Atomic Physics

C. J. FOOT

Department of Physics
University of Oxford

OXFORD UNIVERSITY PRESS
The Oxford Master Series is designed for final year undergraduate and beginning graduate students in physics and related disciplines. It has been driven by a perceived gap in the literature today. While basic undergraduate physics texts often show little or no connection with the huge explosion of research over the last two decades, more advanced and specialized texts tend to be rather daunting for students. In this series, all topics and their consequences are treated at a simple level, while pointers to recent developments are provided at various stages. The emphasis in on clear physical principles like symmetry, quantum mechanics, and electromagnetism which underlie the whole of physics. At the same time, the subjects are related to real measurements and to the experimental techniques and devices currently used by physicists in academe and industry. Books in this series are written as course books, and include ample tutorial material, examples, illustrations, revision points, and problem sets. They can likewise be used as preparation for students starting a doctorate in physics and related fields, or for recent graduates starting research in one of these fields in industry.

CONDENSED MATTER PHYSICS

1. M. T. Dove: Structure and dynamics: an atomic view of materials
2. J. Singleton: Band theory and electronic properties of solids
3. A. M. Fox: Optical properties of solids
4. S. J. Blundell: Magnetism in condensed matter
5. J. F. Annett: Superconductivity
6. R. A. L. Jones: Soft condensed matter

ATOMIC, OPTICAL, AND LASER PHYSICS

7. C. J. Foot: Atomic physics
8. G. A. Brooker: Modern classical optics
9. S. M. Hooker, C. E. Webb: Laser physics

PARTICLE PHYSICS, ASTROPHYSICS, AND COSMOLOGY

10. D. H. Perkins: Particle astrophysics
11. Ta-Pei Cheng: Relativity, gravitation, and cosmology

STATISTICAL, COMPUTATIONAL, AND THEORETICAL PHYSICS

12. M. Maggiore: A modern introduction to quantum field theory
13. W. Krauth: Statistical mechanics: algorithms and computations
14. J. P. Sethna: Entropy, order parameters, and complexity
Preface

This book is primarily intended to accompany an undergraduate course in atomic physics. It covers the core material and a selection of more advanced topics that illustrate current research in this field. The first six chapters describe the basic principles of atomic structure, starting in Chapter 1 with a review of the classical ideas. Inevitably the discussion of the structure of hydrogen and helium in these early chapters has considerable overlap with introductory quantum mechanics courses, but an understanding of these simple systems provides the basis for the treatment of more complex atoms in later chapters. Chapter 7 on the interaction of radiation with atoms marks the transition between the earlier chapters on structure and the second half of the book which covers laser spectroscopy, laser cooling, Bose–Einstein condensation of dilute atomic vapours, matter-wave interferometry and ion trapping. The exciting new developments in laser cooling and trapping of atoms and Bose–Einstein condensation led to Nobel prizes in 1997 and 2001, respectively. Some of the other selected topics show the incredible precision that has been achieved by measurements in atomic physics experiments. This theme is taken up in the final chapter that looks at quantum information processing from an atomic physics perspective; the techniques developed for precision measurements on atoms and ions give exquisite control over these quantum systems and enable elegant new ideas from quantum computation to be implemented.

The book assumes a knowledge of quantum mechanics equivalent to an introductory university course, e.g. the solution of the Schrödinger equation in three dimensions and perturbation theory. This initial knowledge will be reinforced by many examples in this book; topics generally regarded as difficult at the undergraduate level are explained in some detail, e.g. degenerate perturbation theory. The hierarchical structure of atoms is well described by perturbation theory since the different layers of structure within atoms have considerably different energies associated with them, and this is reflected in the names of the gross, fine and hyperfine structures. In the early chapters of this book, atomic physics may appear to be simply applied quantum mechanics, i.e. we write down the Hamiltonian for a given interaction and solve the Schrödinger equation with suitable approximations. I hope that the study of the more advanced material in the later chapters will lead to a more mature and deeper understanding of atomic physics. Throughout this book the experimental basis of atomic physics is emphasised and it is hoped that the reader will gain some factual knowledge of atomic spectra.
The selection of topics from the diversity of current atomic physics is necessarily subjective. I have concentrated on low-energy and high-precision experiments which, to some extent, reflects local research interests that are used as examples in undergraduate lectures at Oxford. One of the selection criteria was that the material is not readily available in other textbooks, at the time of writing, e.g. atomic collisions have not been treated in detail (only a brief summary of the scattering of ultracold atoms is included in Chapter 10). Other notable omissions include: X-ray spectra, which are discussed only briefly in connection with the historically important work of Moseley, although they form an important frontier of current research; atoms in strong laser fields and plasmas; Rydberg atoms and atoms in doubly- and multiply-excited states (e.g. excited by new synchrotron and free-electron laser sources); and the structure and spectra of molecules.

I would like to thank Geoffrey Brooker for invaluable advice on physics (in particular Appendix B) and on technical details of writing a textbook for the Oxford Master Series. Keith Burnett, Jonathan Jones and Andrew Steane have helped to clarify certain points, in my mind at least, and hopefully also in the text. The series of lectures on laser cooling given by William Phillips while he was a visiting professor in Oxford was extremely helpful in the writing of the chapter on that topic. The following people provided very useful comments on the draft manuscript: Rachel Godun, David Lucas, Mark Lee, Matthew McDonnell, Martin Shotter, Claes-Göran Wahlström (Lund University) and the (anonymous) reviewers. Without the encouragement of Sönke Adlung at OUP this project would not have been completed. Irmgard Smith drew some of the diagrams. I am very grateful for the diagrams and data supplied by colleagues, and reproduced with their permission, as acknowledged in the figure captions. Several of the exercises on atomic structure derive from Oxford University examination papers and it is not possible to identify the examiners individually—some of these exam questions may themselves have been adapted from some older sources of which I am not aware.

Finally, I would like to thank Professors Derek Stacey, Joshua Silver and Patrick Sanders who taught me atomic physics as an undergraduate and graduate student in Oxford. I also owe a considerable debt to the book on elementary atomic structure by Gordon Kemble Woodgate, who was my predecessor as physics tutor at St Peter’s College, Oxford. In writing this new text, I have tried to achieve the same high standards of clarity and conciseness of expression whilst introducing new examples and techniques from the laser era.

Background reading

It is not surprising that our language should be incapable of describing the processes occurring with the atoms, for it was invented to describe the experiences of daily life, and these consist only of processes involving exceeding large numbers
of atoms. Furthermore, it is very difficult to modify our language so that it will be able to describe these atomic processes, for words can only describe things of which we can form mental pictures, and this ability, too, in the result of daily experience. Fortunately, mathematics is not subject to this limitation, and it has been possible to invent a mathematical scheme—the quantum theory—which seems entirely adequate for the treatment of atomic processes.

From *The physical principles of the quantum theory*, Werner Heisenberg (1930).

The point of the excerpt is that quantum mechanics is essential for a proper description of atomic physics and there are many quantum mechanics textbooks that would serve as useful background reading for this book. The following short list includes those that the author found particularly relevant: Mandl (1992), Rae (1992) and Griffiths (1995). The book *Atomic spectra* by Softley (1994) provides a concise introduction to this field. The books Cohen-Tannoudji et al. (1977), Atkins (1983) and Basdevant and Dalibard (2000) are very useful for reference and contain many detailed examples of atomic physics. Angular-momentum theory is very important for dealing with complicated atomic structures, but it is beyond the intended level of this book. The classic book by Dirac (1981) still provides a very readable account of the addition of angular momenta in quantum mechanics. A more advanced treatment of atomic structure can be found in Condon and Odabasi (1980), Cowan (1981) and Sobelman (1996).

Oxford

C. J. F.

Web site:
http://www.physics.ox.ac.uk/users/foot

This site has supplementary information and corrections found after going to press.
Contents

1 Early atomic physics
   1.1 Introduction 1
   1.2 Spectrum of atomic hydrogen 1
   1.3 Bohr's theory 3
   1.4 Relativistic effects 5
   1.5 Moseley and the atomic number 7
   1.6 Radiative decay 11
   1.7 Einstein A and B coefficients 11
   1.8 The Zeeman effect 13
      1.8.1 Experimental observation of the Zeeman effect 17
   1.9 Summary of atomic units 18
   Exercises 19

2 The hydrogen atom 22
   2.1 The Schrödinger equation 22
      2.1.1 Solution of the angular equation 23
      2.1.2 Solution of the radial equation 26
   2.2 Transitions 29
      2.2.1 Selection rules 30
      2.2.2 Integration with respect to $\theta$ 32
      2.2.3 Parity 32
   2.3 Fine structure 34
      2.3.1 Spin of the electron 35
      2.3.2 The spin–orbit interaction 36
      2.3.3 The fine structure of hydrogen 38
      2.3.4 The Lamb shift 40
      2.3.5 Transitions between fine-structure levels 41
   Further reading 42
   Exercises 42

3 Helium 45
   3.1 The ground state of helium 45
   3.2 Excited states of helium 46
      3.2.1 Spin eigenstates 51
      3.2.2 Transitions in helium 52
   3.3 Evaluation of the integrals in helium 53
      3.3.1 Ground state 53
      3.3.2 Excited states: the direct integral 54
      3.3.3 Excited states: the exchange integral 55
4 The alkalis
4.1 Shell structure and the periodic table
4.2 The quantum defect
4.3 The central-field approximation
4.4 Numerical solution of the Schrödinger equation
  4.4.1 Self-consistent solutions
4.5 The spin–orbit interaction: a quantum mechanical approach
4.6 Fine structure in the alkalis
  4.6.1 Relative intensities of fine-structure transitions
Further reading
Exercises

5 The LS-coupling scheme
5.1 Fine structure in the LS-coupling scheme
5.2 The jj-coupling scheme
5.3 Intermediate coupling: the transition between coupling schemes
5.4 Selection rules in the LS-coupling scheme
5.5 The Zeeman effect
5.6 Summary
Further reading
Exercises

6 Hyperfine structure and isotope shift
6.1 Hyperfine structure
  6.1.1 Hyperfine structure for s-electrons
  6.1.2 Hydrogen maser
  6.1.3 Hyperfine structure for \( l \neq 0 \)
  6.1.4 Comparison of hyperfine and fine structures
6.2 Isotope shift
  6.2.1 Mass effects
  6.2.2 Volume shift
  6.2.3 Nuclear information from atoms
6.3 Zeeman effect and hyperfine structure
  6.3.1 Zeeman effect of a weak field, \( \mu_B B < A \)
  6.3.2 Zeeman effect of a strong field, \( \mu_B B > A \)
  6.3.3 Intermediate field strength
6.4 Measurement of hyperfine structure
  6.4.1 The atomic-beam technique
  6.4.2 Atomic clocks
Further reading
Exercises

7 The interaction of atoms with radiation
7.1 Setting up the equations
7.1.1 Perturbation by an oscillating electric field 124
7.1.2 The rotating-wave approximation 125
7.2 The Einstein \( B \) coefficients 126
7.3 Interaction with monochromatic radiation 127
  7.3.1 The concepts of \( \pi \)-pulses and \( \pi/2 \)-pulses 128
  7.3.2 The Bloch vector and Bloch sphere 128
7.4 Ramsey fringes 132
7.5 Radiative damping 134
  7.5.1 The damping of a classical dipole 135
  7.5.2 The optical Bloch equations 137
7.6 The optical absorption cross-section 138
  7.6.1 Cross-section for pure radiative broadening 141
  7.6.2 The saturation intensity 142
  7.6.3 Power broadening 143
7.7 The a.c. Stark effect or light shift 144
7.8 Comment on semiclassical theory 145
7.9 Conclusions 146
Further reading 147
Exercises 148

8 Doppler-free laser spectroscopy 151
  8.1 Doppler broadening of spectral lines 151
  8.2 The crossed-beam method 153
  8.3 Saturated absorption spectroscopy 155
    8.3.1 Principle of saturated absorption spectroscopy 156
    8.3.2 Cross-over resonances in saturation spectroscopy 159
  8.4 Two-photon spectroscopy 163
  8.5 Calibration in laser spectroscopy 168
    8.5.1 Calibration of the relative frequency 168
    8.5.2 Absolute calibration 169
    8.5.3 Optical frequency combs 171
Further reading 175
Exercises 175

9 Laser cooling and trapping 178
  9.1 The scattering force 179
  9.2 Slowing an atomic beam 182
    9.2.1 Chirp cooling 184
  9.3 The optical molasses technique 185
    9.3.1 The Doppler cooling limit 188
  9.4 The magneto-optical trap 190
  9.5 Introduction to the dipole force 194
  9.6 Theory of the dipole force 197
    9.6.1 Optical lattice 201
  9.7 The Sisyphus cooling technique 203
    9.7.1 General remarks 203
    9.7.2 Detailed description of Sisyphus cooling 204
    9.7.3 Limit of the Sisyphus cooling mechanism 207
9.8 Raman transitions 208
  9.8.1 Velocity selection by Raman transitions 208
  9.8.2 Raman cooling 210
9.9 An atomic fountain 211
9.10 Conclusions 213
Exercises 214

10 Magnetic trapping, evaporative cooling and
Bose–Einstein condensation 218
10.1 Principle of magnetic trapping 218
10.2 Magnetic trapping 220
  10.2.1 Confinement in the radial direction 220
  10.2.2 Confinement in the axial direction 221
10.3 Evaporative cooling 224
10.4 Bose–Einstein condensation 226
10.5 Bose–Einstein condensation in trapped atomic vapours 228
  10.5.1 The scattering length 229
10.6 A Bose–Einstein condensate 234
10.7 Properties of Bose-condensed gases 239
  10.7.1 Speed of sound 239
  10.7.2 Healing length 240
  10.7.3 The coherence of a Bose–Einstein condensate 240
  10.7.4 The atom laser 242
10.8 Conclusions 242
Exercises 243

11 Atom interferometry 246
11.1 Young’s double-slit experiment 247
11.2 A diffraction grating for atoms 249
11.3 The three-grating interferometer 251
11.4 Measurement of rotation 251
11.5 The diffraction of atoms by light 253
  11.5.1 Interferometry with Raman transitions 255
11.6 Conclusions 257
Further reading 258
Exercises 258

12 Ion traps 259
12.1 The force on ions in an electric field 259
12.2 Earnshaw’s theorem 266
12.3 The Paul trap 267
  12.3.1 Equilibrium of a ball on a rotating saddle 267
  12.3.2 The effective potential in an a.c. field 268
  12.3.3 The linear Paul trap 269
12.4 Buffer gas cooling 268
12.5 Laser cooling of trapped ions 268
12.6 Quantum jumps 269
12.7 The Penning trap and the Paul trap 274
Early atomic physics

1.1 Introduction

The origins of atomic physics were entwined with the development of quantum mechanics itself ever since the first model of the hydrogen atom by Bohr. This introductory chapter surveys some of the early ideas, including Einstein’s treatment of the interaction of atoms with radiation, and a classical treatment of the Zeeman effect. These methods, developed before the advent of the Schrödinger equation, remain useful as an intuitive way of thinking about atomic structure and transitions between the energy levels. The ‘proper’ description in terms of atomic wavefunctions is presented in subsequent chapters.

Before describing the theory of an atom with one electron, some experimental facts are presented. This ordering of experiment followed by explanation reflects the author’s opinion that atomic physics should not be presented as applied quantum mechanics, but it should be motivated by the desire to understand experiments. This represents what really happens in research where most advances come about through the interplay of theory and experiment.

1.2 Spectrum of atomic hydrogen

It has long been known that the spectrum of light emitted by an element is characteristic of that element, e.g. sodium in a street lamp, or burning in a flame, produces a distinctive yellow light. This crude form of spectroscopy, in which the colour is seen by eye, formed the basis for a simple chemical analysis. A more sophisticated approach using a prism, or diffraction grating, to disperse the light inside a spectrograph shows that the characteristic spectrum for atoms is composed of discrete lines that are the ‘fingerprint’ of the element. As early as the 1880s, Fraunhofer used a spectrograph to measure the wavelength of lines, that had not been seen before, in light from the sun and he deduced the existence of a new element called helium. In contrast to atoms, the spectra of molecules (even the simplest diatomic ones) contain many closely-spaced lines that form characteristic molecular bands; large molecules, and solids, usually have nearly continuous spectra with few sharp features. In 1888, the Swedish professor J. Rydberg found that the spectral
lines in hydrogen obey the following mathematical formula:

\[ \frac{1}{\lambda} = R \left( \frac{1}{n^2} - \frac{1}{n'^2} \right), \]  

(1.1)

where \( n \) and \( n' \) are whole numbers; \( R \) is a constant that has become known as the Rydberg constant. The series of spectral lines for which \( n = 2 \) and \( n' = 3, 4, \ldots \) is now called the Balmer series and lies in the visible region of the spectrum.\(^1\) The first line at 656 nm is called the Balmer-\( \alpha \) (or \( H_\alpha \)) line and it gives rise to the distinctive red colour of a hydrogen discharge—a healthy red glow indicates that most of the molecules of \( \text{H}_2 \) have been dissociated into atoms by being bombarded by electrons in the discharge. The next line in the series is the Balmer-\( \beta \) line at 486 nm in the blue and subsequent lines at shorter wavelengths tend to a limit in the violet region.\(^2\) To describe such series of lines it is convenient to define the reciprocal of the transition wavelength as the wavenumber \( \tilde{\nu} \) that has units of \( m^{-1} \) (or often \( cm^{-1} \)),

\[ \tilde{\nu} = \frac{1}{\lambda}. \]  

(1.2)

Wavenumbers may seem rather old-fashioned but they are very useful in atomic physics since they are easily evaluated from measured wavelengths without any conversion factor. In practice, the units used for a given quantity are related to the method used to measure it, e.g., spectrosopes and spectrographs are calibrated in terms of wavelength.\(^3\) A photon with wavenumber \( \tilde{\nu} \) has energy \( E = h\tilde{\nu} \). The Balmer formula implicitly contains a more general empirical law called the Ritz combination principle that states: the wavenumbers of certain lines in the spectrum can be expressed as sums (or differences) of other lines: \( \tilde{\nu}_3 = \tilde{\nu}_1 \pm \tilde{\nu}_2 \), e.g., the wavenumber of the Balmer-\( \beta \) line (\( n = 2 \) to \( n' = 4 \)) is the sum of that for Balmer-\( \alpha \) (\( n = 2 \) to \( n' = 3 \)) and the first line in the Paschen series (\( n = 3 \) to \( n' = 4 \)). Nowadays this seems obvious since we know about the underlying energy-level structure of atoms but it is still a useful principle for analyzing spectra. Examination of the sums and differences of the wavenumbers of transitions gives clues that enable the underlying structure to be deduced, rather like a crossword puzzle—some examples of this are given in later chapters. The observed spectral lines in hydrogen can all be expressed as differences between energy levels, as shown in Fig. 1.1, where the energies are proportional to \( 1/n^2 \). Other series predicted by eqn 1.1 were more difficult to observe experimentally than the Balmer series. The transitions to \( n = 1 \) give the Lyman series in the vacuum ultraviolet region of the spectrum.\(^4\) The series of lines with wavelengths longer than the Balmer series lie in the infra-red region (not visible to the human eye, nor readily detected by photographic film—the main methods available to the early spectroscopists). The following section looks at how these spectra can be explained theoretically.

---

\(^1\) The Swiss mathematician Johann Balmer wrote down an expression which was a particular case of eqn 1.1 with \( n = 2 \); a few years before Johannes (commonly called Janne) Rydberg found the general formula that predicted other series.

\(^2\) A spectrum of the Balmer series of lines is on the cover of this book.

\(^3\) In this book transitions are also specified in terms of their frequency (denoted by \( f \) so that \( f = \omega \)), or in electron volts (eV) where appropriate.

\(^4\) Air absorbs radiation at wavelengths shorter than about 200 nm and so spectrographs must be evacuated, as well as being made with special optics.
1.3 Bohr’s theory

In 1913, Bohr put forward a radical new model of the hydrogen atom using quantum mechanics. It was known from Rutherford’s experiments that inside atoms there is a very small, dense nucleus with a positive charge. In the case of hydrogen this is a single proton with a single electron bound to it by the Coulomb force. Since the force is proportional to $1/r^2$, as for gravity, the atom can be considered in classical terms as resembling a miniature solar system with the electron orbiting around the proton, just like a planet going around the sun. However, quantum mechanics is important in small systems and only certain electron orbits are allowed. This can be deduced from the observation that hydrogen atoms emit light only at particular wavelengths corresponding to transitions between discrete energies. Bohr was able to explain the observed spectrum by introducing the then novel idea of quantisation that goes beyond any previous classical theory. He took the orbits that occur in classical mechanics and imposed quantisation rules onto them.

Bohr assumed that each electron orbits the nucleus in a circle, whose radius $r$ is determined by the balance between centripetal acceleration and the Coulomb attraction towards the proton. For electrons of mass $m_e$ and speed $v$ this gives

$$\frac{m_e v^2}{r} = \frac{e^2}{4\pi\epsilon_0 r^2}.$$  \hspace{1cm} (1.3)

In SI units the strength of the electrostatic interaction between two
Older systems of units give more succinct equations without $4\pi\varepsilon_0$; some of this neatness can be retained by keeping $e^2/4\pi\varepsilon_0$ grouped together.

The combination of charges of magnitude $e$ is characterised by the combination of constants $e^2/4\pi\varepsilon_0$. This leads to the following relation between the angular frequency $\omega = v/r$ and the radius:

$$\omega^2 = \frac{e^2/4\pi\varepsilon_0}{m_e r^3}. \quad (1.4)$$

This is equivalent to Kepler's laws for planetary orbits relating the square of the period $2\pi/\omega$ to the cube of the radius (as expected since all steps have been purely classical mechanics). The total energy of an electron in such an orbit is the sum of its kinetic and potential energies:

$$E = \frac{1}{2}m_e v^2 - \frac{e^2/4\pi\varepsilon_0}{r}. \quad (1.5)$$

Using eqn 1.3 we find that the kinetic energy has a magnitude equal to half the potential energy (an example of the virial theorem). Taking into account the opposite signs of kinetic and potential energy, we find

$$E = -\frac{e^2/4\pi\varepsilon_0}{2r}. \quad (1.6)$$

This total energy is negative because the electron is bound to the proton and energy must be supplied to remove it. To go further Bohr made the following assumption.

**Assumption I** There are certain allowed orbits for which the electron has a fixed energy. The electron loses energy only when it jumps between the allowed orbits and the atom emits this energy as light of a given wavelength.

That electrons in the allowed orbits do not radiate energy is contrary to classical electrodynamics—a charged particle in circular motion undergoes acceleration and hence radiates electromagnetic waves. Bohr's model does not explain why the electron does not radiate but simply takes this as an assumption that turns out to agree with the experimental data. We now need to determine which out of all the possible classical orbits are the allowed ones. There are various ways of doing this and we follow the standard method, used in many elementary texts, that assumes quantisation of the angular momentum in integral multiples of $\hbar$ (Planck's constant over $2\pi$):

$$m_e v r = n\hbar, \quad (1.7)$$

where $n$ is an integer. Combining this with eqn 1.3 gives the radii of the allowed orbits as

$$r = a_0 n^2, \quad (1.8)$$

where the Bohr radius $a_0$ is given by

$$a_0 = \frac{\hbar^2}{(e^2/4\pi\varepsilon_0) m_e}. \quad (1.9)$$

This is the natural unit of length in atomic physics. Equations 1.6 and 1.8 combine to give the famous Bohr formula:

$$E = -\frac{e^2/4\pi\varepsilon_0}{2a_0} \frac{1}{n^2}. \quad (1.10)$$
The positive integer \( n \) is called the principal quantum number.\(^6\)

Bohr's formula predicts that in the transitions between these energy levels the atoms emit light with a wavenumber given by

\[
\tilde{\nu} = R_\infty \left( \frac{1}{n^2} - \frac{1}{n'^2} \right). \tag{1.11}
\]

This equation fits very closely to the observed spectrum of atomic hydrogen described by eqn 1.1. The Rydberg constant \( R_\infty \) in eqn 1.11 is defined by

\[
hcR_\infty = \frac{(e^2/4\pi\varepsilon_0)^2 m_e}{2\hbar^2}. \tag{1.12}
\]

The factor of \( hc \) multiplying the Rydberg constant is the conversion factor between energy and wavenumbers since the value of \( R_\infty \) is given in units of m\(^{-1}\) (or cm\(^{-1}\) in commonly-used units). The measurement of the spectrum of atomic hydrogen using laser techniques has given an extremely accurate value for the Rydberg constant\(^7\) \( R_\infty = 10973.731.568.525 \) m\(^{-1}\). However, there is a subtle difference between the Rydberg constant calculated for an electron orbiting a fixed nucleus \( R_\infty \) and the constant for real hydrogen atoms in eqn 1.1 (we originally wrote \( R \) without a subscript but more strictly we should specify that it is the constant for hydrogen \( R_H \)). The theoretical treatment above has assumed an infinitely massive nucleus, hence the subscript \( \infty \). In reality both the electron and proton move around the centre of mass of the system. For a nucleus of finite mass \( M \) the equations are modified by replacing the electron mass \( m_e \) by its reduced mass

\[
m = \frac{m_e M}{m_e + M}. \tag{1.13}
\]

For hydrogen

\[
R_H = R_\infty \frac{M_p}{m_e + M_p} \approx R_\infty \left( 1 - \frac{m_e}{M_p} \right), \tag{1.14}
\]

where the electron-to-proton mass ratio is \( m_e/M_p \approx 1/1836 \). This reduced-mass correction is not the same for different isotopes of an element, e.g. hydrogen and deuterium. This leads to a small but readily observable difference in the frequency of the light emitted by the atoms of different isotopes; this is called the isotope shift (see Exercises 1.1 and 1.2).

### 1.4 Relativistic effects

Bohr's theory was a great breakthrough. It was such a radical change that the fundamental idea about the quantisation of the orbits was at first difficult for people to appreciate—they worried about how the electrons could know which orbits they were going into before they jumped.

It was soon realised, however, that the assumption of circular orbits is

---

\(^6\)The alert reader may wonder why this is true since we introduced \( n \) in connection with angular momentum in eqn 1.7, and (as shown later) electrons can have zero angular momentum. This arises from the simplification of Bohr's theory. Exercise 1.12 discusses a more satisfactory, but longer and subtler, derivation that is closer to Bohr's original papers. However, the important thing to remember from this introduction is not the formalism but the magnitude of the atomic energies and sizes.

\(^7\)This is the 2002 CODATA recommended value. The currently accepted values of physical constants can be found on the web site of the National Institute of Science and Technology (NIST).
too much of an over-simplification. Sommerfeld produced a quantised mechanical theory of electrons in elliptical orbits that was consistent with special relativity. He introduced quantisation through a general rule that stated ‘the integral of the momentum associated with a coordinate around one period of the motion associated with that coordinate is an integral multiple of Planck’s constant’. This general method can be applied to any physical system where the classical motion is periodic. Applying this quantisation rule to momentum around a circular orbit gives the equivalent of eqn 1.7:

\[ m_e v \times 2\pi r = n\hbar. \]  

In addition to quantising the motion in the coordinate \( \theta \), Sommerfeld also considered quantisation of the radial degree of freedom \( r \). He found that some of the elliptical orbits expected for a potential proportional to \( 1/r \) are also stationary states (some of the allowed orbits have a high eccentricity, more like those of comets than planets). Much effort was put into complicated schemes based on classical orbits with quantisation, and by incorporating special relativity this ‘old quantum theory’ could explain accurately the fine structure of spectral lines. The exact details of this work are now mainly of historical interest but it is worthwhile to make a simple estimate of relativistic effects. In special relativity a particle of rest mass \( m \) moving at speed \( v \) has an energy

\[ E(v) = \gamma mc^2, \]  

where the gamma factor is \( \gamma = 1/\sqrt{1 - v^2/c^2} \). The kinetic energy of the moving particle is \( \Delta E = E(v) - E(0) = (\gamma - 1) mc^2 \). Thus relativistic effects produce a fractional change in energy:

\[ \frac{\Delta E}{E} \sim \frac{v^2}{c^2}. \]  

This leads to energy differences between the various elliptical orbits of the same gross energy because the speed varies in different ways around the elliptical orbits, e.g. for a circular orbit and a highly elliptical orbit of the same gross energy. From eqns 1.3 and 1.7 we find that the ratio of the speed in the orbit to the speed of light is

\[ \frac{v}{c} = \frac{\alpha}{n}, \]  

where the fine-structure constant \( \alpha \) is given by

\[ \alpha = \frac{e^2/4\pi\varepsilon_0}{\hbar c}. \]  

This fundamental constant plays an important role throughout atomic physics. Numerically its value is approximately \( \alpha \approx 1/137 \) (see inside the back cover for a list of constants used in atomic physics). From eqn 1.17 we see that relativistic effects lead to energy differences of order \( \alpha^2 \) times the gross energy. (This crude estimate neglects some
dependence on principal quantum number and Chapter 2 gives a more quantitative treatment of this fine structure.) It is not necessary to go into all the refinements of Sommerfeld's relativistic theory that gave the energy levels in hydrogen very precisely, by imposing quantisation rules on classical orbits, since ultimately a paradigm shift was necessary. Those ideas were superseded by the use of wavefunctions in the Schrödinger equation. The idea of elliptical orbits provides a connection with our intuition based on classical mechanics and we often retain some traces of this simple picture of electron orbits in our minds. However, for atoms with more than one electron, e.g. helium, classical models do not work and we must think in terms of wavefunctions.

1.5 Moseley and the atomic number

At the same time as Bohr was working on his model of the hydrogen atom, H. G. J. Moseley measured the X-ray spectra of many elements. Moseley established that the square root of the frequency of the emitted lines is proportional to the atomic number \( Z \) (that he defined as the position of the atom in the periodic table, starting counting at \( Z = 1 \) for hydrogen), i.e.

\[
\sqrt{f} \propto Z. \quad (1.20)
\]

Moseley's original plot is shown in Fig. 1.2. As we shall see, this equation is a considerable simplification of the actual situation but it was remarkably powerful at the time. By ordering the elements using \( Z \) rather than relative atomic mass, as was done previously, several inconsistencies in the periodic table were resolved. There were still gaps that were later filled by the discovery of new elements. In particular, for the rare-earth elements that have similar chemical properties and are therefore difficult to distinguish, it was said 'in an afternoon, Moseley could solve the problem that had baffled chemists for many decades and establish the true number of possible rare earths' (Segrè 1980). Moseley's observations can be explained by a relatively simple model for atoms that extends Bohr's model for hydrogen.\(^{11}\)

A natural way to extend Bohr's atomic model to heavier atoms is to suppose that the electrons fill up the allowed orbits, starting from the bottom. Each energy level only has room for a certain number of electrons so they cannot all go into the lowest level and they arrange themselves in shells, labelled by the principal quantum number, around the nucleus. This shell structure arises because of the Pauli exclusion principle and the electron spin, but for now let us simply consider it as an empirical fact that the maximum number of electrons in the \( n = 1 \) shell is 2, the \( n = 2 \) shell has 8 and the \( n = 3 \) shell has 18, etc. For historical reasons, X-ray spectroscopists do not use the principal quantum number but label the shells by letters: K for \( n = 1 \), L for \( n = 2 \), M for \( n = 3 \) and so on alphabetically.\(^{12}\) This concept of electronic shells explains the emission of X-rays from atoms in the following way. Moseley produced X-rays by bombarding samples of the given element with electrons that

\(^{11}\)Tragically, Henry Gwyn Jeffreys Moseley was killed when he was only 28 while fighting in the First World War (see the biography by Heilbron (1974)).

\(^{12}\)The chemical properties of the elements depend on this electronic structure, e.g. the inert gases have full shells of electrons and these stable configurations are not willing to form chemical bonds. The explanation of the atomic structure underlying the periodic table is discussed further in Section 4.1. See also Atkins (1994) and Grant and Phillips (2001).
The handwriting in the bottom right corner states that this diagram is the original for Moseley's famous paper in *Phil. Mag.*, 27, 703 (1914).

Fig. 1.2 Moseley's plot of the square root of the frequency of X-ray lines of elements against their atomic number. Moseley's work established the atomic number $Z$ as a more fundamental quantity than the 'atomic weight' (now called relative atomic mass). Following modern convention the units of the horizontal scales would be $10^8 \sqrt{\text{Hz}}$ at the bottom and $10^{-10} \text{m}$ for the log scale at the top. (Archives of the Clarendon Laboratory, Oxford; also shown on the Oxford physics website.)
had been accelerated to a high voltage in a vacuum tube. These fast electrons knock an electron out of an atom in the sample leaving a vacancy or hole in one of its shells. This allows an electron from a higher-lying shell to ‘fall down’ to fill this hole emitting radiation of a wavelength corresponding to the difference in energy between the shells.

To explain Moseley’s observations quantitatively we need to modify the equations in Section 1.3, on Bohr’s theory, to account for the effect of a nucleus of charge greater than the +1e of the proton. For a nuclear charge Ze we replace $e^2/4\pi\varepsilon_0$ by $Ze^2/4\pi\varepsilon_0$ in all the equations, resulting in a formula for the energies like that of Balmer but multiplied by a factor of $Z^2$. This dependence on the square of the atomic number means that, for all but the lightest elements, transitions between low-lying shells lead to emission of radiation in the X-ray region of the spectrum. Scaling the Bohr theory result is accurate for hydrogenic ions, i.e. systems with one electron around a nucleus of charge Ze. In neutral atoms the other electrons (that do not jump) are not simply passive spectators but partly screen the nuclear charge; for a given X-ray line, say the K- to L-shell transition, a more accurate formula is

$$\frac{1}{\lambda} = R_\infty \left\{ \frac{(Z - \sigma_K)^2}{1^2} - \frac{(Z - \sigma_L)^2}{2^2} \right\}.$$  (1.21)

The screening factors $\sigma_K$ and $\sigma_L$ are not entirely independent of $Z$ and the values of these screening factors for each shell vary slightly (see the exercises at the end of this chapter). For large atomic numbers this formula tends to eqn 1.20 (see Exercise 1.4). This simple approach does not explain why the screening factor for a shell can exceed the number of electrons inside that shell, e.g. $\sigma_K = 2$ for $Z = 74$ although only one electron remains in this shell when a hole is formed. This does not make sense in a classical model with electrons orbiting around a nucleus, but can be explained by atomic wavefunctions—an electron with a high principal quantum number (and little angular momentum) has a finite probability of being found at small radial distances.

The study of X-rays has developed into a whole field of its own within atomic physics, astrophysics and condensed matter, but there is only room to mention a few brief facts here. When an electron is removed from the K-shell the atom has an amount of energy equal to its binding energy, i.e. a positive amount of energy, and it is therefore usual to draw the diagram with the K-shell at the top, as in Fig. 1.3. These are the energy levels of the hole in the electron shells. This diagram shows why the creation of a hole in a low-lying shell leads to a succession of transitions as the hole works its way outwards through the shells. The hole (or equivalently the falling electron) can jump more than one shell at a time; each line in a series from a given shell is labelled using Greek letters (as in the series in hydrogen), e.g. K$_\alpha$, K$_\beta$, ... The levels drawn in Fig. 1.3 have some sub-structure and this leads to transitions with slightly different wavelengths, as shown in Moseley’s plot. This is fine structure caused by relativistic effects that we considered for Sommerfeld’s theory; the substitution $e^2/4\pi\varepsilon_0 \rightarrow Ze^2/4\pi\varepsilon_0$, as above, (or
Fig. 1.3 The energy levels of the inner shells of the tungsten atom \((Z = 74)\) and the transitions between them that give rise to X-rays. The level scheme has several important differences from that for the hydrogen atom (Fig. 1.1).

Firstly, the energies are tens of keV, as compared to eV for \(Z = 1\), because they scale as \(Z^2\) (approximately).

Secondly, the energy levels are plotted with \(n = 1\) at the top because when an electron is removed from the K-shell the system has more energy than the neutral atom; energies are shown for an atom with a vacancy (missing electron) in the K-, L-, M- and N-shells. The atom emits X-ray radiation when an electron drops down from a higher shell to fill a vacancy in a lower shell—this process is equivalent to the vacancy, or 'hole', working its way outwards. This way of plotting the energies of the system shows clearly that the removal of an electron from the K-shell leads to a cascade of X-ray transitions, e.g. a transition between the \(n = 1\) and 2 shells gives a line in the K-series which is followed by a line in another series (L-, M-, etc.). When the vacancy reaches the outermost shells of electrons that are only partially filled with valence electrons with binding energies of a few eV (the O- and F-shells in the case of tungsten), the transition energies become negligible compared to those between the inner shells. This level scheme is typical for electrons in a moderately heavy atom, i.e. one with filled K-, L-, M- and N-shells. (The lines of the L-series shown dotted are allowed X-ray transitions, but they do not occur following K\(_\alpha\) emission—for reasons explained in Section 2.2.)

Equivalently \(\alpha \to Z\alpha\) shows that fine structure is of order \((Z\alpha)^2\) times the gross structure, which itself is proportional to \(Z^2\). Thus relativistic effects grow as \(Z^4\) and become very significant for the inner electrons of heavy atoms, leading to the fine structure of the L- and M-shells, seen in Fig. 1.3. This relativistic splitting of the shells explains why in Moseley’s plot (Fig. 1.2) there are two closely-spaced curves for the K\(_\alpha\)-line, and several curves for the L-series.

Nowadays much of the X-ray work in atomic physics is carried out using sources such as synchrotrons; these devices accelerate electrons by the techniques used in particle accelerators. A beam of high-energy electrons circulates in a ring and the circular motion causes the electrons to
radiate X-rays. Such a source can be used to obtain an X-ray absorption spectrum.\(^\text{14}\) There are many other applications of X-ray emission, e.g. as a diagnostic tool for the processes that occur in plasmas in fusion research and in astrophysical objects. Many interesting processes occur at ‘high energies’ in atomic physics but the emphasis in this book is mainly on lower energies.

### 1.6 Radiative decay

An electric dipole moment \(-eD\) oscillating at angular frequency \(\omega\) radiates a power\(^\text{15}\)

\[
P = \frac{e^2 D^2 \omega^4}{12\pi \varepsilon_0 c^5}.
\]

(1.22)

An electron in harmonic motion has a total energy\(^\text{16}\) of \(E = m_e \omega^2 D^2/2\), where \(D\) is the amplitude of the motion. This energy decreases at a rate equal to the power radiated:

\[
\frac{dE}{dt} = -\frac{e^2 \omega^2}{6\pi \varepsilon_0 m_e c^3} E = -\frac{E}{\tau},
\]

(1.23)

where the classical radiative lifetime \(\tau\) is given by

\[
\frac{1}{\tau} = \frac{e^2 \omega^2}{6\pi \varepsilon_0 m_e c^3}.
\]

(1.24)

For the transition in sodium at a wavelength of 589 nm (yellow light), this equation predicts a value of \(\tau = 16\) ns \(\simeq 10^{-8}\) s. This is very close to the experimentally measured value and typical of allowed transitions that emit visible light. Atomic lifetimes, however, vary over a very wide range,\(^\text{17}\) e.g. for the Lyman-\(\alpha\) transition (shown in Fig. 1.1) the upper level has a lifetime of only a few nanoseconds.\(^\text{18,19}\)

The classical value of the lifetime gives the fastest time in which the atom could decay on a given transition and this is often close to the observed lifetime for strong transitions. Atoms do not decay faster than a classical dipole radiating at the same wavelength, but they may decay more slowly (by many orders of magnitude in the case of forbidden transitions).\(^\text{20}\)

### 1.7 Einstein A and B coefficients

The development of the ideas of atomic structure was linked to experiments on the emission, and absorption, of radiation from atoms, e.g. X-rays or light. The emission of radiation was considered as something that just has to happen in order to carry away the energy when an electron jumps from one allowed orbit to another, but the mechanism was not explained.\(^\text{21}\) In one of his many strokes of genius Einstein devised a way of treating the phenomenon of spontaneous emission quantitatively,\(^\text{14}\) Absorption is easier to interpret than emission since only one of the terms in eqn 1.21 is important, e.g. \(E_K = \hbar c R_\infty (Z - \sigma_K)^2\).

\(^\text{15}\) This total power equals the integral of the Poynting vector over a closed surface in the far-field of radiation from the dipole. This is calculated from the oscillating electric and magnetic fields in this region (see electromagnetism texts or Corney (2000)).

\(^\text{16}\) The sum of the kinetic and potential energies.

\(^\text{17}\) The classical lifetime scales as \(\omega^2\). However, we will find that the quantum mechanical result is different (see Exercise 1.7).

\(^\text{18}\) Higher-lying levels, e.g. \(n = 30\), live for many microseconds (Gallagher 1994).

\(^\text{19}\) Atoms can be excited up to configurations with high principal quantum numbers in laser experiments; such systems are called Rydberg atoms and have small intervals between their energy levels. As expected from the correspondence principle, these Rydberg atoms can be used in experiments that probe the interface between classical and quantum mechanics.

\(^\text{20}\) The ion-trapping techniques described in Chapter 12 can probe transitions with spontaneous decay rates less than \(15^{-1}\), using single ions confined by electric and magnetic fields—something that was only a ‘thought experiment’ for Bohr and the other founders of quantum theory. In particular, the effect of individual quantum jumps between atomic energy levels is observed. Radiative decay resembles radioactive decay in that individual atoms spontaneously emit a photon at a given time but taking the average over an ensemble of atoms gives exponential decay.

\(^\text{21}\) A complete explanation of spontaneous emission requires quantum electrodynamics.
This treatment of the interaction of atoms with radiation forms the foundation for the theory of the laser, and is used whenever radiation interacts with matter (see Fox 2001). A historical account of Einstein’s work and its profound implications can be found in Pais (1982).

The frequency dependence of the interaction is considered in Chapter 7.

The word laser is an acronym for light amplification by stimulated emission of radiation.

Einstein considered atoms with two levels of energies, $E_1$ and $E_2$, as shown in Fig. 1.4; each level may have more than one state and the number of states with the same energy is the degeneracy of that level, represented by $g_1$ and $g_2$. Einstein considered what happens to an atom interacting with radiation of energy density $\rho(\omega)$ per unit frequency interval. The radiation causes transitions from the lower to the upper level at a rate proportional to $\rho(\omega)$, where the constant of proportionality is $B_{12}$. The atom interacts strongly only with that part of the distribution $\rho(\omega)$ with a frequency close to $\omega_{12} = (E_2 - E_1)/\hbar$, the atom’s resonant frequency. By symmetry it is also expected that the radiation will cause transitions from the upper to lower levels at a rate dependent on the energy density but with a constant of proportionality $B_{21}$ (the subscripts are in a different order for emission as compared to absorption). This is a process of stimulated emission in which the radiation at angular frequency $\omega$ causes the atom to emit radiation of the same frequency. This increase in the amount of light at the incident frequency is fundamental to the operation of lasers. The symmetry between up and down is broken by the process of spontaneous emission in which an atom falls down to the lower level, even when no external radiation is present. Einstein introduced the coefficient $A_{21}$ to represent the rate of this process. Thus the rate equations for the populations of the levels, $N_1$ and $N_2$, are

$$\frac{dN_2}{dt} = N_1 B_{12} \rho(\omega_{12}) - N_2 B_{21} \rho(\omega_{12}) - N_2 A_{21} \quad (1.25)$$

and

$$\frac{dN_1}{dt} = -\frac{dN_2}{dt}. \quad (1.26)$$

The first equation gives the rate of change of $N_2$ in terms of the absorption, stimulated emission and spontaneous emission, respectively. The second equation is a consequence of having only two levels so that atoms leaving level 2 must go into level 1; this is equivalent to a condition that $N_1 + N_2 = \text{constant}$. When $\rho(\omega) = 0$, and some atoms are initially in the upper level ($N_2(0) \neq 0$), the equations have a decaying exponential solution:

$$N_2(t) = N_2(0) \exp(-A_{21}t), \quad (1.27)$$

where the mean lifetime$^{25}$ is

$$\frac{1}{\tau} = A_{21}. \quad (1.28)$$

Fig. 1.4 The interaction of a two-level atom with radiation leads to stimulated transitions, in addition to the spontaneous decay of the upper level.
Einstein devised a clever argument to find the relationship between the \( A_{21} \) and \( B \)-coefficients and this allows a complete treatment of atoms interacting with radiation. Einstein imagined what would happen to such an atom in a region of black-body radiation, e.g. inside a box whose surface acts as a black body. The energy density of the radiation \( \rho(\omega) \, d\omega \) between angular frequency \( \omega \) and \( \omega + d\omega \) depends only on the temperature \( T \) of the emitting (and absorbing) surfaces of the box; this function is given by the Planck distribution law:\(^{26}\)

\[
\rho(\omega) = \frac{\hbar \omega^3}{\pi^2 c^3} \frac{1}{\exp(\hbar \omega/k_B T) - 1} .
\] (1.29)

Now we consider the level populations of an atom in this black-body radiation. At equilibrium the rates of change of \( N_1 \) and \( N_2 \) (in eqn 1.26) are both zero and from eqn 1.25 we find that

\[
\rho(\omega_{12}) = \frac{A_{21}}{B_{12}} \frac{1}{(N_1/N_2)(B_{12}/B_{21}) - 1} .
\] (1.30)

At thermal equilibrium the population in each of the states within the levels are given by the Boltzmann factor (the population in each state equals that of the energy level divided by its degeneracy):

\[
\frac{N_2}{g_2} = \frac{N_1}{g_1} \exp \left( - \frac{\hbar \omega}{k_B T} \right) .
\] (1.31)

Combining the last three equations (1.29, 1.30 and 1.31) we find\(^{27}\)

\[
A_{21} = \frac{\hbar \omega^3}{\pi^2 c^3} B_{21}
\] (1.32)

and

\[
B_{12} = \frac{g_2}{g_1} B_{21} .
\] (1.33)

The Einstein coefficients are properties of the atom.\(^{28}\) Therefore these relationships between them hold for any type of radiation, from narrow-bandwidth radiation from a laser to broadband light. Importantly, eqn 1.32 shows that strong absorption is associated with strong emission. Like many of the topics covered in this chapter, Einstein’s treatment captured the essential features of the physics long before all the details of the quantum mechanics were fully understood.\(^{29}\)

1.8 The Zeeman effect

This introductory survey of early atomic physics must include Zeeman’s important work on the effect of a magnetic field on atoms. The observation of what we now call the Zeeman effect and three other crucial experiments were carried out just at the end of the nineteenth century, and together these discoveries mark the watershed between classical and quantum physics.\(^{30}\) Before describing Zeeman’s work in detail, I shall

\(^{26}\)Planck was the first to consider radiation quantised into photons of energy \( \hbar \omega \). See Pais (1986).

\(^{27}\)These equations hold for all \( T \), so we can equate the parts that contain \( \exp(\hbar \omega/k_B T) \) and the temperature-independent factors separately to obtain the two equations.

\(^{28}\)This is shown explicitly in Chapter 7 by a time-dependent perturbation theory calculation of \( B_{12} \).

\(^{29}\)To excite a significant fraction of the population into the upper level of a visible transition would require black-body radiation with a temperature comparable to that of the sun, and this method is not generally used in practice—such transitions are easily excited in an electrical discharge where the electrons impart energy to the outermost electrons in an atom. (The voltage required to excite weakly-bound outer electrons is much less than for X-ray production.)

\(^{30}\)Pais (1986) and Segrè (1980) give historical accounts.
briefly mention the other three great breakthroughs and their significance for atomic physics. Röntgen discovered mysterious X-rays emitted from discharges, and sparks, that could pass through matter and blacken photographic film. At the same time, Bequerel’s discovery of radioactivity opened up the whole field of nuclear physics. Another great breakthrough was J. J. Thomson’s demonstration that cathode rays in electrical discharge tubes are charged particles whose charge-to-mass ratio does not depend on the gas in the discharge tube. At almost the same time, the observation of the Zeeman effect of a magnetic field showed that there are particles with the same charge-to-mass ratio in atoms (that we now call electrons). The idea that atoms contain electrons is very obvious now but at that time it was a crucial piece in the jigsaw of atomic structure that Bohr put together in his model. In addition to its historical significance, the Zeeman effect provides a very useful tool for examining the structure of atoms, as we shall see at several places in this book. Somewhat surprisingly, it is possible to explain this effect by a classical-mechanics line of reasoning (in certain special cases). An atom in a magnetic field can be modelled as a simple harmonic oscillator. The restoring force on the electron is the same for displacements in all directions and the oscillator has the same resonant frequency $\omega_0$ for motion along the $x$-, $y$- and $z$-directions (when there is no magnetic field). In a magnetic field $\mathbf{B}$ the situation is that for an electron with charge $-e$, position $\mathbf{r}$ and velocity $\mathbf{v} = \dot{\mathbf{r}}$ is

$$m_e \frac{d\mathbf{v}}{dt} = -m_e \omega_0^2 \mathbf{r} - e \mathbf{v} \times \mathbf{B}.$$  \hspace{1cm} (1.34)

In addition to the restoring force (assumed to exist without further explanation), there is the Lorentz force that occurs for a charged particle moving through a magnetic field. Taking the direction of the field to be the $z$-axis, $\mathbf{B} = B\mathbf{\hat{z}}$, leads to

$$\dot{\mathbf{r}} + 2\Omega_L \mathbf{r} \times \mathbf{\hat{z}} + \omega_0^2 \mathbf{r} = 0.$$ \hspace{1cm} (1.35)

This contains the Larmor frequency

$$\Omega_L = \frac{eB}{2m_e}.$$ \hspace{1cm} (1.36)

We use a matrix method to solve the equation and look for a solution in the form of a vector oscillating at $\omega$:

$$\mathbf{r} = \text{Re} \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \exp(-i\omega t) \right\}.$$ \hspace{1cm} (1.37)

Written in matrix form, eqn 1.35 reads

$$\begin{pmatrix} \omega_0^2 & -2i\omega \Omega_L & 0 \\ 2i\omega \Omega_L & \omega_0^2 & 0 \\ 0 & 0 & \omega_0^2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \omega^2 \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$ \hspace{1cm} (1.38)
The eigenvalues \( \omega^2 \) are found from the following determinant:

\[
\begin{vmatrix}
\omega_0^2 - \omega^2 & -2i\omega \Omega_L & 0 \\
2i\omega \Omega_L & \omega_0^2 - \omega^2 & 0 \\
0 & 0 & \omega_0^2 - \omega^2
\end{vmatrix} = 0.
\]

(1.39)

This gives \( \{ \omega^4 - (2\omega_0^2 + 4\Omega_L^2) \omega^2 + \omega_0^4 \} (\omega^2 - \omega_0^2) = 0 \). The solution \( \omega = \omega_0 \) is obvious by inspection. The other two eigenvalues can be found exactly by solving the quadratic equation for \( \omega^2 \) inside the curly brackets.

For an optical transition we always have \( \Omega_L \ll \omega_0 \) so the approximate eigenfrequencies are \( \omega \simeq \omega_0 \pm \Omega_L \). Substituting these values back into eqn 1.38 gives the eigenvectors corresponding to \( \omega = \omega_0 - \Omega_L \), \( \omega_0 \) and \( \omega_0 + \Omega_L \), respectively, as

\[
r = \begin{pmatrix}
\cos (\omega_0 - \Omega_L) t \\
-\sin (\omega_0 - \Omega_L) t \\
0
\end{pmatrix}, \\
\begin{pmatrix}
0 \\
0 \\
\cos \omega_0 t
\end{pmatrix}, \\
\begin{pmatrix}
\cos (\omega_0 + \Omega_L) t \\
\sin (\omega_0 + \Omega_L) t \\
0
\end{pmatrix}
\]

and

The magnetic field does not affect motion along the \( z \)-axis and the angular frequency of the oscillation remains \( \omega_0 \). Interaction with the magnetic field causes the motions in the \( x \)- and \( y \)-directions to be coupled together (by the off-diagonal elements \( \pm 2i \omega \Omega_L \) of the matrix in eqn 1.38).\(^{34}\) The result is two circular motions in opposite directions in the \( xy \)-plane, as illustrated in Fig. 1.5. These circular motions have frequencies shifted up, or down, from \( \omega_0 \) by the Larmor frequency. Thus the action of the external field splits the original oscillation at a single frequency (actually three independent oscillations all with the same frequency, \( \omega_0 \)) into three separate frequencies. An oscillating electron acts as a classical dipole that radiates electromagnetic waves and Zeeman observed the frequency splitting \( \Omega_L \) in the light emitted by the atom.

This classical model of the Zeeman effect explains the polarization of the light, as well as the splitting of the lines into three components. The calculation of the polarization of the radiation at each of the three different frequencies for a general direction of observation is straightforward using vectors;\(^{35}\) however, only the particular cases where the radiation propagates parallel and perpendicular to the magnetic field are considered here, i.e. the longitudinal and transverse directions of observation, respectively. An electron oscillating parallel to \( \mathbf{B} \) radiates an electromagnetic wave with linear polarization and angular frequency \( \omega_0 \). This \( \pi \)-component of the line is observed in all directions except along the magnetic field;\(^{36}\) in the special case of transverse observation (i.e. in the \( xy \)-plane) the polarization of the \( \pi \)-component lies along \( \hat{e}_z \). The circular motion of the oscillating electron in the \( xy \)-plane at angular frequencies \( \omega_0 + \Omega_L \) and \( \omega_0 - \Omega_L \) produces radiation at these frequencies. Looking transversely, this circular motion is seen edge-on so that it looks like linear sinusoidal motion, e.g. for observation along

\(^{34}\) The matrix does not have off-diagonal elements in the last column or bottom row, so the \( x \)- and \( y \)-components are not coupled to the \( z \)-component, and the problem effectively reduces to solving a \( 2 \times 2 \) matrix.

\(^{35}\) Some further details are given in Section 2.2 and in Woodgate (1980).

\(^{36}\) An oscillating electric dipole proportional to \( \hat{e}_z \cos \omega_0 t \) does not radiate along the \( z \)-axis—observation along this direction gives a view along the axis of the dipole so that effectively the motion of the electron cannot be seen.
The $x$-axis only the $y$-component is seen, and the radiation is linearly polarized perpendicular to the magnetic field—see Fig. 1.6. These are called the $\sigma$-components and, in contrast to the $\pi$-component, they are also seen in longitudinal observation—looking along the $z$-axis one sees the electron’s circular motion and hence light that has circular polarization. Looking in the opposite direction to the magnetic field (from the positive $z$-direction, or $\theta = 0$ in polar coordinates) the circular motion in the anticlockwise direction is associated with the frequency $\omega_0 + \Omega_L$. In addition to showing that atoms contain electrons by measuring the magnitude of the charge-to-mass ratio $e/m_e$, Zeeman also deduced the sign of the charge by considering the polarization of the emitted light. If the sign of the charge was not negative, as we assumed from the start, light at $\omega_0 + \Omega_L$ would have the opposite handedness—from this Zeeman could deduce the sign of the electron’s charge.

For situations that only involve orbital angular momentum (and no spin) the predictions of this classical model correspond exactly to those of quantum mechanics (including the correct polarizations), and the intuition gained from this model gives useful guidance in more complicated cases. Another reason for studying the classical treatment of the Zeeman effect is that it furnishes an example of degenerate perturbation theory in classical mechanics. We shall encounter degenerate perturbation theory in quantum mechanics in several places in this book and an understanding of the analogous procedure in classical mechanics is very helpful.

37 This is left-circularly-polarized light (Corney 2000).
1.8.1 Experimental observation of the Zeeman effect

Figure 1.7(a) shows an apparatus suitable for the experimental observation of the Zeeman effect and Fig. 1.7(b–e) shows some typical experimental traces. A low-pressure discharge lamp that contains the atom to be studied (e.g. helium or cadmium) is placed between the pole pieces of an electromagnet capable of producing fields of up to about 1 T. In the arrangement shown, a lens collects light emitted perpendicular to the field (transverse observation) and sends it through a Fabry–Perot étalon. The operation of such étalons is described in detail by Brooker (2003), and only a brief outline of the principle of operation is given here.

![Diagram](image)

Fig. 1.7 (a) An apparatus suitable for the observation of the Zeeman effect. The light emitted from a discharge lamp, between the pole pieces of the electromagnet, passes through a narrow-band filter and a Fabry–Perot étalon. Key: L1, L2 are lenses; F - filter; P - polarizer to discriminate between π- and σ-polarizations (optional); Fabry–Perot étalon made of a rigid spacer between two highly-reflecting mirrors (M1 and M2); D - detector. Other details can be found in Brooker (2003). A suitable procedure is to (partially) evacuate the étalon chamber and then allow air (or a gas with a higher refractive index such as carbon dioxide) to leak in through a constant-flow-rate valve to give a smooth linear scan. Plots (b) to (e) show the intensity I of light transmitted through the Fabry–Perot étalon. (b) A scan over two free-spectral ranges with no magnetic field. Both (c) and (d) show a Zeeman pattern observed perpendicular to the applied field; the spacing between the π- and σ-components in these scans is one-quarter and one-third of the free-spectral range, respectively—the magnetic field in scan (c) is weaker than in (d). (e) In longitudinal observation only the σ-components are observed—this scan is for the same field as in (c) and the σ-components have the same position in both traces.
• Light from the lamp is collected by a lens and directed on to an interference filter that transmits only a narrow band of wavelength corresponding to a single spectral line.

• The étalon produces an interference pattern that has the form of concentric rings. These rings are observed on a screen in the focal plane of the lens placed after the étalon. A small hole in the screen is positioned at the centre of the pattern so that light in the region of the central fringe falls on a detector, e.g. a photodiode. (Alternatively, the lens and screen can be replaced by a camera that records the ring pattern on film.)

• The effective optical path length between the two flat highly-reflecting mirrors is altered by changing the pressure of the air in the chamber; this scans the étalon over several free-spectral ranges while the intensity of the interference fringes is recorded to give traces as in Fig. 1.7(b–e).

1.9 Summary of atomic units

This chapter has used classical mechanics and elementary quantum ideas to introduce the important scales in atomic physics: the unit of length \( a_0 \) and a unit of energy \( h c R_\infty \). The natural unit of energy is \( \frac{e^2}{4\pi\varepsilon_0 a_0} \) and this unit is called a hartree.\(^{38}\) This book, however, expresses energy in terms of the energy equivalent to the Rydberg constant, 13.6 eV; this equals the binding energy in the first Bohr orbit of hydrogen, or 1/2 a hartree. These quantities have the following values:

\[
a_0 = \frac{\hbar^2}{(e^2/4\pi\varepsilon_0)m_e} = 5.29 \times 10^{-11} \text{ m},
\]

\[
hcR_\infty = \frac{m_e(e^2/4\pi\varepsilon_0)^2}{2\hbar^2} = 13.6 \text{ eV}.
\]

The use of these atomic units makes the calculation of other quantities simple, e.g. the electric field in a hydrogen atom at radius \( r = a_0 \) equals \( e/(4\pi\varepsilon_0 a_0^2) \). This corresponds to a potential difference of 27.2 V over a distance of \( a_0 \), or a field of \( 5 \times 10^{11} \text{ V m}^{-1} \).

Relativistic effects depend on the dimensionless fine-structure constant \( \alpha \):

\[
\alpha = \frac{(e^2/4\pi\varepsilon_0)}{\hbar c} \simeq \frac{1}{137}.
\]

The Zeeman effect of a magnetic field on atoms leads to a frequency shift of \( \Omega_L \) in eqn 1.36.\(^{39}\) In practical units the size of this frequency shift is

\[
\frac{\Omega_L}{2\pi B} = \frac{e}{4\pi m_e} = 14 \text{ GHz T}^{-1}.
\]

Equating the magnetic energy \( \hbar \Omega_L \) with \( \mu_B B \), the magnitude of the energy for a magnetic moment \( \mu_B \) in a magnetic flux density \( B \), shows that the unit of atomic magnetic moment is the Bohr magneton

\[
\mu_B = \frac{e\hbar}{2m_e} = 9.27 \times 10^{-24} \text{ J T}^{-1}.
\]

\(^{38}\) It equals the potential energy of the electron in the first Bohr orbit.

\(^{39}\) This Larmor frequency equals the splitting between the \( \pi \)- and \( \sigma \)-components in the normal Zeeman effect.
This magnetic moment depends on the properties of the unpaired electron (or electrons) in the atom, and has a similar magnitude for all atoms. In contrast, other atomic properties scale rapidly with the nuclear charge; hydrogenic systems have energies proportional to $Z^2$, and the same reasoning shows that their size is proportional to $1/Z$ (see eqns 1.40 and 1.41). For example, hydrogenic uranium $U^{+91}$ has been produced in accelerators by stripping 91 electrons off a uranium atom to leave a single electron that has a binding energy of $92^2 \times 13.6\,\text{eV} = 115\,\text{keV}$ (for $n = 1$) and an orbit of radius $a_0/92 = 5.75 \times 10^{-13}\,\text{m} \equiv 575\,\text{fm}$. The transitions between the lowest energy levels of this system have short wavelengths in the X-ray region.\textsuperscript{40}

The reader might think that it would be a good idea to use the same units across the whole of atomic physics. In practice, however, the units reflect the actual experimental techniques used in a particular region of the spectrum, e.g. radio-frequency, or microwave synthesisers, are calibrated in Hz (kHz, MHz and GHz); the equation for the angle of diffraction from a grating is expressed in terms of a wavelength; and for X-rays produced by tubes in which electrons are accelerated by high voltages it is natural to use keV.\textsuperscript{41} A table of useful conversion factors is given inside the back cover.

The survey of classical ideas in this chapter gives a historical perspective on the origins of atomic physics but it is not necessary, or indeed in some cases downright confusing, to go through a detailed classical treatment—the physics at the scale of atomic systems can only properly be described by wave mechanics and this is the approach used in the following chapters.\textsuperscript{42}

\textbf{Exercises}

(1.1) \textit{Isotope shift}

The deuteron has approximately twice the mass of the proton. Calculate the difference in the wavelength of the Balmer-\(\alpha\) line in hydrogen and deuterium.

(1.2) \textit{The energy levels of one-electron atoms}

The table gives the wavelength\textsuperscript{43} of lines observed in the spectrum of atomic hydrogen and singly-ionized helium. Explain as fully as possible the similarities and differences between the two spectra.

<table>
<thead>
<tr>
<th>H (nm)</th>
<th>He(^{+}) (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>656.28</td>
<td>656.01</td>
</tr>
<tr>
<td>486.13</td>
<td>541.16</td>
</tr>
<tr>
<td>434.05</td>
<td>485.93</td>
</tr>
<tr>
<td>410.17</td>
<td>454.16</td>
</tr>
<tr>
<td></td>
<td>433.87</td>
</tr>
<tr>
<td></td>
<td>419.99</td>
</tr>
<tr>
<td></td>
<td>410.00</td>
</tr>
</tbody>
</table>

\textsuperscript{40}Energies can be expressed in terms of the rest mass energy of the electron $m_e c^2 = 0.511\,\text{MeV}$. The gross energy is $(Z\alpha)^2 \frac{3}{4} m_e c^2$ and the fine structure is of order $(Z\alpha)^4 \frac{1}{2} m_e c^2$.

\textsuperscript{41}Laser techniques can measure transition frequencies of around $10^{15}\,\text{Hz}$ directly as a frequency to determine a precise value of the Rydberg constant, and there are no definite rules for whether a transition should be specified by its energy, wavelength or frequency.

\textsuperscript{42}X-ray spectra are not discussed again in this book and further details can be found in Kuhn (1969) and other atomic physics texts.

\textsuperscript{43}These are the wavelengths in air with a refractive index of 1.0003 in the visible region.
(1.3) **Relativistic effects**
Evaluate the magnitude of relativistic effects in the \( n = 2 \) level of hydrogen. What is the resolv-
ing power \( \lambda/\Delta \lambda_{\text{min}} \) of an instrument that could observe these effects in the Balmer-\( \alpha \) line?

(1.4) **X-rays**
Show that eqn 1.21 approximates to eqn 1.20 when the atomic number \( Z \) is much greater than the screening factors.

(1.5) **X-rays**
It is suspected that manganese (\( Z = 25 \)) is very poorly mixed with iron (\( Z = 26 \)) in a block of al-
loy. Predict the energies of the K-absorption edges of these elements and determine an X-ray photon en-
ty that would give good contrast (between regions of different concentrations) in an X-ray of the block.

(1.6) **X-ray experiments**
Sketch an apparatus suitable for X-ray spectroscopy of elements, e.g. Moseley's experiment. Describe the principle of its operation and the method of measuring the energy, or wavelength, of X-rays.

(1.7) **Fine structure in X-ray transitions**
Estimate the magnitude of the relativistic effects in the M-shell of lead (\( Z = 82 \)) in keV. Also ex-
press you answer as a fraction of the K\( \alpha \) transition.

(1.8) **Radiative lifetime**
For an electron in a circular orbit of radius \( r \) the electric dipole moment has a magnitude of \( D = -e r \) and radiates energy at a rate given by eqn 1.22. Find the time \( \tau \) taken to lose an energy of \( \hbar \omega \).

Use your expression to estimate \( \tau \) for the \( n = 3 \) to \( n = 2 \) transition in hydrogen that emits light of wavelength 656 nm.

**Comment.** This method gives \( 1/\tau \propto (er)^2 \omega^3 \), which corresponds closely to the quantum mechanical result in eqn 7.23.

(1.9) **Black-body radiation**
Two-level atoms with a transition at wavelength \( \lambda = 600 \) nm, between the levels with degeneracies \( g_1 = 1 \) and \( g_2 = 3 \), are immersed in black-body radiation. The fraction in the excited state is 0.1. What is the temperature of the black body and the energy density per unit frequency interval \( \rho(\omega_{12}) \) of the radiation at the transition frequency?

(1.10) **Zeeman effect**
What is the magnitude of the Zeeman shift for an atom in (a) the Earth's magnetic field, and (b) a magnetic flux density of 1 T? Express your answer in both MHz, and as a fraction of the transition frequency \( \Delta f/f \) for a spectral line in the visible.

(1.11) **Relative intensities in the Zeeman effect**
Without an external field, an atom has no preferred direction and the choice of quantisation axis is arbitrary. In these circumstances the light emitted cannot be polarized (since this would establish a preferred orientation). As a magnetic field is gradually turned on we do not expect the intensities of the different Zeeman components to change discontinuously because the field has little effect on transition rates. This physical argument implies that oppositely-polarized components emit-
ted along the direction of the field must have equal intensities, i.e. \( I_{+} = I_{-} \) (notation defined in Fig. 1.6).

What can you deduce about
(a) the relative intensities of the components emitted perpendicularly to the field?
(b) the ratio of the total intensities of light emit-
ted along and perpendicular to the field?

(1.12) **Bohr theory and the correspondence principle**
This exercise gives an alternative approach to the theory of the hydrogen atom presented in Sec-
tion 1.3 that is close to the spirit of Bohr's original papers. It is somewhat more subtle than that usu-
ally given in elementary textbooks and illustrates Bohr's great intuition. Rather than the ad hoc as-
sumption that angular momentum is an integral multiple of \( \hbar \) (in eqn 1.7), Bohr used the corre-
spondence principle. This principle relates the beha-

**Assumption II** The correspondence principle states that in the limit of large quantum numbers a quantum system tends to the same limit as the corresponding classical system.

Bohr formulated this principle in the early days of quantum theory. To apply this principle to hy-
drogen we first calculate the energy gap between adjacent electron orbits of radii \( r \) and \( r' \). For large radii, the change \( \Delta r = r' - r \ll r \).

(a) Show that the angular frequency \( \omega = \Delta E/\hbar \) of radiation emitted when an electron makes a quantum jump between these levels is
\[
\omega \approx \frac{e^2}{4\pi\epsilon_0} \frac{\Delta r}{2\hbar r^2}.
\]

(b) An electron moving in a circle of radius \( r \) acts as an electric dipole radiating energy at the
orbital frequency $\omega$ given by eqn 1.4. Verify that this equation follows from eqn 1.3.

(c) In the limit of large quantum numbers, the quantum mechanical and classical expressions give the same frequency $\omega$. Show that equating the expressions in the previous parts yields $\Delta r = 2(\alpha_0 r)^{1/2}$.

(d) The difference in the radii between two adjacent orbits can be expressed as a difference equation.\footnote{A difference equation is akin to a differential equation but without letting the differences become infinitesimal.} In this case $\Delta n = 1$ and

$$\frac{\Delta r}{\Delta n} \propto r^{1/2}. \quad (1.45)$$

This equation can be solved by assuming that the radius varies as some power $x$ of the quantum number $n$, e.g. if one orbit is labelled by an integer $n$ and the next by $n+1$, then $r = an^x$ and $r' = a(n+1)^x$. Show that $\Delta r = axn^{x-1} \propto n^{x/2}$. Determine the power $x$ and the constant $a$.

Comment. We have found eqn 1.3 from the correspondence principle without considering angular momentum. The allowed energy levels are easily found from this equation as in Section 1.3. The remarkable feature is that, although the form of the equation was derived for high values of the principal quantum number, the result works down to $n = 1$.

(1.13) Rydberg atoms

(a) Show that the energy of the transitions between two shells with principal quantum numbers $n$ and $n' = n + 1$ is proportional to $1/n^3$ for large $n$.

(b) Calculate the frequency of the transition between the $n' = 51$ and $n = 50$ shells.

(c) What is the size of an atom in these Rydberg states? Express your answer both in atomic units and in metres.
The hydrogen atom

The simple hydrogen atom has had a great influence on the development of quantum theory, particularly in the first half of the twentieth century when the foundations of quantum mechanics were laid. As measurement techniques improved, finer and finer details were resolved in the spectrum of hydrogen until eventually splittings of the lines were observed that cannot be explained even by the fully relativistic formulation of quantum mechanics, but require the more advanced theory of quantum electrodynamics. In the first chapter we looked at the Bohr–Sommerfeld theory of hydrogen that treated the electron orbits classically and imposed quantisation rules upon them. This theory accounted for many of the features of hydrogen but it fails to provide a realistic description of systems with more than one electron, e.g. the helium atom. Although the simple picture of electrons orbiting the nucleus, like planets round the sun, can explain some phenomena, it has been superseded by the Schrödinger equation and wavefunctions. This chapter outlines the application of this approach to solve Schrödinger’s equation for the hydrogen atom; this leads to the same energy levels as the Bohr model but the wavefunctions give much more information, e.g. they allow the rates of the transitions between levels to be calculated (see Chapter 7). This chapter also shows how the perturbations caused by relativistic effects lead to fine structure.

2.1 The Schrödinger equation

The solution of the Schrödinger equation for a Coulomb potential is in every quantum mechanics textbook and only a brief outline is given here.\(^1\) The Schrödinger equation for an electron of mass \(m_e\) in a spherically-symmetric potential is

\[
\begin{cases}
-\hbar^2 & \nabla^2 + V(r) \\
2m_e
\end{cases} \psi = E\psi.
\]

(2.1)

This is the quantum mechanical counterpart of the classical equation for the conservation of total energy expressed as the sum of kinetic and potential energies.\(^2\) In spherical polar coordinates we have

\[
\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) - \frac{1}{r^2} l^2,
\]

(2.2)

\(^1\) The emphasis is on the properties of the wavefunctions rather than how to solve differential equations.

\(^2\) The operator for linear momentum is \(p = -i\hbar \nabla\) and for angular momentum it is \(\hbar \mathbf{l} = \mathbf{r} \times \mathbf{p}\). This notation differs in two ways from that commonly used in quantum texts. Firstly, \(\hbar\) is taken outside the angular momentum operators, and secondly, the operators are written without ‘hats’. This is convenient for atomic physics, e.g. in the vector model for the addition of angular momenta.
where the operator \( l^2 \) contains the terms that depend on \( \theta \) and \( \phi \), namely
\[
l^2 = -\left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right\},
\]
(2.3)
and \( \hbar^2 l^2 \) is the operator for the orbital angular momentum squared. Following the usual procedure for solving partial differential equations, we look for a solution in the form of a product of functions \( \psi = R(r)Y(\theta, \phi) \). The equation separates into radial and angular parts as follows:
\[
\frac{1}{R} \frac{\partial}{\partial r} \left( r^2 \frac{\partial R}{\partial r} \right) - \frac{2m_e e^2}{\hbar^2} \{ V(r) - E \} = \frac{1}{Y} l^2 Y.
\]
(2.4)

Each side depends on different variables and so the equation is only satisfied if both sides equal a constant that we call \( b \). Thus
\[
l^2 Y = b Y.
\]
(2.5)

This is an eigenvalue equation and we shall use the quantum theory of angular momentum operators to determine the eigenfunctions \( Y(\theta, \phi) \).

### 2.1.1 Solution of the angular equation

To continue the separation of variables we substitute \( Y = \Theta(\theta) \Phi(\phi) \) into eqn 2.5 to obtain
\[
\sin \theta \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \Theta}{\partial \theta} \right) + b \sin^2 \theta = \frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \phi^2} = \text{const}.
\]
(2.6)

The equation for \( \Phi(\phi) \) is the same as in simple harmonic motion, so
\[
\Phi = Ae^{im\phi} + Be^{-im\phi}.
\]
(2.7)

The constant on the right-hand side of eqn 2.6 has the value \( m^2 \). Physically realistic wavefunctions have a unique value at each point and this imposes the condition \( \Phi(\phi + 2\pi) = \Phi(\phi) \), so \( m \) must be an integer. The function \( \Phi(\phi) \) is the sum of eigenfunctions of the operator for the \( z \)-component of orbital angular momentum
\[
\hbar l_z = -i\hbar \frac{\partial}{\partial \phi}.
\]
(2.8)

The function \( e^{im\phi} \) has magnetic quantum number \( m \) and its complex conjugate \( e^{-im\phi} \) has magnetic quantum number \( -m \).

A convenient way to find the function \( Y(\theta, \phi) \) and its eigenvalue \( b \) in eqn 2.5 is to use the ladder operators \( l_+ = l_x + il_y \) and \( l_- = l_x - il_y \). These operators commute with \( l^2 \), the operator for the total angular momentum squared (because \( l_x \) and \( l_y \) commute with \( l^2 \)); therefore, the three functions \( Y \), \( l_+ Y \) and \( l_- Y \) are all eigenfunctions of \( l^2 \) with the same eigenvalue \( b \) (if they are non-zero, as discussed below). The ladder

3. A and \( B \) are arbitrary constants. Alternatively, the solutions can be written in terms of real functions as \( A'\sin(m\phi) + B'\cos(m\phi) \).

4. The operator \(-\partial^2/\partial \phi^2 \equiv l_z^2 \) and consequently \( \Phi(\phi) \) is an eigenfunction of \( l_z^2 \) with eigenvalue \( m^2 \).

5. The solution of equations involving the angular part of \( \nabla^2 \) arises in many situations with spherical symmetry, e.g., in electrostatics, and the same mathematical tools could be used here to determine the properties of the spherical harmonic functions, but angular momentum methods give more physical insight for atoms.
operators can be expressed in polar coordinates as:

\[ l_+ = e^{i\phi} \left( \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) \]
\[ l_- = e^{-i\phi} \left( -\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right). \tag{2.9} \]

The operator \( l_+ \) transforms a function with magnetic quantum number \( m \) into another angular momentum eigenfunction that has eigenvalue \( m + 1 \). Thus \( l_+ \) is called the raising operator.\(^6\) The lowering operator \( l_- \) changes the magnetic quantum number in the other direction, \( m \to m - 1 \). It is straightforward to prove these statements and other properties of these operators;\(^7\) however, the purpose of this section is not to present the general theory of angular momentum but simply to outline how to find the eigenfunctions (of the angular part) of the Schrödinger equation.

Repeated application of the raising operator does not increase \( m \) indefinitely—for each eigenvalue \( b \) there is a maximum value of the magnetic quantum number\(^8\) that we shall call \( l \), i.e. \( m_{\text{max}} = l \). The raising operator acting on an eigenfunction with \( m_{\text{max}} \) gives zero since by definition there are no eigenfunctions with \( m > m_{\text{max}} \). Thus solving the equation \( l_+ Y = 0 \) (Exercise 2.11) we find that the eigenfunctions with \( m_{\text{max}} = l \) have the form

\[ Y \propto \sin^l \theta e^{il\phi}. \tag{2.10} \]

Substitution back into eqn 2.5 shows that these are eigenfunctions\(^1\) with eigenvalue \( b = l(l + 1) \), and \( l \) is the orbital angular momentum quantum number. The functions \( Y_l m(\theta, \phi) \) are labelled by their eigenvalues in the conventional way.\(^9\) For \( l = 0 \) only \( m = 0 \) exists and \( Y_{0,0} \) is a constant with no angular dependence. For \( l = 1 \) we can find the eigenfunctions by starting from the one with \( l = 1 = m \) (in eqn 2.10) and using the lowering operator to find the others:

\[ Y_{1,1} \propto \sin \theta e^{i\phi}, \]
\[ Y_{1,0} \propto l_- Y_{1,1} \propto \cos \theta, \]
\[ Y_{1,-1} \propto l_- Y_{1,0} \propto \sin \theta e^{-i\phi}. \]

This gives all three eigenfunctions expected for \( l = 1.\(^10\) For \( l = 2 \) this procedure gives

\[ Y_{2,2} \propto \sin^2 \theta e^{i2\phi}, \]
\[ \vdots \]
\[ Y_{2,-2} \propto \sin^2 \theta e^{-i2\phi}. \]

These are the five eigenfunctions with \( m = 2, 1, 0, -1, -2.\(^11\) Normalised angular functions are given in Table 2.1.

Any angular momentum eigenstate can be found from eqn 2.10 by
Table 2.1 Orbital angular momentum eigenfunctions.

\[
\begin{align*}
Y_{0,0} &= \sqrt{\frac{1}{4\pi}} \\
Y_{1,0} &= \sqrt{\frac{3}{4\pi}} \cos \theta \\
Y_{1,\pm 1} &= \pm \sqrt{\frac{3}{8\pi}} \sin \theta \, e^{\pm i\phi} \\
Y_{2,0} &= \sqrt{\frac{5}{16\pi}} \left(3\cos^2 \theta - 1\right) \\
Y_{2,\pm 1} &= \pm \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta \, e^{\pm i\phi} \\
Y_{2,\pm 2} &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta \, e^{\pm 2i\phi} \\
\end{align*}
\]

Normalisation: \( \int_0^{2\pi} \int_0^\pi |Y_{l,m}|^2 \sin \theta \, d\theta \, d\phi = 1 \)

repeated application of the lowering operator:\(^{12}\)

\[
Y_{l,m} \propto (l_-)^{l-m} \sin^l \theta \, e^{il\phi}.
\] (2.11)

To understand the properties of atoms, it is important to know what the wavefunctions look like. The angular distribution needs to be multiplied by the radial distribution, calculated in the next section, to give the square of the wavefunction as

\[
|\psi (r, \theta, \phi)|^2 = R_{n,l}^2 (r) |Y_{l,m} (\theta, \phi)|^2.
\] (2.12)

This is the probability distribution of the electron, or \(-e|\psi|^2\) can be interpreted as the electronic charge distribution. Many atomic properties, however, depend mainly on the form of the angular distribution and Fig. 2.1 shows some plots of \(|Y_{l,m}|^2\). The function \(|Y_{0,0}|^2\) is spherically symmetric. The function \(|Y_{1,0}|^2\) has two lobes along the \(z\)-axis. The squared modulus of the other two eigenfunctions of \(l = 1\) is proportional to \(\sin^2 \theta\). As shown in Fig. 2.1(c), there is a correspondence between these distributions and the circular motion of the electron around the \(z\)-axis that we found as the normal modes in the classical theory of the Zeeman effect (in Chapter 1).\(^{13}\) This can be seen in Cartesian coordinates where

\[
\begin{align*}
Y_{1,0} &\propto \frac{z}{r}, \\
Y_{1,1} &\propto \frac{x + iy}{r}, \\
Y_{1,-1} &\propto \frac{x - iy}{r}.
\end{align*}
\] (2.13)

\(^{12}\)This eigenfunction has magnetic quantum number \(l - (l - m) = m\).

\(^{13}\)Stationary states in quantum mechanics correspond to the time-averaged classical motion. In this case both directions of circular motion about the \(x\)-axis give the same distribution.
Fig. 2.1 Polar plots of the squared modulus of the angular wavefunctions for the hydrogen atom with \( l = 0 \) and 1. For each value of the polar angle \( \theta \) a point is plotted at a distance proportional to \( |Y(\theta, \phi)|^2 \) from the origin. Except for (d), the plots have rotational symmetry about the \( z \)-axis and look the same for any value of \( \phi \). (a) \( |Y_{0,0}|^2 \) is spherical. (b) \( |Y_{1,0}|^2 \propto \cos^{2\theta} \) has two lobes along the \( z \)-axis. (c) \( |Y_{1,1}|^2 \propto \sin^2 \theta \) has an 'almost' toroidal shape—this function equals zero for \( \theta = 0 \). (d) \( |Y_{1,-1}|^2 \) looks the same. (e) \( |Y_{2,1}|^2 \propto |x/r|^2 \) has rotational symmetry about the \( z \)-axis and this polar plot is drawn for \( \phi = \theta \) it looks like (b) but rotated through an angle of \( \pi/2 \). Any linear combination of these is also an eigenfunction of \( \ell^2 \), e.g.

\[
Y_{1,-1} - Y_{1,1} \propto \frac{z}{r} = \sin \theta \cos \phi,
\]

\[
Y_{1,-1} + Y_{1,1} \propto \frac{y}{r} = \sin \theta \sin \phi.
\]

These two real functions have the same shape as \( Y_{1,0} \propto z/r \) but are aligned along the \( z \)- and \( y \)-axes, respectively.\(^{14}\) In chemistry these distributions for \( l = 1 \) are referred to as \( p \)-orbitals. Computer programs can produce plots of such functions from any desired viewing angle (see Blundell 2001, Fig. 3.1) that are helpful in visualising the functions with \( l > 1 \). (For \( l = 0 \) and 1 a cross-section of the functions in a plane that contains the symmetry axis suffices.)

2.1.2 Solution of the radial equation

An equation for \( R(r) \) is obtained by setting eqn 2.4 equal to the constant \( b = l(l + 1) \) and putting in the Coulomb potential \( V(r) = -e^2/4\pi\varepsilon_0 r \). It

\[\text{2.1.2 Solution of the radial equation}\]

An equation for \( R(r) \) is obtained by setting eqn 2.4 equal to the constant \( b = l(l + 1) \) and putting in the Coulomb potential \( V(r) = -e^2/4\pi\varepsilon_0 r \). It
can be cast in a convenient form by the substitution $P(r) = rR(r)$:

$$
-\frac{\hbar^2}{2m_e} \frac{d^2 P}{dr^2} + \left\{ \frac{\hbar^2}{2m_e} \frac{l(l+1)}{r^2} - \frac{e^2}{4\pi \epsilon_0} \frac{1}{r} - E \right\} P = 0.
$$

(2.16)

The term proportional to $l(l+1)/r^2$ is the kinetic energy associated with the angular degrees of freedom; it appears in this radial equation as an effective potential that tends to keep wavefunctions with $l \neq 0$ away from the origin. Dividing through this equation by $E = -|E|$ (a negative quantity since $E \leq 0$ for a bound state) and making the substitution

$$
\rho^2 = \frac{2m_e |E| r^2}{\hbar^2}
$$

(2.17)

reduces the equation to the dimensionless form

$$
\frac{d^2 P}{d\rho^2} + \left\{ -\frac{l(l+1)}{\rho^2} + \frac{\lambda}{\rho} - 1 \right\} P = 0.
$$

(2.18)

The constant that characterises the Coulomb interaction strength is

$$
\lambda = \frac{e^2}{4\pi \epsilon_0} \sqrt{\frac{2m_e}{\hbar^2 |E|}}.
$$

(2.19)

The standard method of solving such differential equations is to look for a solution in the form of a series. The series solutions have a finite number of terms and do not diverge when $\lambda = 2n$, where $n$ is an integer.\(^{15}\)

Thus, from eqn 2.19, these wavefunctions have eigenenergies given by\(^{16}\)

$$
E = -\frac{2m_e}{\hbar^2} \left( \frac{e^2}{4\pi \epsilon_0} \right)^2 \frac{1}{\lambda^2} = -\hbar c R_\infty \frac{1}{n^2}.
$$

(2.20)

This shows that the Schrödinger equation has stationary solutions at energies given by the Bohr formula. The energy does not depend on $l$; this accidental degeneracy of wavefunctions with different $l$ is a special feature of Coulomb potential. In contrast, degeneracy with respect to the magnetic quantum number $m_l$ arises because of the system’s symmetry, i.e., an atom’s properties are independent of its orientation in space, in the absence of external fields.\(^{17}\)

The solution of the Schrödinger equation gives much more information than just the energies; from the wavefunctions we can calculate other atomic properties in ways that were not possible in the Bohr–Sommerfeld theory.

We have not gone through the gory details of the series solution, but we should examine a few examples of radial wavefunctions (see Table 2.2). Although the energy depends only on $n$, the shape of the wavefunctions depends on both $n$ and $l$ and these two quantum numbers are used to label the radial functions $R_{n,l}(r)$. For $n = 1$ there is only the $l = 0$ solution, namely $R_{1,0} \propto e^{-\rho}$. For $n = 2$ the orbital angular momentum quantum number is $l = 0$ or 1, giving

$$
R_{2,0} \propto (1 - \rho) e^{-\rho}, \quad R_{2,1} \propto \rho e^{-\rho}.
$$

15The solution has the general form $P(\rho) = C e^{-\rho v(\rho)}$, where $v(\rho)$ is another function of the radial coordinate, for which there is a polynomial solution (see Woodgate 1980 and Rae 1992).

16Using eqn 1.41.

17This is true for any spherically-symmetric potential $V(r)$.\)
Table 2.2 Radial hydrogenic wavefunctions $R_{n,l}$ in terms of the variable $\rho = Zr/(na_0)$, which gives a scaling that varies with $n$. The Bohr radius $a_0$ is defined in eqn 1.40.

$$
R_{1,0} = \left( \frac{Z}{a_0} \right)^{3/2} 2 e^{-\rho} \\
R_{2,0} = \left( \frac{Z}{2a_0} \right)^{3/2} 2 (1 - \rho) e^{-\rho} \\
R_{2,1} = \left( \frac{Z}{2a_0} \right)^{3/2} \frac{2}{\sqrt{3}} \rho e^{-\rho} \\
R_{3,0} = \left( \frac{Z}{3a_0} \right)^{3/2} 2 \left( 1 - 2\rho + \frac{2}{3} \rho^2 \right) e^{-\rho} \\
R_{3,1} = \left( \frac{Z}{3a_0} \right)^{3/2} \frac{4\sqrt{2}}{3} \rho \left( 1 - \frac{1}{2} \rho \right) e^{-\rho} \\
R_{3,2} = \left( \frac{Z}{3a_0} \right)^{3/2} \frac{2\sqrt{2}}{3\sqrt{5}} \rho^2 e^{-\rho}
$$

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

These show a general feature of hydrogenic wavefunctions, namely that the radial functions for $l = 0$ have a finite value at the origin, i.e. the power series in $\rho$ starts at the zeroth power. Thus electrons with $l = 0$ (called s-electrons) have a finite probability of being found at the position of the nucleus and this has important consequences in atomic physics.

Inserting $|E|$ from eqn 2.20 into eqn 2.17 gives the scaled coordinate

$$
\rho = \frac{Zr}{n a_0} ,
$$

(2.21)

where the atomic number has been incorporated by the replacement $e^2/4\pi \varepsilon_0 \to Ze^2/4\pi \varepsilon_0$ (as in Chapter 1). There are some important properties of the radial wavefunctions that require a general form of the solution and for future reference we state these results. The probability density of electrons with $l = 0$ at the origin is

$$
|\psi_{n,l=0}(0)|^2 = \frac{1}{\pi} \left( \frac{Z}{na_0} \right)^3 .
$$

(2.22)

For electrons with $l \neq 0$ the expectation value of $1/r^3$ is

$$
\left\langle \frac{1}{r^3} \right\rangle = \int_0^\infty \frac{1}{r^3} R_{n,l}^2 (r) \, r^2 \, dr = \frac{1}{l \left( l + \frac{1}{2} \right) (l + 1)} \left( \frac{Z}{na_0} \right)^3 .
$$

(2.23)

These results have been written in a form that is easy to remember; they must both depend on $1/a_0^3$ in order to have the correct dimensions and the dependence on $Z$ follows from the scaling of the Schrödinger
2.2 Transitions

The wavefunction solutions of the Schrödinger equation for particular energies are standing waves and give a distribution of electronic charge \(-e|\psi(r)|^2\) that is constant in time. We shall now consider how transitions between these stationary states occur when the atom interacts with electromagnetic radiation that produces an oscillating electric field\(^{19}\)

\[
E(t) = |E_0| \text{Re} \left( e^{-i\omega t} \hat{e}_{\text{rad}} \right)
\]  
(2.25)

with constant amplitude \(|E_0|\) and polarization vector \(\hat{e}_{\text{rad}}\).\(^{20}\) If \(\omega\) lies close to the atomic resonance frequency then the perturbing electric field puts the atom into a superposition of different states and induces an oscillating electric dipole moment on the atom (see Exercise 2.10). The calculation of the stimulated transition rate requires time-dependent perturbation theory (TDPT), as described in Chapter 7. However, the treatment from first principles is lengthy and we shall anticipate some of the results so that we can see how spectra relate to the underlying structure of the atomic energy levels. This does not require an exact calculation of transition rates, but we only need to determine whether the transition rate has a finite value or whether it is zero (to first order), i.e., whether the transition is allowed and gives a strong spectral line, or is forbidden.

The result of time-dependent perturbation theory is encapsulated in the golden rule (or Fermi’s golden rule);\(^{21}\) this states that the rate of transitions is proportional to the square of the matrix element of the perturbation. The Hamiltonian that describes the time-dependent interaction with the field in eqn 2.25 is \(H' = e \mathbf{r} \cdot \mathbf{E}(t)\), where the electric dipole operator is \(-e \mathbf{r}\).\(^{22}\) This interaction with the radiation stimulates transitions from state 1 to state 2 at a rate\(^{23}\)

\[
\text{Rate} \propto |eE_0|^2 \int \psi_2^* (\mathbf{r} \cdot \hat{e}_{\text{rad}}) \psi_1 \, d^3r \equiv |eE_0|^2 \times |\langle 2 | \mathbf{r} \cdot \hat{e}_{\text{rad}} | 1 \rangle|^2 .
\]

(2.26)

The concise expression in Dirac notation is convenient for later use. This treatment assumes that the amplitude of the electric field is uniform over the atom so that it can be taken outside the integral over the atomic wavefunctions, i.e., that \(E_0\) does not depend on \(r\).\(^{24}\) We write the dipole matrix element as the product

\[
\langle 2 | \mathbf{r} \cdot \hat{e}_{\text{rad}} | 1 \rangle = D_{12} \mathcal{I}_{\text{ang}} .
\]

The radial integral is\(^{25}\)

\[18\] This quantity is related to the quantum mechanical expectation value of the potential energy \(\langle p.e. \rangle\); as in the Bohr model the total energy is \(E = \langle p.e. \rangle / 2\).

\[19\] The interaction of atoms with the oscillating magnetic field in such a wave is considerably weaker; see Appendix C.

\[20\] The unit vector \(\hat{e}_{\text{rad}}\) gives the direction of the oscillating electric field. For example, for the simple case of linear polarization along the \(x\)-axis \(\hat{e}_{\text{rad}} = \hat{e}_x\) and the real part of \(e^{-i\omega t}\) is \(\cos(\omega t)\); therefore \(E(t) = |E_0| \cos(\omega t) \hat{e}_x\).

\[21\] See quantum mechanics texts such as Mandl (1992).

\[22\] This is analogous to the interaction of a classical dipole with an electric field. Atoms do not have a permanent dipole moment, but one is induced by the oscillating electric field. For a more rigorous derivation, see Woodgate (1980) or Loudon (2000).

\[23\] The maximum transition rate occurs when \(\omega\), the frequency of the radiation, matches the transition frequency \(\omega_{12}\), as discussed in Chapter 7. Note, however, that we shall not discuss the so-called 'density of states' in the golden rule since this is not straightforward for monochromatic radiation.

\[24\] In eqn 2.25 the phase of the wave is actually \((\omega t - k \cdot r)\), where \(r\) is the coordinate relative to the atom's centre of mass (taken to be the origin) and \(k\) is the wavevector. We assume that the variation of phase \(k \cdot r\) is small over the atom \((k a_0 \ll 2\pi)\). This is equivalent to \(\lambda \gg a_0\), i.e., the radiation has a wavelength much greater than the size of the atom. This is called the dipole approximation.

\[25\] Note that \(D_{12} = D_{21}\).
\[ D_{12} = \int_{0}^{\infty} R_{n_2,l_2}(r) \, r \, R_{n_1,l_1}(r) \, r^2 \, dr. \] (2.28)

The angular integral is
\[ I_{\text{ang}} = \int_{0}^{2\pi} \int_{0}^{\pi} Y_{l_2,m_2}^* (\theta, \phi) \, \hat{r} \cdot \hat{e}_{\text{rad}} \, Y_{l_1,m_1} (\theta, \phi) \, \sin \theta \, d\theta \, d\phi, \] (2.29)

where \( \hat{r} = r/r \). The radial integral is not normally zero although it can be small for transitions between states whose radial wavefunctions have a small overlap, e.g. when \( n_1 \) is small and \( n_2 \) is large (or the other way round). In contrast, the \( I_{\text{ang}} = 0 \) unless strict criteria are satisfied—these are the selection rules.

### 2.2.1 Selection rules

The selection rules that govern allowed transitions arise from the angular integral in eqn 2.29 which contains the angular dependence of the interaction \( \hat{r} \cdot \hat{e}_{\text{rad}} \) for a given polarization of the radiation. The mathematics requires that we calculate \( I_{\text{ang}} \) for an atom with a well-defined quantization axis (invariably chosen to be the \( z \)-axis) and radiation that has a well-defined polarization and direction of propagation. This corresponds to the physical situation of an atom experiencing the Zeeman effect of an external magnetic field, as described in Section 1.8; that treatment of the electron as a classical oscillator showed that the components of different frequencies within the Zeeman pattern have different polarizations.

We use the same nomenclature of \( \pi \)- and \( \sigma \)-transitions here; transverse observation refers to radiation emitted perpendicular to the magnetic field, and longitudinal observation is along the \( z \)-axis.\(^{26}\)

To calculate \( I_{\text{ang}} \) we write the unit vector \( \hat{r} \) in the direction of the induced dipole as:
\[ \hat{r} = \frac{1}{r} \left( x\hat{e}_x + y\hat{e}_y + z\hat{e}_z \right) \]
\[ = \sin \theta \cos \phi \hat{e}_x + \sin \theta \sin \phi \hat{e}_y + \cos \theta \hat{e}_z. \] (2.30)

Expressing the functions of \( \theta \) and \( \phi \) in terms of spherical harmonic functions as
\[ \sin \theta \cos \phi = \frac{\sqrt{2\pi}}{3} (Y_{1,-1} - Y_{1,1}), \]
\[ \sin \theta \sin \phi = i \frac{\sqrt{2\pi}}{3} (Y_{1,-1} + Y_{1,1}), \]
\[ \cos \theta = \frac{\sqrt{4\pi}}{3} Y_{1,0}, \] (2.31)

leads to
\[ \hat{r} \propto Y_{1,-1} \frac{\hat{e}_x + i\hat{e}_y}{\sqrt{2}} + Y_{1,0} \hat{e}_x + Y_{1,1} \frac{-\hat{e}_x + i\hat{e}_y}{\sqrt{2}}. \] (2.32)

We write the general polarization vector as
\[ \hat{e}_{\text{rad}} = A_{\pi} \frac{\hat{e}_x - i\hat{e}_y}{\sqrt{2}} + A_{\sigma} \hat{e}_x + A_{\sigma} \left( \frac{-\hat{e}_x + i\hat{e}_y}{\sqrt{2}} \right), \] (2.33)

\(^{26}\) If either the atoms have random orientations (e.g. because there is no external field) or the radiation is unpolarized (or both), then an average over all angles must be made at the end of the calculation.
where \( A_\pi \) depends on the component of the electric field along the \( z \)-axis and the component in the \( xy \)-plane is written as a superposition of two circular polarizations with amplitudes \( A_{\pi^+} \) and \( A_{\pi^-} \) (rather than in terms of linear polarization in a Cartesian basis). Similarly, the classical motion of the electron was written in terms of three eigenvectors in Section 1.8: an oscillation along the \( z \)-axis and circular motion in the \( xy \)-plane, both clockwise and anticlockwise.

From the expression for \( \mathbf{\hat{r}} \) in terms of the angular functions \( Y_{l,m}(\theta, \phi) \) with \( l = 1 \) we find that the dipole induced on the atom is proportional to

\[
\mathbf{\hat{r}} \cdot \mathbf{\hat{e}}_{\text{rad}} = A_{\sigma^-} Y_{1,-1} + A_{\sigma^0} Y_{1,0} + A_{\sigma^+} Y_{1,+1}.
\]  

(2.34)

The following sections consider the transitions that arise from these three terms.

### \( \pi \)-transitions

The component of the electric field along the \( z \)-axis \( A_\pi \) induces a dipole moment on the atom proportional to \( \mathbf{\hat{e}}_{\text{rad}} \cdot \mathbf{\hat{e}}_z = \cos \theta \) and the integral over the angular parts of the wavefunctions is

\[
\mathcal{I}^\pi = \int_0^{2\pi} \int_0^\pi Y^*_{l_2,m_2}(\theta, \phi) \cos \theta Y_{l_1,m_1}(\theta, \phi) \sin \theta \, d\theta \, d\phi.
\]  

(2.35)

To determine this integral we exploit the symmetry with respect to rotations about the \( z \)-axis. The system has cylindrical symmetry, so the value of this integral is unchanged by a rotation about the \( z \)-axis through an angle \( \phi_0 \):

\[
\mathcal{I}^\pi_{\text{ang}} = e^{i(m_1 - m_2)\phi_0} \mathcal{I}^\pi_{\text{ang}}.
\]  

(2.36)

This equation is satisfied if either \( \mathcal{I}^\pi_{\text{ang}} = 0 \) or \( m_{l_1} = m_{l_2} \). For this polarization the magnetic quantum number does not change, \( \Delta m_l = 0 \).

### \( \sigma \)-transitions

The component of the oscillating electric field in the \( xy \)-plane excites \( \sigma \)-transitions. Equation 2.34 shows that the circularly-polarized radiation with amplitude \( A_{\pi^\pm} \) excites an oscillating dipole moment on the atom proportional \( Y_{1,1} \propto \sin \theta e^{i\phi} \), for which the angular integral is

\[
\mathcal{I}^{\sigma^\pm} = \int_0^{2\pi} \int_0^\pi Y^*_{l_2,m_2}(\theta, \phi) \sin \theta e^{i\phi} Y_{l_1,m_1}(\theta, \phi) \sin \theta \, d\theta \, d\phi.
\]  

(2.37)

Again, consideration of symmetry with respect to rotation about the \( z \)-axis through an arbitrary angle shows that \( \mathcal{I}^{\sigma^+}_{\text{ang}} = 0 \) unless \( m_{l_1} - m_{l_2} + 1 = 0 \). The interaction of an atom with circularly-polarized radiation of the opposite handedness leads to a similar integral but with \( e^{i\phi} \rightarrow e^{-i\phi} \); this integral \( \mathcal{I}^{\sigma^-}_{\text{ang}} = 0 \) unless \( m_{l_1} - m_{l_2} - 1 = 0 \). Thus the selection rule for the \( \sigma \)-transitions is \( \Delta m_l = \pm 1 \).

We have found the selection rules that govern \( \Delta m_l \) for each of the three possible polarizations of the radiation separately. These apply
when the polarized light interacts with an atom that has a well-defined orientation, e.g. an atom in an external magnetic field. If the light is unpolarized or there is no defined quantisation axis, or both, then \( \Delta m_l = 0, \pm 1 \).

**Example 2.1 Longitudinal observation**

Electromagnetic radiation is a transverse wave with its oscillating electric field perpendicular to the direction of propagation, \( \mathbf{e}_{\text{rad}} \cdot \mathbf{k} = 0 \). Thus radiation with wavevector \( \mathbf{k} = k\mathbf{e}_x \) has \( A_x = 0 \) and \( \pi \)-transitions do not occur.\(^{32}\) Circularly-polarized radiation (propagating along the \( z \)-axis) is a special case for which transitions occur with either \( \Delta m_l = +1 \) or \( \Delta m_l = -1 \), depending on the handedness of the radiation, but not both.

### 2.2.2 Integration with respect to \( \theta \)

In the angular integral the spherical harmonic functions with \( l = 1 \) (from eqn 2.34) are sandwiched between the angular momentum wavefunctions of the initial and final states so that

\[
I_{\text{ang}} \propto \int_0^{2\pi} \int_0^{\pi} Y_{l_2,m_2}^{*} Y_{l_1,m_1} \sin \theta \, d\theta \, d\phi.
\]  

(2.38)

To calculate this angular integral we use the following formula:\(^{33}\)

\[
Y_{l_1,m_1} Y_{l_1,m_1} = A Y_{l_1+1,m_1+m} + B Y_{l_1-1,m_1+m},
\]

(2.39)

where \( A \) and \( B \) are constants whose exact values need not concern us. Thus from the orthogonality of the spherical harmonics\(^{34}\) we find

\[
I_{\text{ang}} \propto A \, \delta_{l_2,l_1+1} \delta_{m_2,m_1+m} + B \, \delta_{l_2,l_1-1} \delta_{m_2,m_1+m}.
\]

The delta functions give the selection rule found previously, namely \( \Delta m_l = m \), where \( m = 0, \pm 1 \) depending on the polarization, and also \( \Delta l = \pm 1 \). In the mathematics the functions with \( l = 1 \) that represent the interaction with the radiation are sandwiched between the orbital angular momentum eigenfunctions of the initial and final states. Thus the rule \( \Delta l = \pm 1 \) can be interpreted as conservation of angular momentum for a photon carrying one unit of angular momentum, \( \hbar \) (Fig. 2.8 illustrates this reasoning for the case of total angular momentum).\(^{35}\) The changes in the magnetic quantum number are also consistent with this picture—the component of the photon's angular momentum along the \( z \)-axis being \( \Delta m_l = 0, \pm 1 \). Conservation of angular momentum does not explain why \( \Delta l \neq 0 \)—this comes about because of parity, as explained below.

### 2.2.3 Parity

Parity is an important symmetry property throughout atomic and molecular physics and its general use will be explained before applying it to...
selection rules. The parity transformation is an inversion through the origin given by $r \rightarrow -r$. This is equivalent to the following transformation of the polar coordinates:

\[
\begin{align*}
\theta &\rightarrow \pi - \theta : \quad \text{a reflection,} \\
\phi &\rightarrow \phi + \pi : \quad \text{a rotation.}
\end{align*}
\]

The reflection produces a mirror image of the original system and parity is also referred to as mirror symmetry. The mirror image of a hydrogen atom has the same energy levels as those in the original atom since the Coulomb potential is the same after reflection. It turns out that all the electric and magnetic interactions 'look the same' after reflection and all atoms have parity symmetry.\(^{36}\) To find the eigenvalues for parity we use the full quantum mechanical notation, with hats to distinguish the operator $\hat{P}$ from its eigenvalue $P$ in the equation

\[\hat{P} \psi = P \psi, \tag{2.40}\]

from which it follows that $\hat{P}^2 \psi = P^2 \psi$. Two successive parity operations correspond to there being no change (the identity operator), i.e. $r \rightarrow -r \rightarrow r$. Thus $P^2 = 1$. Therefore the parity operator has eigenvalues $P = 1$ and $-1$ that correspond to even and odd parity wavefunctions, respectively:

\[\hat{P} \psi = \psi \quad \text{or} \quad \hat{P} \psi = -\psi.\]

Both eigenvalues occur for the spherical harmonic functions,

\[\hat{P} Y_{l,m} = (-1)^l Y_{l,m}. \tag{2.41}\]

The value of the angular integral does not change in a parity transformation\(^{37}\) so

\[I_{\text{ang}} = (-1)^{l_2+l_1+1} I_{\text{ang}}. \tag{2.42}\]

Thus the integral is zero unless the initial and final states have opposite parity (see Exercise 2.12). In particular, electric dipole transitions require an odd change in the orbital angular momentum quantum number ($\Delta l \neq 0$).\(^{38}\)

The treatment above of the parity operator acting on a wavefunction is quite general and even in complex atoms the wavefunctions have a definite parity. The selection rules we have discussed in this section and others are tabulated in Appendix C. If the electric dipole matrix element is zero between two states then other types of transition may occur but at a rate many orders of magnitude slower than allowed transitions.

The allowed transitions between the $n = 1$, 2 and 3 shells of atomic hydrogen are shown in Fig. 2.2, as an example of the selection rules. The $2s$ configuration has no allowed transitions downwards; this makes it metastable, i.e. it has a very long lifetime of about 0.125 s.\(^{39}\)

Finally, a comment on the spectroscopic notation. It can be seen in Fig. 2.2 that the allowed transitions give rise to several series of

---

\(^{36}\)This can be proved formally in quantum mechanics by showing that the Hamiltonians for these interactions commute with the parity operator. The weak interaction in nuclear physics does not have mirror symmetry and violates parity conservation. The extremely small effect of the weak interaction on atoms has been measured in exceedingly careful and precise experiments.

\(^{37}\)See, for example, Mandl (1992).

\(^{38}\)The radial integral is not changed by the parity transformation.

\(^{39}\)This special feature is used in the experiment described in Section 2.3.4.
For hydrogen this is the Lyman series, as marked on Fig. 1.1; however, p-series is a general name.

These names reflect the appearance of the lines in the first experimental observations.

By considering elliptical orbits, rather than just circular ones, Sommerfeld refined Bohr's theory to obtain a relativistic expression for the energy levels in hydrogen that gave very accurate predictions of the fine structure; however, details of that approach are not given here.

2.3 Fine structure

Relativistic effects lead to small splittings of the atomic energy levels called fine structure. We estimated the size of this structure in Section 1.4 by comparing the speed of electrons in classical orbits with the speed of light. In this section we look at how to calculate fine structure by treating relativistic effects as a perturbation to the solutions of the Schrödinger equation. This approach requires the concept that electrons have spin.
2.3.1 Spin of the electron

In addition to the evidence provided by observations of the fine structure itself, that is described in this section, two other experiments showed that the electron has spin angular momentum, not just orbital angular momentum. One of these pieces of experimental evidence for spin was the observation of the so-called anomalous Zeeman effect. For many atoms, e.g. hydrogen and sodium, the splitting of their spectral lines in a magnetic field does not have the pattern predicted by the normal Zeeman effect (that we found classically in Section 1.8). This anomalous Zeeman effect has a straightforward explanation in terms of electron spin (as shown in Section 5.5). The second experiment was the famous Stern–Gerlach experiment that will be described in Section 6.4.1.43

Unlike orbital angular momentum, spin does not have eigenstates that are functions of the angular coordinates. Spin is a more abstract concept and it is convenient to write its eigenstates in Dirac’s ket notation as $|s m_s\rangle$. The full wavefunction for a one-electron atom is the product of the radial, angular and spin wavefunctions: $\Psi = R_{n,l}(r) Y_{l,m_l} (\theta, \phi) |s m_s\rangle$.

Or, using ket notation for all of the angular momentum, not just the spin,

$$\Psi = R_{n,l}(r) |l m_l s m_s\rangle.$$  \hspace{1cm} (2.43)

These atomic wavefunctions provide a basis in which to calculate the effect of perturbations on the atom. However, some problems do not require the full machinery of (degenerate) perturbation theory and for the time being we shall treat the orbital and spin angular momenta by analogy with classical vectors. To a large extent this vector model is intuitively obvious and we start to use it without formal derivations. But note the following points. An often-used shorthand for the spin eigenfunctions is spin-up:

$$|s = \frac{1}{2}, m_s = \frac{1}{2}\rangle \equiv \uparrow ,$$  \hspace{1cm} (2.44)

and similarly $|\downarrow\rangle$ for the $m_s = -\frac{1}{2}$ state (spin-down). However, in quantum mechanics the angular momentum cannot be completely aligned ‘up’ or ‘down’ with respect to the $z$-axis, otherwise the $x$- and $y$-components would be zero and we would know all three components simultaneously.44 The vector model mimics this feature with classical vectors drawn with length $|s| = \sqrt{s(s+1)} = \sqrt{3}/2$. (Only the expectation value of the square of the angular momentum has meaning in quantum mechanics.) The spin-up and spin-down states are as illustrated in Fig. 2.3 with components along the $z$-axis of $\pm \frac{1}{2}$. We can think of the vector as rotating around the $z$-axis, or just having an undefined direction in the $xy$-plane corresponding to a lack of knowledge of the $x$- and $y$-components (see also Grant and Phillips 2001).

The name ‘spin’ invokes an analogy with a classical system spinning on its axis, e.g. a sphere rotating about an axis through its centre of mass, but this mental picture has to be treated with caution; spin cannot be equal to the sum of the orbital angular momenta of the constituents since that will always be an integer multiple of $\hbar$. In any case, the electron is

43 The fine structure, anomalous Zeeman effect and Stern–Gerlach experiment all involve the interaction of the electron’s magnetic moment with a magnetic field—the internal field of the atom in the case of fine structure. Stern and Gerlach detected the magnetic interaction by its influence on the atom’s motion, whereas the Zeeman effect and fine structure are observed by spectroscopy.

44 This is not possible since the operators for the $x$, $y$ and $z$-components of angular momentum do not commute (save in a few special cases; we can know that $s_x = s_y = s_z = 0$ if $s = 0$).
Fig. 2.3 The representation of (a) spin-up and (b) spin-down states as vectors precessing around the \( z \)-axis.

a structureless elementary particle with no measurable size. So we are left with the experimental fact that the electron has an intrinsic spin angular momentum of \( \hbar/2 \) and these half-integer values are perfectly acceptable within the general theory of angular momentum in quantum mechanics.

### 2.3.2 The spin–orbit interaction

The Schrödinger equation is non-relativistic, as can readily be seen by looking at the kinetic-energy operator that is equivalent to the non-relativistic expression \( p^2/2m_e \). Some of the relativistic effects can be taken into account as follows. An electron moving through an electric field \( \mathbf{E} \) experiences an effective magnetic field \( \mathbf{B} \) given by

\[
\mathbf{B} = -\frac{1}{c^2} \mathbf{v} \times \mathbf{E}.
\]  

(2.45)

This is a consequence of the way an electric field behaves under a Lorentz transformation from a stationary to a moving frame in special relativity. Although a derivation of this equation is not given here, it is certainly plausible since special relativity and electromagnetism are intimately linked through the speed of light \( c = 1/\sqrt{\varepsilon_0\mu_0} \). This equation for the speed of electromagnetic waves in a vacuum comes from Maxwell’s equations; \( \varepsilon_0 \) being associated with the electric field and \( \mu_0 \) with the magnetic field. Rearrangement to give \( \mu_0 = 1/(\varepsilon_0 c^2) \) suggests that magnetic fields arise from electrodynamics and relativity.\(^{45}\)

We now manipulate eqn 2.45 into a convenient form, by substituting for the electric field in terms of the gradient of the potential energy \( V \) and unit vector in the radial direction:

\[
\mathbf{E} = \frac{1}{e} \frac{\partial V}{\partial r} \mathbf{r}.
\]  

(2.46)

The factor of \( e \) comes in because the electron’s potential energy \( V \) equals its charge \( -e \) times the electrostatic potential. From eqn 2.45 we have

\[
\mathbf{B} = \frac{1}{m_e c^2} \left( \frac{1}{er} \frac{\partial V}{\partial r} \right) \mathbf{r} \times m_e \mathbf{v} = \frac{\hbar}{m_e c^2} \left( \frac{1}{er} \frac{\partial V}{\partial r} \right) \mathbf{1},
\]  

(2.47)

where the orbital angular momentum is \( \hbar \mathbf{r} \times m_e \mathbf{v} \). The electron has an intrinsic magnetic moment \( \mathbf{\mu} = -g_s \mu_B \mathbf{s} \), where the spin has a magnitude of \( |s| = s = 1/2 \) (in units of \( \hbar \)) and \( g_s \simeq 2 \), so the moment...
has a magnitude close to one Bohr magneton \( (\mu_B = e\hbar/2m_e) \). The interaction of the electron’s magnetic moment with the orbital field gives the Hamiltonian

\[
H = - \mu \cdot B = g_s \mu_B s \cdot \frac{\hbar}{m_e c^2} \left( \frac{1}{r} \frac{\partial V}{\partial r} \right) l.
\] (2.48)

However, this expression gives energy splittings about twice as large as observed. The discrepancy comes from the Thomas precession—a relativistic effect that arises because we are calculating the magnetic field in a frame of reference that is not stationary but rotates as the electron moves about the nucleus. The effect is taken into account by replacing \( g_s \) with \( g_s - 1 \approx 1 \). Finally, we find the spin–orbit interaction, including the Thomas precession factor, is

\[
H_{s-o} = (g_s - 1) \frac{\hbar^2}{2m_e c^2} \left( \frac{1}{r} \frac{\partial V}{\partial r} \right) s \cdot l.
\] (2.49)

For the Coulomb potential in hydrogen we have

\[
\frac{1}{r} \frac{\partial V}{\partial r} = \frac{e^2/4\pi\epsilon_0}{r^3}.
\] (2.50)

The expectation value of this Hamiltonian gives an energy change of

\[
E_{s-o} = \frac{\hbar^2}{2m_e c^2} \frac{e^2}{4\pi\epsilon_0} \left( \frac{1}{r^3} \right) \langle s \cdot l \rangle.
\] (2.51)

The separation into a product of radial and angular expectation values follows from the separability of the wavefunction. The integral \( \langle 1/r^3 \rangle \) is given in eqn 2.23. However, we have not yet discussed how to deal with interactions that have the form of dot products of two angular momenta; let us start by defining the total angular momentum of the atom as the sum of its orbital and spin angular momenta,

\[
j = l + s.
\] (2.52)

This is a conserved quantity for a system without any external torque acting on it, e.g., an atom in a field-free region of space. This is true both in classical and quantum mechanics, but we concentrate on the classical explanation in this section. The spin–orbit interaction between \( l \) and \( s \) causes these vectors to change direction, and because their sum is constrained to be equal to \( j \) they move around as shown in Fig. 2.4.49 Squaring and rearranging eqn 2.52, we find that \( 2s \cdot l = j^2 - l^2 - s^2 \). Hence we can find the expectation value in terms of the known values for \( \langle j^2 \rangle, \langle l^2 \rangle \) and \( \langle s^2 \rangle \) as

\[
\langle s \cdot l \rangle = \frac{1}{2} \left\{ j(j+1) - l(l+1) - s(s+1) \right\}.
\] (2.53)

Thus the spin–orbit interaction produces a shift in energy of

\[
E_{s-o} = -\frac{\beta}{2} \left\{ j(j+1) - l(l+1) - s(s+1) \right\},
\] (2.54)

46 This is almost equivalent to using \( g_s/2 \approx 1 \), but \( g_s - 1 \) is more accurate at the level of precision where the small deviation of \( g_s \) from 2 is important (Haar and Curtis 1987). For further discussion of Thomas precession see Cowan (1981), Elsberg and Resnick (1985) and Munroe (2001).

47 We have derived this classically, e.g., by using \( \mathbf{H} = \mathbf{r} \times \mathbf{m}_e \mathbf{v} \). However, the same expression can be obtained from the fully relativistic Dirac equation for an electron in a Coulomb potential by making a low-velocity approximation, see Sakurai (1967). That quantum mechanical approach justifies treating \( l \) and \( s \) as operators.

48 Using the approximation \( g_s - 1 \approx 1 \).

Fig. 2.4 The orbital and spin angular momenta add to give a total atomic angular momentum of \( j \).

49 In this precession about \( j \) the magnitudes of \( l \) and \( s \) remain constant. The magnetic moment (proportional to \( s \)) is not altered in an interaction with a magnetic field, and because of the symmetrical form of the interaction in eqn 2.49, we do not expect \( l \) to behave any differently. See also Blundell (2001) and Section 5.1.
where the spin–orbit constant \( \beta \) is (from eqns 2.51 and 2.23)

\[
\beta = \frac{\hbar^2}{2m_e^2 c^2} \frac{e^2}{4\pi \varepsilon_0} \frac{1}{(n \alpha_o)^3} \frac{1}{l (l + \frac{1}{2}) (l + 1)} .
\]

(2.55)

A single electron has \( s = \frac{1}{2} \), so, for each \( l \), its total angular momentum quantum number \( j \) has two possible values:

\[
j = l + \frac{1}{2} \quad \text{or} \quad l - \frac{1}{2}.
\]

From eqn 2.54 we find that the energy interval between these levels, \( \Delta E_{s-o} = E_{j=l+\frac{1}{2}} - E_{j=l-\frac{1}{2}} \), is

\[
\Delta E_{s-o} = \beta \left( l + \frac{1}{2} \right) = \frac{\alpha^2 \hbar c R_{\infty}}{n^3 l (l + 1)} .
\]

(2.56)

Or, expressed in terms of the gross energy \( E(n) \) in eqn 1.10,

\[
\Delta E_{s-o} = \frac{\alpha^2}{n l (l + 1)} E(n) .
\]

(2.57)

This agrees with the qualitative discussion in Section 1.4, where we showed that relativistic effects cause energy changes of order \( \alpha^2 \) times the gross structure. The more complete expression above shows that the energy intervals between levels decrease as \( n \) and \( l \) increase. The largest interval in hydrogen occurs for \( n = 2 \) and \( l = 1 \); for this configuration the spin–orbit interaction leads to levels with \( j = 1/2 \) and \( j = 3/2 \).

The full designation of these levels is \( 2p^2P_{1/2} \) and \( 2p^2P_{3/2} \), in the notation that will be introduced for the \( LS \)-coupling scheme. But some of the quantum numbers (defined in Chapter 5) are superfluous for atoms with a single valence electron and a convenient short form is to denote these two levels by \( 2P_{1/2} \) and \( 2P_{3/2} \); these correspond to \( np_j \), where \( P \) represents the (total) orbital angular momentum for this case. (The capital letters are consistent with later usage.) Similarly, we may write \( 2S_{1/2} \) for the \( 2s^2S_{1/2} \) level; \( 3D_{3/2} \) and \( 3D_{5/2} \) for the \( j = 3/2 \) and \( 5/2 \) levels, respectively, that arise from the \( 3d \) configuration. But the full notation must be used whenever ambiguity might arise.

### 2.3.3 The fine structure of hydrogen

As an example of fine structure, we look in detail at the levels that arise from the \( n = 2 \) and \( n = 3 \) shells of hydrogen. Equation 2.54 predicts that, for the \( 2p \) configuration, the fine-structure levels have energies of

\[
E_{s-o} (2P_{1/2}) = -\beta_{2p} ,
E_{s-o} (2P_{3/2}) = \frac{1}{2} \beta_{2p} ,
\]

as shown in Fig. 2.5(a). For the \( 3d \) configuration

\[
E_{s-o} (3D_{3/2}) = -\frac{3}{2} \beta_{3d} ,
E_{s-o} (3D_{5/2}) = \beta_{3d} ,
\]
as shown in Fig. 2.5(b). For both configurations, it is easy to see that the spin–orbit interaction does not shift the mean energy

$$\bar{E} = (2j + 1) E_j(n, l) + (2j' + 1) E_{j'}(n, l),$$  

where $j' = l - 1/2$ and $j = l + 1/2$ for the two levels. This calculation of the 'centre of gravity' for all the states takes into account the degeneracy of each level.

The spin–orbit interaction does not affect the $2S_1/2$ or $3S_{1/2}$ so we might expect these levels to lie close to the centre of gravity of the configurations with $l > 0$. This is not the case. Fig. 2.6 shows the energies of the levels for the $n = 3$ shell given by a fully relativistic calculation. We can see that there are other effects of similar magnitude to the spin–orbit interaction that affect these levels in hydrogen. Quite remarkably, these additional relativistic effects shift the levels by just the right amount to make $n P_{1/2}$ levels degenerate with the $n S_{1/2}$ levels, and $n P_{3/2}$ degenerate with $n D_{3/2}$. This structure does not occur by chance, but points to a deeper underlying cause. The full explanation of these observations requires relativistic quantum mechanics and the technical details of such calculations lie beyond the scope of this book.\footnote{See graduate-level quantum mechanics texts, e.g. Sakurai (1967) and Series (1988).} We shall simply quote the solution of the Dirac equation for an electron in a Coulomb potential; this gives a formula for the energy $E_{\text{Dirac}}(n, j)$ that depends only on $n$ and $j$, i.e. it gives the same energy for levels of the same $n$ and $j$ but different $l$, as in the cases above. In a comparison of the

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig2.5}
\caption{The fine structure of hydrogen. The fine structure of (a) the $2p$ and (b) the $3d$ configurations are drawn on different scales: $\beta_{2p}$ is considerably greater than $\beta_{3d}$. All $p$- and $d$-configurations look similar apart from an overall scaling factor.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig2.6}
\caption{The theoretical positions of the energy levels of hydrogen calculated by the fully relativistic theory of Dirac depend on $n$ and $j$ only (not $l$), as shown in this figure for the $n = 3$ shell. In addition to the spin–orbit interaction, the effects that determine the energies of these levels are: the relativistic mass correction and, for $s$-electrons only, the Darwin term (that accounts for relativistic effects that occur at small $r$, where the electron's momentum becomes comparable to $m_e c$).}
\end{figure}
exact relativistic solution of the Dirac equation and the non-relativistic energy levels, three relativistic effects can be distinguished.

(a) There is a straightforward relativistic shift of the energy (or equivalently mass), related to the binomial expansion of \( \gamma = (1-v^2/c^2)^{-1/2} \), in eqn 1.16. The term of order \( v^2/c^2 \) gives the non-relativistic kinetic energy \( p^2/2m_e \). The next term in the expansion is proportional to \( v^4/c^4 \) and gives an energy shift of order \( v^2/c^2 \) times the gross structure — this is the effect that we estimated in Section 1.4.

(b) For electrons with \( l \neq 0 \), the comparison of the Dirac and Schrödinger equations shows that there is a spin–orbit interaction of the form given above, with the Thomas precession factor naturally included.\(^{53}\)

(c) For electrons with \( l = 0 \) there is a Darwin term proportional to \( |\psi(r = 0)|^2 \) that has no classical analogue (see Woodgate (1980) for further details).

That these different contributions conspire together to perturb the wavefunctions such that levels of the same \( n \) and \( j \) are degenerate seems improbable from a non-relativistic point of view. It is worth reiterating the statement above that this structure arises from the relativistic Dirac equation; making an approximation for small \( v^2/c^2 \) shows that these three corrections, and no others, need to be applied to the (non-relativistic) energies found from the Schrödinger equation.

2.3.4 The Lamb shift

Figure 2.7 shows the actual energy levels of the \( n = 2 \) and \( n = 3 \) shells. According to relativistic quantum theory the \( 2S_{1/2} \) level should be exactly degenerate with \( 2P_{1/2} \) because they both have \( n = 2 \) and \( j = 1/2 \), but in reality there is an energy interval between them, \( E(2S_{1/2}) - E(2P_{1/2}) \approx 1 \text{ GHz} \). The shift of the \( 2S_{1/2} \) level to a higher energy (lower binding energy) than the \( E_{\text{Dirac}}(n = 2, j = 1/2) \) is about one-tenth of the interval between the two fine-structure levels, \( E(2P_{3/2}) - E(2P_{1/2}) \approx 11 \text{ GHz} \). Although small, this discrepancy in hydrogen was of great historical importance in physics. For this simple one-electron atom the predictions of the Dirac equation are very precise and that theory cannot account for Lamb and Retherford’s experimental measurement that the \( 2S_{1/2} \) level is indeed higher than the \( 2P_{1/2} \) level.\(^{54}\)

The explanation of this Lamb shift goes beyond relativistic quantum mechanics and requires quantum electrodynamics (QED)—the quantum field theory that describes electromagnetic interactions. Indeed, the observation of the Lamb shift experiment was a stimulus for the development of this theory.\(^{55}\) An intriguing feature of QED is so-called vacuum fluctuations—regions of free space are not regarded as being completely empty but are permeated by fluctuating electromagnetic fields.\(^{56}\) The QED effects lead to a significant energy shift for electrons with \( l = 0 \) and hence break the degeneracy of \( 2S_{1/2} \) and \( 2P_{1/2} \).\(^{57}\) The largest QED shift occurs for the \( 1S_{1/2} \) ground level of hydrogen but there is no other level nearby and so a determination of its energy requires a precise mea-
2.3 Fine structure

Fig. 2.7 The fine structure of the \( n = 2 \) and \( n = 3 \) shells of hydrogen and the allowed transitions between the levels. According to the Dirac equation, the \( 2S_{1/2} \) and \( 2P_{1/2} \) levels should be degenerate, but they are not. The measured positions show that the \( 2S_{1/2} \) level is shifted upwards relative to the position \( E_{\text{Dirac}}(n = 2, j = 1/2) \) and is therefore not degenerate with the \( 2P_{1/2} \) level. Such a shift occurs for all the s-electrons (but the size of the energy shift decreases with increasing \( n \)). The explanation of this shift takes us beyond relativistic quantum mechanics into the realm of quantum electrodynamics (QED)—the quantum field theory that describes electromagnetic interactions.

Fig. 2.8 The conservation of total angular momentum in electric dipole transitions that gives the selection rule in eqn 2.59 can be represented as vector addition. The photon has one unit of angular momentum, and so to go from level \( j_1 \) to \( j_2 \) the vectors must form a triangle, as shown for the case of (a) \( j_1 = 1/2 \) to \( j_2 = 1/2 \), (b) \( j_1 = 1/2 \) to \( j_2 = 3/2 \) and (c) \( j_1 = 3/2 \) to \( j_2 = 3/2 \).

measurement of a large frequency. Nowadays this can be achieved by laser spectroscopy (Chapter 8) but the near degeneracy of the two \( j = 1/2 \) levels with \( n = 2 \) was crucial in Lamb’s experiment.\(^{58}\) Another important feature in that experiment was the metastability of the \( 2S_{1/2} \) level, whose lifetime was given in Section 2.2.3. That level decays \( \sim 10^8 \) times more slowly than that of \( 2P_{1/2} \). In an atomic beam of hydrogen (at room temperature) the atoms have typical velocities of about 3000 m s\(^{-1}\) and atoms excited into the \( 2p \) configuration travel an average distance of only \( 5 \times 10^{-6} \) m before decaying with the emission of Lyman-\( \alpha \) radiation. In contrast, metastable atoms travel the full length of the apparatus (\( \simeq 1 \) m) and are de-excited when they collide with a detector (or the wall of the vacuum chamber). Hydrogen, and hydrogenic systems, are still used for experimental tests of fundamental theory because their simplicity allows very precise predictions.

2.3.5 Transitions between fine-structure levels

Transitions in hydrogen between the fine-structure levels with principal quantum numbers \( n = 2 \) and 3 give the components of the Balmer-\( \alpha \) line shown in Fig. 2.7; in order of increasing energy, the seven allowed

\(^{58}\)Higher shells have smaller shifts between the \( j = 1/2 \) levels.
transitions between the levels with different $j$ are as follows:

\[ 2 \text{P}_{3/2} - 3 \text{S}_{1/2}, \]
\[ 2 \text{P}_{3/2} - 3 \text{D}_{3/2}, \]
\[ 2 \text{P}_{3/2} - 3 \text{D}_{5/2}, \]
\[ 2 \text{S}_{1/2} - 3 \text{P}_{1/2}, \]
\[ 2 \text{P}_{1/2} - 3 \text{S}_{1/2}, \]
\[ 2 \text{S}_{1/2} - 3 \text{P}_{3/2}, \]
\[ 2 \text{P}_{1/2} - 3 \text{D}_{3/2}. \]

These obey the selection rule $\Delta l = \pm 1$ but an additional rule prevents a transition between $3 \text{D}_{5/2}$ and $2 \text{P}_{1/2}$, namely that the change of the total angular momentum quantum number in an electric dipole transition obeys

\[ \Delta j = 0, \pm 1. \quad (2.59) \]

This selection rule may be explained by angular momentum conservation (as mentioned in Section 2.2.2). This rule can be expressed in terms of vector addition, as shown in Fig. 2.8; the conservation condition is equivalent to being able to form a triangle from the three vectors representing $j$ of the initial state, the final state, and a unit vector for the (one unit of) angular momentum carried by the photon. Hence, this selection rule is sometimes referred to as the triangle rule. The projection of $j$ along the $z$-axis can change by $\Delta m_j = 0, \pm 1$. (Appendix C gives a summary of all selection rules.)

**Further reading**

Much of the material covered in this chapter can be found in the introductory quantum mechanics and atomic physics texts listed in the References. For particular topics the following are useful: Segrè (1980) gives an overview of the historical development, and Series (1988) reviews the work on hydrogen, including the important Lamb shift experiment.

---

### Exercises

(2.1) **Angular-momentum eigenfunctions**

(a) Verify that all the eigenfunctions with $l = 1$ are orthogonal to $Y_{0,0}$.

(b) Verify that all the eigenfunctions with $l = 1$ are orthogonal to those with $l = 2$.

(2.2) **Angular-momentum eigenfunctions**

(a) Find the eigenfunction with orbital angular momentum quantum number $l$ and magnetic quantum number $m = l - 1$.

(b) Verify that $Y_{l,l-1}$ is orthogonal to $Y_{l-1,l-1}$. 
(2.3) Radial wavefunctions
Verify eqn 2.23 for \( n = 2, \ l = 1 \) by calculating the radial integral (for \( Z = 1 \)).

(2.4) Hydrogen
For a hydrogen atom the normalised wavefunction of an electron in the 1s state, assuming a point nucleus, is

\[
\psi(r) = \left( \frac{1}{\pi a_0^2} \right)^{1/2} e^{-r/a_0},
\]

where \( a_0 \) is the Bohr radius. Find an approximate expression for the probability of finding the electron in a small sphere of radius \( r_B \ll a_0 \) centred on the proton. What is the electronic charge density in this region?

(2.5) Hydrogen
The Balmer-\( \alpha \) spectral line is observed from a (weak) discharge in a lamp that contains a mixture of hydrogen and deuterium at room temperature. Comment on the feasibility of carrying out an experiment using a Fabry–Perot étalon to resolve (a) the isotope shift, (b) the fine structure and (c) the Lamb shift.

(2.6) Transitions
Estimate the lifetime of the excited state in a two-level atom when the transition wavelength is (a) 100 nm and (b) 1000 nm. In what spectral regions do these wavelengths lie?

(2.7) Selection rules
By explicit calculation of integrals over \( \theta \), for the case of \( \pi \)-polarization only, verify that \( p \rightarrow d \) transitions are allowed, but not \( s \rightarrow d \).

(2.8) Spin–orbit interaction
The spin–orbit interaction splits a single-electron configuration into two levels with total angular momentum quantum numbers \( j = l + 1/2 \) and \( j' = l - 1/2 \). Show that this interaction does not shift the mean energy (centre of gravity) of all the states given by \( (2j + 1) E_j + (2j' + 1) E_{j'} \).

(2.9) Selection rule for the magnetic quantum number
Show that the angular integrals for \( \sigma \)-transitions contain the factor

\[
\int_0^{2\pi} e^{i(m_1 - m_2)\phi} \, d\phi.
\]

Hence derive the selection rule \( \Delta m_i = \pm 1 \) for this polarization. Similarly, derive the selection rule for the \( \pi \)-transitions.

(2.10) Transitions
An atom in a superposition of two states has the wavefunction

\[
\Psi(t) = A\psi_1(r) e^{-iE_1 t/\hbar} + B\psi_2(r) e^{-iE_2 t/\hbar}.
\]

The distribution of electronic charge is given by

\[
-e |\Psi(t)|^2 = -e\left[ |A\psi_1|^2 + |B\psi_2|^2 + 2A^*B\psi_1^*\psi_2 |\cos(\omega_1 t - \phi)\right].
\]

Part of this oscillates at the (angular) frequency of the transition \( \omega_{12} = \omega_2 - \omega_1 = (E_2 - E_1)/\hbar \).

(a) A hydrogen atom is in a superposition of the 1s ground state, \( \psi_1 = R_{1,0}(r) Y_{0,0}(\theta, \phi) \), and the \( m_i = 0 \) state of the 2p configuration, \( \psi_2 = R_{2,1}(r) Y_{1,0}(\theta, \phi) \); \( A \approx 0.995 \) and \( B = 0.1 \) (so the term containing \( B^2 \) can be ignored). Sketch the form of the charge distribution for one cycle of oscillation.

(b) The atom in a superposition state may have an oscillating electric dipole moment

\[
-eD(t) = -e \left( \Psi^*(t) r \Psi(t) \right).
\]

What are the conditions on \( \psi_1 \) and \( \psi_2 \) for which \( D(t) \neq 0 \).

(c) Show that an atom in a superposition of the same states as in (a) has a dipole moment of

\[
-eD(t) = -e \left| 2A^*B \right| I_{\text{ang}} \times \left\{ \int R_{2,1}(r) R_{1,0}(r) \sin^2 \theta \cos(\omega_{12} t) \, d^2 r \right\} \cos(\omega_{12} t t_{\text{ang}})
\]

where \( I_{\text{ang}} \) is an integral of the angular parts of the wavefunction with respect to \( \theta \) and \( \phi \). Calculate the amplitude of this dipole, in units of \( e a_0 \), for \( A = B = 1/\sqrt{2} \).

(d) A hydrogen atom is in a superposition of the 1s ground state and the \( m_i = 1 \) state of the 2p configuration, \( \psi_1 = R_{2,1}(r) Y_{1,1}(\theta, \phi) \). Sketch the form of the charge distribution at various points in its cycle of oscillation.

(e) Comment on the relationship between the time dependence of the charge distributions sketched in this exercise and the motion of the electron in the classical model of the Zeeman effect (Section 1.8).

(2.11) Angular eigenfunctions
We shall find the angular momentum eigenfunctions using ladder operators, by assuming that for some value of \( l \) there is a maximum value of the magnetic quantum number \( m_{\text{max}} \). For this case \( Y_{l,m_{\text{max}}} \propto \Theta(\theta)e^{im_{\text{max}}\phi} \) and the function \( \Theta(\theta) \) can be found from

\[
l_{+}\Theta(\theta) \exp(i m_{\text{max}} \phi) = 0.
\]
(a) Show that $\Theta(\theta)$ satisfies the equation
\[
\frac{1}{\Theta(\theta)} \frac{\partial \Theta(\theta)}{\partial \theta} = m_{\text{max}} \frac{\cos \theta}{\sin \theta}.
\]

(b) Find the solution of the equation for $\Theta(\theta)$. (Both sides have the form $f'(\theta)/f(\theta)$ whose integral is $\ln\{f(\theta)\}$.) By substituting this solution into eqn 2.5 to show that $b = m_{\text{max}}(m_{\text{max}} + 1)$, or otherwise, obtain eqn 2.10.

(2.12) Parity and selection rules
Show that eqn 2.42 implies that $l_z - l_z$ is odd.

Hence, or otherwise, prove that $I_{\text{ang}}$ is zero unless the initial and final states have opposite parity.

(2.13) Selection rules in hydrogen
Hydrogen atoms are excited (by a pulse of laser light that drives a multi-photon process) to a specific configuration and the subsequent spontaneous emission is resolved using a spectrograph. Infra-red and visible spectral lines are detected only at the wavelengths 4.05 $\mu$m, 1.87 $\mu$m and 0.656 $\mu$m. Explain these observations and give the values of $n$ and $l$ for the configurations involved in these transitions.
Helium

Helium has only two electrons but this simplicity is deceptive. To treat systems with two particles requires new concepts that also apply to multi-particle systems in many branches of physics, and it is very worthwhile to study them carefully using the example of helium. There is truth in the saying that atomic physicists count ‘one, two, many’ and a detailed understanding of the two-electron system is sufficient for much of the atomic structure in this book.\(^1\)

3.1 The ground state of helium

Two electrons in the Coulomb potential of a charge Ze, e.g. the nucleus of an atom, obey a Schrödinger equation of the form

\[
\left\{ \frac{-\hbar^2}{2m} \nabla_1^2 + \frac{-\hbar^2}{2m} \nabla_2^2 + \frac{e^2}{4\pi\varepsilon_0} \left( \frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}} \right) \right\} \psi = E\psi.
\]

(3.1)

Here \(r_{12} = |r_1 - r_2|\) is the distance between electron 1 and electron 2 and the electrostatic repulsion of electrons is proportional to \(1/r_{12}\). Neglecting this mutual repulsion for the time being, we can write the equation as

\[
(H_1 + H_2) \psi = E^{(0)} \psi,
\]

(3.2)

where

\[
H_1 \equiv \frac{-\hbar^2}{2m} \nabla_1^2 - \frac{Ze^2}{4\pi\varepsilon_0 r_1}
\]

(3.3)

and \(H_2\) is a similar expression for electron 2. Writing the atomic wavefunction as a product of the wavefunctions for each electron, \(\psi = \psi(1) \psi(2)\), allows us to separate eqn 3.2 into two single-electron Schrödinger equations:

\[
H_1 \psi(1) = E_1 \psi(1)
\]

(3.4)

and a similar equation for \(\psi(2)\) with energy \(E_2\). The solutions of these one-electron equations are hydrogenic wavefunctions with energies given by the Rydberg formula. Helium has \(Z = 2\) and in its ground state both electrons have energy \(E_1 = E_2 = -4\hbar c R_\infty = -54.4\,\text{eV}\). Thus the total energy of the atom (neglecting repulsion) is

\[
E^{(0)} = E_1 + E_2 = -109\,\text{eV}.
\]

(3.5)
Now we need to calculate the perturbation produced by the electron-electron repulsion. The system has the spatial wavefunction

\[
\psi_{1s} = R_{1s}^2(r_1) R_{1s}^2(r_2) \times \frac{1}{4\pi},
\]

where radial wavefunctions are defined in Table 2.2. The expectation value of the repulsion is (see Section 3.3)

\[
\frac{e^2}{4\pi\varepsilon_0} \int_0^\infty \int_0^\infty \psi_{1s}^2 \frac{1}{r_{12}} \psi_{1s}^2 r_1^2 dr_1 r_2^2 dr_2 = 34 \text{ eV}.
\]

Adding this to the (zeroth-order) estimate \(E^{(0)}\) gives an energy of \(E(1s^2) = -109 + 34 = -75 \text{ eV}\). It takes an energy of 75 eV to remove both electrons from a helium atom leaving a bare helium nucleus \(\text{He}^+\); the second ionization energy. To go from \(\text{He}^+\) to \(\text{He}^{++}\) takes 54.4 eV, so this estimate suggests that the first ionization energy (required to remove one electron from He to create \(\text{He}^+\)) is \(\text{IE(He)} \approx 75 - 54 \approx 21 \text{ eV}\). But the expectation value in eqn 3.7 is not small compared to the binding energy and therefore the perturbation has a significant effect on the wavefunctions. The necessary adjustment of the wavefunctions can be accounted for by the variational method. This method gives a value close to the measured ionization energy 24.6 eV. Helium has the highest first ionization energy of all elements because of its closed \(n = 1\) shell. For a plot of the ionization energies of the elements see Grant and Phillips (2001, Chapter 11, Fig. 18).

According to the Pauli exclusion principle, two electrons cannot have the same set of quantum numbers. Therefore there must be some additional quantum number associated with the two 1s-electrons in the ground state of helium—this is their spin (introduced in Section 2.3.1). The observed filling-up of the atomic (sub-)shells in the periodic table implies that two spin states are associated with each set of spatial quantum numbers \(n, \ell, m_l\). However, electrostatic energies do not depend on spin and we can find the spatial wavefunctions separately from the problem of finding the spin eigenfunctions.

### 3.2 Excited states of helium

To find the energy of the excited states we use the same procedure as for the ground state—at first we neglect the mutual repulsion term and separate eqn 3.1 into two one-electron equations that have solutions

\[
\psi_{1s}(1) = R_{1s}(r_1) \times \frac{1}{\sqrt{4\pi}},
\]

\[
\psi_{nl}(2) = R_{nl}(r_2) Y_{l,m} (\theta_2, \phi_2)
\]

for the configuration 1s\(n\ell \). The spatial part of the atomic wavefunction is the product

\[
\psi_{\text{space}} = \psi_{1s}(1) \psi_{nl}(2).
\]
Another wavefunction has the same energy, namely
\[ \psi_{\text{space}} = u_{1s}(2)u_{nl}(1). \]  
(3.9)

These two states are related by a permutation of the labels on the electrons, $1 \leftrightarrow 2$; the energy cannot depend on the labeling of identical particles so there is exchange degeneracy. To consider the effect of the repulsive term on this pair of wavefunctions with the same energy (degenerate states) we need degenerate perturbation theory. There are two approaches. The look-before-you-leap approach is first to form eigenstates of the perturbation from linear combinations of the initial states.\footnote{This is guided by looking for eigenstates of symmetry operators that commute with the Hamiltonian for the interaction, as in Section 4.5.} In this new basis the determination of the eigenenergies of the states is simple. It is instructive, however, simply to press ahead and go through the algebra once.\footnote{In the light of this experience one can take the short cut in future.}

We rewrite the Schrödinger equation (eqn 3.1) as
\[ (H_0 + H') \psi = E \psi, \]  
(3.10)

where $H_0 = H_1 + H_2$, and we consider the mutual repulsion of the electrons $H' = e^2/4\pi \varepsilon_0 r_{12}$ as a perturbation. We also rewrite eqn 3.2 as
\[ H_0 \psi = E^{(0)} \psi, \]  
(3.11)

where $E^{(0)} = E_1 + E_2$ is the unperturbed energy. Subtraction of eqn 3.11 from eqn 3.10 gives the energy change produced by the perturbation, $\Delta E = E - E^{(0)}$, as
\[ H' \psi = \Delta E \psi. \]  
(3.12)

A general expression for the wavefunction with energy $E^{(0)}$ is a linear combination of expressions 3.8 and 3.9, with arbitrary constants $a$ and $b$,
\[ \psi = a \, u_{1s}(1)u_{nl}(2) + b \, u_{1s}(2)u_{nl}(1). \]  
(3.13)

Substitution into eqn 3.12, multiplication by either $u_{1s}(1)^* u_{nl}(2)^*$ or $u_{1s}(2)^* u_{nl}(1)^*$, and then integration over the spatial coordinates for each electron ($r_1, \theta_1, \phi_1$ and $r_2, \theta_2, \phi_2$) gives two coupled equations that we write as
\[ \begin{pmatrix} J & K \\ K & J \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \Delta E \begin{pmatrix} a \\ b \end{pmatrix}. \]  
(3.14)

This is eqn 3.12 in matrix form. The direct integral is
\[ J = \frac{1}{4\pi \varepsilon_0} \int \int |u_{1s}(1)|^2 \frac{e^2}{r_{12}} |u_{nl}(2)|^2 \, dr_1^3 \, dr_2^3 \]  
\[ = \frac{1}{4\pi \varepsilon_0} \int \int \rho_{1s}(r_1) \rho_{nl}(r_2) \, dr_1^3 \, dr_2^3, \]  
(3.15)

where $\rho_{1s}(1) = -e |u_{1s}(1)|^2$ is the charge density distribution for electron 1, and similarly for $\rho_{nl}(2)$. This direct integral represents the Coulomb repulsion of these charge clouds (Fig. 3.1). The exchange integral is
\[ K = \frac{1}{4\pi \varepsilon_0} \int \int u_{1s}^*(1) u_{nl}^*(2) \frac{e^2}{r_{12}} u_{1s}(2) u_{nl}(1) \, dr_1^3 \, dr_2^3. \]  
(3.16)
Fig. 3.1 The direct integral in a 1s\( n \)s configuration of helium corresponds to the Coulomb repulsion between two spherically-symmetric charge clouds made up of shells of charge like those shown.

\[
\begin{array}{c}
1L \\
\uparrow \\
2K \\
\downarrow \\
J \\
\downarrow \\
1snl \\
\text{Unperturbed configuration}
\end{array}
\]

Unlike the direct integral, this does not have a simple classical interpretation in terms of charge (or probability) distributions—the exchange integral depends on interference of the amplitudes. The spherical symmetry of the 1s wavefunction makes the integrals straightforward to evaluate (Exercises 3.6 and 3.7).

The eigenvalues \( \Delta E \) in eqn 3.14 are found from

\[
\begin{vmatrix}
J - \Delta E & K \\
K & J - \Delta E
\end{vmatrix} = 0. \tag{3.17}
\]

The roots of this determinantal equation are \( \Delta E = J \pm K \). The direct integral shifts both levels together but the exchange integral leads to an energy splitting of \( 2K \) (see Fig. 3.2). Substitution back into eqn 3.14 gives the two eigenvectors in which \( \hat{b} = a \) and \( \hat{b} = -a \). These correspond to symmetric (S) and antisymmetric (A) wavefunctions:

\[
\psi^S_{\text{space}} = \frac{1}{\sqrt{2}} \left\{ u_{1s}(1)u_{nl}(2) + u_{1s}(2)u_{nl}(1) \right\},
\]

\[
\psi^A_{\text{space}} = \frac{1}{\sqrt{2}} \left\{ u_{1s}(1)u_{nl}(2) - u_{1s}(2)u_{nl}(1) \right\}.
\]

The wavefunction \( \psi^A_{\text{space}} \) has an eigenenergy of \( E^{(0)} + J - K \), and this is lower than the energy \( E^{(0)} + J + K \) for \( \psi^S_{\text{space}} \). (For the 1s\( n \)l configurations in helium \( K \) is positive.) This is often interpreted as the electrons ‘avoiding’ each other, i.e. \( \psi^A_{\text{space}} = 0 \) for \( r_1 = r_2 \), and for this wavefunction, the probability of finding electron 1 close to electron 2 is small (see Exercise 3.3). This anticorrelation of the two electrons makes the expectation of the Coulomb repulsion between the electrons smaller than for \( \psi^S_{\text{space}} \).

The occurrence of symmetric and antisymmetric wavefunctions has a classical analogue illustrated in Fig. 3.3. A system of two oscillators, with the same resonance frequency, that interact (e.g. they are joined together by a spring) has antisymmetric and symmetric normal modes as illustrated in Fig. 3.3(b) and (c). These modes and their frequencies are found in Appendix B as an example of the application of degenerate perturbation theory in Newtonian mechanics.\(^{10}\)

The exchange integral decreases as \( n \) and \( l \) increase because of the reduced overlap between the wavefunctions of the excited electron and

\(^9\)It is easy to check which wavefunction corresponds to which eigenvalue by substitution into the original equation.

\(^{10}\)Another example is the classical treatment of the normal Zeeman effect.
the 1s-electron. These trends are an obvious consequence of the form of
the wavefunctions: the excited electron’s average orbit radius increases
with energy and hence with \( n \); the variation with \( l \) arises because the
effective potential from the angular momentum (‘centrifugal’ barrier)
leads to the wavefunction of the excited electron being small at small
\( r \). However, in the treatment as described above, the direct integral
does not tend to zero as \( n \) and \( l \) increase, as shown by the following
physical argument. The excited electron ‘sees’ the nuclear charge of +2e
surrounded by the 1s electronic charge distribution, i.e. in the region far
from the nucleus where \( nl \)-electron’s wavefunction has a significant value
it experiences a Coulomb potential of charge +1e. Thus the excited
electron has an energy similar to that of an electron in the hydrogen
atom, as shown in Fig. 3.4. But we have started with the assumption
that both the 1s- and \( nl \)-electrons have an energy given by the Rydberg
formula for \( Z = 2 \). The direct integral \( J \) equals the difference between
these energies.\(^{11}\) This work was an early triumph for wave mechanics
since previously it had not been possible to calculate the structure of
helium.\(^{12}\)

In this section we found the wavefunctions and energy levels in helium
by direct calculation but looking back we can see how to anticipate the
answer by making use of symmetry arguments. The Hamiltonian for the
electrostatic repulsion, proportional to \( 1/r_{12} \equiv 1/(r_1 - r_2) \), commutes
with the operator that interchanges the particle labels 1 and 2, i.e. the
swap operation \( 1 \leftrightarrow 2 \). (Although we shall not give this operator a
symbol it is obvious that it leaves the value of \( 1/r_{12} \) unchanged.) Comm-
uting operators have simultaneous eigenfunctions. This prompts us

\( ^{11}\) This can also be seen from eqn 3.15.
The integration over \( r_1 \), \( \psi_1 \) and \( \varphi_1 \)
leads to a repulsive Coulomb potential
\( \sim e/4\pi\varepsilon_0 r_2 \) that cancels part of the at-
ttractive potential of the nucleus, when
\( r_2 \) is greater than the values of \( r_1 \) where
\( \psi_1 \) is appreciable.

\( ^{12}\) For hydrogen, the solution of
Schrödinger’s equation reproduced the
energy levels calculated by the
Bohr–Sommerfeld theory. However,
wave mechanics does give more in-
formation about hydrogen than the
old quantum theory, e.g. it allows the
detailed calculation of transition rates.
Fig. 3.4 The energy levels of the helium atom with those of hydrogen for comparison. The $1s^2$ ground configuration is tightly bound. For the excited configurations of helium the $1s$-electron screens the outer electron from the nuclear charge so that the $1s\ell d$ configurations in helium have similar energy to the shell with principal quantum number $n$ in hydrogen. The hydrogenic levels are indicated on the right. The interval between the $^1L$ and $^3L$ terms (equal to twice the exchange integral) is clear for the $1s2s, 1s2p, 1s3s, 1s3p$ and $1s4s$ configurations but it is smaller for higher $n$ and $l$. 
to construct the symmetrised wavefunctions $\psi^A_{\text{space}}$ and $\psi^S_{\text{space}}$. In this basis of eigenstates it is simple to calculate the effect of the electrostatic repulsion.

### 3.2.1 Spin eigenstates

The electrostatic repulsion between the two electrons leads to the wavefunctions $\psi^S_{\text{space}}$ and $\psi^A_{\text{space}}$ in the excited states of the helium atom. The ground state is a special case where both electrons have the same spatial wavefunction, so only a symmetric solution exists. We did not consider spin since electrostatic interactions depend on the charge of the particle, not their spin. Neither $H_0$ nor $H'$ contains any reference to the spin of the electrons. Spin does, however, have a profound effect on atomic wavefunctions. This arises from the deep connection between spin and the symmetry of the wavefunction of indistinguishable particles. Note that here we are considering the total wavefunction in the systems that includes both the spatial part (found in the previous section) and the spin. Fermions have wavefunctions that are antisymmetric with respect to particle-label interchange, and bosons have symmetric ones. As a consequence of this symmetry property, fermions and bosons fill up the levels of a system in different ways, i.e. they obey different quantum statistics.

Electrons are fermions so atoms have total wavefunctions that are antisymmetric with respect to permutation of the electron labels. This requires $\psi_{\text{space}}$ to associate with an antisymmetric spin function $\psi^A_{\text{spin}}$, and the other way round:

$$\psi = \psi^S_{\text{space}} \psi^A_{\text{spin}} \quad \text{or} \quad \psi^A_{\text{space}} \psi^S_{\text{spin}}.$$  \hspace{1cm} (3.18)

These antisymmetrised wavefunctions that we have constructed fulfil the requirement of having particular symmetry with respect to the interchange of indistinguishable particles. Now we shall find the spin eigenfunctions explicitly. We use the shorthand notation where $\uparrow$ and $\downarrow$ represent $m_s = 1/2$ and $-1/2$, respectively. Two electrons have four possible combinations: the three symmetric functions,

$$\psi^S_{\text{spin}} = |\uparrow\uparrow\rangle = \frac{1}{\sqrt{2}} \{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\} \hspace{1cm} (3.19)$$

$$= |\downarrow\downarrow\rangle,$$

corresponding to $S = 1$ and $M_S = +1, 0, -1$; and an antisymmetric function

$$\psi^A_{\text{spin}} = \frac{1}{\sqrt{2}} \{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\}, \hspace{1cm} (3.20)$$

corresponding to $S = 0$ (with $M_S = 0$). Spectroscopists label the eigenstates of the electrostatic interactions with the symbol $2S+1L$, where $S$ is the total spin and $L$ is the total orbital angular momentum quantum number. The $1s\pi l$ configurations in helium $L = l$, so the allowed terms

$^{13}$For two electrons, swapping the particle labels twice brings us back to where we started, so $\psi(1,2) = \pm \psi(2,1)$. Therefore the two possible eigenvalues are 1 for $\psi^S_{\text{space}}$ and $-1$ for $\psi^A_{\text{space}}$.

$^{14}$Indistinguishable means that the particles are identical and have the freedom to exchange positions, e.g. atoms in a gas which obey Fermi–Dirac or Bose–Einstein statistics depending on their spin. In contrast, atoms in a solid can be treated as distinguishable, even if they are identical, because they have fixed positions—we could label the atoms 1, 2, etc. and still know which is which at some later time.

$^{15}$These statements about the result of adding two $s = 1/2$ angular momenta can be proved by formal angular momentum theory. Simplified treatments describe $S = 0$ as having one electron with 'spin-up' and the other with 'spin-down'; but both $M_S = 0$ states are linear combinations of the states $|m_{s1} = +1/2, m_{s2} = -1/2\rangle$ and $|m_{s1} = -1/2, m_{s2} = +1/2\rangle$. 

are $^1L$ and $^3L$, e.g. the $1s2s$ configuration in helium gives rise to the terms $^1S$ and $^3S$, where $S$ represents $L = 0$.\textsuperscript{16}

In summary, we have calculated the structure of helium in two distinct stages.

1) Energies Degenerate perturbation theory gives the space wavefunctions $\psi^S_{\text{space}}$ and $\psi^A_{\text{space}}$ with energies split by twice the exchange integral. In helium the degeneracy arises because the two electrons are identical particles so there is exchange degeneracy, but the treatment is similar for systems where a degeneracy arises by accident.

2) Spin We determined the spin associated with each energy level by constructing symmetrised wavefunctions. The product of the spatial functions and the spin eigenstates gives the total atomic wavefunction that must be antisymmetric with respect to particle-label interchange.

Exchange degeneracy, exchange integrals, degenerate perturbation theory and symmetrised wavefunctions all occur in helium and their interrelationship is not straightforward so that misconceptions abound. A common misinterpretation is to infer that because levels with different total spin, $S = 0$ and 1, have different energies then there is a spin-dependent interaction—this is not correct, but sometimes in condensed matter physics it is useful to pretend that it is! (See Blundell 2001.) The interactions that determine the gross structure of helium are entirely electrostatic and depend only on the charge and position of the particles. Also, degenerate perturbation theory is sometimes regarded as a mysterious quantum phenomenon. Appendix A gives further discussion and shows that symmetric and antisymmetric normal modes occur when two classical systems, with similar energy, interact, e.g. two coupled oscillators.

3.2.2 Transitions in helium

To determine which transitions are allowed between the energy levels of helium we need a selection rule for spin: the total spin quantum number does not change in electric dipole transitions. In the matrix element $\langle \psi_{\text{final}} | r | \psi_{\text{initial}} \rangle$ the operator $r$ does not act on spin; therefore, if the $\psi_{\text{final}}$ and $\psi_{\text{initial}}$ do not have the same value of $S$, then their spin functions are orthogonal and the matrix element equals zero.\textsuperscript{17} This selection rule gives the transitions shown in Fig. 3.5.

\textsuperscript{17}This anticipates a more general discussion of this and other selection rules for the $LS$-coupling scheme in a later chapter.
3.3 Evaluation of the integrals in helium

In this section we shall calculate the direct and exchange integrals to make quantitative predictions for some of the energy levels in the helium atom, based on the theory described in the previous sections. This provides an example of the use of atomic wavefunctions to carry out a calculation where there are no corresponding classical orbits and gives an indication of the complexities that arise in systems with more than one electron. The evaluation of the integrals requires care and some further details are given in Appendix B. The important point to be learnt from this section, however, is not the mathematical techniques but rather to see that the integrals arise from the Coulomb interaction between electrons treated by straightforward quantum mechanics.

3.3.1 Ground state

To calculate the energy of the 1s² configuration we need to find the expectation value of \( e^2/4\pi\epsilon_0 r_{12} \) in eqn 3.1—this calculation is the same as the evaluation of the mutual repulsion between two charge distributions in classical electrostatics, as in eqn 3.15 with \( \rho_{1s}(r_1) \) and \( \rho_{1s}(r_2) = \rho_{1s}(r_2) \). The integral can be considered in different ways. We could calculate the energy of the charge distribution of electron 1 in the potential created by electron 2, or the other way around. This section does neither; it uses a method that treats each electron symmetrically (as in Appendix B), but of course each approach gives the same numerical result. Electron 1 produces an electrostatic potential at radial distance \( r_2 \) given by

\[
V_{12}(r_2) = \int_0^{r_2} \frac{1}{4\pi\epsilon_0 r_{12}} \rho(r_1) \, d^3r_1. \tag{3.21}
\]

Fig. 3.5 The allowed transitions between the terms of helium are governed by the selection rule \( \Delta S = 0 \) in addition to the rule \( \Delta l = \pm 1 \) found previously. Since there are no transitions between singlets and triplets it is convenient to draw them as two separate systems. Notice that in the radiative decay of helium atoms excited to high-lying levels there are bottlenecks in the metastable 1s2s \(^3\)S and 1s2s \(^1\)S terms.
The spherical symmetry of s-electrons means that the charge in the region $r_1 < r_2$ acts like a point charge at the origin, so that

$$V_{12}(r_2) = \frac{Q(r_2)}{4\pi \epsilon_0 r_2},$$

where $Q(r_2)$ is the charge within a radius of $r_2$ from the origin, which is given by

$$Q(r_2) = \int_0^{r_2} \rho(r_1) 4\pi r_1^2 dr_1. \quad (3.22)$$

The electrostatic energy that arises from the repulsion equals

$$E_{12} = \int_0^{\infty} V_{12}(r_2) \rho(r_2) 4\pi r_2^2 dr_2. \quad (3.23)$$

For the $1s^2$ configuration there is an exactly equal contribution to the energy from $V_{21}(r_1)$, the (partial) potential at $r_1$ produced by electron 2. Thus the total energy of the repulsion between the electrons is twice that in eqn 3.23. Using the radial wavefunction for a 1s-electron, we find

$$J_{1s^2} = 2\times \frac{e^2}{4\pi \epsilon_0} \int_0^{\infty} \left\{ \int_0^{r_1} 4Z^2 e^{-(Z/a_0)2r_1} r_1^2 dr_1 \right\} 4Z^2 e^{-(Z/a_0)2r_2} r_2^2 dr_2$$

$$= \frac{e^2}{4\pi \epsilon_0} \frac{5}{2a_0} Z = (13.6 \text{ eV}) \times \frac{5}{4} Z. \quad (3.24)$$

For helium this gives $J_{1s^2} = 34 \text{ eV}$. 

### 3.3.2 Excited states: the direct integral

A $1s\nu l$ configuration of helium has an energy close to that of an $n$-electron in hydrogen, e.g. in the $1s2p$ configuration the 2p-electron has a similar binding energy to the $n = 2$ shell of hydrogen. The obvious explanation, in Bohr’s model, is that the 2p-electron lies outside the 1s-orbit so that the inner electron screens the outer one from the full nuclear charge. Applying an analogous argument to the quantum treatment of helium leads to the Hamiltonian $H = H_{0a} + H'_a$, where

$$H_{0a} = -\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) - \frac{e^2}{4\pi \epsilon_0} \left( \frac{2}{r_1} + \frac{1}{r_2} \right) \quad (3.25)$$

and

$$H'_a = \frac{e^2}{4\pi \epsilon_0} \left( \frac{1}{r_{12}} - \frac{1}{r_2} \right). \quad (3.26)$$

In the expression for $H_{0a}$, electron 2 experiences the Coulomb attraction of a charge +1e. In $H'_a$ the subtraction of $e^2/4\pi \epsilon_0 r_2$ from the mutual repulsion means that the perturbation tends to zero at a large distance from the nucleus (which is intuitively reasonable). This decomposition differs from that in Section 3.1. The different treatment of the two electrons makes the perturbation theory a little tricky, but Heisenberg
did the calculation as described in Bethe and Salpeter (1957) or Bethe and Salpeter (1977); he found the direct integral

\[ J_{1sni} = \frac{e^2}{4\pi \varepsilon_0} \int_{0}^{\infty} \left( \int \frac{1}{r_{12}} - \frac{1}{r_2} \right) |u_{1s}(1)|^2 |u_{nlm}(2)|^2 \, d^3r_1 \, d^3r_2. \]  

(3.27)

This must be evaluated with the appropriate wavefunctions, i.e. \( u_{nlm}^{Z=1} \) rather than \( u_{nlm}^{Z=2} \), and \( u_{nlm}^{Z=2} \) as before.\(^{21}\) For the excited electron \( u_{nlm} = R_{nl}(r) Y_m(\theta, \phi) \), where \( R_{nl}(r) \) is the radial function for \( Z = 1 \). We write the direct integral as

\[ J_{1sni} = \frac{e^2}{4\pi \varepsilon_0} \int_{0}^{\infty} \int_{0}^{\infty} J(r_1, r_2) R_{10}^2(r_1) R_{nl}^2(r_2) r_1^2 r_2^2 \, dr_1 \, dr_2, \]  

(3.28)

where the angular parts are contained in the function\(^{22}\)

\[ J(r_1, r_2) = \int_{0}^{2\pi} \int_{0}^{\pi} \int_{0}^{2\pi} \left( \int \frac{1}{r_{12}} - \frac{1}{r_2} \right) \frac{1}{4\pi} |Y_m(\theta_2, \phi_2)|^2 \times \sin \theta_1 \, d\theta_1 \, d\phi_1 \, \sin \theta_2 \, d\theta_2 \, d\phi_2. \]  

(3.29)

The calculation of this integral requires the expansion of \( 1/r_{12} \) in terms of spherical harmonics:\(^{23}\)

\[ \frac{1}{r_{12}} = \frac{1}{r_2} \sum_{k=0}^{\infty} \left( \frac{r_1}{r_2} \right)^k \frac{4\pi}{2k+1} \sum_{q=-k}^{k} Y_{k,q}^* (\theta_1, \phi_1) Y_{k,q} (\theta_2, \phi_2) \]  

(3.30)

for \( r_2 > r_1 \) (and \( r_1 \leftrightarrow r_2 \) when \( r_1 > r_2 \)). Only the term for \( k = 0 \) survives in the integration over angles in eqn 3.29 to give\(^{24}\)

\[ J(r_1, r_2) = \begin{cases} 0 & \text{for } r_1 < r_2, \\ 1/r_1 - 1/r_2 & \text{for } r_1 > r_2. \end{cases} \]

When \( r_1 < r_2 \) the original screening argument applies and eqn 3.25 gives a good description. When \( r_1 > r_2 \) the appropriate potential is proportional to \(-2/r_2 - 1/r_1\) and \( J(r_1, r_2) \) accounts for the difference between this and \(-2/r_1 - 1/r_2\) used in \( H_{\text{eo}} \). Thus we find

\[ J_{1sni} = \frac{e^2}{4\pi \varepsilon_0} \int_{0}^{\infty} \left( \int_{r_2}^{\infty} \left( \frac{1}{r_1} - \frac{1}{r_2} \right) R_{10}^2(r_1) r_1^2 \, dr_1 \right) R_{nl}^2(r_2) r_2^2 \, dr_2. \]  

(3.31)

Evaluation of this integral for the \( 1s2p \) configuration (in Exercise 3.6) gives \( J_{1s2p} = -2.8 \times 10^{-2} \, \text{eV} \)—three orders of magnitude smaller than \( J_{1s2s} \) in eqn 3.7 (evaluated from eqn 3.24). The unperturbed wavefunction for \( Z = 1 \) has energy equal to that of the corresponding level in hydrogen and the small negative direct integral accounts for the incompleteness of the screening of the \( n \ell \)-electron by the inner electron.

### 3.3.3 Excited states: the exchange integral

The exchange integral has the same form as eqn 3.16 but with \( u_{nlm}^{Z=1} \) rather than \( u_{nlm}^{Z=2} \) (and \( u_{1s}^{Z=2} \) as before). Within the spatial wavefunction...
\[ u_{nlm} = R_{nl}(r)Y_{lm}(\theta, \phi) \] only the radial part depends on \( Z \). We write the exchange integral as (cf. eqn 3.28)

\[
K_{1s1s} = \frac{e^2}{4\pi\epsilon_0} \int \int K(r_1, r_2) R_{1s}(r_1) R_{1s}(r_2) R_{nl}(r_2) \frac{r_1^2}{r_2^2} \, dr_1 \frac{r_1^2}{r_2^2} \, dr_2.
\]

(3.32)

The function \( K(r_1, r_2) \) containing the angular integrals is (cf. eqn 3.29)

\[
K(r_1, r_2) = \int \int \int \frac{1}{r_{12}} Y_{lm}^* (\theta_1, \phi_1) \frac{1}{4\pi} Y_{lm} (\theta_2, \phi_2)
\times \sin \theta_1 \, d\theta_1 \, d\phi_1 \sin \theta_2 \, d\theta_2 \, d\phi_2.
\]

(3.33)

For the 1snp configuration only the second term of the expansion in eqn 3.30, with \( k = 1 \), survives in the integration because of the orthogonality of the spherical harmonic functions (see Exercise 3.7), to give

\[
K(r_1, r_2) = \begin{cases} 
\frac{r_1}{3r_2} & \text{for } r_1 < r_2, \\
\frac{r_2}{3r_1} & \text{for } r_2 < r_1.
\end{cases}
\]

(3.34)

Carrying out the integration over the radial wavefunctions in eqn 3.32 for the 1s2p configuration gives the splitting between \( ^3P \) and \( ^1P \) as \( 2K_{1s2p} \approx 0.21 \text{ eV} \) (close to the measured value of 0.25 eV).

The assumption that the excited electron lies outside the 1s wavefunction does not work so well for 1sns configurations since \( \psi_{ns} (0) \) has a finite value and the above method of calculating \( J \) and \( K \) is less accurate.\footnote{The overlap of the 1s and \( nl \) wavefunctions becomes smaller as \( n \) and \( l \) increase. In Heisenberg’s treatment where screening is taken into account, the direct integral gives the deviation from the hydrogenic levels (which could be characterised by a quantum defect as in the alkalis, see Chapter 4). For electrons with \( l \neq 0 \) the term \( -\hbar^2 l(l+1)/2mr^2 \) in the Schrödinger equation causes the electron’s wavefunction to lie almost entirely outside the region where \( u_{nlm}^2 = R_{ls} (r) / \sqrt{4\pi} \) has a significant value.}

\footnote{The 1s2s configuration of helium has a singlet–triplet separation of \( E_1 (S) - E_3 (S) = 2K_{1s2s} \approx 0.80 \text{ eV} \) and the direct integral is also larger than that for 1s2p—these trends are evident in Fig. 3.4 (see also Exercise 3.7).}

In some respects, helium is a more typical atom than hydrogen. The Schrödinger and Dirac equations can be solved exactly for the one-electron system, but not for helium or other atoms with more electrons. Thus in a careful study of helium we encounter the approximations needed to treat multi-electron systems, and this is very important for understanding atomic structure in general. Helium also gives a good example of the influence of identical particles on the occupation of the states in quantum systems. The energy levels of the helium atom (and the existence of exchange integrals) do not depend on the fact that the two electrons are identical, as demonstrated in Exercises 3.3 and 3.4; however, this is a common point of confusion. The books recommended for further reading give clear and accurate descriptions of helium that reward careful study.

### Further reading

The recommended books are divided into two categories corresponding to the two main themes in this chapter: (a) a description of how to calculate the electrostatic energy in an atom with more than one electron, which introduces principles that can be used in atoms with more
3.3 Evaluation of the integrals in helium

electrons; and (b) a discussion of the influence of identical particles on the statistics of a quantum system that is important throughout physics. The influence of identical particles on the occupation of the quantum levels of a system with many particles, i.e. Bose-Einstein and Fermi-Dirac statistics, is discussed in statistical mechanics texts. Clear descriptions of helium may be found in the following textbooks: Cohen-Tannoudji et al. (1977), Woodgate (1980) and Mandl (1992). The calculation of the direct and exchange integrals in Section 3.3 is based on the definitive work by Bethe and Salpeter (1957), or see Bethe and Jackiw (1986).

A very instructive comparison can be made between the properties of the two electrons in helium and the nuclear spin statistics of homonuclear diatomic molecules described in Atkins (1983, 1994). There are diatomic molecules with nuclei that are identical bosons, identical fermions and cases of two similar but not identical particles, and their study gives a wider perspective than the study of helium alone. The nuclei of the two atoms in a hydrogen molecule are protons which are fermions (like the two electrons in helium). For reasons explained in the above references, we can consider only those parts of the molecular wavefunction that describe the rotation \( \psi_{\text{rot}} \) and the nuclear spin states \( \psi_I \)—these are spatial and spin wavefunctions, respectively. For \( \text{H}_2 \) the wavefunction must have overall antisymmetry with respect to an interchange of particle labels since the nuclei are protons, each with a spin of 1/2. This requires that a rotational must pair with a spin function of the opposite symmetry:

\[
\psi_{\text{molecule}} = \psi_{\text{rot}}^S \psi_I^A \quad \text{or} \quad \psi_{\text{rot}}^A \psi_I^S .
\]

This is analogous to eqn 3.18 for helium; as described in Section 3.2.1, the two spin-1/2 nuclei in a hydrogen molecule give a total (nuclear) spin of 0 and 1, with one state and three states, respectively. The 1 to 3 ratio of the number of nuclear spins associated with the energy levels for \( \psi_{\text{rot}}^S \) and \( \psi_{\text{rot}}^A \), respectively, influences the populations of these rotational energy levels in a way that is directly observed in molecular spectra (the intensity of the lines in spectra depends on the population of the initial level). The molecule \( \text{H}_D \) made from hydrogen and deuterium does not have identical nuclei so there is no overall symmetry requirement, but it has similar energy levels to those of \( \text{H}_2 \) apart from the mass dependence. This gives a real physical example where the statistics depends on whether the particles are identical or not, but the energy of the system does not. Exercise 3.4 discusses an artificial example: a helium-like system that has the same energy levels as a helium atom and hence the same direct and exchange integrals, even though the constituent particles are not identical.

27 Molecules made up of two atoms with identical nuclei.

28 These books also summarise the helium atom and the quantum mechanics of these molecular systems is very closely related to atomic physics.

29 The wavefunction of the hydrogen molecule has exchange symmetry—crudely speaking, the molecule looks the same when rotated through 180°.
Exercises

More advanced problems are indicated by a *.

(3.1) Estimate of the binding energy of helium
(a) Write down the Schrödinger equation for the helium atom and state the physical significance of each of the terms.
(b) Estimate the equilibrium energy of an electron bound to a charge \( Ze \) by minimising
\[
E(r) = \frac{\hbar^2}{2mr^2} - \frac{Ze^2}{4\pi\varepsilon_0 r}.
\]
(c) Calculate the repulsive energy between the two electrons in helium assuming that \( r_{12} \approx r \). Hence estimate the ionization energy of helium.
(d) Estimate the energy required to remove a further electron from the helium-like ion \( \text{Si}^{12+} \), taking into account the scaling with \( Z \) of the energy levels and the expectation value for the electrostatic repulsion. The experimental value is \( 2400 \text{ eV} \). Compare the accuracy of your estimates for \( \text{Si}^{12+} \) and helium. (IE(He) = 24.6 eV.)

(3.2) Direct and exchange integrals for an arbitrary system
(a) Verify that for
\[
\psi^A(r_1, r_2) = \frac{1}{\sqrt{2}} \{ u_\alpha(r_1)u_\beta(r_2) - u_\alpha(r_2)u_\beta(r_1) \}
\]
and \( H' = e^2/4\pi\varepsilon_0 r_{12} \) the expectation value \( \langle \psi^A | H' | \psi^A \rangle \) has the form \( J - K \) and give the expressions for \( J \) and \( K \).
(b) Write down the wavefunction \( \psi^S \) that is orthogonal to \( \psi^A \).
(c) Verify that \( \langle \psi^A | H' | \psi^S \rangle = 0 \) so that \( H' \) is diagonal in this basis.

(3.3) Exchange integrals for a delta-function interaction
A particle in a square-well potential, with \( V(x) = 0 \) for \( 0 < x < \ell \) and \( V(x) = \infty \) elsewhere, has normalised eigenfunctions \( u_0(x) = \sqrt{2/\ell} \sin(\pi x / \ell) \) and \( u_1(x) = \sqrt{2/\ell} \sin(2\pi x / \ell) \).
(a) What are the eigenenergies \( E_0 \) and \( E_1 \) of these two wavefunctions for a particle of mass \( m \)?
(b) Two particles of the same mass \( m \) are both in the ground state so that the energy of the whole system is \( 2E_0 \). Calculate the perturbation produced by a point-like interaction described by the potential \( a \delta(x_1 - x_2) \), with \( a \) a constant.
(c) Show that, when the two interacting particles occupy the ground and first excited states, the direct and exchange integrals are equal. Also show that the delta-function interaction produces no energy shift for the antisymmetric spatial wavefunction and explain this in terms of correlation of the particles. Calculate the energy of the other level of the perturbed system.
(d) For the two energy levels found in part (c), sketch the spatial wavefunction as a function of the coordinates of the two particles \( x_1 \) and \( x_2 \). The particles move in one dimension but the two-particle wavefunction exists in a two-dimensional Hilbert space—draw either a contour plot in the \( x_1, x_2 \)-plane or attempt a three-dimensional sketch (by hand or computer).
(e) The two particles are identical and have spin 1/2. What is the total spin quantum number \( S \) associated with each of the energy levels found in part (c)?
* (f) Discuss qualitatively the energy levels of this system for two particles that have slightly different masses \( m_1 \neq m_2 \), so that they are distinguishable? [Hint. The spin has not been given because it is not important for non-identical particles.]

Comment. The antisymmetric spatial wavefunction in part (c) clearly has different properties from a straightforward product \( u_0u_1 \). The exchange integral is a manifestation of the entanglement of the multiple-particle system.

(3.4) A helium-like system with non-identical particles
Imagine that there exists an exotic particle with the same mass and charge as the electron but spin \( 3/2 \) (so it is not identical to the electron). This particle and an electron form a bound system with a helium nucleus. Compare the energy levels of this system with those of the helium atom. Describe the energy levels of a system with two of the ex-