have themselves a pattern of two separate sub-groups, so that when the spectrum is slightly blurred the lines merge to form two ‘lumps’, therefore to first approximation we have a doublet. For an example see figure 5.1 which shows the fine structure of the Balmer-alpha ( $n = 2 \rightarrow 3$ ) line. The fine structure is approximately ten-thousand times smaller than the gross structure. For example for the Balmer-alpha line the observed splitting is

10 GHz    in frequency units  
0.3  $\text{cm}^{-1}$     in wavenumber units  
40  $\mu\text{eV}$

None of this structure was predicted by our treatment of hydrogen up till now. However, it should not come as a surprise that our treatment of hydrogen so far has ‘only’ been accurate to relative precision  $10^{-4}$  because we have seen that the speed of the electron is of order  $\alpha c \simeq 10^{-2}c$ . Corrections from special relativity always appear as a power series in  $(v/c)^2$  so we can expect the predictions so far to be inaccurate by of order a part in  $10^4$ .

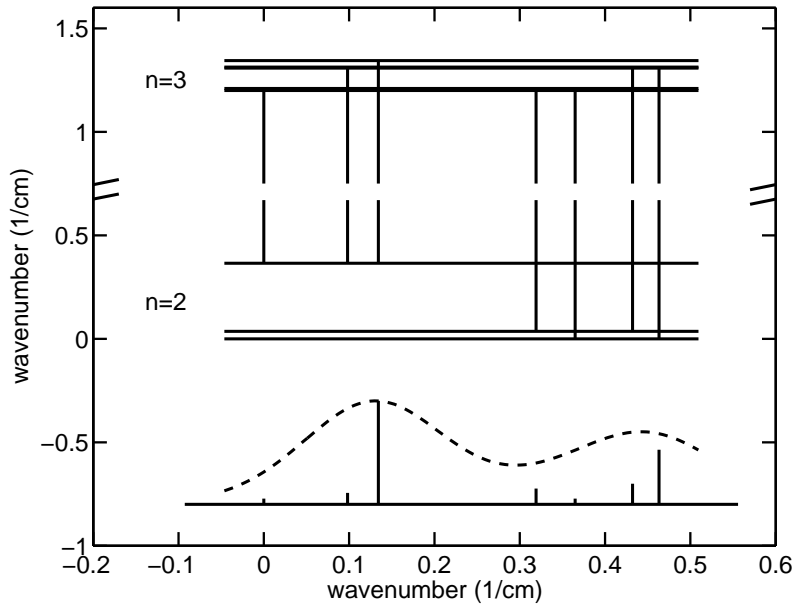


Figure 5.1: Fine structure of the  $n = 2 \rightarrow n = 3$  Balmer alpha line of hydrogen. The energy level splittings and the observed emission spectrum are shown in wavenumber units ( $\text{cm}^{-1}$ ). The zeros of the scales have been set arbitrarily at the lowest energy level of those shown, and at the smallest wavenumber of the transitions in this group. The top group of 5 levels (two pairs and a singleton) belongs to the  $n = 3$  configuration, the bottom group of 3 levels belongs to  $n = 2$ ; the break in the vertical lines is to show that the gross structure splitting between  $n = 3$  and  $n = 2$  ( $1/656 \times 10^{-7} \simeq 15000 \text{ cm}^{-1}$ ) has been suppressed. The dashed line shows the effect of blurring of the spectrum by the Doppler shift from random motion of the atoms in a gas of hydrogen at room temperature. Then the detailed structure is smoothed away and only a pair of lines (a doublet) is resolved.

In many branches of physics a precision better than one in a thousand would already be regarded as very good, but here at the fundamentals we have to do better than that. Furthermore, our treatment of other hydrogen-like systems such as highly charged ions has been less accurate still, because the electron moves quicker when it is near a more highly-charged nucleus. For hydrogen-like Uranium,  ${}_{92}\text{U}^{91+}$ , our treatment is essentially unusable because the electron is moving at speeds of order half the speed of light. [Problem (easy): use the hydrogen results, including their scaling with  $Z$ , to confirm this statement.]

## 5.2 The Dirac result

To cope with relativity (to be technically precise, to cope with Lorentz covariance) properly we have to reject the Schrödinger equation altogether, and replace it with an equation which behaves the right way under changes of reference frame. Such an equation was discovered by Dirac. It had all the right properties, only it had the surprising feature that instead of treating one wavefunction at a time, it used a vector of four wavefunctions to describe a single particle! The genius of Dirac was not only to discover the equation, but also to reason that the four-wavefunction vector described the two spin states of the electron, and predicted a previously unknown form of the electron, which we now know as its anti-particle, the positron.

The Dirac equation for two point-particles interacting by a  $1/r$  potential can be solved exactly, if we assume one particle is fixed (i.e. it has infinite mass). The answer for nuclear charge  $Z = 1$  is

$$E_{nj} = mc^2 \left[ 1 + \alpha^2 \left( \frac{1}{n - (j + \frac{1}{2}) + [(j + \frac{1}{2})^2 - \alpha^2]^{1/2}} \right)^2 \right]^{-1/2} \quad (5.1)$$

where  $j$  is the total angular momentum quantum number ( $\hat{\mathbf{j}} = \hat{\mathbf{l}} + \hat{\mathbf{s}}$ ) for the electron. The combination of orbital and spin angular momenta will be discussed shortly. To allow for a nucleus of charge  $Z$ , just replace  $\alpha$  by  $Z\alpha$ . Note that this energy includes the rest energy  $mc^2$ , so if we want to recover the Schrödinger prediction we should expand the square root in powers of  $\alpha$ , and then take off the  $mc^2$  contribution. [Problem: confirm that you then get the hydrogen energy  $-(1/2)m\alpha^2c^2/n^2$  to 2nd order in  $\alpha$ ].

The Dirac equation itself is beyond this level of the course, but you should be aware that it is the right relativistic equation, and you should notice two things about the prediction for the energy levels:

1. The formula contains  $\alpha$  only in the form  $\alpha^2$ , so when expanded in a power series, the next term after the hydrogen gross structure  $-(1/2)m\alpha^2c^2$  will be of order  $\alpha^4$ . Allowing for a nucleus of charge  $Z$ , this becomes order  $Z^4\alpha^4$ .
2. The formula depends on  $j$ , and therefore predicts slightly different energies for two states of the same  $n$  but different  $j$ . This difference is the *fine structure*.

I can't resist adding that even the Dirac equation is slightly off: in 1947 Lamb and Retherford<sup>1</sup> (not Rutherford), in a triumph of experimental physics, precisely measured a small splitting between states of different  $l$  but the same  $j$ , which is not predicted by Dirac on its own. The splitting is now called the Lamb shift (another Nobel prize here). This greatly stimulated the discovery of Quantum Electrodynamics which accounts for this further effect and pretty much everything else about electrons and photons.

### Dirac summary

The summary of the Dirac equation prediction, up to order  $Z^4\alpha^4$ , is

power of $Z\alpha$	i.e.	coefficient	interpretation	
0	1	$mc^2$	rest energy	(5.2)
2	$Z^2\alpha^2$	$-mc^2/2n^2$	gross structure	
4	$Z^4\alpha^4$	$+\frac{mc^2}{8n^4}(3 - \frac{4n}{j+1/2})$	fine structure	

**This power series gives an easy way to remember that the fine structure in hydrogen-like ions scales as  $Z^4$ .**

## 5.3 Schrödinger method to account for fine structure

The fine structure of hydrogen and of other atoms can be calculated in an approximate way by treating the relativistic effects as small corrections to the basic Schrödinger equation which we have used up till now, equation (3.12). This approach has the advantage that it gives some good physical insights, and for atoms other than hydrogen the Dirac equation quickly leads to intractable calculational problems, while the method of using Schrödinger's equation with changes in the Hamiltonian can be calculated with greater success. The method is an example of an important general technique called *perturbation theory* which will be the subject of the next two chapters. We shall then examine its application to the fine structure of hydrogen in chapter 8.

<sup>1</sup>W. E. Lamb and R. C. Retherford, Physical Review **72**, 241(1947).

We have already seen that the Dirac result refers to the total angular momentum of the electron, obtained from the sum of its orbital and spin angular momenta. These were introduced in the previous part of the course, but I shall comment further on them here as a revision.

## 5.4 Physical nature of orbital and spin angular momenta

Orbital angular momentum and spin angular momentum have some similarities but also some important differences. Most of the similarities are in the mathematics of the *operators* which describe them, and their associated eigenvalues. The most clear differences are in the physical embodiment of these two types of angular momentum, and the associated quantum *states*.

### Similarities

1. Both are forms of angular momentum.
2. Both are described by operators which obey the set of commutation rules for angular momentum operators:

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z \quad (5.3)$$

for  $\hat{\mathbf{J}}$  composed of entirely orbital or entirely spin angular momentum, or of a combination of the two types.

3. The eigenvalues of the operator  $\hat{J}^2$  are of the form  $J(J+1)\hbar^2$  independent of whether orbital or spin angular momentum is involved, and the eigenvalues of  $\hat{J}_z$  are  $M\hbar$  where the possible values of  $M$  are  $-J, -J+1, \dots, J-1, J$ .
4. Total angular momentum is conserved in an isolated physical system, irrespective of what type of angular momentum it may be composed of.
5. Many physical interactions do not couple spatial and spin degrees of freedom. If only such interactions are involved, then orbital and spin angular momentum are each individually conserved in isolated physical systems. If coupling between spin and spatial degrees of freedom is present (e.g. spin-orbit interaction in atoms, see later) then the orbital and spin angular momenta may not be conserved individually, but their sum is.

The upshot of all this is that for many purposes we can work with angular momentum without caring which type it is. However, there are important differences also.

### Differences

1. Orbital angular momentum is completely described by and contained in the spatial wavefunction. Spin angular momentum is not: it is a completely new degree of freedom which requires an *extension to the mathematical structure of quantum mechanics* (see below).
2. The orbital angular momentum of a single particle can have any size and any direction, depending on its state of motion; the spin angular momentum of a single particle is fixed in size, and always the same for the same type of particle, though it can have any direction (and for zero rest-mass particles such as photons the spin direction is related to the direction of propagation).
3. Orbital angular momentum always involves a movement of matter, i.e. a non-zero current density; spin angular momentum does not.
4. The quantum number  $L$  associated with purely orbital angular momentum is always an integer; the quantum number  $S$  associated with spin angular momentum can be an integer or a half-integer.
5. The gyromagnetic ratio (ratio of magnetic dipole moment to angular momentum, see later) is 1 (in standard units) for orbital angular momentum, but need not be, and usually is not, 1 for spin angular momentum.

I have written these differences in order of importance, with the most important first. The extension to quantum mechanics which is mentioned in point 1 did not appear in the list of similarities, because it is to do with the states rather than the operators.

The full extension is the one introduced by Dirac, in which an electron is described not by a single wavefunction, but by a set of four, and this treatment can handle relativistic calculations and anti-particles. However, in the limit of kinetic energies small compared to the rest energy, we can ignore pair creation and then it is sufficient to use two wavefunctions to describe a single electron. The most convenient way to think about this pair is to regard it as a two-component vector. Some examples will help to clarify this.

Before we became aware of the spin degree of freedom, we would have described a single electron moving along the  $y$  axis by a wavefunction such as

$$\psi(x, y, z, t) = Ne^{i(ky - \omega t)}.$$

We now need to include the spin state by replacing this with a more complete statement. The spin is described by an operator  $\hat{\mathbf{S}}$ . For a single electron, we know the mathematics of this operator is the same as the mathematics of a certain set of  $2 \times 2$  matrices (the Pauli spin matrices), so the most convenient way to represent the spin state is to use a two-component vector. We *define* the entity

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

to be a mathematical representation of the spin state called ‘spin up’ (i.e. eigenstate of  $\hat{S}_z$  with eigenvalue  $(+1/2)\hbar$ ) and we *define* the entity

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

to be a mathematical representation of the spin state called ‘spin down’ (i.e. eigenstate of  $\hat{S}_z$  with eigenvalue  $(-1/2)\hbar$ ). These definitions will prove to be sensible ones, because the matrix methods turn out to model the observed physical behaviour of electron spin perfectly, and that is all we require of a physical description.

For a spin-half particle, a general spin state can be written as a combination of spin-up and spin-down:

$$a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \quad (5.4)$$

normalized such that  $|a|^2 + |b|^2 = 1$ . For spin greater than one half (e.g. a photon of spin 1, or two electrons together, whose total spin could be 0 or 1) we need spin vectors of more components.

Note that the spin vectors are abstract mathematical entities. The vector space they live in is not part of our everyday space of position and time, but is rather a ‘spin space’ in which only quantum mechanical spin can live and move and have its being. In case this makes you worried that a whole new dimension has been added to the world which you didn’t know was there, you should not worry, because since you are composed of electrons as well as other things, you already live in this spin space as well as in the everyday space which you have already noticed. (Efforts to relate spin to ‘ordinary’ spatial degrees of freedom are active in areas of theoretical physics such as string theory, and it does appear that spin can be regarded as arising from certain subtleties of motion of strings at a minute scale, but then these theories involve plenty of extra dimensions in any case, so the world is still more complicated than you thought.)

Now we need to combine the spatial wavefunction with the spin state. In many situations, a simple product is sufficient (to be precise, this is a tensor product). So if our electron moving along the  $y$  axis has spin state ‘up’ along the  $z$  axis, then the total state is

$$\Psi(x, y, z, \zeta, t) = N e^{i(ky - \omega t)} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (5.5)$$

where the symbol  $\zeta$  simply serves to remind us that the total state depends on a spin ‘coordinate’ as well as on the familiar  $x, y, z$  (we are now using the word ‘coordinate’ in a generalized sense to include spin as well as spatial degrees of freedom).

If the electron’s spin state is along the  $x$  direction, the total state is

$$\Psi_i(x, y, z, \zeta, t) = N e^{i(ky - \omega t)} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (5.6)$$

which is again a simple product. However, we need to be aware that states can also involve a more intimate mixture of spatial and spin degrees of freedom. For example, suppose the electron just described now enters a Stern-Gerlach apparatus oriented with a gradient of magnetic field in the  $z$  direction. Then the spin-up component will experience a force upwards, while the spin-down component experiences a force downwards. Upon emerging from the magnet, the electron is in a superposition of one beam propagating along and upwards, the other propagating along and downwards, where the final motional state depends on the spin state because of the action of the magnet. The result is

$$\Psi_f(x, y, z, \zeta, t) = N \frac{1}{\sqrt{2}} \left[ e^{i(ky + \Delta z - \omega t)} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + e^{i(ky - \Delta z - \omega t)} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] \quad (5.7)$$

where  $\hbar\Delta/m$  is the size of the upward or downward velocity imparted to the electron by the magnet.

Notice that the state of the electron cannot now be written as a product of spatial and spin parts:

$$\Psi_f(x, y, z, \zeta, t) \neq \psi(x, y, z, t) \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{for any } \psi, a, b \quad (5.8)$$

This is a subtle idea, it means we can no longer say what the spin state is, in and of itself, because it is *entangled with* the spatial state, and similarly we cannot say what the spatial state is, in and of itself, because it is entangled with the spin state.

### 5.4.1 Dirac notation

Throughout these notes I am taking pains not to assume the reader is familiar with Dirac notation. However, Dirac notation is powerful and elegant and I recommend it. In Dirac notation we regard the quantum state as a vector in an abstract space called Hilbert space. It does not matter whether the underlying physical observables are of spatial or spin character, since the only observable quantities are overlaps (dot products) between one Dirac ket and another. In Dirac notation a spin state can be written simply by giving the quantum numbers for the total spin and its projection along the quantisation axis, e.g.

$$|s = 1/2, m_s = +1/2\rangle \quad \text{for 'spin up' state of a spin-half particle}$$

$$|s = 1/2, m_s = -1/2\rangle \quad \text{for 'spin down' state of a spin-half particle}$$

Spin-half is sufficiently common that for this special case a short-hand is also adopted:

$$|\uparrow\rangle \equiv |s = 1/2, m_s = +1/2\rangle \quad (5.9)$$

$$|\downarrow\rangle \equiv |s = 1/2, m_s = -1/2\rangle \quad (5.10)$$

Simply as a reminder that spin states other than spin-half are possible, I shall now quote a result which you have met in previous lectures, concerning possible total spin states of a pair of spin-half particles:

$$|S = 0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle_1 |\downarrow\rangle_2 - |\downarrow\rangle_1 |\uparrow\rangle_2) \quad (5.11)$$

$$(5.12)$$

$$|S = 1, M_S = 1\rangle = |\uparrow\rangle_1 |\uparrow\rangle_2$$

$$|S = 1, M_S = 0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle_1 |\downarrow\rangle_2 + |\downarrow\rangle_1 |\uparrow\rangle_2)$$

$$|S = 1, M_S = -1\rangle = |\downarrow\rangle_1 |\downarrow\rangle_2 \quad (5.13)$$

We shall return to this subject when we come to consider a pair of electrons (the helium atom).

### 5.4.2 Main points

The main points of this chapter have been:

the fact that the Dirac equation gives a full relativistic treatment; the formula for the energy levels can be written as a power series in  $Z^2\alpha^2$ ; hence the  $Z^4$  scaling of fine structure; the nature of spin and its mathematical representation.





# Chapter 6

## Perturbation Theory

Perturbation theory is an important and powerful method in physics. It enables us to make progress when a physical system is too complicated to be analysed exactly. The essential idea is to solve the behaviour in steps. First we approximate the system by some simple Hamiltonian whose Schrödinger equation we know how to solve. Then we add the bit missed out, and use perturbation theory to calculate how our previous results (energy eigenfunctions and eigenvalues) must be changed.

### 6.1 First order non-degenerate time-independent perturbation theory

The main result is as follows. If a system has a Hamiltonian of the form

$$H = H^0 + \delta H \tag{6.1}$$

where the effects of  $\delta H$  are small compared with those of  $H^0$ , then we can solve the Schrödinger equation

$$H\psi_k = E_k\psi_k \tag{6.2}$$

by using the form

$$\begin{aligned} \psi_k &= \psi_k^0 + \delta\psi_k \\ E_k &= E_k^0 + \delta E_k \end{aligned}$$

where  $\psi_k^0$  and  $E_k^0$  are the eigenfunctions and eigenvalues of  $H_0$ , i.e.

$$H^0\psi_k^0 = E_k^0\psi_k^0, \tag{6.3}$$

and  $\delta E_k$  and  $\delta\psi_k$  express the difference between the energies and eigenstates of  $H$  and those of  $H^0$ . It is always possible to do this. The notation has the following logic. We use the subscript  $k$  to count the different eigenstates, and the superscript zero is applied to quantities related to the zeroth order Hamiltonian  $H^0$  which is present before the perturbation is applied. Our aim will be to find  $\delta E_k$  and  $\delta\psi_k$ .

Plugging all this into (6.2) we have

$$(H^0 + \delta H)(\psi_k^0 + \delta\psi_k) = (E_k^0 + \delta E_k)(\psi_k^0 + \delta\psi_k)$$

Multiplying out the brackets we obtain:

$$H^0\psi_k^0 + H^0\delta\psi_k + \delta H\psi_k^0 + \delta H\delta\psi_k = E_k^0\psi_k^0 + E_k^0\delta\psi_k + \delta E_k\psi_k^0 + \delta E_k\delta\psi_k \tag{6.4}$$

So far everything we have done is exact, there are no approximations involved.

At this stage we make an approximation: we ignore 2nd order small quantities, and thus obtain

$$\begin{aligned} H^0 \psi_k^0 + H^0 \delta\psi_k + \delta H \psi_k^0 &\simeq E_k^0 \psi_k^0 + E_k^0 \delta\psi_k + \delta E_k \psi_k^0 \\ \Rightarrow H^0 \delta\psi_k + \delta H \psi_k^0 &= E_k^0 \delta\psi_k + \delta E_k \psi_k^0, \end{aligned} \quad (6.5)$$

where we used (6.3) to cancel two terms. Now multiply by  $(\psi_k^0)^*$  and integrate:

$$\begin{aligned} \Rightarrow \langle \psi_k^0 | H^0 | \delta\psi_k \rangle + \langle \psi_k^0 | \delta H | \psi_k^0 \rangle &= \langle \psi_k^0 | E_k^0 | \delta\psi_k \rangle + \langle \psi_k^0 | \delta E_k | \psi_k^0 \rangle \\ \Rightarrow \langle \psi_k^0 | \delta H | \psi_k^0 \rangle &= \delta E_k. \end{aligned} \quad (6.6)$$

The last step cancelled two terms by using the result

$$\langle \psi_k^0 | H^0 | \delta\psi_k \rangle = E_k^0 \langle \psi_k^0 | \delta\psi_k \rangle. \quad (6.7)$$

To prove this, use the fact that  $H^0$  is Hermitian, which implies that  $\langle \psi_k^0 | H^0 | \delta\psi_k \rangle = \langle \delta\psi_k | H^0 | \psi_k^0 \rangle^*$ , and then examine the integral, use (6.3) and to finish recall that the energy  $E_k^0$  is real. The more expert quantum mechanic can see (6.7) immediately by operating  $H^0$  to the left inside the bracket, but if you don't know what I mean by that, then ignore it.

So far we have derived in equation (6.6) the important result that a small perturbing Hamiltonian will shift the energy of the  $k$ 'th state by an amount equal to the expectation value of the perturbation in the original state:

$$\delta E_k^{(1)} = \langle \psi_k^0 | \delta H | \psi_k^0 \rangle. \quad (6.8)$$

I have inserted the superscript (1) to act as a reminder that this result is only valid to first order—the 2nd order small quantities have been neglected.

Next we would like to know the effect on the wavefunction. To find this, return to equation (6.5) and multiply by one of the eigenstates different from  $\psi_k^0$  before integrating:

$$\langle \psi_n^0 | H^0 | \delta\psi_k \rangle + \langle \psi_n^0 | \delta H | \psi_k^0 \rangle = \langle \psi_n^0 | E_k^0 | \delta\psi_k \rangle + \langle \psi_n^0 | \delta E | \psi_k^0 \rangle \quad (6.9)$$

$$\Rightarrow \langle \psi_n^0 | \delta H | \psi_k^0 \rangle = \langle \psi_n^0 | \delta\psi_k \rangle (E_k^0 - E_n^0) \quad (6.10)$$

using the same reasoning as before. Next, use the fact that  $\delta\psi_k$ , the change in wavefunction, can always be written in terms of the zeroth order wavefunctions, since they form a complete set:

$$\delta\psi_k = \sum_n c_n \psi_n^0 \quad \text{where } c_n = \langle \psi_n^0 | \delta\psi_k \rangle.$$

The perturbation theory result (6.10) tells us the coefficients, so we have

$$\delta\psi_k^{(1)} = \sum_{n \neq k} \frac{\langle \psi_n^0 | \delta H | \psi_k^0 \rangle}{E_k^0 - E_n^0} \psi_n^0. \quad (6.11)$$

I have again inserted the superscript (1) to remind us that the result is a first order approximation to the complete  $\delta\psi_k$ . Note that the term  $n = k$  is omitted from this sum: this is ok since any contribution from  $\psi_k^0$  to  $\delta\psi_k$  is taken care of at the end by imposing an overall normalisation of the new state  $\psi_k^0 + \delta\psi_k$  (to first order approximation).

As long as, for the energy level  $k$  we are considering, no other level has the same energy, then all is well. However, if there is degeneracy  $E_n^0 = E_k^0$  then the answer (6.11) blows up. The case of degenerate states will be considered in chapter 7.

### 6.1.1 Simple examples

These were treated in the lecture:

A particle is in a one-dimensional box (infinite square well potential) of width  $a$ . A small potential step is introduced at the centre of the box, raising the potential energy in the right hand half of the box by  $V_0$ . Calculate the energy eigenstates and their energies using first order perturbation theory, and comment on the accuracy of the calculation.

A one-dimensional simple harmonic oscillator has potential energy function  $V = (1/2)m\omega^2x^2$ . A small additional potential energy of the form  $V' = ax^2$  is introduced, where  $a \ll m\omega^2$ . Calculate the energies of the lowest two energy eigenstates using first order perturbation theory, and comment on the accuracy of the calculation.

## 6.2 2nd and higher orders

So far we have obtained the first order correction to the energies, equation (6.8) and the wavefunctions (6.11). We can now do better. Examining the exact equation (6.4), we can now substitute in the values of  $\delta E_k$  and  $\delta\psi_k$ , but in doing so we need to keep in mind that we have not calculated them precisely, we have just got a result which is valid when 2nd order quantities are neglected. In order to make further progress, it is useful to make the notation more powerful, in order to clarify what is approximate and what is not.

To this end, we now write

$$\delta E_k = \delta E_k^{(1)} + \delta E_k^{(2)} + \delta E_k^{(3)} \dots \quad (6.12)$$

$$\delta\psi_k = \delta\psi_k^{(1)} + \delta\psi_k^{(2)} + \delta\psi_k^{(3)} \dots \quad (6.13)$$

The superscript indicates the degree of approximation involved, as follows:  $\delta E_k^{(1)}$  is an approximation to  $\delta E$  which is valid when 2nd and higher order quantities are neglected;  $(\delta E_k^{(1)} + \delta E_k^{(2)})$  is an approximation to  $\delta E$  which is valid when 3rd and higher order quantities are neglected, and so on. A similar statement applies to each term in the sequence for  $\delta\psi_k$ .

The quantities which we have calculated so far, equations (6.8) and (6.11), are  $\delta E_k^{(1)}$  and  $\delta\psi_k^{(1)}$  where the superscript (1) was already introduced with the new notation in mind.

Substituting (6.12) and (6.13) into the exact equation (6.4) gives

$$H_0\psi_k^0 = E_k^0\psi_k^0 \quad (6.14a)$$

$$+H^0\delta\psi_k^{(1)} + \delta H\psi_k^0 = +E_k^0\delta\psi_k^{(1)} + \delta E_k^{(1)}\psi_k^0 \quad (6.14b)$$

$$+H^0\delta\psi_k^{(2)} + \delta H\delta\psi_k^{(1)} = +E_k^0\delta\psi_k^{(2)} + \delta E_k^{(1)}\delta\psi_k^{(1)} + \delta E_k^{(2)}\psi_k^0 \quad (6.14c)$$

$$+\dots = +\dots \quad (6.14d)$$

where I have gathered all the zeroth order terms on line a, all the first order terms on line b, and so on.

We have already seen that the zeroth order terms cancel, and we have expressions for the first order terms. To obtain the 2nd order terms, notice that the first order result is such that line b cancels. Therefore we are just left with line c (in 2nd order we are ignoring the further lines of the equation):

$$H^0 \delta\psi_k^{(2)} + \delta H \psi_k^{(1)} = E_k^0 \delta\psi_k^{(2)} + \delta E_k^{(1)} \delta\psi_k^{(1)} + \delta E_k^{(2)} \psi_k^0. \quad (6.15)$$

Multiply by  $(\psi_k^0)^*$  and integrate, and use the same observation we employed in equation (6.7) to cancel the first term on each side. What remains is

$$\langle \psi_k^0 | \delta H | \delta\psi_k^{(1)} \rangle = \langle \psi_k^0 | \delta E_k^{(1)} | \delta\psi_k^{(1)} \rangle + \langle \psi_k^0 | \delta E_k^{(2)} | \psi_k^0 \rangle. \quad (6.16)$$

Now recall that  $\psi_k^0$  does not appear at all in the expression (6.11) for  $\delta\psi_k^{(1)}$ , and therefore the first term on the right hand side is zero. Hence

$$\delta E_k^{(2)} = \langle \psi_k^0 | \delta H | \delta\psi_k^{(1)} \rangle \quad (6.17)$$

which is nice and memorable. Now substituting in (6.11) we obtain the 2nd order energy correction:

$$\begin{aligned} \delta E_k^{(2)} &= \langle \psi_k^0 | \delta H \left[ \sum_{n \neq k} \frac{\langle \psi_n^0 | \delta H | \psi_k^0 \rangle}{E_k^0 - E_n^0} | \psi_n^0 \rangle \right] \\ &= \sum_{n \neq k} \frac{|\langle \psi_n^0 | \delta H | \psi_k^0 \rangle|^2}{E_k^0 - E_n^0}. \end{aligned} \quad (6.18)$$

You should now be able to see how the method proceeds to all orders. The next step is to get the 2nd order correction to the wavefunctions, then with that information we can get the 3rd order correction to the energies, and so on. Provided the series of terms  $E_k^0 + \delta E_k^{(1)} + \delta E_k^{(2)} + \dots$  converges, then the series sum will be exactly the correct energy  $E_k$  under the perturbed Hamiltonian. In practice, we usually are content to calculate just the first order, or a few low-order terms, and accept the degree of approximation involved. However, it is important to have some idea of what that degree of approximation is.

Here we shall just consider how accurate the first order result is. To this end, reason that when the perturbation is small, then the series converges quickly, so we expect the inaccuracy of the first order approximation to be given mostly by the 2nd order term. Examining (6.18) and (6.8) we find

$$\frac{|\delta E_k^{(2)}|}{|\delta E_k^{(1)}|} \approx \frac{|\delta E_k^{(1)}|}{|E_n^0 - E_k^0|} \quad (6.19)$$

where the symbol  $\approx$  is here used to mean ‘is the same order of magnitude as’. This result says the 2nd order correction is small compared to the 1st order one, when the first order shift is small compared to the energy level spacing of the original (unperturbed) Hamiltonian.

### 6.2.1 Energy levels ‘repel’

We just remarked on the size of the 2nd order correction  $\delta E_k^{(2)}$ . It is useful to note its sign as well. The series in (6.18) is dominated by the terms for which the denominator is small, in other words by the contribution from energy levels close to the one whose shift is being calculated. Since the numerator of each term is always positive, the sign of the contribution is given by the denominator. For a given energy level  $k$ , the 2nd order shift is upwards (positive sign) whenever the other energy level  $n$  has less energy ( $E_n^0 < E_k^0$ ), and downwards whenever the other energy level has more energy ( $E_n^0 > E_k^0$ ). The net effect is always such that levels ‘repel’, i.e. the 2nd order shift tends to push energy levels apart from one another.

For example, the 2nd order contribution to the shift of the ground state energy is always such as to lower it.

## 6.3 Examples in hydrogen ground state

We shall next apply non-degenerate perturbation theory to two problems of interest in the hydrogen atom and hydrogen-like ions. Since at this stage we can only treat the non-degenerate case, we will be limited to calculations for the ground state. The other states of hydrogen will be treated later.

### 6.3.1 Nuclear volume correction

Up till now we have treated the proton at the centre of the hydrogen atom as if it were a point-like particle. In fact it is a system of quarks and gluons with a size of order  $10^{-15}$  m. The details of the complex set of interactions between the electron in hydrogen and the quarks in the nucleus need not concern us, but we can easily consider the main effect: this is that the nuclear charge is spread out over a finite volume of space.

Let us model the nucleus (a single proton for hydrogen, or a system of protons and neutrons for a hydrogen-like ion or an isotope) as a sphere of radius  $R$  having constant charge density smeared out over its volume  $(4/3)\pi R^3$ . The value of  $R$  will be in the region of  $10^{-15}$  m, it can be deduced from scattering experiments, or else we can calculate a formula for the effect on the spectrum of hydrogen, and then use the observed spectrum to deduce  $R$ . We shall do this using perturbation theory.

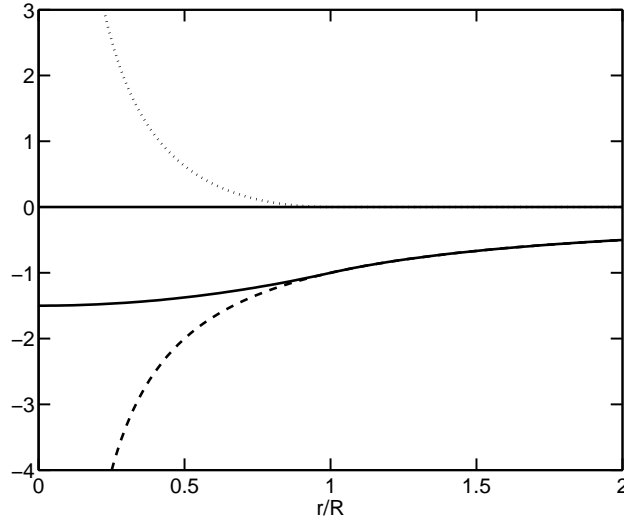


Figure 6.1: Potential energy for volume effect calculation, in units of  $e^2/(4\pi\epsilon_0 R)$ . Full curve: full potential energy  $V$  (eq. (6.23)). Dashed curve: zero order potential energy  $V^0$ . Dotted curve: perturbation  $\delta V = V - V_0$  (eq. (6.25)).

The zeroth order Hamiltonian is the one we have used up till now:

$$H^0 = \frac{\hat{p}^2}{2\mu} - \frac{e^2}{4\pi\epsilon_0 r}. \quad (6.20)$$

To find the perturbing Hamiltonian, we need to do a bit of electrostatics to deduce the electric potential from a sphere of charge. *Outside* the sphere it follows from Gauss's theorem that the electric field is the same as the one from a point charge, leading to potential energy

$$V = -\frac{e^2}{4\pi\epsilon_0 r} \quad \text{for } r \geq R. \quad (6.21)$$

*Inside* the sphere, the electric field is still spherically symmetric and outwards, but now Gauss's theorem tells us it is equal to that of the charge enclosed by a sphere of radius  $r$ , which is a fraction  $(r/R)^3$  of the total amount  $e$ , hence the magnitude of the electric field  $\mathcal{E}$  inside the nucleus is

$$\mathcal{E} = \frac{(r/R)^3 e}{4\pi\epsilon_0 r^2} \quad \text{for } r \leq R. \quad (6.22)$$

This equation says the field increases linearly with  $r$ . Integrating it with respect to  $r$  we find the electric potential is quadratic in  $r$ . The constant of integration must be such that the electric potential is continuous at  $r = R$  (since there cannot be an infinite electric field there), hence the potential energy of the electron when it is inside the nucleus is

$$V = -\frac{e^2}{4\pi\epsilon_0 R} \left( \frac{3}{2} - \frac{r^2}{2R^2} \right) \quad \text{for } r \leq R. \quad (6.23)$$

Equations (6.21) and (6.23) together tell us the correct Hamiltonian for our problem:

$$H = \frac{\hat{p}^2}{2\mu} + V(r). \quad (6.24)$$

Therefore the perturbation is

$$\delta H = H - H^0 = \begin{cases} \frac{e^2}{4\pi\epsilon_0 R} \left( -\frac{3}{2} + \frac{r^2}{2R^2} + \frac{R}{r} \right) & \text{for } r \leq R \\ 0 & \text{for } r \geq R \end{cases} \quad (6.25)$$

This perturbation is shown in figure 6.1. Notice that  $\delta H$  tends to infinity as  $r \rightarrow 0$ , but this does not imply it is a 'large' perturbation, since its *effects* (the change in energy levels and eigenfunctions) will be small. Notice also that we avoided the mistake of thinking equation (6.23) itself is the perturbation, which of course would be quite wrong.

Applying the theory to the ground state of hydrogen, we find the energy shift

$$\begin{aligned}\delta E^{(1)} &= \langle \psi^0 | \delta H | \psi^0 \rangle \\ &= \int_0^R \int_0^\pi \int_0^{2\pi} \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} \frac{e^2}{4\pi\epsilon_0 R} \left( -\frac{3}{2} + \frac{r^2}{2R^2} + \frac{R}{r} \right) \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} r^2 \sin\theta dr d\theta d\phi\end{aligned}$$

using equation (3.27) for the wavefunction and (6.25) for the perturbation. Notice that the radial integral stops at  $r = R$  because the perturbation is zero for  $r > R$ . Our integral looks a little tricky, but as good atomic physicists we should think about the size of the quantities involved. The nuclear radius  $R$  is of order 1 fermi, i.e.  $10^{-15}$  m, but the electron wavefunction size  $a_0$  is of order 1 Angstrom, i.e.  $10^{-10}$  m for hydrogen, therefore to *very* good approximation  $\exp(-r/a_0) \simeq 1$  for all values of  $r$  which appear in the integral. Using this insight, the integral is trivial, and we have

$$\delta E^{(1)} = \frac{4}{10a_0^3} \frac{e^2}{4\pi\epsilon_0} R^2. \quad (6.26)$$

The order of magnitude of this quantity is  $10^{-8}$  eV, so it is very small compared to the zero-order energy separations which are of order a few eV, therefore this 1st order perturbation calculation is very accurate. However, we are about to consider relativistic considerations which will show that in truth the hydrogen atom is not as simple as we have assumed here, so in fact we can only trust the general form of this calculation of the nuclear volume effect. The main point is that the shift is found to be proportional to the mean square nuclear radius, and has a size of order  $(R/a_0)^2$  times the zero-order energies.

Such a small shift is not easy to observe directly, but it can be seen by comparing the spectrum of one isotope with another. However in such a comparison we need to keep in mind that the energy levels are shifted both by the nuclear volume effect and by the change in reduced mass of the electron. If more than two isotopes are available, then the two contributions can be distinguished by taking advantage of their different dependence on the relative atomic mass. Alternatively, we can bring in other information, such as the nuclear masses which can often be accurately measured another way (e.g. by comparing the motional frequencies of ions of the relevant isotopes confined in an ion trap).

### 6.3.2 Relativistic kinetic energy correction

According to special relativity, the kinetic energy of a particle of rest mass  $m$  and momentum  $p$  is

$$\begin{aligned}T &= (p^2 c^2 + m^2 c^4)^{1/2} - mc^2 \quad [\text{relativistic kinetic energy}] \\ &\simeq \frac{p^2}{2m} - \frac{1}{2mc^2} \left( \frac{p^2}{2m} \right)^2 + \dots\end{aligned} \quad (6.27)$$

we shall apply this result to the quantum theory of hydrogen, replacing  $p$  by its operator  $\hat{p}$ . It is, however, questionable whether this replacement is strictly correct, since we are combining a classical relativity result with a quantum non-relativistic equation of motion, whereas we ought to replace the whole equation of motion (and then we would be dealing with the Dirac equation). However, it is found that this approach does reproduce the results of the Dirac equation correctly, and it yields physical insight, so we shall stay with it.

We next consider whether to replace  $m$  in (6.27) by the reduced mass  $\mu$ . For the first term,  $p^2/2m$ , this is the right thing to do since it then reproduces the low-velocity limit correctly. However it should be kept in mind that the dynamics of a two-body system cannot be simplified by the reduced mass concept in the relativistic limit, therefore it is questionable whether the reduced mass should be used in the relativistic corrections. However, when the nuclear mass is large compared to the electron mass (i.e. all cases except for positronium) this is a correction to a correction, and we shall not concern ourselves with it at the level of these notes. We shall therefore quote the small energy shift using the rest mass  $m$  of the electron.

In order to simplify the further calculations, we now use the neat trick of expressing the kinetic energy  $p^2/2m$  as total energy minus potential energy, and hence the perturbing Hamiltonian is

$$\delta H = -\frac{1}{2mc^2}(E_n - V)^2. \quad (6.28)$$

This example is treated in the problems. The result is

$$\delta E_1^{(1)} = -\frac{5}{8}Z^4\alpha^4 mc^2. \quad (6.29)$$

The size of the shift is smaller than the unperturbed energies by a factor of order  $Z^2\alpha^2 \simeq v^2/c^2$ , as we should expect for a relativistic correction. This is a small perturbation for hydrogen, though it is large compared to the nuclear volume effect.



## Chapter 7

# Degenerate perturbation theory

When two or more energy eigenstates are degenerate, then perturbation theory as we have discussed it so far breaks down, because it is impossible for the shift caused by the perturbation to be small compared to the separation between the energy levels before the perturbation was applied. This is found in practice to be a very common situation, so it is important to know how to handle it. This chapter is dedicated to a careful consideration of what we need to do.

### 7.1 A summary

Before going into the reasons why, I am going to tell you at the outset both our problem and its solution.

**Problem:** equation (6.11), and hence equation (6.18) etc. all blow up when  $E_n = E_k$  for one or more terms in the sum.

**Solution:** we arrange that whenever  $E_n^0 = E_k^0$ , then  $\langle \psi_n^0 | \delta H | \psi_k^0 \rangle = 0$ . Then we use the expressions already calculated, dropping the terms whose numerator is zero.

The essence of the solution is as simple as that! What we now need to do is (1) find out how to arrange  $\langle \psi_n^0 | \delta H | \psi_k^0 \rangle = 0$  when we need it, and (2) in so doing, prove that it is always possible to arrange this for the case of degenerate states.

Notice that when the matrix element  $\langle \psi_n^0 | \delta H | \psi_k^0 \rangle$  is zero for the case of degenerate energy levels, then in the sum (6.11) we have a “zero divided by zero” and hence mathematically ill-defined quantity. Therefore we need to return to the previous equation (6.10). This equation says  $0 = 0$  for this case, which is mathematically well-defined (and correct!), but it means we now have no information about the value of  $\langle \psi_n^0 | \delta H | \psi_k^0 \rangle$ . However, by looking into this more fully it is possible to show that whenever  $\langle \psi_n^0 | \delta H | \psi_k^0 \rangle = 0$ , then  $\delta \psi_k$  does not contain any contribution from  $\psi_n^0$ , so it is correct simply to omit the term from the sum.

The mathematical procedure which arranges that  $\langle \psi_n^0 | \delta H | \psi_k^0 \rangle = 0$  (for  $n \neq k$ ) is called ‘diagonalising the perturbing Hamiltonian’. This phrase refers to the matrix which represents the perturbing Hamiltonian in the basis provided by the zeroth order states. We are arranging that off-diagonal elements of this matrix are zero, hence the phrase. However, notice that we don’t in fact need all the off-diagonal elements to be zero, and in practice they won’t be all zero: only the ones for which the states were degenerate before the perturbation is applied are made zero by the ‘diagonalization’ (which is therefore, strictly speaking, only a partial diagonalization).

### 7.1.1 Degenerate sub-space

To get to grips with the concepts here, we need to think hard about energy eigenstates in the presence of degeneracy. Let  $E_k^0$  be a set of energy levels, and  $\psi_k^0$  the associated energy eigenfunctions (all normalized and mutually orthogonal), of some Hamiltonian  $H^0$ . Consider the state

$$u = a\psi_1^0 + b\psi_2^0 \quad (7.1)$$

This is a perfectly allowable wavefunction as long  $|a|^2 + |b|^2 = 1$ , which we assume. The crucial question is:

is  $u$  an eigenstate of energy?

You must be *absolutely clear in your mind about this!* The answer is:

*if  $\psi_1^0$  and  $\psi_2^0$  are degenerate, then yes, otherwise no.*

(*Proof:* operate  $H^0$  onto  $u$  and see what you get.)

So, in the case of degeneracy we have an interesting situation. Not only  $\psi_1^0$  and  $\psi_2^0$ , but also *any* linear combination of them, are all eigenfunctions of  $H^0$ . This is an infinite number of states, all having the same energy. The better way to think about this is to see that an arbitrary linear combination of two orthogonal quantum states is enough to span a complete ‘space’ of states, just as two orthogonal vectors span a two-dimensional vector space. Therefore, rather than thinking in terms of an infinite number of states, it is more insightful to think of the degenerate states as a two-dimensional space of states.

More generally, if some number  $m$  of mutually orthogonal states are all degenerate, then they form an  $m$ -dimensional space of degenerate states.

This observation is crucial to degenerate perturbation theory, because it enables us to see that there is some flexibility in how we define the set of zero-order states which are used in the perturbation expansion. For example, if for some perturbation  $\langle \psi_1^0 | \delta H | \psi_2^0 \rangle \neq 0$ , which will cause trouble in the perturbation expansion when  $E_1^0 = E_2^0$ , then we can always try some other pair of *mutually orthogonal states in the degenerate space*. For example, try

$$\psi_+^0 = (\psi_1^0 + \psi_2^0)/\sqrt{2} \quad (7.2)$$

$$\psi_-^0 = (\psi_1^0 - \psi_2^0)/\sqrt{2}. \quad (7.3)$$

These are equally good candidates to be used as the zero-order states in the perturbation theory calculation, because they are eigenstates of the zero order Hamiltonian  $H^0$ , and they are orthogonal to each other and to all the other eigenstates of  $H^0$ . If we find  $\langle \psi_+^0 | \delta H | \psi_-^0 \rangle = 0$  then our work is done: the perturbation is ‘diagonal in the new set of states’ and we can use the standard perturbation theory results.

In the example just given, we got lucky: the states  $\psi_+^0$  and  $\psi_-^0$  were simply found by a guess. In general, it can be shown that it is *always* possible to diagonalize a Hamiltonian in a given space. The set of mutually orthogonal states (the diagonal *basis*) which achieves this will not usually be as simple as  $\psi_\pm$ , but it can be found. The proof relies on some matrix methods which I won't go into here, please consult the text books if you are interested. It is essentially the standard method to diagonalize a matrix (find the eigenvectors, form a matrix  $U$  out of them, and then use  $U$  to change basis).

### 7.1.2 The general method

We can now see how the method works in general:

*Degenerate perturbation theory proceeds exactly as non-degenerate perturbation theory, as long as we ensure that the zero-order states satisfy  $\langle \psi_n^0 | \delta H | \psi_k^0 \rangle = 0$  whenever  $E_n^0 = E_k^0$ .*

When  $\langle \psi_n^0 | \delta H | \psi_k^0 \rangle = 0$  we say the perturbation  $\delta H$  does not “connect” the two states. The degenerate states can always be manipulated into a form which ensures this. The manipulation is referred to as “diagonalizing the perturbing Hamiltonian”, and can be achieved by an automatic though somewhat involved mathematical procedure. However, we can almost always avoid the full complexities of that procedure, by using methods which carry more physical insight. The methods to be tried first are

- make use of some symmetry of the problem, which enables one to guess the appropriate basis
- identify *constants of the motion*

## 7.2 Constants of the motion

The phrase ‘constant of the motion’ can be used in either classical or quantum mechanics, it refers to some physical property which does not change in time, although the physical system in question may be changing in time in other ways. For example, for a ball thrown across a tennis court, the horizontal component of momentum is a constant of the motion (in the absence of air resistance) while the vertical component of momentum, and height, and the horizontal position, are not. For a thrown gyroscope (or a spinning tennis ball) the angular momentum as well as the horizontal velocity is a constant of the motion.

In the quantum context, a constant of the motion is a physical observable  $\hat{Q}$  such as position, momentum, angular momentum (orbital or spin), etc. whose value does not change in time for the system under consideration. That is, if the observable were to be measured, the set of values which could be obtained, and their probability distribution, are independent of the time at which the measurement is made. You should know by now that if an observable  $\hat{Q}$  does not evolve in time, then its operator must commute with the Hamiltonian. But this in turn implies that the eigenstates of  $\hat{Q}$  can simultaneously be eigenstates of the Hamiltonian and hence that any state of definite energy can also have a definite value of  $\hat{Q}$ , so the eigenvalues of  $\hat{Q}$  are *good quantum numbers*.

term	meaning	mathematical implication
constant of the motion	quantity which does not vary in time	the operator commutes with the Hamiltonian: $[\hat{Q}, \hat{H}] = 0$
good quantum number	quantum number associated with an operator which is a constant of the motion	eigenstate of $\hat{Q}$ is simultaneously an eigenstate of $\hat{H}$

These concepts are useful in perturbation theory when we can identify an observable which is a constant of the motion under the original Hamiltonian  $H^0$  and also under the perturbed Hamiltonian  $H^0 + \delta H$ . Suppose  $\hat{Q}$  is such an observable. Then we know two useful things:

- the zeroth order energy eigenstates can be (they don't *have to be*, but they *can be*) eigenstates of  $\hat{Q}$ :

$$\hat{Q}\psi_k^0 = q_k\psi_k^0 \quad \text{for energy eigenstates } \psi_k^0$$

- $\hat{Q}$  commutes with the perturbation:

$$[\hat{Q}, \delta H] = 0 \quad \Rightarrow \quad \hat{Q}\delta H = \delta H\hat{Q}$$

From these we can prove the crucial condition  $\langle \psi_n^0 | \delta H | \psi_k^0 \rangle = 0$  for zero-order energy eigenstates  $\psi_n^0$  as follows:

$$\begin{aligned} \langle \psi_n^0 | \delta H | \psi_k^0 \rangle &= \frac{1}{q_k} \langle \psi_n^0 | \delta H q_k | \psi_k^0 \rangle \\ &= \frac{1}{q_k} \langle \psi_n^0 | \delta H \hat{Q} | \psi_k^0 \rangle \quad [\text{since } \psi_k^0 \text{ is eigenstate of } \hat{Q} \text{ (as well as of } H^0)] \\ &= \frac{1}{q_k} \langle \psi_n^0 | \hat{Q} \delta H | \psi_k^0 \rangle \quad [\text{since } \hat{Q} \text{ commutes with } \delta H] \\ &= \frac{1}{q_k} \langle \psi_n^0 | q_n \delta H | \psi_k^0 \rangle \quad [\text{Hermitian } \hat{Q} \text{ and } \psi_n^0 \text{ eigenstate of } \hat{Q}] \\ &= \frac{q_n}{q_k} \langle \psi_n^0 | \delta H | \psi_k^0 \rangle \\ \Rightarrow \quad \langle \psi_n^0 | \delta H | \psi_k^0 \rangle \left( 1 - \frac{q_n}{q_k} \right) &= 0 \end{aligned}$$

therefore<sup>1</sup>

$$\text{either } \langle \psi_n^0 | \delta H | \psi_k^0 \rangle = 0 \quad \text{or} \quad q_n = q_k$$

i.e. the perturbation does not connect states of different  $\hat{Q}$ .

---

<sup>1</sup>For a slicker proof, just consider  $0 = \langle \psi_n^0 | [\hat{Q}, \delta H] | \psi_k^0 \rangle = \langle \psi_n^0 | \delta H | \psi_k^0 \rangle (q_n - q_k)$ .

The method to apply degenerate perturbation theory by means of constants of the motion is as follows.

1. Find an observable which is a constant of the motion under both the original and the perturbing Hamiltonian. Call this observable  $\hat{Q}$ .
2. We know it is possible to find states which are eigenstates both of the zero-order Hamiltonian and of  $\hat{Q}$ , so find them.
3. As long as we use these zero-order states, the standard results of perturbation theory will apply without problem (as proved above).

Sometimes the states identified in step (2) still have the problem of simultaneous degeneracy with respect to both energy and  $\hat{Q}$ , in which case either a further constant of the motion must be found or else the brute force ‘diagonalisation’ procedure (which I have not described) has to be resorted to. Assuming this does not happen (it will not for the cases we meet in this introduction to atomic physics), the only subtlety is that the states to be identified in step (2) might not be exactly the same as the states you were used to for the physical system under consideration. For example, in the hydrogen atom, we found so far that the energy eigenstates are a product of a radial part  $R(r)$ , a spin part  $\chi(\zeta)$  (not treated in detail yet, but we know it is there), and an angular part in the form of a spherical harmonic  $Y_{l,m}(\theta, \phi)$ . However, once we include relativistic corrections to the Hamiltonian, it will emerge that we should not use zeroth order eigenstates in the form  $R(r)Y_{l,m}(\theta, \phi)\chi(\zeta)$ . Instead, we will have to combine together states of different  $m$  and spin direction, but of the same  $n$  and  $l$ . This is an example of the coupling of two angular momenta (orbital and spin in this example). The case of coupled angular momentum is very prevalent in atomic physics. I will finish this chapter by discussing it in detail, first for classical mechanics, and then with the main quantum ideas.

### 7.3 Example: charged gyroscopes

Suppose we have two classical gyroscopes in a space station, so that they are in free fall and we can ignore gravity. The gyroscopes are spinning fast, and their axes are not parallel. Each gyroscope will have a fixed orientation in space because the angular momentum about any axis in space does not change in time (conservation of angular momentum). Now suppose that each gyroscope has a small lump of electric charge placed somewhere on the flywheel—think of an insulating plastic patch charged by a few electrons for example. The gyroscopes will now influence each other by Coulomb repulsion, and the exact form of this influence as a function of position and time is very complicated because the gyroscopes are rotating. However, imagine that the charge really is very small, so the influence of the gyroscopes on each other is tiny compared with the size of each one’s rotation about its own axis. What will happen? We can make three predictions, two of them certain and exact, the other approximate but correct to a high degree of accuracy.

What is certain is that

1. *Something* will happen—the motion will change in some way.
2. The total angular momentum of both gyroscopes together will not change. This is certain because internal forces (i.e. internal to the system consisting of the pair of gyroscopes) cannot produce a net torque on the whole system.

A very reasonable (and in fact quite accurate) approximation is that the speed of rotation of each gyroscope about its own axis will not change. I am not going to prove this here, [Problem: (hard) find a way to treat the small extra torques which arise in this scenario and see if you can prove the claim], but I claim that it is physically reasonable. What happens is that both gyroscopes begin to precess slowly, and they do so in such a way that the total angular momentum of the whole system stays constant. That means they precess *about the direction of the total angular momentum*. Think about it!—draw a vector diagram of the two individual angular momentum vectors adding up to make the total angular momentum, and figure out a way to have the two vectors move with constant length but changing direction in such a way that their sum does not change.

In mathematical terms, the size of the two classical angular momenta  $l_1, l_2$  does not change, while their changing direction means their individual  $z$  components fluctuate in time. Meanwhile the total angular momentum vector  $\mathbf{L} = \mathbf{l}_1 + \mathbf{l}_2$  is fixed, i.e. fixed in both size and direction.  $\mathbf{l}_1$  and  $\mathbf{l}_2$  precess about  $\mathbf{L}$ .

### 7.3.1 Quantum gyroscopes

Now we will analyse the same problem using quantum mechanics. In order to have a reasonable picture, assume the gyroscopes are very small and rotating very fast, and the charge is just one electron's worth—in fact our two gyroscopes will be simply two electrons in an atom! This is a classic example of degenerate perturbation theory. Before we allow for the influence of the electron-electron repulsion, the energy eigenstates  $\psi_k^0$  are angular momentum eigenstates having quantum numbers  $l_1, m_{l1}, l_2, m_{l2}$  where the labels 1, 2 label the two electrons. The fact that these are eigenstates of  $z$ -component of angular momentum  $\hat{l}_{z1}, \hat{l}_{z2}$ , as well as of total angular momentum squared for each electron, indicates that the gyroscopes have a fixed orientation in space. Next we allow for the perturbation caused by their Coulomb repulsion. We find that the energy eigenstates  $\psi_k$  of the complete Hamiltonian  $H^0 + \delta H$  are still, to good approximation, eigenstates of  $\hat{\mathbf{l}}_1^2, \hat{\mathbf{l}}_2^2$ , but they are *no longer* eigenstates of  $\hat{l}_{z1}, \hat{l}_{z2}$ . Therefore the numbers  $m_{l1}, m_{l2}$  are not good quantum numbers any more. This expresses the fact that the  $z$  component of angular momentum of each electron now wanders up and down due to precession (see diagram in lecture).

The total angular momentum of the pair of electrons is represented by the operator  $\hat{\mathbf{L}} = \hat{\mathbf{l}}_1 + \hat{\mathbf{l}}_2$ . Let  $L$  and  $M_L$  be the quantum numbers associated with eigenstates of the total angular momentum operators  $\hat{L}^2$  and  $\hat{L}_z$ .

Allowing for the Coulomb repulsion between the electrons, the energy eigenstates  $\psi_k$  now combine states of the same  $l_i$  but different  $m_{li}$ , so as to produce states of definite  $L$  and  $M_L$ . This corresponds to the statement that the total angular momentum  $\hat{\mathbf{L}}$  remains constant under the perturbation, having both a constant size and a constant direction.

Actually, and this illustrates the importance of the idea of the degenerate space of states, we could have used states of definite  $L$  and  $M_L$  all along. That is, instead of picking  $\psi_k^0$  to be a state given by quantum numbers  $l_1, m_{l1}, l_2, m_{l2}$ , we could have chosen it to be a state of definite  $l_1, l_2, L, M_L$ , because that would have been an eigenstate of  $H^0$  as well. This is the trick of degenerate perturbation theory: we have some choice about exactly which form the zero order states have, because we can take groups of degenerate states and combine together the members within each group. If we can do this in such a way as to obtain degenerate states which are **not connected to one another by the perturbation**, then our work is done since the blowing-up problem of equation (6.11) is avoided, and we simply use the standard results (6.8), (6.11) for the energy shift and change in wavefunction.

In this example the perturbation can't connect states of different  $L, M_L$ , so these are the states which the system naturally adopts as its energy eigenstates (recall the behaviour of the 2D harmonic oscillator and of the vibrating drum skin) and each acquires a well-defined energy shift. The energy shift is given by

$$\delta E = \left\langle \{q.\}, L, M_L \left| \delta H \right| \{q.\}, L, M_L \right\rangle \quad (7.4)$$

where  $\{q.\}$  is the list of all the other relevant quantum numbers, which in this case is  $n_1, n_2, l_1, l_2$  but N.B. not  $m_{l1}, m_{l2}$ .

Some complete worked examples of degenerate perturbation theory will be provided in the next chapter. The appendix below gives an example problem for which an exact solution is possible, so that the perturbation result can be confirmed.

## 7.4 Appendix: An exact analysis of a simple example

To get some insight into perturbation theory in the degenerate case, we will consider an example where we can work out the exact answer without using perturbation theory. The example we will examine is a 2-dimensional harmonic potential well  $V_0 = (1/2)m\omega^2(x^2 + y^2)$ , with a perturbation to be described in a moment.

The situation before the perturbation is applied can be analyzed by separation of variables, giving solutions of the form  $\psi_0(x, y) = u_0(x)v_0(y)$ . The energy levels of the 1-dimensional Schrödinger equations are  $E_{x,m} = (m+1/2)\hbar\omega$  and  $E_{y,n} = (n+1/2)\hbar\omega$ , so the total energy is  $E_{m,n} = (m+n+1)\hbar\omega$ . All the energy levels apart from the ground state are degenerate. For example the two different states

$$f^0(x, y) = Axe^{-x^2/2a^2} e^{-y^2/2a^2}, \quad g^0(x, y) = Aye^{-x^2/2a^2} e^{-y^2/2a^2} \quad (7.5)$$

where  $a = (\hbar/m\omega)^{1/2}$  and  $A = \sqrt{2/\pi a^4}$  both have energy  $E_{1,0} = E_{0,1} = 2\hbar\omega$ .

Suppose now a perturbing potential of the form

$$V' = \epsilon(x + y)^2$$

is applied to the system (for example, if the system is an electron in a trap, this perturbation could be caused by an electric field along the diagonal  $x = y$  and increasing linearly with distance). It is quite easy to see what must happen: the potential energy is no longer circularly symmetric, the well now has an elliptical shape, being more tightly confining along the diagonal  $x = y$ .

In order to analyze the problem, it is convenient to introduce new coordinates  $x' = (x + y)/\sqrt{2}$ ,  $y' = (x - y)/\sqrt{2}$ , so the total potential energy can be written

$$V_0 + V' = \frac{1}{2}m\omega^2(x^2 + y^2) + \epsilon(x + y)^2 = \left(\frac{1}{2}m\omega^2 + 2\epsilon\right)x'^2 + \frac{1}{2}m\omega^2y'^2$$

Define  $\omega'$  by putting  $(1/2)m\omega^2 + 2\epsilon = (1/2)m\omega'^2$  and we obtain the exact solution by separation into two one-dimensional problems as before, in the new coordinate system  $x', y'$ . The energy levels are now  $E = (m + 1/2)\hbar\omega' + (n + 1/2)\hbar\omega$  and the energy eigenstates are of the form  $\psi = u(x + y)v(x - y)$ .

The first significant point is that the new choice of coordinates was open to us all along: we could just as readily have written the eigenstates of the original unperturbed Hamiltonian in this coordinate system. We would have obtained for the first excited states, for example,

$$\psi_+^0(x, y) = A \frac{x+y}{\sqrt{2}} e^{-(x+y)^2/4a^2} e^{-(x-y)^2/4a^2} = \frac{1}{\sqrt{2}}(f^0(x, y) + g^0(x, y)) \quad (7.6)$$

$$\psi_-^0(x, y) = A \frac{x-y}{\sqrt{2}} e^{-(x+y)^2/4a^2} e^{-(x-y)^2/4a^2} = \frac{1}{\sqrt{2}}(f^0(x, y) - g^0(x, y)) \quad (7.7)$$

where the reader should check the final equality by expanding the squared terms in the exponents.

Now let us examine how perturbation theory deals with this example. If we picked the original way of writing the states,  $f^0(x, y)$  and  $g^0(x, y)$ , then we would have an infinity in equation (6.11) and therefore the original approximation in going from (6.4) to (6.5) has broken down. However, if we pick the equally valid zero-order states  $\psi_+^0(x, y)$  and  $\psi_-^0(x, y)$  then in equation (6.10) the left hand side, which is here  $\langle \psi_+^0 | \delta H | \psi_-^0 \rangle$ , is equal to zero, which is consistent with the right hand side, so everything is mutually consistent and correct, and the first order perturbation theory results are valid. To prove  $\langle \psi_+^0 | \delta H | \psi_-^0 \rangle = 0$  simply examine the integral:

$$\langle \psi_+^0 | \delta H | \psi_-^0 \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A^* x' e^{-x'^2/2a^2} e^{-y'^2/2a^2} \epsilon x'^2 A y' e^{-x'^2/2a^2} e^{-y'^2/2a^2} dx' dy' = 0. \quad (7.8)$$

To confirm the perturbation theory value for the energy shifts, evaluate

$$\delta E_+ = \langle \psi_+^0 | \epsilon x'^2 | \psi_+^0 \rangle = \frac{3\epsilon}{m\omega} \quad (7.9)$$

$$\delta E_- = \langle \psi_-^0 | \epsilon x'^2 | \psi_-^0 \rangle = \frac{\epsilon}{m\omega} \quad (7.10)$$

and compare with the exact expression expanded to first order in the perturbation size  $\epsilon$ :

$$E_+ = \frac{3}{2}\hbar\omega' + \frac{1}{2}\hbar\omega \simeq 2\hbar\omega + \frac{3\epsilon}{m\omega} \quad (7.11)$$

$$E_- = \frac{1}{2}\hbar\omega' + \frac{3}{2}\hbar\omega \simeq 2\hbar\omega + \frac{\epsilon}{m\omega}. \quad (7.12)$$



## Chapter 8

# Hydrogen fine structure according to Schrödinger

We have already commented that the Schrödinger equation is not quite perfect when it comes to understanding relativistic effects, but nevertheless we can use it by treating the relativistic effects as small perturbations (ok for hydrogen, but not ok for hydrogen-like uranium). This gives a lot of physical insight which we can then apply to other atoms.

The relativistic corrections naturally divide into 3 parts: relativistic kinetic energy correction (sometimes called a relativistic ‘mass correction’); ‘spin–orbit interaction’; and ‘Darwin term’.

### 8.1 Relativistic kinetic energy correction

This was discussed in section 6.3, and calculated for the ground state. Here we will extend the treatment to all states.

The perturbation Hamiltonian is (eq. (6.28)),

$$\delta H = -\frac{1}{2mc^2}(H^0 - V)^2. \quad (8.1)$$

It so happens that for this perturbation, the zero-order states which we have already used, namely  $R_{n,l}(r)Y_{l,m}(\theta, \phi)$ , are appropriate, because the perturbing Hamiltonian does not connect states of different  $l$ . The proof is given (for you to complete) in the problem set. Therefore the first-order energy shift is simply the expectation value of  $\delta H$  in the familiar zeroth-order states. The integral has three terms which involve the zero order energy and mean values of  $1/r$  and  $1/r^2$ . These are given in table 3.1. The result is

$$\delta E_n^{(1)} = -\frac{Z^2\alpha^2}{4n^2}E_n^0 \left(3 - \frac{4n}{l+1/2}\right) = Z^4\alpha^4\frac{\mu c^2}{8n^4} \left(3 - \frac{4n}{l+1/2}\right). \quad (8.2)$$

Note that (8.2) looks like the full Dirac result (5.2) except that we have  $l$  instead of  $j$ . The further terms which, together with this kinetic energy correction, reproduce Dirac fully are the spin-orbit interaction and the Darwin term, these are discussed next.

## 8.2 Spin-orbit interaction

This is the interaction between the magnetic dipole moment of the electron and the magnetic field it experiences owing to its motion through the electric field of the nucleus. (Alternatively, one may say that the nucleus moves with respect to the electron and so produces a magnetic field at the position of the electron.) We will calculate the contribution by an argument which is fairly simple and has good physical insight but which unfortunately needs fixing at the end by a factor of two. To fix the calculation a further piece of physics is needed, the result of which will merely be quoted here.

### 8.2.1 Magnetic dipoles

To understand magnetic dipoles, first let's examine how they come about in classical physics. For example, consider a classical electron moving in a circular orbit of radius  $r$ . If it has speed  $v$  then the orbital angular momentum  $l = rmv$  and the current around the loop is  $ev/2\pi r$ . The magnetic dipole moment of a current loop is the current times the area, so

$$|\mu| = \frac{\pi r^2 ev}{2\pi r} = \frac{e}{2}rv = \frac{e}{2m}l \quad (8.3)$$

Examining the signs and directions carefully, we find that the full relationship is

$$\boldsymbol{\mu} = -\frac{e}{2m}\mathbf{l}, \quad (8.4)$$

More generally, we need to consider non-circular orbits. It is found that the result extends to this case [Problem (fairly easy): prove it yourself]. Hence we have the quantum operator relationship

$$\hat{\boldsymbol{\mu}} = -\frac{e}{2m}\hat{\mathbf{l}} \quad (8.5)$$

for the magnetic moment produced by the orbital motion of an electron.

It is useful to gather all the dimensional quantities on the right hand side together in one place, giving us a unit of magnetic dipole moment called the *Bohr magneton*:

$$\mu_B = \frac{e\hbar}{2m}. \quad (8.6)$$

We can expect spin angular momentum also to have a magnetic moment associated with it, and it does, but it turns out that the proportionality constant is twice that for orbital angular momentum (Dirac's equation predicts this<sup>1</sup>, I know of no simpler argument for it). To allow for all eventualities we therefore introduce a dimensionless factor called the *g factor*, and write:

$$\boldsymbol{\mu} = -g_j \mu_B \frac{\hat{\mathbf{j}}}{\hbar}, \quad (8.7)$$

with

$$\begin{aligned} g_l &= 1 \\ g_s &= 2. \end{aligned}$$

N.B. you will often see equation (8.7) written  $\boldsymbol{\mu} = -g_j \mu_B \hat{\mathbf{j}}$ . In the latter version it is understood that the angular momentum  $\hat{\mathbf{j}}$  is in units of  $\hbar$ . We shall not make that choice for now.

## 8.2.2 Magnetic field experienced by the electron

The above serves as an introduction to some notations, and a reminder on magnetic dipoles. Now we will consider the magnetic field experienced by the electron in hydrogen. One way to think about this is to say that the electron sees the nuclear charge going around it, so the nuclear charge  $q$  moving at velocity  $\mathbf{v}_n$  with respect to the electron produces a magnetic field at the electron according to the Biot-Savart law:

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{q \mathbf{v}_n \wedge \mathbf{r}}{r^3} \quad (8.8)$$

$$= \mu_0 \epsilon_0 \frac{q}{4\pi \epsilon_0} \frac{\mathbf{v}_n \wedge \hat{\mathbf{r}}}{r^2} \quad (8.9)$$

$$= \frac{-\mathbf{v} \wedge \mathbf{E}}{c^2} \quad (8.10)$$

where  $\mathbf{E}$  is the electric field of the nucleus at the position of the electron, and  $\mathbf{v} = -\mathbf{v}_n$  is the velocity of the electron in the reference frame of the nucleus.

For an atom of many electrons, any given electron experiences a magnetic field from the nucleus and all the other electrons which move with respect to it. Each of the contributions has the form (8.10) with  $\mathbf{E}$  being the contribution to the electric field at the electron in question, so the total magnetic field is still given by (8.10), with  $\mathbf{E}$  now the total electric field at the electron in question.

This suggests that the relation (8.10) can be obtained by a more direct insight into the fields themselves. In fact it is a simple result of special relativity<sup>2</sup>.

<sup>1</sup>In fact Quantum Electrodynamics makes a slight adjustment, and the true *g* factor for spin is greater than 2 by an amount of order  $\alpha$ , see formula in section 1.2

<sup>2</sup>See also Panofsky and Phillips pp. 330-331.

Here is the relativistically insightful way to think about equation (8.10). Consider a particle of charge  $q$  moving at velocity  $\mathbf{v}$  with respect to reference frame  $S$ . If the electric and magnetic field in this reference frame are  $\mathbf{E}$  and  $\mathbf{B}$ , then the Lorentz force on the particle is  $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \wedge \mathbf{B})$ . Now consider the motion in reference frame  $S'$  in which the particle is at rest. Relativistic corrections to mass and acceleration are 2nd order in  $v/c$ , so to first order in  $v$ , the force is still  $q(\mathbf{E} + \mathbf{v} \wedge \mathbf{B})$ . However, in frame  $S'$  the particle is not moving, therefore this force must arise completely from the electric field  $\mathbf{E}'$  in the new reference frame. We conclude that the formula relating the electric field in a moving frame to the electric and magnetic fields in the original frame is

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \wedge \mathbf{B} \quad \text{to first order in } v. \quad (8.11)$$

Now, the equation we are discussing, eq. (8.10), is simply the corresponding result for the magnetic field in a moving frame:

$$\mathbf{B}' = \mathbf{B} - \frac{\mathbf{v} \wedge \mathbf{E}}{c^2} \quad \text{to first order in } v. \quad (8.12)$$

The Biot-Savart formula is an approximation to the full relativistic result, being valid up to first order in  $v/c$ .

Notice that for the case of a radial field,  $\mathbf{E}$  is in the direction of  $\mathbf{r}$  and therefore  $\mathbf{B}$  is in the direction of  $-\mathbf{v} \wedge \mathbf{r} = \mathbf{l}/m$ , i.e. in the direction of the orbital angular momentum vector.

### 8.2.3 The Hamiltonian and perturbation calculation

The magnetic field interacts with the magnetic dipole moment of the electron, to give the perturbation Hamiltonian

$$\begin{aligned} H_{so} &= -\boldsymbol{\mu} \cdot \mathbf{B} \\ &= g_s \frac{e\hbar}{2m} \frac{\hat{\mathbf{s}}}{\hbar} \cdot \frac{-\mathbf{v} \wedge \mathbf{E}}{c^2} \\ &= \frac{e}{mc^2} \hat{\mathbf{s}} \cdot \left( \frac{e}{4\pi\epsilon_0 r^3} \mathbf{r} \wedge \mathbf{v} \right) \\ &\stackrel{?}{=} \frac{e^2}{4\pi\epsilon_0 m^2 c^2} \frac{1}{r^3} \hat{\mathbf{l}} \cdot \hat{\mathbf{s}} \quad [ \text{using } \mathbf{r} \wedge \mathbf{mv} = \mathbf{l} ] \end{aligned}$$

This calculation is correct but the question-mark highlights a subtlety: we have just done a calculation in the reference frame of the electron, which owing to the orbital motion is not an inertial reference frame. When we transform back to the centre of mass frame using special relativity, a further precession is introduced, called the *Thomas precession*, and the result in the centre of mass frame is exactly half what we have calculated:

$$H_{so} = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0 m^2 c^2} \frac{1}{r^3} \hat{\mathbf{l}} \cdot \hat{\mathbf{s}} \quad [ \text{for } V = -e^2/4\pi\epsilon_0 r. \quad [ \text{correct result} ] \quad (8.13)$$

More generally, we can apply the same analysis to the case of an electron orbiting in any radial electric field (thus a central potential energy function  $V(r)$ ), and one obtains (exercise for the reader):

$$H_{so} = \frac{1}{2} \frac{1}{m^2 c^2} \frac{1}{r} \frac{dV}{dr} \hat{\mathbf{l}} \cdot \hat{\mathbf{s}}. \quad (8.14)$$

N.B. you may meet formulae for  $H_{so}$  which appear to look like (8.13) and (8.14) multiplied by a factor  $\hbar^2$ . This is because some authors use dimensionless versions of angular momentum operators, and move the  $\hbar$  factors into the coefficient. However in fact there is nothing especially to do with quantum mechanics (and hence  $\hbar$ ) in the calculation so far: it is just the classical physics analysis of the interaction between a magnetic dipole and the magnetic field it experiences when it moves through a central electric field.

Now let's see what happens when we apply this interaction to the hydrogen atom using perturbation theory—here comes the quantum part. We have an interaction between  $\hat{\mathbf{l}}$  and  $\hat{\mathbf{s}}$  but we can be sure this will not manage to change the total angular momentum  $\hat{\mathbf{j}} = \hat{\mathbf{l}} + \hat{\mathbf{s}}$  because, as we discussed at length in sections 7.3 and 7.3.1, the total angular momentum can't be affected by internal forces. Therefore the quantities  $\hat{\mathbf{j}}^2$  and  $\hat{j}_z$  are constants of the motion. Furthermore, we can express  $\hat{\mathbf{l}} \cdot \hat{\mathbf{s}}$  in terms of  $\hat{\mathbf{j}}^2$ ,  $\hat{\mathbf{l}}^2$  and  $\hat{\mathbf{s}}^2$  quite easily:

$$\hat{\mathbf{j}}^2 = \hat{\mathbf{j}} \cdot \hat{\mathbf{j}} \quad (8.15)$$

$$= (\hat{\mathbf{l}} + \hat{\mathbf{s}}) \cdot (\hat{\mathbf{l}} + \hat{\mathbf{s}}) \quad (8.16)$$

$$= \hat{\mathbf{l}}^2 + \hat{\mathbf{s}}^2 + \hat{\mathbf{l}} \cdot \hat{\mathbf{s}} + \hat{\mathbf{s}} \cdot \hat{\mathbf{l}} \quad (8.17)$$

$$\begin{aligned} &= \hat{\mathbf{l}}^2 + \hat{\mathbf{s}}^2 + 2\hat{\mathbf{l}} \cdot \hat{\mathbf{s}} \quad [\text{since } \hat{\mathbf{s}} \text{ commutes with } \hat{\mathbf{l}}] \\ \implies \hat{\mathbf{l}} \cdot \hat{\mathbf{s}} &= \frac{1}{2} (\hat{\mathbf{j}}^2 - \hat{\mathbf{l}}^2 - \hat{\mathbf{s}}^2) \end{aligned} \quad (8.18)$$

This is an exact relationship between the operators<sup>3</sup>.

This result is very convenient because the magnitudes of all three angular momenta (orbital, spin and total) are constants of the motion under the zeroth-order Hamiltonian, which means the zeroth-order energy eigenstates, i.e. the eigenstates before we take spin-orbit coupling into account, are eigenstates of  $\hat{\mathbf{l}}^2$ ,  $\hat{\mathbf{s}}^2$ , and  $\hat{\mathbf{j}}^2$  (when they are chosen appropriately in the case of degenerate sets of states). These zeroth-order states may conveniently be written  $|n, l, s, j\rangle$  and we have (using (8.18)):

$$\hat{\mathbf{l}} \cdot \hat{\mathbf{s}} |n, l, s, j\rangle = \frac{1}{2} \hbar^2 (j(j+1) - l(l+1) - s(s+1)) |n, l, s, j\rangle. \quad (8.19)$$

It is now easy to apply the perturbation theory: to get the energy shift to first order we just take the expectation value of  $H_{so}$  in a state of well-defined  $j$ ,  $l$  and  $s$ :

$$\delta E_{so} = \langle n, l, s, j | H_{so} | n, l, s, j \rangle = \frac{1}{2} \frac{\hbar^2 e^2}{4\pi\epsilon_0 m^2 c^2} \left\langle \frac{1}{r^3} \right\rangle \frac{1}{2} (j(j+1) - l(l+1) - s(s+1)), \quad (8.20)$$

where we used (8.13) and (8.19).

The vector model picture of this spin-orbit coupling is that  $\mathbf{l}$  and  $\mathbf{s}$  are precessing about  $\mathbf{j}$ .

More generally, when an electron moves in any central potential  $V(r)$ ,

$$\delta E_{so} = \langle n, l, s, j | H_{so} | n, l, s, j \rangle = \frac{1}{2} \frac{\hbar^2}{m^2 c^2} \left\langle \frac{1}{r} \frac{dV}{dr} \right\rangle \frac{1}{2} (j(j+1) - l(l+1) - s(s+1)). \quad (8.21)$$

Since the energy of the levels now depends on  $j$ , we extend the term notation to indicate  $j$  as a subscript, thus

$$\{l, s, j\} \text{ are indicated by } {}^{2s+1}l_j$$

For example, the ground state of hydrogen is  $1s \ {}^2S_{1/2}$ , and the group of states with  $n = 2$  is  $2s \ {}^2S_{1/2}$ ,  $2p \ {}^2P_{1/2}$ ,  $2p \ {}^2P_{3/2}$ .

---

<sup>3</sup>Some textbooks quote a cosine rule involving  $|\mathbf{l}||\mathbf{s}|\cos\theta$ , but this is an unnecessary departure from strictly accurate quantum theory.

Finally, let us evaluate  $\langle 1/r^3 \rangle$  in hydrogen. Recalling  $a_0 m \alpha c = \hbar$  for the Bohr radius  $a_0$ , we can see that  $\langle 1/r^3 \rangle$  should be proportional to  $1/a_0^3 = m^3 \alpha^3 c^3 / \hbar^3$  and hence

$$\delta E_{so} = \langle H_{so} \rangle \propto \frac{1}{2} \frac{\hbar^2 e^2}{4\pi\epsilon_0 m^2 c^2} \frac{m^3 \alpha^3 c^3}{\hbar^3} = \frac{1}{2} m c^2 \alpha^4. \quad (8.22)$$

The full result is given in table 3.1, leading to

$$\delta E_{so} = \frac{m c^2 \alpha^4}{4n^3} \frac{j(j+1) - l(l+1) - s(s+1)}{l(l+1/2)(l+1)} \quad (8.23)$$

for  $l > 0$ . Note that the dependence on  $1/r^3$  leads to a dependence on  $1/n^3$ . When the spin-orbit interaction is combined with the effect of the mass correction, we obtain the correct Dirac result to order  $\alpha^4$ , except for the  $l = 0$  case, where the kinetic energy correction is non-zero but the spin-orbit effect is zero.

### 8.3 Darwin term for $l = 0$

The final relativistic correction to the Schrödinger equation for hydrogen is a special case which only applies to  $l = 0$ . It is to do with the fact that the wavefunction for  $l = 0$  does not vanish at the origin, so the non-relativistic approximation  $Z e^2 / 4\pi\epsilon_0 r \ll m c^2$  is seriously violated. The interaction Hamiltonian is

$$H_D = -\frac{\hbar^2}{4m^2 c^2} e \mathbf{E} \cdot \nabla \quad (8.24)$$

and the associated energy shift is proportional to  $|\psi(0)|^2$  (the electron probability density at the nucleus). By a strange luck the energy shift due to the Darwin term for  $l = 0$  is exactly given by substituting  $l = 0$  in the formula for the energy shift given by the sum of the two previous interactions! (kinetic energy correction and spin-orbit interaction.) Actually this luck is not so strange since this whole perturbative calculation is somewhat artificial, introducing relativistic phenomena such as spin and mass energy without bothering to make the fundamental equation (Schrödinger) relativistic. We are bound to keep introducing terms until we get the correct Dirac result!

Now, the above details of the Darwin term are off the syllabus, but as a matter of general knowledge you should know that

- (1) the states of the same  $j$  (and  $n$ ) are degenerate in hydrogen according to all but QED
- (2) the spin-orbit shift is zero for  $l = 0$
- (3) the mass shift plus spin-orbit is the whole story for  $l \neq 0$ , but not for  $l = 0$ .

Adding the contributions of the three terms outlined above, we get the Dirac energy correct to order  $Z^4 \alpha^4$ , see the Dirac summary (5.2).

You should check your understanding by working out the size of the fine structure from the formula. The observed fine structure of hydrogen is given in many textbooks. To have an order of magnitude in mind, notice that the splitting of the  $2p^2P_{3/2}$  level from the  $2p^2P_{1/2}$  level is about 10 GHz.

### 8.3.1 Main points

The main points of this chapter have been:

the Schrödinger treatment by perturbation theory with three contributions: kinetic energy correction, spin-orbit interaction, Darwin term for  $l = 0$ ; the relationship between spin and magnetic moment; gyromagnetic ratio  $g$  and Bohr magneton  $\mu_B$ ; the magnetic field experienced by an electron in an atom; the general form of  $\delta E_{so}$ .

## 8.4 Appendix on the correct choice of eigenstates for hydrogen

Did you notice that we solved the problem of degenerate perturbation theory in hydrogen without ever explicitly calculating the zero-order states and making sure the perturbation did not connect degenerate pairs of them? This is because we used the powerful method of constants of the motion: this told us enough information about the states (i.e. it specified which observables they had to be eigenstates of) that we didn't have to write the states down explicitly. However, to check our knowledge, it is a good exercise to take a look at some of these states.

We saw that the kinetic energy shift is diagonal in states of definite  $l$ , and the spin-orbit interaction is diagonal in states of definite  $j$ , where the underlying operators are related by  $\hat{\mathbf{j}} = \hat{\mathbf{l}} + \hat{\mathbf{s}}$ . What we need, therefore, is a set of states of definite  $l$  and  $j$ . Such states can exist because  $\hat{l}^2$  and  $\hat{j}^2$  commute (even though  $\hat{\mathbf{j}}$  and  $\hat{\mathbf{l}}$  do not, see below for explanation). They are given by the quantum theory of coupling of angular momentum, whose general treatment is outside the syllabus. However, we will quote some examples without proof. Only the angular and spin parts will be shown. The complete state is a product of this and the radial part  $R_{n,l}(r)$ . All have definite spin quantum number  $s = 1/2$ , but the spin state is also involved because the direction of the spin (associated with quantum number  $m_s$ ) is no longer well-defined: it is not a constant of the motion, so the energy eigenstates are superpositions of states of different  $m_s$ . Similar statements apply to  $\hat{\mathbf{l}}$  and the quantum number  $m_l$ .

Here are the examples for  $l = 0$  and 1. On the left hand side of each equation the angular momentum quantum numbers are specified using spectroscopic notation, and the equation shows how the state is composed from the spherical harmonics and the spin states:

$$\begin{aligned} |^2s_{1/2}, m_j = 1/2\rangle &= |Y_{0,0}\rangle |\uparrow\rangle \\ |^2s_{1/2}, m_j = -1/2\rangle &= |Y_{0,0}\rangle |\downarrow\rangle \end{aligned} \quad (8.25)$$

$$\begin{aligned} |^2p_{1/2}, m_j = 1/2\rangle &= \sqrt{\frac{2}{3}} |Y_{1,1}\rangle |\downarrow\rangle - \frac{1}{\sqrt{3}} |Y_{1,0}\rangle |\uparrow\rangle \\ |^2p_{1/2}, m_j = -1/2\rangle &= -\sqrt{\frac{2}{3}} |Y_{1,-1}\rangle |\uparrow\rangle + \frac{1}{\sqrt{3}} |Y_{1,0}\rangle |\downarrow\rangle \end{aligned} \quad (8.26)$$

$$\begin{aligned} |^2p_{3/2}, m_j = 3/2\rangle &= |Y_{1,1}\rangle |\uparrow\rangle \\ |^2p_{3/2}, m_j = 1/2\rangle &= \sqrt{\frac{2}{3}} |Y_{1,0}\rangle |\uparrow\rangle + \frac{1}{\sqrt{3}} |Y_{1,1}\rangle |\downarrow\rangle \\ |^2p_{3/2}, m_j = -1/2\rangle &= \sqrt{\frac{2}{3}} |Y_{1,0}\rangle |\downarrow\rangle + \frac{1}{\sqrt{3}} |Y_{1,-1}\rangle |\uparrow\rangle \\ |^2p_{3/2}, m_j = -3/2\rangle &= |Y_{1,-1}\rangle |\downarrow\rangle \end{aligned} \quad (8.27)$$

### 8.4.1 Commutation of vector operators

The commutation of vector operators needs a little thought. With classical vectors we have  $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ , but  $\mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}$ , so we should not expect vector operators to commute in general in quantum physics. If in doubt, always work in terms of the components and see what you find. In the discussion of hydrogen fine structure we used the fact that if  $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$  then eigenstates of  $\hat{J}^2$  are not eigenstates of  $\hat{L}_z$  or  $\hat{S}_z$ . To show this we will prove that  $[\hat{J}^2, \hat{L}_z] \neq 0$ :

$$\begin{aligned} [\hat{J}^2, \hat{L}_z] &= [\hat{L}^2 + \hat{S}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \hat{L}_z] \\ &= [2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \hat{L}_z] \\ &= [2(\hat{L}_x \hat{S}_x + \hat{L}_y \hat{S}_y + \hat{L}_z \hat{S}_z), \hat{L}_z] \\ &= 2 \left( [\hat{L}_x, \hat{L}_z] \hat{S}_x + [\hat{L}_y, \hat{L}_z] \hat{S}_y \right) \\ &= 2i\hbar \left( -\hat{L}_y \hat{S}_x + \hat{L}_x \hat{S}_y \right) \\ &= 2i\hbar \hat{C}_z \quad \text{where } \hat{\mathbf{C}} = \hat{\mathbf{L}} \wedge \hat{\mathbf{S}}. \end{aligned}$$

A similar argument would apply to  $[\hat{J}^2, \hat{S}_z]$ .



## Chapter 9

# The Zeeman effect

We have now finished our discussion of a free hydrogen atom or hydrogen-like ion, that is, one with no external influences on it. We now turn to the influence of an externally applied magnetic field. Such a field causes a further splitting of the energy levels, and hence further structure in the emission and absorption spectrum. This is called the Zeeman effect. When the field is large (exactly how large will be seen below) the way the components vary with  $B$  can change and this regime is called the Paschen-Back effect.

At this introductory stage, we will only consider the restricted case of the Zeeman effect for  $s$  states, that is, when there is no orbital angular momentum (but there is spin). This will permit us to examine some of the basic ideas. The full Zeeman effect which involves both orbital and spin contributions is an important tool in the analysis of atomic spectra and in other applications, it will be treated next year.

### 9.1 Weak $B$ field: Zeeman effect

An external magnetic field will interact primarily with the net magnetic dipole moment  $\boldsymbol{\mu}$  of an atom if it has one. From basic electromagnetic theory, the interaction Hamiltonian is

$$H_Z = -\boldsymbol{\mu} \cdot \mathbf{B}. \quad (9.1)$$

#### 9.1.1 Classical treatment and normal Zeeman effect

First let us consider the classical physics of this situation. The external field exerts a torque on the magnetic dipole, tending to make it parallel with the field. Owing to the non-zero angular momentum of the system, this torque causes the magnetic dipole to *precess around the magnetic field direction*. This can be compared to the situation when a toy gyroscope stands on a table. The torque produced by the combination of gravity acting through its centre of mass, and the normal reaction force from the table top on the tip of the gyroscope's axis, tends to give it angular momentum in such a direction as to tip it over. However, when added to the angular momentum the gyroscope already possesses, the net result is to cause the gyroscope not to fall down but to precess. See the problems for more details, where you should deduce that the angular frequency of the precession of a magnetic dipole in a magnetic field  $B$  is the *Larmor frequency*  $\omega_L = g\mu_B B/\hbar$  where  $\mu_B$  is the Bohr magneton introduced in section 8.2.1 and  $g$  is the gyromagnetic ratio.

Now, for an electron moving around an orbit, the magnetic dipole moment associated with the *orbital* motion has gyromagnetic ratio  $g_1 = 1$  (section 8.2.1)—in both a classical as well as a quantum treatment. Hence before spin was discovered, the Larmor frequency was expected to be  $\omega_L = \mu_B B / \hbar$ .

If, while undergoing such a precession, the electron also oscillates so as to emit light, then one expects the frequency of the emitted light to be shifted up or down by  $\omega_L$ , or not changed, depending on its direction of polarization. For example, any oscillatory motion of the electron along the  $\mathbf{B}$  field direction is not affected by the precession, so light polarized parallel to  $\mathbf{B}$  is unshifted in frequency (and light polarized perpendicular to  $\mathbf{B}$  is shifted in frequency by  $\pm\omega_L$ ).

The classical treatment therefore leads to an expected emission spectrum consisting of three components of frequencies  $\omega_0$  and  $\omega_0 \pm \mu_B B / \hbar$ , where  $\omega_0$  is the transition frequency in the absence of a magnetic field. An observed emission spectrum with this spacing and structure is called the “normal Zeeman effect”. However, in practice the Zeeman spectrum is usually found to be different from this, in which case it is called an “anomalous Zeeman effect”. The word “normal” is here used in the sense of “not strange”: the “anomalous” Zeeman effect is in fact more typical.

A normal Zeeman emission spectrum is what is predicted by both a classical and a quantum treatment when the initial and final states both have zero spin. Hence the observation of the anomalous Zeeman effect was evidence (and remains evidence) that something other than movement of charge causes at least part of the magnetic behaviour of atoms.

In the following quantum treatment, we will examine the opposite extreme: when there is spin, but there is no orbital angular momentum. We will find that in this case the prediction remains a pattern of three lines, but now with the spacing  $2\mu_B B / \hbar$ .

### 9.1.2 Quantum treatment: Zeeman effect for $s$ states

We treat the influence of the Zeeman Hamiltonian (9.1) by first order perturbation theory.

Before this external perturbation acts, the Hamiltonian has spherical symmetry, and therefore the total angular momentum is a constant of the motion and so  $j$  and  $m_j$  are both good quantum numbers. States of the same  $j$  but different  $m_j$  are degenerate. When the external magnetic field is turned on, there is a torque which results in the magnetic dipole precessing about the field direction. Without loss of generality we can choose the field direction to be the  $z$  axis. The precession then changes the direction of  $\mathbf{j}$ , but not the value of its projection along the  $z$  direction. Therefore, in the presence of the field,  $\hat{j}_z$  is still a constant of the motion, and  $m_j$  remains a good quantum number (as does  $j$ ). This means that degenerate perturbation theory proceeds straightforwardly by using eigenstates of  $\hat{j}_z$  as the zeroth order states.

We already met the connection between magnetic dipole moments and angular momentum. For a system with orbital angular momentum  $\hat{\mathbf{l}}$  and spin angular momentum  $\hat{\mathbf{s}}$ , the net magnetic dipole moment is

$$\boldsymbol{\mu} = -(g_l \mu_B \hat{\mathbf{l}} + g_s \mu_B \hat{\mathbf{s}}). \quad (9.2)$$

where  $g_l = 1$  and  $g_s = 2$ , and we are using notation in which the operators  $\hat{\mathbf{l}}$  and  $\hat{\mathbf{s}}$  are dimensionless.

To keep things simple, we will only discuss  $s$  states here, i.e. the case where the orbital angular momentum is zero (the phrase ‘ $s$  states’ refers to spectroscopic notation for  $l = 0$ ). In this case the magnetic dipole moment arises wholly from the spin of the system, so we have

$$\boldsymbol{\mu} = -g_s\mu_B\hat{\mathbf{s}} = -g_s\mu_B\hat{\mathbf{j}}, \quad (9.3)$$

where  $\hat{\mathbf{j}} = \hat{\mathbf{l}} + \hat{\mathbf{s}}$  is the operator for total angular momentum. The purpose in using  $\hat{\mathbf{j}}$  rather than  $\hat{\mathbf{s}}$  is because in general, in the presence of spin-orbit coupling,  $\hat{\mathbf{s}}$  is not a constant of the motion, while  $\hat{\mathbf{j}}$  is. Of course,  $l = 0$  states don’t have any spin-orbit coupling, so for them  $\hat{\mathbf{s}}$  is a constant of the motion, but it is useful to adopt a notation which uses the same quantum numbers ( $j$  and  $m_j$ , not  $m_s$ ) as the ones which are needed to describe the atomic structure. Choosing the  $z$  axis to lie along  $\mathbf{B}$ , we obtain

$$-\boldsymbol{\mu} \cdot \mathbf{B} = -\mu_z B = 2\mu_B B \hat{j}_z \quad (9.4)$$

and the first order shift in the energy levels is

$$\Delta E = \langle j, m_j | 2\mu_B B \hat{j}_z | j, m_j \rangle = 2\mu_B B m_j, \quad (9.5)$$

where we have labeled the states by the total angular momentum quantum numbers, and the final evaluation of the matrix element was easy because the states are eigenstates of the operator in question,  $\hat{j}_z$ .

In hydrogen  $s$  states,  $j = 1/2$  and  $m_j = \pm 1/2$ , but other systems can have larger values of spin, and hence more possible values of  $m_j$ . In any case, the energy level which at zero magnetic field had its  $2j + 1$  states all degenerate, is now split into  $2j + 1$  states all of different energy, one for each value of  $m_j$ . For  $j = 1/2$  and  $g = 2$  the states are at energies  $\pm 2\mu_B B(1/2) = \pm\mu_B B$ , hence the separation between them is  $2\mu_B B$ .

### 9.1.3 Effect on the emission spectrum

The easiest way to observe the Zeeman effect is by observing the emission spectrum of the atom in the presence of a magnetic field. The emission spectrum does not reveal the energy levels directly, but rather the differences between them, but in the presence of the Zeeman splitting there are many more such differences. As the magnetic field is increased, each emission line is observed to divide into a group of closely spaced lines.

The strongest transitions (producing the brightest emitted light) in atoms involve so-called “electric dipole radiation”, to be discussed in chapter 13, but this type of radiation is only possible for a restricted set of transitions which obey certain *selection rules*. The selection rules will be discussed in detail in chapters 12 and 13, but for now we will simply comment that one rule is that the parity of the atomic state must change in order for the transition to be possible by the strongest (electric dipole) mechanism. The parity of the spherical harmonics is equal to  $(-1)^l$  so the parity rule implies that  $l$  must change by an odd integer during the transition, and therefore it rules out transitions between  $s$  states (in fact a state with  $l = 0$  can only be connected by electric dipole transitions to a state with  $l = 1$ ). Therefore our restricted study of the Zeeman effect will have to consider non-electric-dipole transitions. Although not the strongest transitions, they still include some interesting and important examples.

One example is the 1s–2s transition in hydrogen. Whereas allowed (electric dipole) transitions involve the emission of a single photon, this transition can only take place by the emission of two photons, in order that angular momentum is conserved (a photon carries 1 unit of intrinsic angular momentum). This has two important implications. First, the transition is weak, so the 2s state only decays slowly, which means it can have a very well-defined energy. This means the transition frequency can in principle be measured very precisely. Its study now provides one of the most stringent tests of our basic physical theories. Secondly, by probing the transition with counter-propagating laser beams, the Doppler shift associated with motion of the hydrogen atoms is cancelled, since the photons from one beam are shifted up in frequency by the same amount as those in the other beam are shifted down, hence one important source of imprecision in the experiment is removed.

For the 1s–2s transition, both the upper and the lower state are split into 2 by the Zeeman effect. Let  $\Delta m_j = m'_j - m_j$  where  $m'_j$  is the magnetic quantum number of the upper state (2s) and  $m_j$  is the magnetic quantum number of the lower state (1s). The transitions in which  $\Delta m_j = 0$  are called  $\pi$  transitions, those in which  $\Delta m_j = \pm 1$  are called  $\sigma_+$  and  $\sigma_-$  transitions, respectively. For 1s–2s there are four cases. The two  $\pi$  transitions ( $\Delta m_j = 0$ ) have the same frequency, equal to the unshifted frequency of the line (i.e. the same as the value at zero  $B$  field). The  $\sigma$  transitions have frequencies shifted from this by  $\pm 2\mu_B B/\hbar$ . Therefore the whole pattern consists of three equally spaced spectral lines, with separations  $2\mu_B B/\hbar$ , centred on the unshifted value.

A few moments' thought will show you that whenever the splittings of upper and lower states are the same, transitions associated with  $\Delta m_j = \pm 1, 0$  will always have this pattern of three equally spaced emission lines. Therefore we have the result that when there is spin but there is no orbital angular momentum, the Zeeman spectrum has the form of three components as in the normal Zeeman effect, but with double the spacing, owing to the  $g = 2$  value for spin.

More generally, when both orbital and spin angular momentum is involved in the states, the pattern of emission lines is more complicated. This will be treated next year. This anomalous effect gives a multitude of different patterns, depending on the values of  $l$ ,  $s$  and  $j$ , hence its observation often permits these quantum numbers to be deduced. Thus it gives much useful information which greatly facilitates the interpretation of atomic spectra.

## 9.2 Strong $B$ field: Paschen-Back effect

The above perturbation theory is valid as long as the Zeeman shift  $\Delta E$  is small compared to the separation of the energy eigenstates before the perturbation was applied. If the Zeeman shift is similar to the fine structure, for example, then the analysis based on first-order perturbation theory can no longer be expected to be exact. The limit where the field-induced splitting is large compared to the fine structure is called the Paschen-Back effect. It is analysed by treating the Zeeman Hamiltonian as the first perturbation to the gross structure of the atom, and then introducing the fine structure contribution (spin-orbit effect, etc.) after that. However, because the  $s$  states don't have any fine structure (they have no spin-orbit contribution and their energies are shifted but not split by kinetic energy corrections), for the restricted case we have been discussing, the Paschen-Back effect leads to the same pattern of splittings as the low-field Zeeman effect.

## 9.3 Uses of the Zeeman effect

With modern computers we can make a fair stab at calculating atomic structure from scratch, but the problem is so complicated for some atoms that even this is only moderately successful. Historically, the problem of working out the identities of all the different atomic energy levels from the observed spectral lines was like trying to deduce the structure of a piano from the noise it makes when thrown down stairs. The general Zeeman effect is a very powerful tool in this analysis. Merely by counting components in the Zeeman spectrum (anomalous effect) one can make a pretty fair guess at the values of  $J$  for the two levels involved in the transition. The relative positions of the Zeeman lines then lead rapidly to confirming  $J$  and identifying orbital and spin angular momentum quantum numbers.

Nowadays, the mere identities of atomic levels are not exactly hot news, but occasionally we need to know the details of some particular transition, in order to use it for some purpose, and the vast majority of transitions in all but the most simple atoms will (rightly) forever remain uncalculated and unobserved until they are needed. The Zeeman effect therefore retains its old usefulness as an identifier of levels.

However, the more common uses of the Zeeman effect are as a diagnostic of magnetic fields in an inaccessible environment such as the atmosphere of a star or the inside of a tokamak fusion reactor, and as a tool for manipulating atoms. The Zeeman effect is used for example to define an axis in space so that atoms in atomic clocks don't wander in direction. It is also a central ingredient in other magnetic resonance experiments. A more ingenious use is in atom traps, where a magnetic field gradient provides a force to confine cold atoms. A Zeeman effect which varies with position is also used in one laser cooling method to compensate for the changing Doppler effect when atoms are slowed down to very low speeds by laser light.

### 9.3.1 Main points

$H = -\boldsymbol{\mu} \cdot \mathbf{B}$ ; precession in the vector model; Zeeman effect giving level splittings and three transition lines; normal Zeeman effect; Paschen-Back effect in strong fields; uses of the Zeeman effect.



# Chapter 10

## Helium: gross structure

The Schrödinger equation for helium is:

$$H \Psi(r_1, \theta_1, \phi_1, \zeta_1; r_2, \theta_2, \phi_2, \zeta_2) = E \Psi(r_1, \theta_1, \phi_1, \zeta_1; r_2, \theta_2, \phi_2, \zeta_2)$$

where the Hamiltonian is, ignoring small corrections,

$$H = \frac{\hat{\mathbf{p}}_1^2}{2m} + \frac{\hat{\mathbf{p}}_2^2}{2m} - \frac{Ze^2}{4\pi\epsilon_0 r_1} - \frac{Ze^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}}. \quad (10.1)$$

The last term is the Coulomb repulsion between the electrons, their distance from one another is  $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ . We are using  $\zeta$  as the spin ‘coordinate’.

Since the Hamiltonian is a *sum* of spatial and spin parts (in which the spin part is zero), it is separable and hence there exist solutions which have the form of a product of spatial and spin parts:

$$\Psi(r_1, \theta_1, \phi_1, \zeta_1; r_2, \theta_2, \phi_2, \zeta_2) = \psi(r_1, \theta_1, \phi_1; r_2, \theta_2, \phi_2) \otimes \chi(\zeta_1, \zeta_2). \quad (10.2)$$

We will treat the spatial part first. Recall that the quantity  $|\psi|^2 dx_1 dy_1 dz_1 dx_2 dy_2 dz_2$  is the probability to find the first electron between  $x_1, y_1, z_1$  and  $x_1 + dx_1, y_1 + dy_1, z_1 + dz_1$  and the second electron between  $x_2, y_2, z_2$  and  $x_2 + dx_2, y_2 + dy_2, z_2 + dz_2$ . We can think of helium as the two electrons moving around the stationary nucleus, though strictly speaking it is the motion in the centre of mass frame which we are discussing.

### 10.1 Main features of the structure, screening

Unfortunately, there is no known way to find an exact solution of the Helium Schrödinger equation by analytical methods. This is the quantum equivalent of the famous 3-body problem which can’t be exactly solved in classical mechanics. However, we can get a good understanding by learning from hydrogen and using some common sense.

It turns out that the higher-lying energy levels of helium are quite close to those of hydrogen (see problems). We can understand this as follows: imagine building a helium atom by starting with the bare nucleus, and adding the two electrons. Clearly, when we add the first electron we have a hydrogen-like system with  $Z = 2$ . Let this first electron occupy the ground state, then it has a spherically symmetric wavefunction. When the second electron is brought in, while it is far from the nucleus it experiences the electric field due to a spherically symmetric charge distribution of both nucleus and inner electron together, making a total charge of 1. Hence we would expect its energy levels to be close to those of the hydrogen atom.

This simple idea is called *screening*—we say the charge of the inner electron partially ‘screens’ the charge of the nucleus.

In more detail, we are arguing that when at least one electron is in the ground state, then the total potential energy of the system can be approximated

$$V = -\frac{Ze^2}{4\pi\epsilon_0 r_1} - \frac{Ze^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}} \simeq V_{\text{eff}}(r_1) + V_{\text{eff}}(r_2) \quad (10.3)$$

where

$$V_{\text{eff}}(r) \simeq \begin{cases} -\frac{Ze^2}{4\pi\epsilon_0 r} & \text{for } r \rightarrow 0 \\ -\frac{(Z-1)e^2}{4\pi\epsilon_0 r} & \text{for } r \rightarrow \infty \end{cases} \quad (10.4)$$

Thus, the potential energy of either electron is dominated by its interaction with the nucleus when it is close to the nucleus, while when it is far from the nucleus it sees the nuclear charge ‘screened’ by the other electron’s charge, i.e. reduced by 1. The true potential energy depends in a complicated way on the two electron positions, but the approximation (10.3) captures the main features and it has two very useful properties.

First, it is simply a sum of a term dependent only on  $r_1$ , and a term dependent only on  $r_2$ , therefore the Schrödinger equation becomes separable in  $r_1$  and  $r_2$ .

Secondly, the approximation (10.3) involves a purely *central* field for either electron, therefore the Schrödinger equation further separates into radial and angular parts, in which the angular solutions are spherical harmonics, and the radial solutions are counted by quantum numbers  $n$  and  $l$ . The radial solutions are not the same as the ones for hydrogen, (though they are similar in the limit  $r \rightarrow \infty$ ), but they are still counted by two quantum numbers  $n$  and  $l$  where  $n$  is in the range  $1 \cdots \infty$  and  $l$  is the orbital angular momentum quantum number, in the range  $0 \cdots n-1$ . To understand this, recall figure (3.1) which showed the effective potential energy curves of hydrogen including the centrifugal barrier. We now have a different shape of the potential energy function (it is no longer simply  $1/r$ ), but the general idea of a set of energy levels for each  $l$  value is retained. The main difference is that *states of the same  $n$  but different  $l$  are now expected to have different energy*—recall that the degeneracy w.r.t.  $l$  in hydrogen was a surprise, it is a special property of the  $1/r$  potential.

Bringing all this together, we find there exist overall solutions, in the *central field approximation* (10.3), which have the form

$$\psi(r_1, \theta_1, \phi_1; r_2, \theta_2, \phi_2) \simeq \phi_{1s}(r_1) \phi_{nlm}(r_2, \theta_2, \phi_2) \quad [ \text{to be adjusted later} \quad (10.5)$$

where to keep things simple we only wrote states in which at least one electron is in the ground state. The wavefunctions  $\phi_{1s}$  and  $\phi_{nlm}$  are labeled by the quantum numbers  $n, l, m$ , but they are not exactly the same as the hydrogen ones, though they have similar features (same number of nodes, vary as  $r^l$  near the origin, etc.).

We will see later that the true energy eigenstates involve sums and differences of these product wavefunctions, but this is the best way to proceed with the analysis at this juncture.

Using the idea of screening, it is easy to see that in helium we should expect states of lower  $l$  to have lower energy, because the wavefunctions of low  $l$  have a greater probability near the nucleus than states of high  $l$ . Therefore the low  $l$  states ‘feel’ the full nuclear charge more, they are less screened, and are hence more tightly bound. You are invited to examine this in the problems, where the energies are compared to hydrogen ones by making an adjustment to the nuclear charge.

Since the wavefunctions for the two electrons are specified by quantum numbers  $n$  and  $l$ , the eigenstates of the helium atom are labeled  $1s2s$ ,  $1s2p$ ,  $1s3s$ ,  $1s3p$ ,  $1s3d$  etc. The ground state is  $1s1s$ , which is written  $1s^2$  for brevity. It will be discussed in the next section.

The list of quantum numbers  $nsn'l$ , given in spectroscopic notation, is called the **configuration**.



In helium we are only concerned with excitation of one electron. That is, we don't have states such as  $2s2p$ ,  $2s3s$  etc. The reason is simply that it takes more energy to excite both electrons, and in helium and other light atoms the energy required is more than the ionization energy. That is, an atom with this much energy rapidly dissociates into an ion plus free electron. In some other elements, however, excitation of two or more electrons in bound states is possible.

### 10.1.1 Ground state

So far we have missed out the lowest lying energy level. The second electron has a lowest energy state with principal quantum number 1, giving the state  $1s^2$  for the whole atom. Note that in this case the approximation of spherical symmetry of the net potential experienced by each electron remains very good, so the quantum numbers  $n, l$  remain correct labels for this case also, though the screening is now very different from before so the energy level is not at all close to the hydrogen value... it is **low** down, at  $-25$  eV, because the nucleus is not well screened. Therefore the spectrum of helium has the following pattern: a bunch of transitions way off in the far ultra-violet, and then all the rest at noticeably longer wavelengths (mid ultra-violet upwards). The former are transitions between ground and excited states, the latter are transitions between one excited state and another.

Helium was in fact first discovered through its distinctive absorption spectrum (rather than by the usual way of finding some lying around, or by chemistry). It was discovered in the outer layer of the sun before it was ever found on earth, and this is the origin of its name (from *helios*, Greek for the sun).

Note that the ionization potential of 24.580 eV is large in helium, indeed it is the largest of any neutral atom. This is because both electrons are close the nucleus, and they each experience an effective nuclear charge significantly greater than 1. The effective nuclear charge (i.e. after screening is taken into account) can be estimated from the measured ionisation energy—see problems.

[Problem: helium is smaller (has a smaller mean radius) than hydrogen. Why?]

[Problem (tricky): estimate how far one would expect  $2s^2$  to be above  $1s^2$  in helium].

To get a good calculated estimate of the ground state energy, we need a method to solve the Schrödinger equation approximately. We could try perturbation theory, starting with the Hamiltonian without the electron-electron repulsion, but then the perturbation would be similar in size to the zero order energy level separations, so perturbation theory does not work well<sup>1</sup>. A better approach is to use the variational method, to be discussed next.

## 10.2 The variational method

The variational method is a method to find approximate solutions to Schrödinger's equation. The starting point is the observation (which we will prove in a moment) that

$$\langle \psi | H | \psi \rangle \geq E_0 \quad (10.6)$$

where  $E_0$  is the lowest eigenvalue of  $H$ , i.e. the ground state energy. The content of (10.6) is that if we try any wavefunction  $\psi$  whatsoever, and calculate the mean energy (i.e. the expectation of value  $H$  in the trial state), then the answer is guaranteed to be greater than or equal to the ground state energy. Therefore it gives us an upper bound on  $E_0$ . The essential idea of the variational method is to try lots of different states  $\psi$ , then the one giving the smallest mean energy gives the closest estimate of  $E_0$ , and therefore is the best approximation (out of the states tried) to the true ground state.

---

<sup>1</sup>To get a quick order-of-magnitude estimate of the electron-electron repulsion energy, one may consider the work required to assemble a sphere of charge whose radius is the Bohr radius and which contains two electrons' worth of charge.

This is best understood by an example. One is discussed in the lecture, and further examples are in the text books. In practice one makes a sensible guess of the general functional form to be expected for  $\psi$ , and one makes the exact shape of the function depend on one or more parameters. For example, in a one-dimensional potential well, an appropriate guess might be a Gaussian function, with a parameter  $a$  which sets its width:

$$\psi_a(x) = \left(\frac{2a}{\pi}\right)^{1/4} e^{-ax^2}. \quad (10.7)$$

We then evaluate the mean energy, which will depend on the parameter  $a$ :

$$E(a) = \langle \psi_a | H | \psi_a \rangle. \quad (10.8)$$

Next, find the value of the parameter  $a$  which gives the least energy, i.e. solve  $dE/da = 0$ . The value of  $a$  which minimizes  $E(a)$  indicates which form of  $\psi$  was the best approximation to the true ground state, and the associated mean energy is the lowest upper bound on  $E_0$ . The method does not give any information on how close an estimate to  $E_0$  one has, but by trying more and more adjustments to the wavefunction, one gets a feel for whether one is approaching a 'hard bound'.

### Proof of (10.6)

Let  $u_n(x)$  be the true eigenfunctions of  $H$ , and  $E_n$  the eigenvalues. Then any trial function  $\psi$  can be expanded in terms of the  $u_n(x)$ :

$$\psi = \sum_n u_n \langle u_n | \psi \rangle \quad (10.9)$$

Putting this in (10.6)

$$\begin{aligned} \langle \psi | H | \psi \rangle &= \langle \psi | H \sum_n |u_n\rangle \langle u_n | \psi \rangle \\ &= \sum_n \langle \psi | E_n |u_n\rangle \langle u_n | \psi \rangle \\ &= \sum_n E_n |\langle \psi | u_n \rangle|^2 \\ &\geq \sum_n E_0 |\langle \psi | u_n \rangle|^2 = E_0 \end{aligned}$$

### Notes for lecture example

We treat a very simple example to illustrate the method. The system is the harmonic oscillator, having potential energy  $V = (1/2)m\omega^2 x^2$ . We propose the guessed wavefunction

$$\psi = \begin{cases} N(a^2 - x^2) & \text{for } -a \leq x \leq a \\ 0 & \text{for } |x| > a \end{cases} \quad (10.10)$$

The wavefunction must be normalised, so  $N = (15/16a^5)^{1/2}$ . Now evaluate

$$\begin{aligned} \langle H \rangle &= N^2 \int_{-a}^a (a^2 - x^2) \left( -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2} m\omega^2 x^2 \right) (a^2 - x^2) dx \\ &= \frac{5\hbar^2}{4ma^2} + \frac{m\omega^2 a^2}{14} \end{aligned}$$

and minimise w.r.t.  $a$ :

$$\frac{\partial \langle H \rangle}{\partial a} = 0 \quad \text{when } a^4 = \frac{35\hbar^2}{2m^2\omega^2}.$$

Substituting back, this gives  $\langle H \rangle = 0.598\hbar\omega$ . The true ground state energy is  $0.5\hbar\omega$ : even our very simple guess at  $\psi$  leads to quite a close estimate of  $E_0$ .

### 10.2.1 Application to helium

In the helium atom, a suitable first guess for the state would be to give both electrons the same wavefunction of exponential form, with the distance scale of the exponential decay as a parameter:

$$\psi(r_1, \theta_1, \phi_1; r_2, \theta_2, \phi_2) = \frac{Z^3}{\pi a_0^3} e^{-Z(r_1+r_2)/a_0}. \quad [\text{trial wavefunction}] \quad (10.11)$$

We have set the distance scale  $a_0/Z$  in terms of a parameter  $Z$  which thus acts like an effective charge of the nucleus in the presence of screening. Make sure you understand the normalization of (10.11): we get 1 upon integrating over all possible positions for *both* electrons. We next evaluate  $\langle \psi | H | \psi \rangle$  using the full Hamiltonian (10.1). The integrals are tricky and will not be discussed here; they lead to the mean energy

$$E(Z) = -2E_R (4Z - Z^2 - 5Z/8)$$

where  $E_R$  is the Rydberg energy<sup>2</sup>. Upon minimizing  $E(Z)$  (for you to do—see problems), we obtain a minimum value  $-77.5$  eV. Note, this is energy of both electrons. If one electron were now removed from the system, the remaining electron would see a hydrogenic potential of nuclear charge 2 and hence the energy of the system would be  $-4E_R = -54.4$  eV. Hence our prediction for the ionization energy is  $77.5 - 54.4 = 23.1$  eV, which isn't too bad compared with the correct value of 25 eV.

The above gives a rough approximation to the ground state wavefunction and energy. To do better, considerable skill is needed in selecting appropriate trial wavefunctions and parameters. A general method called the *self consistent field* or Hartree method can be applied to helium and other atoms, this will be treated next year.

### 10.2.2 Main points

Schrodinger equation for helium and main features of the energy levels; comparison with hydrogen energy levels; configuration notation based on spherical symmetry; screening; variational method.

---

<sup>2</sup>We don't care whether full or reduced mass of the electrons is used, since the calculation is only approximate.

## Chapter 11

# Exchange symmetry and the Pauli exclusion principle

One of the most important, perhaps *the* most important idea in physics is that of *symmetry*. By this we mean not just mirror reflection, but symmetry as the general concept of a physical system staying unchanged when something is done to it. For example, a sphere is symmetric with respect to rotations, and the behaviour of any group of particles doesn't change when we displace the system in space, a property referred to as translational symmetry. The latter idea can be used, through a quantum mechanical argument, to derive that momentum is conserved! (See Feynman Lectures on Physics, vol 1, chapter 52). This is a truly staggering insight. It is also a strong hint at how powerful a concept symmetry is in physics, especially quantum physics.

The concept of symmetry allows us to see, for example, that the electric field produced by a spherically symmetric charge distribution must also be spherically symmetric. It must be, since if there were some non-symmetric kink in the field, *how would the electric field 'know' where the kink should go?* Indeed such a situation would be like a contradiction: if the field produced by some object were not spherically symmetric, it follows that the object can't have been spherically symmetric after all.

One can think of this in terms of information. The laws of the universe are cagey: Nature is like a poker player holding her cards close to her chest, never allowing any physical phenomenon to reveal more information in any situation than the minimum logically required by the conditions. So for example, a kink in an electric field gives information by picking out a direction in space, but if there is no direction picked out by the underlying charge distribution, this information has appeared from nowhere, which does not happen.

## 11.1 Identical particles

The symmetry we will be concerned with is concerned with identical particles. If two particles are identical with one another, then the system is not affected if we swap the particles; this is the symmetry operation we want to discuss. I recommend the excellent discussion in Shankar. These notes do not replace that, or another textbook, but they give the main points.

### 11.1.1 Classical situation

Consider the case of two billiard balls discussed in the lecture. In classical physics, we can imagine two identical billiard balls as an idealization: we know real billiard balls will differ slightly from one another in one way or another, but there is no problem with the abstract concept of a pair of identical balls in classical physics. It means two balls, which, no matter how carefully we examine them, we cannot tell apart.

Suppose we have two such balls near two pockets of a snooker table. We can refer to the balls by giving them labels 1 and 2. These labels do not refer to any *intrinsic* property of either ball, rather they refer to their environment. For example we can say ‘let’s call the ball near the bottom left hand pocket ball 1, and the ball near the bottom right hand pocket 2’. Suppose the balls are now caused to move up the table. We will consider two possible predictions for the outcome:

Prediction A: “Ball 1 ends up in top left pocket, ball 2 ends up in top right pocket.”

Prediction B: “Ball 1 ends up in top right pocket, ball 2 ends up in top left pocket.”

Now, if someone came to the snooker table after the balls have finished moving, and did not witness their trajectories, then they would have no way of telling which of predictions A and B was correct, because they cannot tell the balls apart.

However, in classical physics, each ball must have followed some particular trajectory, so anyone who carefully watched the balls moving along their trajectories would be able to tell which pocket ‘ball 1’ (the ball starting near the bottom left pocket) went to, and similarly for ‘ball 2’ (the one starting near the bottom right pocket). Such a person could tell which of predictions A and B was true, and which false.

Now, this argument does not require that anyone actually does the observations. The essential point is that the balls are distinguished by their *histories*. Therefore in classical physics, two configurations related by exchanging identical particles are physically nonequivalent. By ‘exchanging’ here we mean really physically swapping the particles around by some physical process. Later when we consider quantum systems, we will discuss the effects of exchanging the labels which we artificially assign to particles.

### 11.1.2 Quantum case: Exchange Principle

By contrast, in quantum mechanics, there is strictly no physical basis for distinguishing between identical particles. This is because it is not possible to keep track of particles through their trajectories: when wavefunctions of two identical particles overlap, it is strictly impossible to keep track of which particle is which.

This truth is encapsulated in the

**Exchange Principle:** Two physical configurations related by the exchange of identical particles must be treated as one and the same configuration and described by the same state vector.

More colloquially: if there is nothing to distinguish between two particles, then *the quantum state must not give any information which depends on which particle is which*.

We can mathematically express this idea as follows:

**Exchange principle:** If  $\psi(\{q\}_1, \{q\}_2)$  is the state of two identical particles which have sets of coordinates<sup>1</sup> (such as position and spin)  $\{q\}_1$  and  $\{q\}_2$ , where 1 and 2 label the particles, then

$$|\psi(\{q\}_1, \{q\}_2)|^2 = |\psi(\{q\}_2, \{q\}_1)|^2 \quad (11.1)$$

i.e. the modulus squared of the wavefunction does not change when the labels on the particles are exchanged. Note that we take the modulus squared of the wavefunction because the physically observable quantities depend on  $|\psi|^2$  not on  $\psi$  directly. This is important, because it leads to two possibilities, both of which are observed in Nature:

$$\text{either } \psi(\{q\}_1, \{q\}_2) = \psi(\{q\}_2, \{q\}_1) \quad (\text{Bosons}) \quad (11.2)$$

$$\text{or } \psi(\{q\}_1, \{q\}_2) = -\psi(\{q\}_2, \{q\}_1) \quad (\text{Fermions}) \quad (11.3)$$

If you discover a relationship like this, or a remarkable implication of it, then you might get half the particles in the universe named after you! The first case is called ‘symmetric with respect to exchange of particles’, the second case is called ‘antisymmetric with respect to exchange of particles’.

---

<sup>1</sup>Here the word “coordinates” is being used in a generic way, to refer to any complete set of physical parameters which are needed to specify a state, including both spatial and spin “coordinates”.

*Proof of (11.2), (11.3).* We need to show that the only possibilities in (11.2) and (11.3) are the plus sign or the minus sign, not some general complex number  $\exp(i\theta)$  of modulus 1. Suppose we measure some observable  $\hat{Q}$  on each of a pair of identical particles. We get two eigenvalues  $a$  and  $b$ , but we cannot say which particle had which eigenvalue. Let  $f(\mathbf{r}, \zeta)$  be the eigenstate of  $\hat{Q}$  of eigenvalue  $a$ , and  $g(\mathbf{r}, \zeta)$  be the eigenstate of  $\hat{Q}$  of eigenvalue  $b$ . Then the measured outcome is consistent with the product state

$$\phi_{1,2} = f(\mathbf{r}_1, \zeta_1) g(\mathbf{r}_2, \zeta_2)$$

and also with the product state

$$\phi_{2,1} = f(\mathbf{r}_2, \zeta_2) g(\mathbf{r}_1, \zeta_1).$$

There are no other product states consistent with the measured outcome, therefore the state of the pair of particles must be somewhere in the space spanned by  $\phi_{1,2}$  and  $\phi_{2,1}$ :

$$\psi(\mathbf{r}_1, \zeta_1, \mathbf{r}_2, \zeta_2) = \alpha\phi_{1,2} + \beta\phi_{2,1}. \quad (11.4)$$

The Exchange Principle requires

$$\psi(\mathbf{r}_1, \zeta_1, \mathbf{r}_2, \zeta_2) = \gamma\psi(\mathbf{r}_2, \zeta_2, \mathbf{r}_1, \zeta_1)$$

for some  $\gamma$  to be discovered. Hence

$$\alpha\phi_{1,2} + \beta\phi_{2,1} = \gamma(\alpha\phi_{2,1} + \beta\phi_{1,2}).$$

Equating coefficients of  $\phi_{1,2}$  and  $\phi_{2,1}$  we have

$$\alpha = \gamma\beta, \quad \beta = \gamma\alpha$$

which have the only solutions  $\gamma = \pm 1$ , hence the Bose and Fermi possibilities.

Our analysis has also showed us how to form wavefunctions of a pair of particles which satisfy the symmetry requirement. When  $\gamma = 1$  (Bosons) we have  $\alpha = \beta$  so the state is

$$\psi_S = \frac{1}{\sqrt{2}} (\phi_{1,2} + \phi_{2,1}). \quad (11.5)$$

When  $\gamma = -1$  (Fermions) we have  $\alpha = -\beta$  so the state is

$$\psi_A = \frac{1}{\sqrt{2}} (\phi_{1,2} - \phi_{2,1}). \quad (11.6)$$

You should yourself try swapping the labels for these states, and observe that  $\psi_A$  changes sign but  $\psi_S$  does not.

“Once a boson always a boson”: there does not exist in the universe any physical interaction which can cause a symmetric state (w.r.t. particle exchange) to evolve into an antisymmetric one, or vice versa<sup>2</sup>.

“All photons are photons.” Exchange symmetry is not just an artificial construct, it has real physical consequences, which we will discuss. This means that if any one pair of photons (for example) is found to be in a symmetric state, then it follows that all pairs of photons must be in symmetric states, since otherwise we could distinguish some photons from others.

---

<sup>2</sup>This follows from the fact that Hamiltonian operators can't depend on labeling of identical particles, see problems for proof.



We find in practice that particles of half-integer spin such as electrons and protons are fermions, while particles of integer spin such as photons are bosons. This relationship between spin and exchange symmetry is called the spin-statistics theorem, which may itself be derived from a subtle higher-order symmetry which is beyond the scope of this course.

**Comment: an atom on the far side of the moon**

Let's consider some implications of (11.2) and (11.3). To keep things simple we will concentrate on the spatial part of the total states (the illustrative examples in this section will assume pairs of particles whose spin state is symmetric under exchange of particle labels). Suppose we have a photon in some state  $\phi(\mathbf{r}_1)$ . Suppose another photon is then placed in the same state  $\phi(\mathbf{r}_2)$  (this means the second photon overlaps the other, having the same frequency and direction). The total state of the two photons is then

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \phi(\mathbf{r}_1) \phi(\mathbf{r}_2) \quad (11.7)$$

Now, although we called the photons 1 and 2, there is nothing in this state which depends on which photon is called 1, and which is called 2. It is therefore a possible state and it obeys the boson relation (11.2). Photons from a laser are like this: lots of them in the same quantum state.

Now consider two different states  $f$  and  $g$  for a photon. For example, these could be states of the same frequency and polarization, but different direction (think of two laser beams shining onto a screen with a small angle between them). If the first photon is in state  $f(\mathbf{r}_1)$ , then the second photon is in  $g(\mathbf{r}_2)$ . However, who told us which photon to put in  $f$ , and which in  $g$ ?—the state could equally be  $f(\mathbf{r}_2)$  combined with  $g(\mathbf{r}_1)$ . When there is nothing to choose between two cases, both occur, and the true state of the photons will be

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} (f(\mathbf{r}_1) g(\mathbf{r}_2) + f(\mathbf{r}_2) g(\mathbf{r}_1)). \quad (11.8)$$

Make sure you convince yourself that this state obeys the Bose relationship (11.2).

Now consider the experiment we are discussing: two laser beams illuminating a screen. The light hitting the screen is composed of two plane waves with a small angle between them. This is like the Michelson interferometer with the mirrors at a small angle, or the Young's slits experiment: an interference pattern of straight lines will be observed on the screen. The pattern is still observed even when the intensity of the light emitted by the lasers is so low that less than one photon at a time, on average, is present in the system (we can arrange this without disturbing the lasers' operation by attaching thick absorption filters to both laser outputs). Now, if each of the photons arrived simply in one beam or the other, then we would never have an interference effect. However, fringes are observed. Think hard about the fact that this is a really remarkable observation: if each photon emerges from *either* one laser *or* the other, which is surely what we would guess, then no interference fringes would be seen. Yet they are seen. This seems to imply that each photon was somehow created by both lasers at once, even though the lasers never interact. Indeed, we could imagine in principle that they are situated light-years apart.

In case you are not sufficiently impressed by this observation, perhaps because you think photons are insubstantial, you may be interested to learn that the same experiment can be done, and has been done<sup>3</sup>, with independently prepared beams of Bosonic atoms. A large number of atoms are prepared all in the same state (by Bose-Einstein condensation) in a small trap, and a second group is *independently* prepared and an adjacent trap. When the atoms are released they fall under gravity, spreading out until the two clouds overlap. They arrive one by one at a position-sensitive detector, and their arrival positions are observed to form a pattern of lines, consistent with interference of their de Broglie waves. Once again even if the atoms fall one by one, the interference is still observed (as long as it is not possible to tell which trap a given atom fell from).

<sup>3</sup>M. R. Andrews, C. G. Townsend, H.-J. Miesner, D. S. Durfee, D. M. Kurn, W. Ketterle, "Observation of Interference Between Two Bose Condensates", *Science* **275**, 637-641 (1997).

Exchange symmetry provides an explanation of this phenomenon. It is not true that two non-interacting sources can collaborate in the emission of a single particle, but they can independently emit identical particles. This means that when a particle is detected, it can be the case (and it is the case in this experiment) that it is strictly impossible to identify which source emitted the particle. One might describe the result as implying that a given atom was emitted by both sources, but by saying “a given atom” we have already committed the error of implying that one atom (of given properties) can be distinguished from another. Rather, we ought to keep clear in the discussion the fact that both sources exhibit a propensity to emit atoms, and therefore the total wavefunction of the atoms has terms in it accounting for each atom being emitted from all available sources, merely from the exchange symmetry requirement<sup>4</sup> (c.f. eq. (11.8)). The implication is that the atoms in your coffee cup have some sort of shared existence with the atoms on the far side of the moon<sup>5</sup>.

### 11.1.3 Pauli Exclusion Principle

Consider the case of two fermions such as two electrons. This time the total state is of type (11.3). Suppose two electrons are both in some state  $\phi$ . Then the total state would be  $\psi = \phi(\mathbf{r}_1, \zeta_1)\phi(\mathbf{r}_2, \zeta_2)$ , but this has the wrong symmetry so is not allowed according to (11.3). To understand what is going on, take the case where the two electrons are in different states  $f$  and  $g$ . Then the total state, with the correct symmetry (i.e. antisymmetric with respect to exchange) is

$$\psi(\mathbf{r}_1, \zeta_1, \mathbf{r}_2, \zeta_2) = \frac{1}{\sqrt{2}} (f(\mathbf{r}_1, \zeta_1)g(\mathbf{r}_2, \zeta_2) - f(\mathbf{r}_2, \zeta_2)g(\mathbf{r}_1, \zeta_1)). \quad (11.9)$$

Now, observe what happens if we let  $g$  become the same as  $f$ : in that case  $\psi$  vanishes. The upshot of this is

**The Pauli Exclusion Principle:** *no two identical Fermions can ever be in the same quantum state.*

The Pauli Exclusion Principle underlies atomic structure through and through, and has many consequences in many areas of physics.

## 11.2 Example to illustrate exchange effects

Unfortunately, although exchange symmetry is a beautiful and fascinating concept, many textbooks contain misleading and sometimes even wrong statements about it. For example, sometimes the spin degree of freedom is left out, or only included in an unclear way as an afterthought. Sometimes energy level splittings are ascribed to exchange effects, even though they would still be present if the particles were non-identical.

To clarify these points, let us consider a simple situation and compare the behaviours when there are either two non-identical particles, or two identical Bosons, or two identical Fermions. The situation will be a pair of particles confined in a one-dimensional box, and interacting with one another by a repulsive force which depends on their separation. We assume the repulsive force is such that it increases when the particles are close, and its effects will be described by a potential energy function  $V(|x_1 - x_2|)$ . To be specific, you could imagine a Coulomb potential if you like, but in any case to keep the mathematics simple, we will assume  $V$  only makes a small contribution to the total energy.

<sup>4</sup>N.B. it is a very severe test of our physical intuition to arrive at the clearest way to describe this experiment in terms of particles: your tutors will probably not all agree on which statements are least misleading, though they will agree on the fringe pattern observed.

<sup>5</sup>This implication may be questioned if it is not possible to bring these atoms together in a controlled way that leaves strictly no evidence of their previous history, since then they will remain effectively labelled by their history, so that exchange symmetry effects will be unobservable. A full discussion of this brings us to the interpretational problems of quantum theory which remain unresolved and which are outside the remit of these notes.

We assume the two particles have the same mass  $M$ . The Schrödinger equation for our problem is

$$-\frac{\hbar^2}{2M} \left( \frac{\partial^2 \psi}{\partial x_1^2} + \frac{\partial^2 \psi}{\partial x_2^2} \right) + (U(x_1) + U(x_2) + V(|x_1 - x_2|)) \psi = E\psi \quad (11.10)$$

where  $U(x)$  is the potential energy for the box, which we take to be an infinite potential well of width  $a$ .

### 11.2.1 Non-identical particles

If the two particles are not identical (e.g. think of a neutron and a proton, which have almost the same mass, but which are not identical), then we can simply ignore the requirements of the Exchange Principle: they do not arise.

Solving Schrödinger's equation, we see immediately we can separate spatial and spin parts, so the work to be done is to solve the spatial part. In the limit where  $V$  is small, the problem is separable and reduces to the familiar infinite potential well, with eigenfunctions and energies

$$u_n^0(x) = \sqrt{\frac{2}{a}} \sin(n\pi x/a), \quad E_n^0 = \frac{\hbar^2 \pi^2}{2Ma^2} n^2. \quad (11.11)$$

For the pair of particles in the well, if  $V$  were zero, a complete set of solutions would be

$$\psi_{n,m}^0(x_1, x_2) = u_n^0(x_1)u_m^0(x_2), \quad E_{n,m}^0 = \frac{\hbar^2 \pi^2}{2Ma^2} (n^2 + m^2). \quad (11.12)$$

In this complete set, notice that when  $n \neq m$  then there is a pair of degenerate states, namely  $u_n^0(x_1)u_m^0(x_2)$  and  $u_m^0(x_1)u_n^0(x_2)$ , both of which have energy  $(n^2 + m^2)\hbar^2\pi^2/2Ma^2$ .

Now allow  $V$  to be non-zero. Since  $V$  is positive we expect all the energy levels to be increased, and we are about to find that in the presence of  $V$ , the degeneracy is lifted. Analyzing this by degenerate perturbation theory, we need to select zero-order states in each degenerate space so that they are not connected by  $V$ . The appropriate combinations are

$$\psi_{n,m,+}^0 = (u_n^0(x_1)u_m^0(x_2) + u_m^0(x_1)u_n^0(x_2)) / \sqrt{2} \quad (11.13)$$

$$\text{and } \psi_{n,m,-}^0 = (u_n^0(x_1)u_m^0(x_2) - u_m^0(x_1)u_n^0(x_2)) / \sqrt{2}. \quad (11.14)$$

(You should confirm that  $\langle \psi_{n,m,+}^0 | V | \psi_{n,m,-}^0 \rangle = 0$ .) In the presence of  $V$ , therefore, the energy eigenstates are  $\psi_{n,m,+}^0$  and  $\psi_{n,m,-}^0$  plus small corrections.

Now consider the energies. Each pair  $\psi_{n,m,+}^0$  and  $\psi_{n,m,-}^0$  were degenerate before the perturbation was introduced. To calculate their first order energy shift we need to calculate

$$\delta E_+ = \langle \psi_{n,m,+}^0 | V | \psi_{n,m,+}^0 \rangle \quad (11.15)$$

$$\text{and } \delta E_- = \langle \psi_{n,m,-}^0 | V | \psi_{n,m,-}^0 \rangle. \quad (11.16)$$

Both of these are shifts upwards in energy, because  $V > 0$ , but by different amounts. Without troubling to do the integrals, we can see which will be larger by a physical argument. We know that  $V$  gets larger when the particles are close together. Therefore, whichever state has the particles closer together on average will have the greater energy shift contributed by  $V$ . Now, the particles are close together when  $x_1 \simeq x_2$ , so we need to examine the probability densities represented by  $|\psi_{n,m,+}^0|^2$  and  $|\psi_{n,m,-}^0|^2$  in the region where  $x_1 \simeq x_2$ . For example, when  $x_1 = x_2$ , we can choose to call both positions simply  $x$ , and we have

$$\psi_{n,m,+}^0(x_1 = x_2 = x) = \sqrt{2} u_n^0(x)u_m^0(x) \quad (11.17)$$

$$\psi_{n,m,-}^0(x_1 = x_2 = x) = \frac{1}{\sqrt{2}} (u_n^0(x)u_m^0(x) - u_m^0(x)u_n^0(x)) = 0 \quad (11.18)$$

Thus in state '+', the particles can readily be close together, while in state '-', they are highly unlikely to be.

The overall conclusion is that

- the particles are on average closer together when they are in state  $\psi_{n,m,+}^0$  than when they are in state  $\psi_{n,m,-}^0$ .

This means the energy shift is larger for  $\psi_{n,m,+}^0$  than for  $\psi_{n,m,-}^0$ , hence in the presence of  $V$ , state  $\psi_{n,m,+}^0$  has the greater energy.

Now, it won't have escaped your notice that the states we have been talking about have definite symmetry with respect to exchange of particle labels. However, they arose simply as solutions to Schrödinger's equation, not from the Exchange Principle. In fact, we should now recall that we have not discussed the spin state of our pair of particles. Since the particles are not identical, and the spin does not appear in the Hamiltonian, it follows that any spin state is possible, so in fact the full state including spin might *not* have definite symmetry w.r.t. exchange of particle labels.

The argument relied on perturbation theory, and you may be wondering what happens to the state when we include all orders of perturbation theory, and thus get an exact solution. Notice that the interaction  $V(|x_1 - x_2|)$  is strictly unchanged when we swap particle labels. It doesn't even change sign. From this it is easy to show that  $V$  does not connect states of different exchange symmetry (try it yourself now, or see problems for general proof). Therefore, in all the higher order corrections to the states in perturbation theory, only symmetric contributions are added to  $\psi_{n,m,+}^0$  and antisymmetric contributions are added to  $\psi_{n,m,-}^0$ . The exact solutions therefore come in pairs, one symmetric, one antisymmetric. The changes introduced by the perturbation do mean the states are not exactly equal to  $\psi_{n,m,+}^0$  and  $\psi_{n,m,-}^0$ . Instead they become

$$\begin{aligned} \psi_{n,m,+} &= (u_n(x_1)u_m(x_2) + u_m(x_1)u_n(x_2)) / \sqrt{2} \\ \text{and } \psi_{n,m,-} &= (v_n(x_1)v_m(x_2) - v_m(x_1)v_n(x_2)) / \sqrt{2}. \end{aligned} \quad (11.19)$$

where the functions  $v$  need not be the same as the functions  $u$  (and neither are the same as  $u^0$ ).

The exchange symmetry of energy eigenstates can be analysed in a more general way by introducing an *exchange operator*  $\mathcal{X}$  which is defined such that its influence on any state is to swap the labels attached to a given pair of particles. It is easy to show  $\mathcal{X}^2 = 1$  and from this that  $\mathcal{X}$  has eigenvalues  $\pm 1$ . The eigenstates of eigenvalue  $+1$  are symmetric w.r.t. exchange, those of eigenvalue  $-1$  are antisymmetric w.r.t. exchange. It can be further be shown that when the Hamiltonian is unchanged by swapping particle labels, then  $\mathcal{X}$  commutes with the Hamiltonian,  $[\mathcal{X}, H] = 0$ . In this case  $\mathcal{X}$  and  $H$  can have joint eigenstates (see problems).

### 11.2.2 Identical Bosons

Now suppose the two particles in our box are identical bosons, such as alpha-particles. In this case the solution of the Schrödinger equation proceeds *exactly* as before. Once we have our set of solutions, the Exchange Principle can be invoked to see which of them are physically possible.

At first sight, since we are now discussing identical bosons, we might feel that the eigenstates  $\psi_{n,m,-}^0$  are not allowed, because they have the wrong exchange symmetry. However, we should recall the spin degree of freedom as well. The total state of the pair of particles is

$$\Psi = (\text{spatial state}) \otimes (\text{spin state})$$

We have seen that the spatial states under consideration each have definite symmetry w.r.t exchange of particles, some of them symmetric, and some antisymmetric. For Bosons we require the overall state, when all labels are exchanged (i.e. including both spatial and spin degrees of freedom), to be symmetric. There are two possibilities:

either (symmetric spatial state)  $\otimes$  (symmetric spin state)  
 or (antisymmetric spatial state)  $\otimes$  (antisymmetric spin state)

Now, in practice alpha particles will usually have spin zero, in which case the spin state can only be symmetric, and then only the first possibility occurs. However, there exist Bosons of non-zero spin, such as photons and various atoms and nuclei, and for them both possibilities will be observed.

In more general physical problems, it is found that any Hamiltonian which describes identical particles must have exchange symmetry, and therefore  $[\mathcal{X}, H] = 0$ . This implies the energy eigenstates can have definite symmetry w.r.t. exchange, so at least some of them are allowed under the requirements of the Exchange Principle. However, we don't always require that the spatial and spin parts of the state appear in the form of a product. If the Hamiltonian involves a coupling between spatial and spin degrees of freedom, as in spin-orbit coupling, for example, then the eigenstates will not usually separate. Instead the overall state is a combination of spatial and spin parts which cannot be written as a product, but which still retains overall symmetry w.r.t. exchange of particle labels. To have an example for illustration, here is such a state:

$$\frac{1}{\sqrt{2}} (f(\mathbf{r}_1) |+\rangle_1 g(\mathbf{r}_2) |-\rangle_2 + f(\mathbf{r}_2) |+\rangle_2 g(\mathbf{r}_1) |-\rangle_1) \quad (11.20)$$

where the  $|\pm\rangle$  states are the spin-one,  $m_s = \pm 1$  spin states.

### 11.2.3 Identical Fermions

The treatment of our example problem for a pair of Fermions is now straightforward. The Schrödinger equation has again the same solutions for the spatial part of the state and for the energies. The overall result must be antisymmetric, so the possibilities are

either (symmetric spatial state)  $\otimes$  (antisymmetric spin state)  
 or (antisymmetric spatial state)  $\otimes$  (symmetric spin state)

In the case where the two fermions have spin half, which is the most common situation, then the spin states are easy to analyse. They were given in section 5.4.1. It is seen that the 'singlet' (total spin  $S = 0$ ) is antisymmetric w.r.t. exchange, and the 'triplet' (total spin  $S = 1$ ) is symmetric w.r.t. exchange. Therefore for two identical spin-half particles, whenever the space and spin parts are separable, we have

either (symmetric spatial state)  $\otimes$  (singlet ( $S = 0$ ) spin state)  
 or (antisymmetric spatial state)  $\otimes$  (triplet ( $S = 1$ ) spin state)

In more general problems, the further comments made above about bosons apply. The state does not always have to be separable into spatial and spin parts, but for fermions it must be overall antisymmetric w.r.t. exchange of particle labels.

### 11.2.4 Discussion

We examined a physical problem where two particles were in a common potential well, and the particles had an interaction with one another. There are many examples in physics of this generic situation, such as the electrons in the helium atom, the electrons in other atoms with larger atomic number, nucleons in a nucleus, etc. It is sometimes asserted that the observation of a splitting of the energy levels into pairs, associated with eigenstates whose spatial part has definite symmetry, is ‘because of the exchange interaction’. However, this is quite wrong, because the splitting, and the symmetric and antisymmetric (spatial part of the) energy eigenstates, is observed for non-identical particles too. There is *no* ‘exchange interaction’ as such. That is, there is no source of potential energy associated with the exchange requirements for identical particles. The only interactions in the universe are the familiar Coulomb interaction, the gravitational interaction, the strong interaction, and the weak interaction.

While the Exchange Principle does *not* affect the energy values, what it does do is *restrict the available states*. Thus, for example, we saw that for non-identical particles in the situation we have discussed, there was no restriction on the spin states, but for identical particles there was: a spatial state of given symmetry could only be associated with a particular class of spin states.

Since there is no ‘exchange interaction’, there are no ‘exchange forces’. However, both terms have acquired a certain currency in physics. They are used to refer to the following idea. Owing to the restrictions on the type of states which are imposed by symmetry requirements, the mean separation between pairs of electrons *of given spin state* will depend on the spin state. That is, other things being equal (i.e. in a given environment) the states with symmetric spin part will have a spatial contribution such that the electrons are on average further apart, and the states with antisymmetric spin will have a spatial contribution such that the electrons are on average closer together. This is not owing to any change in the mean position of either electron alone, it is because of a *correlation* between their positions.

For a physical picture of the correlation, imagine two classical particles moving to and fro in the same one-dimensional potential well and with the same amount of energy. Each particle has the same average position, and the same excursion in position. However, if the particles are moving in opposite directions, then much of the time they are far apart, but if they are moving in the same direction, and begin close to one another, then on average they remain close. Thus the nature of the *correlation* between the motions of the two particles affects the mean value of the separation.

Owing to the different mean separation, the Coulomb repulsion will cause the state having symmetric spin part to have less energy. As a short-hand way of speaking, the contribution to the total Coulomb energy which depends on the symmetry of the state is said to arise from an ‘exchange force’. When the two single-particle wavefunctions involved in the superposition  $\psi^+$  or  $\psi^-$  do not appreciably overlap one another, this contribution is negligible. In many situations, wavefunctions fall off exponentially with distance: in this case so will the ‘exchange force’.

### 11.2.5 More than two particles

States of three or more identical particles must be either symmetric (bosons) or antisymmetric (fermions) with respect to exchange of any pair. Suppose energy measurements of three identical particles show that they have energies  $E_f$ ,  $E_g$  and  $E_h$ . The total state must then involve products of the form  $f(x_i)g(x_j)h(x_k)$  where  $f, g, h$  are the energy eigenstates and  $i, j, k \in \{1, 2, 3\}$  label the particles, but there is no way to say which particle has which energy. The task of finding a state with all the right exchange symmetry might be quite difficult for many particles, but fortunately the properties of determinants do the job for us. The joint state of three fermions is given by

$$\psi_{a,b,c}(x_1, x_2, x_3) = \frac{1}{(3!)^{1/2}} \begin{vmatrix} f(x_1) & g(x_1) & h(x_1) \\ f(x_2) & g(x_2) & h(x_2) \\ f(x_3) & g(x_3) & h(x_3) \end{vmatrix} \quad (11.21)$$

$$= (3!)^{-1/2} (f(x_1)g(x_2)h(x_3) - f(x_1)h(x_2)g(x_3) + g(x_1)h(x_2)f(x_3) - g(x_1)f(x_2)h(x_3) + h(x_1)f(x_2)g(x_3) - h(x_1)g(x_2)f(x_3)) \quad (11.22)$$

In atomic physics, such determinants are called Slater determinants. You should check that the resulting state is indeed antisymmetric w.r.t. exchange of any pair—in fact it follows immediately from standard properties of determinants.

To find a totally symmetric state, write the same sequence of product states as given by the determinant, but don’t include the sign changes associated with the determinant.

To include spin, a method which always succeeds in generating states of the right symmetry is simply to include it in each state in the determinant. For example

$$\frac{1}{(2!)^{1/2}} \begin{vmatrix} f(x_1) \uparrow_1 & g(x_1) \downarrow_1 \\ f(x_2) \uparrow_2 & g(x_2) \downarrow_2 \end{vmatrix} = \frac{1}{2^{1/2}} (f(x_1) \uparrow_1 g(x_2) \downarrow_2 - f(x_2) \uparrow_2 g(x_1) \downarrow_1) \quad (11.23)$$

is an allowed state for two spin-half particles, in (what I hope is) an obvious notation. This state is not separable into a product of spatial and spin parts. Space/spin-separable states can be found for two particles, we already met some examples above. For example

$$\frac{1}{2} \begin{vmatrix} f(x_1) \uparrow_1 & g(x_1) \downarrow_1 \\ f(x_2) \uparrow_2 & g(x_2) \downarrow_2 \end{vmatrix} + \frac{1}{2} \begin{vmatrix} f(x_1) \downarrow_1 & g(x_1) \uparrow_1 \\ f(x_2) \downarrow_2 & g(x_2) \uparrow_2 \end{vmatrix} = \frac{1}{2} (f(x_1)g(x_2) - f(x_2)g(x_1)) (\uparrow_1 \downarrow_2 + \uparrow_2 \downarrow_1). \quad (11.24)$$

However, such a space/spin separation is rare for allowed states of three or more identical particles, even though the Hamiltonian may be separable.

### 11.2.6 Exchange symmetry of eigenstates of summed operators

We are often interested in physical properties which have the form of a sum of single-particle observables. For example, we may be interested in the total kinetic energy  $\sum_i \hat{p}_i^2/2m$  or the total orbital angular momentum

$$\hat{\mathbf{L}} = \hat{\mathbf{l}}_1 + \hat{\mathbf{l}}_2 + \hat{\mathbf{l}}_3 + \cdots. \quad (11.25)$$

I have written this out explicitly in order to make it obvious that this operator, and more generally all operators in the form of such a simple sum, do not change when the particle labels are swapped around. From this it follows that they commute with the exchange operator and therefore their eigenstates have definite exchange symmetry. This means that states which are allowed according to the exchange symmetry rules can be eigenstates of the total properties of the system such as  $\hat{L}^2$  and  $\hat{L}_z$ , while usually they are not eigenstates of individual operators such as  $\hat{l}_{z1}$ .

This feature is important, because it underlies the notion of a composite system: it confirms that for many purposes a composite system can be treated as a single entity with overall properties, without worrying about its internal structure. It is hard to imagine a physical world without the property of allowing well-defined composite systems. Our ability to make sense of our surroundings even in the most basic way for everyday life depends on it.

### 11.2.7 When can identical particles be treated as if they were non-identical?

Until this subject of exchange symmetry came up, we happily did our classical mechanics without ever worrying about it. We were happy to say things like ‘an atom was emitted by the oven and passed through the magnetic field and hit the detector’, and we did not find that this way of thinking gave misleading predictions. Furthermore, even in the quantum mechanical discussion of the electron in hydrogen, we did not worry about all the other electrons in the universe. What these examples illustrate is that for many purposes identical particles can be treated as if they could be told apart. This fact is highly important to our ability to make sense of our surroundings: it would be very confusing if every time we picked up an object on the earth, its behaviour could only be made sense of by taking into consideration other objects on the moon, or elsewhere in the universe.



For a more detailed discussion of this important point, I refer you once again to Shankar or other textbooks. The essence of the result is

- If, during the course of a given sequence of events, two or more particles always occupy states with negligible overlap, then the particles are said to be *distinguishable* and their behaviour can be accurately modelled by treating them as if they were non-identical.

To clarify what is meant by “occupy states with negligible overlap”, let us consider an example. Consider a pair of spin zero particles, one on the earth, the other on the moon. We could in principle measure the system and learn that near the earth there was just one particle, and near the moon there was just one particle, without learning which particle was which. Assuming this has been done, a possible state is

$$\psi_{EM,S} = 2^{-1/2}(\phi_E(\mathbf{r}_1)\phi_M(\mathbf{r}_2) + \phi_E(\mathbf{r}_2)\phi_M(\mathbf{r}_1)) \quad (11.26)$$

where  $\phi_E$  and  $\phi_M$  are two example wavefunctions, such that  $\phi_E$  is peaked near the earth and  $\phi_M$  is peaked near the moon.

Now, if the particles were not identical, then a measurement could in principle tell us not only that one was near the earth and one near the moon, but also which was which. Instead of labelling them 1 and 2, we could call them Fred and Kate. A possible state would then be

$$\psi_{EM} = \phi_E(\mathbf{r}_F)\phi_M(\mathbf{r}_K) \quad (11.27)$$

in the case where Fred is found on the earth, and Kate on the moon.

Now, the essential idea is that *as long as the wavefunctions peaked near E and M do not overlap with one another*, then the predictions using  $\psi_{EM,S}$  (equation (11.26)) are the same as the predictions using  $\psi_{EM}$  (equation (11.27)). For, all the observable quantities can be written as matrix elements  $\langle \psi | \hat{Q} | \psi \rangle$ . Upon evaluating such matrix elements, state (11.26) includes ‘cross terms’ of the form

$$\langle \phi_E(\mathbf{r}_1)\phi_M(\mathbf{r}_2) | \hat{Q} | \phi_M(\mathbf{r}_1)\phi_E(\mathbf{r}_2) \rangle \quad (11.28)$$

while state (11.27) does not. Assuming the operator  $\hat{Q}$  does not dramatically alter  $\phi_M$  so that it overlaps with  $\phi_E$ , the cross term is close to zero and can be neglected. When we recall that wavefunctions centred in potential wells typically *decay exponentially* with distance outside the well, we see that cross terms really can be absurdly small and to neglect them is not just ok, but quite correct. To include them would be ridiculous because physics is not an exact science anyway—I mean not *exactly* exact.

The concept of distinguishability is most important in statistical mechanics, where you have already met it. There, its main use is as an aid to counting orthogonal states of a system of several particles. If the particles are distinguishable, then it can be shown that the number of properly symmetrized states is the same as the number of product states. The latter are easier to count, so this is a useful observation. For an example, consider the available spin states of two spin-half nuclei in two different atoms. We argue that the nuclei are distinguishable because they are located in different atoms, and therefore we may count product spin states. These are up-up, up-down, down-up and down-down: a total of four. This is the correct total, although the states available to the system are really

$$\begin{aligned} & \psi_S(\text{spatial}) \otimes (\text{singlet}), \\ & \psi_A(\text{spatial}) \otimes (\text{triplet}, M=1), \\ & \psi_A(\text{spatial}) \otimes (\text{triplet}, M=0), \\ \text{and } & \psi_A(\text{spatial}) \otimes (\text{triplet}, M=-1). \end{aligned}$$

### 11.3 Exchange symmetry in helium

In chapter 10 we analysed the Helium atom using a central potential model which included the attraction to the nucleus and which included in an approximate way most of the electron-electron repulsion energy. Now that we are familiar with exchange symmetry, we can go further.

First, notice that the exact Hamiltonian must be symmetric w.r.t. exchange of the two electrons, therefore the true energy eigenstates don't have the form (10.5) but rather they come in pairs of the form

$$\begin{aligned} \psi_S &= \frac{1}{\sqrt{2}} (\phi_{1s}(r_1) \phi_{nlm}(r_2, \theta_2, \phi_2) + \phi_{1s}(r_2) \phi_{nlm}(r_1, \theta_1, \phi_1)) \\ \text{and } \psi_A &= \frac{1}{\sqrt{2}} (\phi_{1s}(r_1) \phi_{nlm}(r_2, \theta_2, \phi_2) - \phi_{1s}(r_2) \phi_{nlm}(r_1, \theta_1, \phi_1)). \end{aligned}$$

Also, owing to the smaller mean separation between the electrons in state  $\psi_S$  compared to state  $\psi_A$ , the former will have more energy (from Coulomb repulsion between electrons) than the latter<sup>6</sup>.

Everything we have said so far would be so for non-identical particles. However, electrons are in fact identical, and the Exchange Principle restricts the states such that  $\psi_S$  is only associated with an antisymmetric spin state, while  $\psi_A$  is only associated with a symmetric spin state. Spin states for a pair of spin-half particles were listed in section 5.4.1. It is seen that the 'singlet' is antisymmetric w.r.t. exchange, and the 'triplet' is symmetric w.r.t. exchange.

Hence the overall conclusion is that, for each pair of states in helium,

- the one with lower energy must be a triplet, the one with higher energy must be a singlet.

It is possible to confirm this experimentally by using the Zeeman effect.

For the ground state of helium  $1s^2$ , we have a special case because the two wavefunctions  $\phi_{1s}$  and  $\phi_{nl}$  are one and the same wavefunction. This means the antisymmetric form  $\psi_A$  vanishes, and we only have the symmetric form

$$\psi_S(\text{ground state}) = \phi_{1s}(r_1) \phi_{1s}(r_2).$$

The Exchange Principle then requires that the associated spin state must be the singlet. Another way of putting it is to say the  $1s^2$  triplet state is ruled out by the Pauli Exclusion Principle.

<sup>6</sup> Actually, there is a subtlety here. Whereas the two states must each have definite symmetry, they don't necessarily have to have the form  $f(1)g(2) \pm f(2)g(1)$ . Rather, if the symmetric state is  $f(1)g(2) + f(2)g(1)$ , then the antisymmetric one can be  $u(1)v(2) - u(2)v(1)$  where  $u$  and  $v$  are functions different from  $f$  and  $g$ , recall equations (11.19). This is what really happens in helium, and the lower energy of the triplet state is mostly because the electrons are closer to the nucleus, rather than further from each other. This point is not treated in textbooks, and is off-syllabus for us.

These conclusions are remarkable, because they say the energy of the helium atom depends on its spin state, even though the Hamiltonian does not depend on spin! The first time one meets the explanation for this remarkable fact, the argument seems long and tortuous. However, once we have got the hang of it, it is not so bad. The summary is as follows.

1. States which are degenerate under the central field Hamiltonian are shifted and split by the electron-electron interaction Hamiltonian.
2. The spatial parts of the eigenstates of the total Hamiltonian have definite symmetry w.r.t. exchange of the electrons, the symmetric spatial state having more energy than the antisymmetric one.
3. The requirement of asymmetry for identical fermion states means that each spatial state must be associated with a specific spin state.
4. The combination of (1)-(3) leads to an energy splitting between spin singlet and triplet states, even though no spin-dependent interactions are involved.
5. The singlet terms are the higher in energy since they are antisymmetric in spin, so must have symmetric spatial part, and when the spatial state is symmetric, the mean separation of the electrons is smaller, leading to a higher energy of repulsion between them.

One can readily see that the above summary will carry over to the case of many electrons. This explains the splitting between singlet and triplet states observed in all atoms from group 2 of the periodic table, and related observations in all atoms of more than one electron.

### Main points

The main points of this chapter have been:

identical particles; Exchange Principle; symmetry w.r.t. particle exchange: Bosons and Fermions; Pauli Exclusion Principle; symmetric and antisymmetric spatial states and their energy splitting; implications for spin state from Exchange Principle; application to helium atom: each triplet state lies lower in energy than the corresponding singlet; brief comments on the extension to many particles; concept of distinguishability and wavefunction overlap.

## 11.4 Appendix: detailed derivation of states and splitting

I will do the calculation for helium, starting by neglecting the electron-electron repulsion all together to begin with. This serves to bring out the features of the calculation, which we will then make more precise at the end. The level of this material is just about on the syllabus: you would not be expected to reproduce the full derivation in detail, but you should be familiar with the main points.

So, considering just the attraction to the nucleus, we have the Hamiltonian

$$H_0 = \left( \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} - \frac{Ze^2}{4\pi\epsilon_0 r_1} - \frac{Ze^2}{4\pi\epsilon_0 r_2} \right) \quad (11.29)$$

where  $Z = 2$ . Let  $f$  and  $g$  be eigenstates of the single-electron Hamiltonian  $\mathbf{p}^2/2m - Ze^2/4\pi\epsilon_0 r$ , with eigenvalues  $E_a$  and  $E_b$  respectively. Then, by the fact that  $H_0$  is separable, the product states

$$\begin{aligned} u_{fg} &\equiv f(\mathbf{r}_1)g(\mathbf{r}_2), \\ u_{gf} &\equiv g(\mathbf{r}_1)f(\mathbf{r}_2), \end{aligned}$$

are eigenstates of  $H_0$ , both having the eigenvalue  $E_0 = E_f + E_g$ . Note that the separation of variables method has converted our single Schrödinger equation into the pair of equations

$$\left. \begin{aligned} H_0 u_{fg} &= E_0 u_{fg} \\ H_0 u_{gf} &= E_0 u_{gf} \end{aligned} \right\} \quad (11.30)$$

If we write a general combination of  $u_{fg}$  and  $u_{gf}$  as  $\alpha u_{fg} + \beta u_{gf}$  where  $\alpha$  and  $\beta$  are any complex numbers satisfying  $|\alpha|^2 + |\beta|^2 = 1$ , then the pair of Schrödinger equations can be expressed as a single matrix equation

$$\begin{pmatrix} \langle fg|H_0|fg\rangle & \langle fg|H_0|gf\rangle \\ \langle gf|H_0|fg\rangle & \langle gf|H_0|gf\rangle \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E_0 \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad (11.31)$$

which is just the simple expression

$$\begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E_0 \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \quad (11.32)$$

We have a degeneracy (two states of the same energy), which will be removed by the electron-electron interaction  $H^r = e^2/4\pi\epsilon_0 r_{12}$ . Let's assume that this interaction is small enough that the energy shifts of the initially degenerate states  $u_{fg}$  and  $u_{gf}$  are primarily due to the coupling between them, rather than to the influence of further states of energy  $\neq (E_a + E_b)$ . Hence we can approximate the problem as a two-state system, with the Schrödinger equation

$$\begin{pmatrix} E_0 + H_{11}^r & H_{12}^r \\ H_{21}^r & E_0 + H_{22}^r \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

where

$$\begin{aligned} H_{11}^r &= \langle fg|H^r|fg\rangle = \int f^*(\mathbf{r}_1)g^*(\mathbf{r}_2)H^r f(\mathbf{r}_1)g(\mathbf{r}_2)dV_1dV_2 \\ &\equiv J \quad [\text{'direct' integral}] \\ H_{22}^r &= \langle gf|H^r|gf\rangle = \int g^*(\mathbf{r}_1)f^*(\mathbf{r}_2)H^r g(\mathbf{r}_1)f(\mathbf{r}_2)dV_1dV_2 \\ &= H_{11}^r \quad [\text{since } H^r \text{ only depends on relative coordinate } r_{12}] \\ H_{12}^r &= \langle fg|H^r|gf\rangle = \int f^*(\mathbf{r}_1)g^*(\mathbf{r}_2)H^r g(\mathbf{r}_1)f(\mathbf{r}_2)dV_1dV_2 \\ &\equiv K \quad [\text{'exchange' integral}] \\ &= H_{21}^r \quad [\text{real and Hermitian } H^r] \end{aligned}$$

The names 'direct' and 'exchange' integrals come from the fact that the direct integral can be interpreted as the mutual energy caused by repulsion between two overlapping charge distributions  $|f|^2$  and  $|g|^2$ , while the exchange integral does not have such a simple interpretation. This is simply an example of the problems which can arise when trying to interpret a two-particle quantum state in terms of single-particle states. It has become common practice to call the off-diagonal part of  $H^r$  an 'exchange interaction', but this is to some extent misleading because there is no new interaction here: it is just the *electrostatic interaction* acting on a two-particle state.

Our Schrödinger equation now has the very simple form

$$\begin{pmatrix} E_0 + J & K \\ K & E_0 + J \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

which is just that of the familiar situation of coupled pendula. By now we know how to solve this in our sleep,  $(E_0 + J - \lambda)^2 - K^2 = 0 \Rightarrow E_0 + J - \lambda = \pm K$ , and the answer is:

eigenvalues:  $E_0 + J + K$ ,  $E_0 + J - K$

eigenstates:  $\beta = \alpha$ ,  $\beta = -\alpha$ .

Therefore the two-electron energy eigenstates are  $u_{fg} \pm u_{gf}$ , i.e.

$$\frac{f(1)g(2) + g(1)f(2)}{\sqrt{2}}, \quad \frac{f(1)g(2) - g(1)f(2)}{\sqrt{2}}, \quad (11.33)$$

and the first of these has more energy than the second. The physical reason why the first state has more energy is that in this symmetric state, those parts of  $f$  and  $g$  which overlap interfere constructively, making an enhanced probability for the electrons to be close to one another (and hence have a higher energy of interaction) while in the antisymmetric state the overlapping parts interfere destructively, so the electrons are further apart and have less energy of interaction.

In the case of coupled classical pendulums we would have similar results: one normal mode has a higher energy than the other because the different *correlations* in the motions of the pendulums cause a different energy from the coupling.

Note that we have obtained the answer *without appealing either to the spin or to the exchange symmetry of the states*. As far as spin goes, at the moment it looks as though each spatial wavefunction can be associated with any spin state. As far as symmetry goes, the electrons could be distinguishable, for all we care (e.g. one painted blue, somehow, and the other painted red). N.B. many (in fact the majority) of textbooks are quite confused on this point, or downright wrong.

Now, as it happens, we are dealing with two identical electrons, rather than one ‘red’ and one ‘blue’. Therefore, by the principles of quantum mechanics as they apply to identical particles, the total state must be one of *definite symmetry* with respect to exchange of the particles. Indeed, since electrons are fermions the state must be antisymmetric. Therefore at first sight only the second of our two solutions (11.33) looks physically possible. In fact both solutions are possible because each electron has a further degree of freedom, namely its spin, but the implication is that not every spin state can go with every spatial state, as discussed in section 11.3.

An exact (within numerical error) treatment would involve a full (numerical) solution of the Schrödinger equation including all the parts of the Hamiltonian from the start. The solutions are then found to have the same general features as given here, but the precise values of energy and the shape of the wavefunctions would differ from those given by the approximate method.



## Chapter 12

# Time-dependent perturbation theory

Up till now, all the problems we have treated have had Hamiltonians which do not depend on time. Time-dependent perturbation theory provides a general method to calculate the effects of a small time-dependent contribution to the Hamiltonian.

If the Hamiltonian varies in time, it means the energy of the system varies in time, so one of the most obvious results is that the system can gain or lose energy.

### 12.1 First order time-dependent perturbation theory

(This section is copied from Shankar with a few notation changes). Suppose a system has a Hamiltonian of the form

$$H = H^0 + \delta H(t) \quad (12.1)$$

where  $H^0$  is time-independent, but  $\delta H$  depends on time. Since the eigenstates of  $H^0$  form a complete set, we can always write the exact wavefunction at any time in terms of them:

$$\psi(t) = \sum_n c_n(t) \psi_n^0 \quad (12.2)$$

where the zeroth-order wavefunctions  $\psi_n^0$  do not depend on time. All the time-dependence is taken up by the coefficients  $c_n$ . Now, if  $\delta H$  were not there, we know these would vary as

$$c_n(t) = c_n(0) e^{-iE_n^0 t/\hbar}. \quad [\text{when } \delta H = 0] \quad (12.3)$$

So let us use this information, and write

$$\psi(t) = \sum_n d_n(t) e^{-iE_n^0 t/\hbar} \psi_n^0, \quad (12.4)$$

then if  $d_n$  changes with time, it is because of  $\delta H$ . We expect therefore that the evolution of  $d_n$  can be written as a nice power series in  $\delta H$ .

Now operate on both sides of (12.4) with  $i\hbar\partial/\partial t - H$  to get

$$0 = \sum_n \left( i\hbar\dot{d}_n - d_n\delta H(t) \right) e^{-iE_n^0 t/\hbar} \psi_n^0 \quad (12.5)$$

$$\implies \sum_n i\hbar\dot{d}_n e^{-iE_n^0 t/\hbar} \psi_n^0 = \sum_n d_n \delta H(t) e^{-iE_n^0 t/\hbar} \psi_n^0. \quad (12.6)$$

Multiply by  $(\psi_f^0)^*$  and integrate:

$$i\hbar\dot{d}_f e^{-iE_f^0 t/\hbar} = \sum_n d_n e^{-iE_n^0 t/\hbar} \langle \psi_f^0 | \delta H(t) | \psi_n^0 \rangle. \quad (12.7)$$

To tidy this up a little, define

$$\hbar\omega_{fn} = E_f^0 - E_n^0 \quad (12.8)$$

then

$$i\hbar\dot{d}_f = \sum_n d_n e^{i\omega_{fn} t} \langle \psi_f^0 | \delta H(t) | \psi_n^0 \rangle. \quad [\text{exact}] \quad (12.9)$$

This result is exact and fully equivalent to the original Schrödinger equation  $H\psi = i\hbar\dot{\psi}$ .

Let us consider the case where, at  $t = 0$ , the system is in state  $\psi_i^0$ . Then

$$d_n(0) = \delta_{ni}.$$

What happens next? To zero order approximation, we ignore  $\delta H$  altogether, and equation (12.9) tells us  $\dot{d}_f = 0$ , i.e. the state doesn't change with time. To get the first order result, notice that in (12.9) the  $\delta H$  is already of first order, so we should substitute in the zero order expression for  $d_n$ . This gives us the first-order result

$$i\hbar\dot{d}_f = e^{i\omega_{fi} t} \langle \psi_f^0 | \delta H(t) | \psi_i^0 \rangle. \quad (12.10)$$

whose solution is

$$d_f(t) = \delta_{fi} - \frac{i}{\hbar} \int_0^t e^{i\omega_{fi} t'} \langle \psi_f^0 | \delta H(t') | \psi_i^0 \rangle dt'. \quad (12.11)$$

You should note that the integral here is a Fourier transform<sup>1</sup>.

The next step is to apply this result to a few examples and so gain familiarity with it: please consult textbooks and problems.

### 12.1.1 Limit of slow and fast perturbations

You need to be aware of what happens in the cases of a very slow and a very fast perturbation. I will just quote results here rather than derive them: consult the textbooks for more information.

<sup>1</sup>A full Fourier analysis integral must extend from  $t' = -\infty$  to  $t' = \infty$  but at any time  $t$  the perturbation  $\delta H$  can only have acted between times 0 and  $t$ , so as far as the Fourier integral is concerned,  $\delta H$  is zero for  $t' < 0$  and for  $t' > t$ .



How slow is slow, and how fast is fast? Every quantum system has a set of natural time-scales, given by the inverse of the frequencies  $\omega_{fi}$  which come from its energy level separations. If a change in Hamiltonian occurs on a time-scale fast compared to the highest of these frequencies, then it can be regarded as sudden, and if it is slow compared to the lowest frequency, then it can be regarded as adiabatic. A more detailed statement, in which the matrix elements  $\langle \psi_f^0 | \delta H(t) | \psi_i^0 \rangle$  are examined along with the frequencies  $\omega_{fi}$ , is needed in the case of very high or very low  $\omega_{fi}$  (e.g. degeneracy).

### Sudden perturbation

If the Hamiltonian changes abruptly in a short interval of time, say  $t = 0$  to  $t = \epsilon$ , then the state of the system immediately after the abrupt change is the same as it was before:  $\psi(t = \epsilon) = \psi(t = 0)$ . After the change, the system has a new Hamiltonian, so its subsequent evolution for times  $t > \epsilon$  may well be different from what it was for times  $t < 0$ , but there was no appreciable change in the *state* during the short interval  $0 < t < \epsilon$ . You should recall the well-known behaviour of a collection of crockery on a table when the table-cloth is abruptly whisked away.

Example: a mass on a spring, of potential energy  $(1/2)m\omega^2x^2$ , is in the ground state of motion. The spring constant is suddenly (in a time small compared to  $1/\omega$ ) doubled. What is the wavefunction of the mass immediately after this change? Answer: the same as it was before the change, i.e.

$$\psi = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar}.$$

Note, however, that this is not the ground state of the new oscillator potential, in fact it is not an energy eigenstate of the new Hamiltonian at all, so the state will subsequently evolve away from  $\psi$ .

### Adiabatic perturbation

**Adiabatic theorem:** If the Hamiltonian  $H(t)$  changes slowly enough from  $H^A$  to  $H^B$  in a time  $T$ , then if the system starts out at time  $t = 0$  in an eigenstate  $\psi_n^A$  of  $H^A$ , it will finish in the corresponding eigenstate  $\psi_n^B$  of  $H^B$ .

The meaning of the ‘corresponding’ eigenstate is as follows. As the Hamiltonian slowly changes, all the energy levels slowly change. We keep track of the movement of the energy levels, and the state  $\psi_n^B$  is the one whose energy under  $H^B$  was linked continuously to the energy of  $\psi_n^A$  under  $H^A$ . In the case where energy levels cross (degeneracy) or come very close to one another, there is a danger the change may not be adiabatic. We then have to keep track of the changes in the wavefunctions to avoid any ambiguity.

Example: a mass on a spring (with vibrational frequency  $\omega$ ) is in the ground state, before the spring constant changes. The spring constant is then slowly doubled, on a time scale long compared to  $1/\omega$ . What is the wavefunction after this change? Answer: the mass ends up in the ground state of the new Hamiltonian. Doubling the spring constant will increase the vibrational frequency by  $\sqrt{2}$ , so the new state is

$$\psi = e^{i\theta} \left(\frac{m\sqrt{2}\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/\sqrt{2}\hbar}.$$

This is an eigenstate of the new Hamiltonian, and so does not further evolve.

Note, the adiabatic theorem says what state the system ends up in, but during the change the state will pick up a phase, owing to its energy and to phases associated with the change itself. This is why we need to include the factor  $\exp(i\theta)$  in the answer. For a system in a single energy eigenstate, this phase is unobservable, but if the system started in a superposition of energy eigenstates, then it will end in a superposition, and the phase *difference* between the two or more components can be revealed by interference effects (c.f. ‘‘Aharonov-Bohm effect’’ and ‘‘Berry phase’’).

## 12.2 Periodic perturbation: Fermi Golden Rule

The case of a periodic perturbation is quite common: think of an atom being illuminated by a light wave, for example. Such perturbations will in reality vary as sines or cosines, but for mathematical convenience we consider the case

$$\delta H(t) = H' e^{-i\omega t}. \quad (12.12)$$

Note,  $H'$  is the amplitude of the oscillation.  $H'$  is small and time-independent.

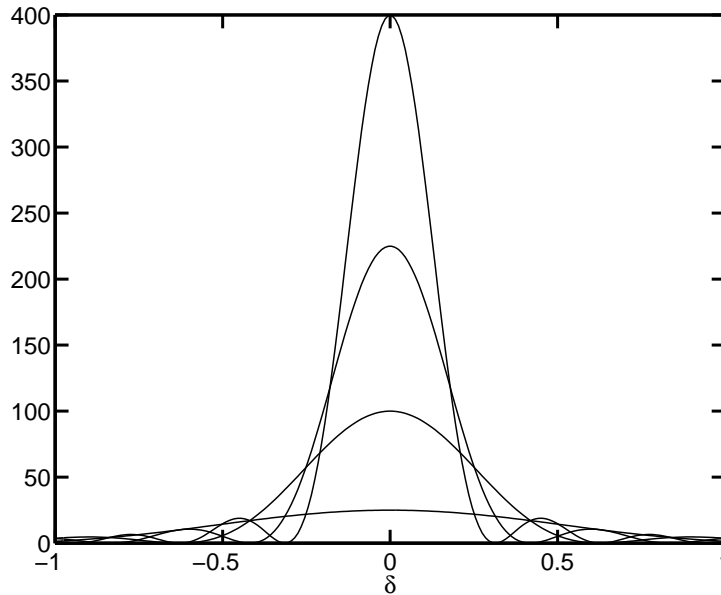


Figure 12.1: The function  $t^2 \text{sinc}^2(t\delta/2)$  plotted for times  $t = 5, 10, 15, 20$ . The transition probability from state  $i$  to state  $f$  is proportional to this, where  $\delta = \omega_{fi} - \omega$  is the detuning from resonance. The function gets more and more sharply peaked as  $t$  increases. The height is  $t^2$ , the first zeros occur at  $\delta = \pm 2\pi/t$ .

Applying our general first-order result (12.11), we have for the amplitude to arrive in a state other than the one the system started in:

$$d_f(t) = -\frac{i}{\hbar} \int_0^t e^{i\omega_{fi}t'} \langle \psi_f^0 | H' | \psi_i^0 \rangle e^{-i\omega t'} dt'. \quad (12.13)$$

$$= -\frac{i}{\hbar} \langle \psi_f^0 | H' | \psi_i^0 \rangle \frac{e^{i(\omega_{fi}-\omega)t} - 1}{i(\omega_{fi} - \omega)} \quad (12.14)$$

$$= -\frac{i e^{i(\omega_{fi}-\omega)t/2}}{\hbar} \langle \psi_f^0 | H' | \psi_i^0 \rangle \text{sinc}[(\omega_{fi} - \omega)t/2] t \quad [\text{recall } \text{sinc}(x) = \sin(x)/x] \quad (12.15)$$

The perturbation has appeared as a ‘square pulse’ acting from time 0 to  $t$ , and the system has ‘Fourier analysed’ it into a sinc function. The probability for the transition  $i \rightarrow f$  is

$$P_{i \rightarrow f} = |d_f|^2 = \frac{1}{\hbar^2} |\langle \psi_f^0 | H' | \psi_i^0 \rangle|^2 \text{sinc}^2[(\omega_{fi} - \omega)t/2] t^2. \quad (12.16)$$

This is shown in figure 12.1 for various times  $t$ . The transition probability is appreciable for final states such that  $|(\omega_{fi} - \omega)t/2| \leq \pi$ , which means for energies in the range

$$E_f^0 \sim E_i^0 + \hbar\omega \pm 2\pi\hbar/t. \quad (12.17)$$

As time goes on, the system gets more and more ‘choosy’ as to which final states it will go to, until for large times the only final state with a non-negligible probability is the one with  $E_f^0 = E_i^0 + \hbar\omega$  (if there

In the limit  $t \rightarrow \infty$  the system will only go to a final state of perfectly correct energy, but of course there is no such thing as either infinity or perfection in the physical world. To get some useful information, instead of assessing the probability to go to a final state of specific energy, it is better to ask for the probability of going to any state at all. Thus we want to calculate

$$P = \sum_{f \neq i} P_{i \rightarrow f}. \quad (12.18)$$

If the quantum system has  $dN$  quantum states in each energy range  $E$  to  $E + dE$ , then we define the

$$\text{density of states} \quad \rho(E) = \frac{dN}{dE}. \quad (12.19)$$

This permits us to convert the sum (12.18) into an integral over energy. The number of states at energy  $E$  is  $\rho(E)dE$ , so

$$P \simeq \int_{-\infty}^{\infty} P_{i \rightarrow f} \rho(E) dE \quad (12.20)$$

$$= \frac{1}{\hbar^2} \int_{-\infty}^{\infty} |\langle \psi_f^0 | H' | \psi_i^0 \rangle|^2 \text{sinc}^2 [(E - (E_i^0 + \hbar\omega))t/2\hbar] t^2 \rho(E) dE \quad (12.21)$$

where in the sinc function the energy of the final state is  $E_f^0 = E$ . If we assume  $t$  is long enough that the  $\text{sinc}^2$  function is much more sharply peaked as a function of energy than either the matrix element or the density of states, then these are both approximately constant in the interval over which the integrand is non-zero, so they can be brought out the front:

$$P \simeq \frac{1}{\hbar^2} |\langle \psi_f^0 | H' | \psi_i^0 \rangle|^2 \rho(E_i^0 + \hbar\omega) \int_{-\infty}^{\infty} \text{sinc}^2 [(E_f^0 - (E_i^0 + \hbar\omega))t/2\hbar] t^2 dE \quad (12.22)$$

To do the integral, change variable to  $x = (E - (E_i^0 + \hbar\omega))t/2\hbar$ , then  $dE/dx = 2\hbar/t$ , and use the result

$$\int_{-\infty}^{\infty} \text{sinc}^2(x) dx = \pi. \quad (12.23)$$

Hence

$$P \simeq \frac{2\pi}{\hbar} |\langle \psi_f^0 | H' | \psi_i^0 \rangle|^2 \rho(E) t. \quad (12.24)$$

We find the probability now increases linearly with time, whereas in equation (12.16) the increase was quadratic in time. What is happening is that the  $\text{sinc}^2$  function picks out a narrower and narrower range of final energies as time goes on, and hence the time-dependence of the total transition probability is brought 'down' from a quadratic to a linear function of time. The linear result means that now we have a constant *transition rate*

$$R = \frac{dP}{dt} \simeq \frac{2\pi}{\hbar} |\langle \psi_f^0 | H' | \psi_i^0 \rangle|^2 \rho(E_f^0) \quad [\text{Fermi Golden Rule}] \quad (12.25)$$

Equation (12.25) is called the Fermi Golden Rule. It is useful whenever we are interested in transitions from one state to another, so it pops up in many areas of physics. It says that with a periodic perturbation, the transition *rate* is constant, and is proportional to the density of states at energy  $E_f^0 = E_i^0 + \hbar\omega$  and to the square of the matrix element<sup>2</sup>.

### Which expression is relevant?

We remarked that the time-dependence of the excitation probability went over from a quadratic to a linear function in passing from (12.16) to (12.24). The change is related to the integration over states, (12.18) and (12.20). Understanding the proper use of these two results involves a consideration of time-scales, as follows.

<sup>2</sup>The final state  $f_0$  appearing in the matrix element can be taken to be any state whose energy is in the range  $E_i^0 + \hbar\omega \pm 2\pi\hbar/t$ , since we assumed the matrix element was the same for all such states.

To get the general idea, it is useful to express everything in time and frequency units. We already expressed the energy change  $E_f - E_i$  in terms of a transition frequency  $\omega_{fi} = (E_f - E_i)/\hbar$ . We also noted that the length of time  $t$  for which the oscillating perturbation acts introduces an energy selectivity  $\pm\hbar/t$ , and hence a frequency selectivity  $\pm 1/t$  via the  $\text{sinc}^2$  function. We argued that for long times this  $\text{sinc}^2$  function is sharp compared to the density of states function, but to quantify how long, we need some idea of how  $\rho(E)$  depends on energy.

In a system such as an atom, up till now we have used the idea of energy eigenstates, and assumed that these were precisely defined in energy. However, that is an idealisation which arises from treating an atom as an isolated system. In fact, atoms are weakly coupled to their environment, and as a result no atom stays in an excited state forever: it decays by emitting one or more photons, a process called *spontaneous emission*. Each excited state therefore has a *natural lifetime*  $\tau$ , and its energy is only well-defined to within  $\Delta E = \hbar/\tau$ . The density of states function allows us to incorporate this fact: in the vicinity of an energy level of natural lifetime  $\tau$ ,  $\rho(E)$  has a peak of width  $\hbar/\tau$ , and the height of the peak is “one state per energy range  $\Delta E$ ”, i.e.

$$\rho(E = E_f) \simeq 1/\Delta E = \tau/\hbar. \quad (12.26)$$

The Fermi Golden Rule involved the approximation that the  $\text{sinc}^2$  function is sharply peaked compared to the density of states function. In the case we just assumed, where the transition is to a single final state of natural lifetime  $\tau$ , we deduce that this requires

$$t \gg \tau. \quad (12.27)$$

There is another timescale set by how long we have to wait for the transition to happen: this is to do with the matrix element  $\langle \psi_f^0 | H' | \psi_i^0 \rangle$ . We put this into frequency units by defining the “Rabi (pronounced ‘Rah-bee’) frequency”:

$$\Omega = \frac{2}{\hbar} |\langle \psi_f^0 | H' | \psi_i^0 \rangle|. \quad (12.28)$$

The reason for including the factor 2 will emerge in a moment.

We can now see that there are two regimes where transitions driven by a periodic perturbation take a simple form, and a more complicated intermediate regime.

1. Strongly driven transition:  $\Omega \gg 1/\tau$ .

In this case the quadratic dependence predicted by (12.16) is relevant. This is because even for times  $t < \tau$ , i.e. long before the Fermi Golden Rule is accurate, there is an appreciable probability of transition. We have from (12.16)

$$P_{i \rightarrow f} = |d_f|^2 = \Omega^2 t^2 / 4 \quad (12.29)$$

when the perturbation is “on resonance”, i.e.  $\omega = \omega_{fi}$ . This gives at  $t = \tau$  the transition probability  $P_{i \rightarrow f} = \Omega^2 \tau^2 / 4 \gg 1$  which is of course impossible, it simply means that for strong driving the first-order perturbation theory is only valid for times  $t \ll \tau$ . For a system where only one transition is near resonance (i.e. there are no further close-lying energy levels), the full time dependence is in fact  $P_{i \rightarrow f} = \sin^2(\Omega t / 2)$ , i.e. an oscillation of angular frequency  $\Omega$ , which explains the factor 2 in the definition of  $\Omega$ . The result (12.29) is the small- $t$  approximation to this complete dependence.

2. Weakly driven transition:  $\Omega \ll 1/\tau$ .

In this case the transition probability only becomes appreciable for times  $t \gg \tau$ , and by this time the Fermi Golden Rule applies, so a constant transition rate is expected.

Most atomic excited states have natural lifetimes of order a few ns, so for transitions involving an excited state, usually the transition is found to be weakly driven and the Fermi Golden Rule applies<sup>3</sup>. States with very long lifetimes (which can be from seconds to millions of years!) are called *metastable*. Transitions between the ground state (which is stable) and a metastable state, or between metastable states, can often in practice be strongly driven by an applied laser beam or microwave field. Examples are given in the next chapter.

### 12.2.1 Forbidden transitions and selection rules

One important implication of the Fermi Golden Rule is that some transitions don't happen! That is, we can deduce that if the matrix element is zero, i.e. the perturbing Hamiltonian does not connect a given pair of initial and final states, then the system won't be driven between those states by that interaction. The transition is said to be 'forbidden'. We can often show mathematically or physically that certain classes of matrix element are zero, and hence a whole class of transitions is forbidden. Such deductions are called *selection rules*. Whenever a selection rule prevents a transition  $i \rightarrow f$ , that does not mean the quantum system cannot ever go from state  $i$  to state  $f$ , it simply means the Hamiltonian  $H'$  under consideration cannot drive it on its journey: it has to go by some other physical process.

For an example of a selection rule, suppose  $H' \propto z = r \cos \theta$ , and consider a system with spherical symmetry:

$$\langle R_{n',l'} Y_{l',m'} | z | R_{n,l} Y_{l,m} \rangle = \int_0^\infty R_{n',l'}^*(r) r R_{n,l}(r) r^2 dr \int_0^\pi W_{l',m'}^*(\theta) \cos \theta W_{l,m}(\theta) \sin \theta d\theta \int_0^{2\pi} e^{-im'\phi} e^{im\phi} d\phi \quad (12.30)$$

where  $W_{l,m}(\theta)$  is the  $\theta$ -dependent part of  $Y_{l,m}(\theta, \phi)$ . From the  $\phi$  integral, we deduce the selection rule

$$m' = m \quad \text{for interaction } H' = z \quad (12.31)$$

That is, an interaction  $\propto z$  can only drive transitions between states having the same  $z$ -component of angular momentum ( $m$ ).

#### Main points

The main points of this chapter have been:

basic result of first-order time-dependent perturbation theory; the sudden and the adiabatic cases; the effect of a periodic perturbation; the Fermi Golden Rule; transition matrix element; selection rules.

---

<sup>3</sup>The spontaneous emission process requires special consideration, because the system providing the perturbation to the atom, namely the surrounding radiation field in its vacuum state, is itself strongly affected by the transition, since it gains a photon where previously it had none. Therefore the effect has to be treated as a coupling between two quantum systems, and the density of states function accounts for the available states of the radiation field as well as of the atom. It is found to be extremely broad as a function of frequency. This, combined with the small matrix element, means that spontaneous emission in atoms is described to extremely good approximation by a constant transition rate.



## Chapter 13

# Interaction of atoms with light

We have already remarked on the fact that atoms can emit and absorb electromagnetic radiation. Now that we have examined the Fermi Golden Rule we also have some idea of the quantum mechanics of this: the electric field of the light wave interacts with the charge in the atom and thus provides a periodic perturbation, resulting in transitions. Before looking at this in a bit (but not much) more detail, we should take a moment to think in physical terms about what is going on.

### 13.1 Emission of light waves from an atom

One of the most immediately obvious and important features of atoms is that they interact strongly with light. A single sodium atom illuminated by a modest laser beam will scatter 30 million photons per second. This is a lot of photons for such a small emitter! A human dark-adapted eye looking at this atom from a distance of 10 cm would receive around 6000 photons per second. [Problem: check this calculation.] This is enough to make the light emitted by a single atom detectable to the naked eye.

The primary way in which electromagnetic waves are produced is by **accelerating charge**. Before quantum mechanics was understood, it was thought that the electrons in an atom could not be orbiting the nucleus because if they were then the high accelerations involved would produce strong electromagnetic radiation and the electrons would rapidly lose energy and plunge into the nucleus in a fraction of a nanosecond. There is some correct insight in this argument: any accelerating charge will indeed emit electromagnetic waves, quantum mechanics is fully in agreement with that. The reason that the energy eigenstates in an atom are much more stable than the classical argument suggests is that the charge is not accelerating: it is smeared out around the atom and *does not move* when the atom is in an energy eigenstate. To be precise, the probability distribution  $|\psi|^2$  of the electrons is completely independent of time. This is in fact just the standard property of energy eigenstates which you should know already: it is why they are called stationary states.

Absorption or emission of electromagnetic radiation happens when an atom is stimulated to absorb or decay by some external time-dependent perturbation. In the middle of such a transition process, the atom is for a while in a **superposition of energy eigenstates**. Consider two energy eigenstates of hydrogen, for example  $\psi_{1s}$  and  $\psi_{2p}$  having energies  $E_1$  and  $E_2$  respectively. According to the time-dependent Schrödinger equation, the wavefunction of each of these states has a time-dependence  $\exp(-iEt/\hbar)$ . Therefore if the atom is in a superposition of the two, then its wavefunction depends on time in the following way:

$$\psi_{\mathbf{r},t} = c_1\psi_{1s}(\mathbf{r})e^{-iE_1t/\hbar} + c_2\psi_{2p}(\mathbf{r})e^{-iE_2t/\hbar} \quad (13.1)$$

The charge density of the electron is equal to

$$-e|\psi|^2 = -e \left( |c_1\psi_{1s}|^2 + |c_2\psi_{2p}|^2 + 2\text{Re} \left[ c_1^*c_2\psi_{1s}^*\psi_{2p}e^{-(E_2-E_1)t/\hbar} \right] \right) \quad (13.2)$$

To be absolutely clear on what is going on, we will take the case that  $c_1\psi_{1s}$  and  $c_2\psi_{2p}$  are real, though the argument will not depend on this. Then we have

$$-e|\psi|^2 = -e \left( c_1^2\psi_{1s}^2 + c_2^2\psi_{2p}^2 + 2c_1c_2\psi_{1s}\psi_{2p}\cos\omega_{21}t \right) \quad (13.3)$$

where  $\omega_{21} = (E_2 - E_1)/\hbar$ . We now have a charge density distribution which is oscillating at the frequency  $\omega_{21}$ . That is how an atom can emit light, and that is why the frequency of the light waves coming out is equal to the frequency difference of two energy eigenstates of the atom.

Once the atom is in such a superposition, therefore, light waves<sup>1</sup> are emitted. Therefore the atom must be losing energy. In terms of the wavefunction, what happens is that the coefficients  $c_1$  and  $c_2$  vary (smoothly) in time:  $c_2$  decreases and  $c_1$  increases. Typical timescales for a strong optical transition at 600 nm wavelength are as follows:

period of one oscillation of the emitted light	$2 \text{ fs} = 2 \times 10^{-15} \text{ s}$
time for $c_2$ to fall to 0 and $c_1$ to rise to 1	$20 \text{ ns} = 2 \times 10^{-8} \text{ s}$
$\Rightarrow$ number of oscillation periods in the emitted pulse	$10^7$

The final number is also the number of wavelengths long which the pulse of light is; we find a length of 6 metres in this example (but note the pulse length can be very different for other transitions).

There are two things which remain to be worked out: first, how does the atom get into a superposition of energy eigenstates in the first place, and secondly what determines the speed with which the coefficients  $c_1$  and  $c_2$  change. More fully, we would like to know how the coefficients change as a function of time. You should not be surprised to learn that for the case of spontaneous emission, the answer is exponential decay of  $|c_2|^2$ , and  $|c_1|^2 = 1 - |c_2|^2$ .

The atom gets into a superposition of energy eigenstates because it is jiggled by a light wave of the right frequency. This is called stimulation. However, it is remarkable that even in the absence of any light, atoms still decay. We now understand this in terms of the quantum vacuum state: the vacuum state has zero mean energy, but its r.m.s. energy is not zero: there are continual fluctuations at all frequencies, and these stimulate atoms to decay.

---

<sup>1</sup>Throughout this section when I refer to light waves, of course in general I mean electromagnetic waves of frequency  $\omega$ , which might be anywhere from radio waves to X rays etc.



The calculation of the atomic decay in terms of these vacuum fluctuations is very complicated, and beyond undergraduate level. However, the case of absorption is more straightforward: we consider a single travelling-wave of light (e.g. a laser beam) interacting with an atom, and apply time-dependent perturbation theory (see below). There is a very beautiful argument due to Einstein which links the spontaneous decay process to the absorption process and to stimulated emission. This will be treated next year. The essential point is that strong absorbers are strong emitters, which you should recall from your study of thermodynamics and Kirchoff's theorem relating emissivity and absorptivity .

### 13.1.1 Electric dipole radiation

The way an incident light beam couples to an atom is primarily through the electric field  $\mathbf{E}$  of the light coupling to the charges in the atom. We will consider hydrogen in order to have a simple example, the argument can then be generalised to other atoms.

Let  $\mathbf{E}$  be the electric field of the light wave, at some instant of time. The forces due to this field on the proton and the electron are  $\pm\mathbf{E}e$  respectively. When the light wavelength is large compared to the dimensions of the atom, at any given instant the field is the same over the whole of the atom to good approximation. Therefore the potential energy function whose gradient gives the forces on the particles is

$$V = -e(\mathbf{E} \cdot \mathbf{r}_p - \mathbf{E} \cdot \mathbf{r}_e) = e\mathbf{E} \cdot \mathbf{r}. \quad (13.4)$$

where  $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_p$  is the relative coordinate.

The quantity  $\mathbf{d} = -e\mathbf{r}$  is the electric dipole moment of the atom (make sure this is clear to you). Therefore the interaction is a coupling  $\delta H = -\mathbf{d} \cdot \mathbf{E}$  between the electric field of the light and the electric dipole moment of the atom. For this reason the associated pattern of radiation is called *electric dipole radiation*.

In an energy eigenstate, the atom has no dipole moment, but an oscillating electric dipole appears during a transition, owing to the oscillating charge distribution shown by equation (13.3).

Invoking the Fermi Golden Rule (12.25), we find the transition rate is proportional to the *electric dipole matrix element* squared:

$$R \propto |\langle \psi_2 | e\mathbf{E} \cdot \mathbf{r} | \psi_1 \rangle|^2 \quad (13.5)$$

where  $\psi_1, \psi_2$  are the atomic wavefunctions for the two states involved in the transition. An example complete calculation of a transition rate is given in the lecture, see also the problems.

### 13.1.2 Selection rules for electric dipole radiation

The selection rules (recall section 12.2.1) which can be deduced for electric dipole radiation are as follows.

1. total angular momentum:
a. $J = 0 \not\leftrightarrow J = 0$
b. $M_J = 0 \not\leftrightarrow M_J = 0$ when $\Delta J = 0$
c. $\Delta J = 0, \pm 1$
d. $\Delta M_J = 0, \pm 1$
2. parity:
a. parity must change
b. $\implies \Delta l = \pm 1$
3. Configuration: only one electron jumps
4. when $L$ and $S$ are good quantum numbers:
a. $\Delta S = 0,$
b. $\Delta L = 0, \pm 1,$ and $L = 0 \not\leftrightarrow L = 0$

These rules have been given in order of how strict they are. The rules in the first box are strict. Rules 1a.–d. follow immediately from the principle of conservation of angular momentum, once we realize that the electric dipole interaction operator can impart at most one unit of angular momentum to any quantum system. One way to derive this fact is to allow the operator  $e\mathbf{E} \cdot \mathbf{r} = e(E_x x + E_y y + E_z z)$  to act on a spherical harmonic  $Y_{l,m}$ , and examine the overlap integral

$$\int \int Y_{l+\Delta l, m+\Delta m}^* e(E_x x + E_y y + E_z z) Y_{l,m} \sin \theta d\theta d\phi. \quad (13.6)$$

It is found that if the integral is to be non-zero,  $l$  and  $m$  can change by at most 1 (and if  $l = 0$ ,  $\Delta l = 1$ ). The reason why the rule is nevertheless quoted in terms of total angular momentum  $J$  is that the essential point is the conservation of total angular momentum. For example, in hydrogen there are transitions with  $\delta J = 0$ , even though  $\delta l = 0$  is ruled by another selection rule (2b.). All the rules 1a.–d. can be grouped together as a single

*‘triangle rule’*: to ensure angular momentum is conserved, it must be possible to form a triangle out of the three angular momentum vectors whose lengths are  $J$ ,  $J + \Delta J$  and 1.

Rule 2a. is strict: because the operator  $e(E_x x + E_y y + E_z z)$  has odd parity (this means it changes sign when  $x \rightarrow -x$ ,  $y \rightarrow -y$ ,  $z \rightarrow -z$ ), it follows that the matrix element can only be non-zero when the final state has opposite parity to the initial state.

The rules in the second box apply when the approximation of spherical symmetry of the Hamiltonian holds. Thus rule 2b. applies the parity rule to the case where the wavefunctions are spherical harmonics: then to change parity,  $l$  must change by an odd number (and we have already ruled out changes by more than 1).

Rules 3–4b are only relevant to atoms of more than one electron. I include them here only for completeness, they will be discussed next year.  $L$  and  $S$  are quantum numbers associated with the total orbital and total spin angular momentum of all the electrons in an atom.

**Violations**

When a transition occurs which is not allowed by the electric dipole selection rules, it typically means it happened by a process other than electric dipole radiation. Examples are a coupling between the electric field of the light and the quadrupole moment of the atom, or between the magnetic field of the light and the magnetic dipole moment of the atom. However, these interactions are intrinsically much weaker than the electric dipole interaction, so the associated transition rates are much slower.

**Do electrons jump?**

It is common parlance to talk about an electron ‘jumping’ in an atomic transition, but in fact neither the emission nor the absorption process is abrupt. As we have seen, absorption or emission is a continuous process whose time scale can be short (10 ns for allowed optical transition) or long (millions of years for microwave transitions). When a photon detector registers a ‘click’, that does not imply the atomic emission was an abrupt process, it implies that the mechanism of the detector had a short time-scale.

**13.1.3 An example of non-Fermi Golden Rule behaviour: magnetic resonance**

If the electric dipole matrix element between two states is non-zero, then it almost always happens (I have not been able to think of a counter-example) that one or both of the states has a short natural lifetime because it can decay by electric dipole radiation (whether via the transition in question or another one). As a result, physical phenomena involving electric dipole transitions are almost always in practice in the regime where the transition rate is constant and well described by the Fermi Golden Rule.

An important example of non-Fermi Golden Rule behaviour involves metastable states and another way of driving transitions: magnetic resonance. Here the system is subject to an oscillating magnetic field  $\mathbf{B}(t) = \mathbf{B}_0 \exp(i\omega t)$ , and the field couples to the magnetic dipole moment  $\boldsymbol{\mu}$  of the system:  $\delta H = -\boldsymbol{\mu} \cdot \mathbf{B}_0 \exp(i\omega t)$ . Both initial and final states can be metastable with very long natural lifetimes if, for example, they are different Zeeman levels of the ground state of the system. This example is discussed in the lecture and the problems.

**13.1.4 Main points**

The main points of this chapter have been: absorption or emission of light requires accelerating charge, which happens when an atom is in a superposition of energy eigenstates; electric dipole matrix element and selection rules.