Electronic Engineering Semiconductors and Devices

A second edition of Electronic Engineering Materials and Devices

JOHN ALLISON
Reader
Department of Electronic and Electrical Engineering
University of Sheffield

McGRAW-HILL BOOK COMPANY

London • New York • St Louis • San Francisco • Auckland
Bogotá • Guatemala • Hamburg • Lisbon • Madrid • Mexico
Montreal • New Delhi • Panama • Paris • San Juan • São Paulo
Singapore • Sydney • Tokyo • Toronto

ELECTRONIC ENGINEERING SEMICONDUCTORS AND DEVICES
INTERNATIONAL EDITION

1 2 3 4 5 6 7 8 9 0 CMO UPE 9 5 4 3 2 1 0

British Library Cataloguing in Publication Data
Allison, John, (date)
Electronic engineering semiconductors and devices.
2nd ed.
1. Electronic equipment
I. Title II. Allison, John, (date). Electronic
engineering materials and devices
621.381

ISBN 0-07-084194-2

British Library of Congress Cataloging-in-Publication Data
Allison, John.
Electronic engineering semiconductors and devices/John Allison.
2nd ed.
p. cm.
Rev. ed. of: Electronic engineering materials and devices. (1971) ISBN 0-07-084194-2
1. Electronics-Materials. 2.Semiconductors. I. Allison, John.
Electronic engineering materials and devices. II. Title.
TK7871.A44 1989
621.381—dc20 89-12813

When ordering this title use ISBN 0-07-100627-3

Printed in Singapore

TO
VICTORIA

We hope she is amused

'Perhaps a frail memorial, but sincere,
Not scorn'd in heav'n, though little notic'd here.'

William Cowper (1731–1800)

Contents

Preface xi

Physical constants xiv

Properties of some common semiconductors at room temperature xv

1 The quantum behaviour of waves and particles 1

1. 1 Introduction 1
1. 2 Black-body radiation 2
1. 3 The photoelectric effect 3
1. 4 The Bohr atom 7
1. 5 Particle–wave duality 11
1. 6 Wavepackets: group and phase velocities of particles 16
1. 7 The Schrödinger wave equation 21
1. 8 Interpretation of the wavefunction ψ 24
1. 9 The uncertainty principle 25
1.10 Beams of particles and potential barriers 27
Problems 33

2 The electronic structure of atoms 35

2.1 Introduction 35
2.2 A particle in a one-dimensional potential well 35
2.3 The hydrogen atom 40
2.4 The exclusion principle and the periodic table of elements 45
Problems 47

3 Collections of particles in gases and solids 48

3.1 Introduction 48
3.2 Assemblies of classical particles—ideal gases 49
3.3 Collections of particles that obey the exclusion principle 56
Problems 59

4 Conduction in metals 62

4.1 Introduction 62
4.2 A simple model of a conductor 62
4.3 Electrons trapped in a three-dimensional potential box 64
4.4 Maximum number of possible energy states 66
4.5 The energy distribution of electrons in a metal 68
4.6 The Fermi level in a metal 71
4.7 Conduction processes in metals 72
Problems 77

5 Electrons in solids—an introduction to band theory 78

5.1 Introduction 78
5.2 Allowed energy bands of electrons in solids 79
5.3 The velocity and effective mass of electrons in a solid 90
5.4 Conductors, semiconductors and insulators 95
5.5 Electrical resistance of solids 97
Problems 100

6 Semiconductors 102

6. 1 Introduction 102
6. 2 The crystal structure of semiconductors 102
6. 3 Conduction processes in semiconductors 107
6. 4 Density of carriers in intrinsic semiconductors 110
6. 5 Extrinsic or impurity semiconductors 116
6. 6 Electron processes in real semiconductors 121
6. 7 The density of carriers and the position of the Fermi level in extrinsic semiconductors 124
6. 8 Compensation doping 128
6. 9 Electrical conduction in semiconductors 129
6.10 The continuity equation 134
6.11 Semiconductor measurements 137
Problems 144

7 Junction diodes 150

7. 1 Introduction 150
7. 2 The pn junction in equilibrium with no applied voltage 150
7. 3 Current flow in a pn junction with forward bias 157
7. 4 Current flow in a pn junction with reverse bias 164
7. 5 Current–voltage characteristics of a junction diode 166
7. 6 Electron and hole injection efficiencies 166
7. 7 The pn junction with finite dimensions 168
7. 8 The geometry of the depletion layer and depletion-layer capacitance 169

7. 9 Diffusion capacitance and the small-signal equivalent circuit of a pn junction 174
7.10 Switching characteristics 176
7.11 Metal–semiconductor junctions 178
7.12 Breakdown in pn junctions; the Zener diode 184
7.13 The tunnel diode 190
7.14 Fabrication technology 191
Problems 194

8 The bipolar junction transistor 201

8.1 Introduction 201
8.2 Phenomenological description of current transport in the abrupt-junction bipolar transistor 201
8.3 Gain parameters of the bipolar transistor 206
8.4 Non-ideal transistor structures 210
8.5 Small-signal equivalent circuit 221
8.6 Fabrication of junction transistors 224
8.7 Silicon controlled rectifier 225
Problems 232

9 Field-effect devices 235

9.1 General properties and types of field-effect transistor (FET) 235
9.2 The junction field-effect transistor (JFET) 236
9.3 Drain and transfer characteristics of the JFET 238
9.4 The metal–oxide–silicon field-effect transistor (MOSFET) 244
9.5 The small-signal equivalent circuit and frequency response of FETs 252
9.6 Charge-coupled devices (CCDs) 254
Problems 257

10 Integrated circuits 261

10.1 Introduction 261
10.2 Methods of isolation 262
10.3 Monolithic integrated circuits 265
10.4 A brief outline of silicon planar integrated-circuit fabrication 266
10.5 Bipolar integrated components 276
10.6 MOS integrated circuits and components 284
10.7 Recent technological developments 288
10.8 Comments on integrated-circuit types 295
Problems 298

11 Optoelectronic devices 301

11.1 Optical display devices 301
11.2 Optical detectors 319
11.3 Solar cells 331
Problems 336

12 Compound semiconductors and devices 339

12.1 Introduction 339
12.2 Compound semiconductors 340
12.3 Doping compound semiconductors 342
12.4 GaAs Schottky diodes and MESFETs 345
12.5 High-electron-mobility transistors 352
12.6 Heterostructure bipolar transistors 353
12.7 Gunn diode oscillator 355

13 New technologies and devices 359

13.1 Multicomponent semiconductors—energy band engineering 359
13.2 Multiple-quantum-well structures—wavefunction engineering 359
13.3 Amorphous semiconductor devices 361

Index 372

Preface

A tremendous flurry of exciting progress in electronic engineering has occurred in the two decades since the first edition of this book appeared. We have seen, for example, the development of rudimentary integrated circuits containing a few components into today's version containing a complete computer, incorporating millions of active devices, on a single chip. The apparently daunting task for the undergraduate student and the electronic device engineer (or the prospective author!) of keeping abreast of such an exponentially expanding technological development is somewhat ameliorated only because the principles of operation of constituent semiconductor components have remained constant. In spite of dramatic changes in the manner in which devices are constructed and interconnected, a flavour of which can be obtained from a quick glance at the later chapters, the essential device science is consequently as relevant today as it was 20 years ago. So, fortunately, is the basic philosophy on which the earlier book was founded, namely to concentrate essentially on timeless electronic fundamentals, such as the description of charge transport in semiconductor junctions, using the operating principles of recent devices as engineering examples of the application of such quintessential electronic concepts.

In the early days of electronic engineering, the essentials of vacuum electronic devices could be understood without too much difficulty and many of their characteristics could be derived theoretically by classical methods. Today, not only is the physical nature of the transport of charge in modern electronic devices more complicated, but many of the more recent devices have properties that cannot be explained satisfactorily without recourse to quantum electronics.

Many textbooks are available that describe, for example, the theoretical physics of the solid state, often in great detail and not without some degree of mathematical complexity. Other, usually engineering, texts often merely provide a cursory description of device behaviour as a preliminary to a detailed discussion of their circuit application. This book is an attempt to close the gap between the two extremes. It provides a physical description of the properties of electronic materials that is sufficiently detailed to allow complete characterization of the electrical performance of modern electronic devices in a manner that can be fully understood by the engineer.

The author is acutely aware of the rapid advances being made in electronic engineering. There is nothing to suggest that the rate of technological development observed over the last few decades will not continue; indeed the pace will probably increase. New devices are continually being developed and absorbed into the technology, causing some textbooks on the subject to become obsolete almost as soon as they are published. It is hoped that, by concentrating on fundamental processes occurring in electronic materials and by discussing contemporary devices as specific examples of these processes, this book will avoid such a fate. Although lack of space prevents complete in-depth coverage of all devices, sufficient insight into the basic properties of elementary devices is given to enable the reader to progress to the study of whatever new or perhaps as yet undeveloped device may be his or her own particular interest.

The text is based in the main on a series of lecture courses given to university electronic and electrical engineering undergraduates in their first, second and (part of) final years. For this reason a choice had to be made between subdividing the subject matter so as to make the level of treatment progressively more difficult or arranging the material in a more logically acceptable way. The former method is usual primarily in teaching, where the presentation has to be geared to the mathematical ability of the student, but it suffers the disadvantage of lack of continuity, each device being described on several occasions, each time with an increasing depth of treatment. In this book the latter course has been adopted in an attempt to provide a text that unifies much of the electronic materials and device teaching over the complete subject range, so emphasizing the relevance of each topic. This arrangement need not be a disadvantage, as few textbooks are read, in the first instance, straight through from beginning to end. The main advantage is that, while the students are able to cover the contents by any one of many routes, dictated by their own ability or as directed by their tutors, they will at the same time possess a book that amalgamates all the material into a coherent whole.

It might be supposed that this textbook is aimed solely at electronic and electrical engineering students, but practicing engineers who feel the need for retraining should also find it helpful, as well as students of associated disciplines in applied technology, physics and chemistry.

I am most grateful to my colleagues at Sheffield for many enjoyable and useful discussions, in particular Dr Clive Woods, whose meticulous reasoning, based on sound engineering principles, has been a constant source of inspiration. I am also most appreciative of friends in industry and in-house for providing additional information and illustrations and to the help of generations of students for their patient forbearance. I would also like to express my sincere thanks to Miss Margaret Eddell and her staff for their unstinting help in preparing the manuscript, especially Miss Elaine Jessop for her ability to read the author's mind, as well as decipher his hieroglyphics so efficiently.

Finally, my warmest thanks are extended to my wife, Judith, for her support and for providing endless coffee and a few cherished moments of silence, and also to Victoria, who was only a twinkle in her father's eye at the beginning of this project but is now able to offer her assistance, without which this manuscript might have been produced much earlier!

May I venture to hope that readers of this book will find it as interesting and informative as I have found it enjoyable and exciting to write.

John Allison

Physical constants

	Symbol		Units
Permittivity of free space	ϵ_0	8.854×10^{-12}	$F\,m^{-1}$
Permeability of free space	μ_0	$4\pi \times 10^{-7}$	$H\,m^{-1}$
Electronic charge	e	1.602×10^{-19}	C
Electronic rest mass	m	9.108×10^{-31}	kg
Electronic charge/mass ratio	e/m	1.759×10^{11}	$C\,kg^{-1}$
Proton rest mass		$1836m$	kg
Planck's constant	h	6.625×10^{-34}	J s
Boltzmann's constant	k	1.380×10^{-23}	$J\,K^{-1}$
kT at room temperature		0.0259	eV

Properties of some common semiconductors at room temperature

	Si	*Ge*	*GaAs*	*InSb*
Atomic weight	28.09	72.59	–	–
Atomic density (m^{-3})	5.02×10^{28}	4.42×10^{28}	–	–
Lattice constant, a(nm)	0.543	0.565	0.563	0.645
Relative permittivity, ϵ_r	11.8	16.0	13.5	11.5
Energy gap, E_g(eV)	1.08	0.66	1.58	0.23
Electron mobility, μ_e($m^2\,V^{-1}\,s^{-1}$)	0.13	0.38	0.85	7.0
Hole mobility, μ_h($m^2\,V^{-1}\,s^{-1}$)	0.05	0.18	0.04	0.10
Intrinsic concentration, n_i(m^{-3})	1.38×10^{16}	$2{\cdot}5 \times 10^{19}$	9×10^{12}	1.6×10^{22}
Electron diffusion constant, D_e($m^2\,s^{-1}$)	0.0031	0.0093	0.020	0.0093
Hole diffusion constant, D_h($m^2\,s^{-1}$)	0.0007	0.0044	–	–
Density of states at conduction band edge, N_c(m^{-3})	2.8×10^{25}	1.0×10^{25}	4.7×10^{23}	–
Density of states at valence band edge, N_v(m^{-3})	1.0×10^{25}	6.0×10^{24}	7.0×10^{24}	–
Melting point (°C)	1420	936	1250	523

1 The quantum behaviour of waves and particles

1.1 Introduction

Throughout our discussion of the electronic properties of materials and the application of these properties to a physical understanding of the operation of electronic devices, we shall constantly be referring to the interaction of particles of atomic size, for example electrons, with other particles or with waves. We shall discover that in some instances elementary mechanics is no longer adequate to describe the dynamics of microscopic particles and that this theory has to be supplemented by one that is more generally applicable, the so-called *quantum* or *wave* mechanics.

Classical mechanics is based on laws developed by Newton, for example

$$F = \mathrm{d}p/\mathrm{d}t \qquad (1.1)$$

Newton's laws, together with a classical electromagnetic theory as summarized by Maxwell's equations, proved adequate for the quantitative explanation of most experiments done before the beginning of the twentieth century. Equation (1.1) was found to be quite satisfactory for predicting the dynamics of large-scale systems. This is, of course, also true today. The term 'large-scale systems' in this context applies equally well to normal engineering laboratory experiments and to the more obviously large systems such as collections of planets.

A series of experiments conducted at the beginning of this century exposed a basic limitation of classical Newtonian mechanics, which is its inability to predict correctly events that take place on a microscopic or atomic scale. In the sections that follow we shall discuss the experiments that led to this failure in classical mechanics, while at the same time we will attempt to lay a general foundation for later discussions of a more general quantum-mechanical theory. It should be stressed that quantum mechanics does not entirely supplant Newtonian mechanics but rather augments it in that it is more widely applicable. We shall show that, within the limits of laboratory-sized objects, however, the newer mechanics reduces to the classical theory, which, because it is simpler to apply, is still to be preferred.

1.2 Black-body radiation

One of the earliest experiments to defy analysis by classical methods was the determination of the frequency spectrum of emitted radiation of an incandescent radiator or 'black body'. In this experiment the intensity of emitted radiation is measured as a function of frequency or wavelength for a fixed temperature, with typical results as indicated diagrammatically in Fig. 1.1.

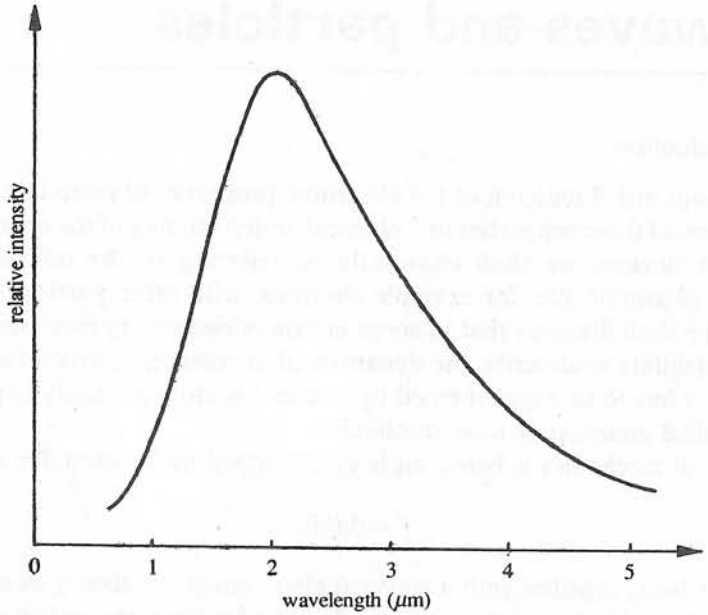


Fig. 1.1 Relative intensity of radiation emitted from a black body as a function of wavelength.

The various earlier theories attempting to explain this experimental evidence were based on classical mechanical ideas incorporated in thermodynamic theory. These theories were never successful in agreeing with the experiment, particularly in the short-wavelength limit. We know now that they broke down because of a fundamental misconception that atomic oscillators are capable of emitting or absorbing energy in continuously variable amounts. It was not until 1901 that Max Planck discredited this false notion by correctly predicting the intensity of radiation at all frequencies. His theory involved the hypothesis that energy could only be absorbed or emitted by the black body in discrete amounts. He assumed that the energy of light waves, for example, is transported in packets or bundles, called *photons* or *quanta*. He further assumed the energy of a photon to be given by

$$E = hf \text{ joules} \tag{1.2}$$

where f is the frequency of the radiation. Planck's constant, h, is a universal constant, which is found for the black-body radiator and other experiments to have a value $h = 6.626 \times 10^{-34}$ J s. Equation (1.2) is sometimes more conveniently written as

$$E = \hbar\omega \tag{1.3}$$

where ω is the angular frequency of the radiation, $2\pi f$, and $\hbar$ equals $h/2\pi$. Thus, the total energy of the black-body radiator was envisaged to exist only in discrete allowed energy states:

$$0, \quad \hbar\omega, \quad 2\hbar\omega, \quad 3\hbar\omega, \quad \ldots, \quad n\hbar\omega$$

transition between these states being brought about by absorption or emission of one or more photons of radiation, each of energy $\hbar\omega$.

It is not surprising that such a theory, being completely opposed to the existing continuously variable energy theories, was not readily accepted, even though it explained the experimental findings most satisfactorily. However, Planck's quantum hypothesis, which forms the basis of modern quantum mechanics, was further vindicated by later experimental evidence, as will be discussed in subsequent sections.

1.3 The photoelectric effect

If light of sufficiently short wavelength impinges on the surface of certain solids, then it is possible for electrons to be emitted from the solid. This is called the photoelectric effect. In the early twentieth century, Einstein reinforced Planck's photon concept of light by providing a satisfactory quantitative explanation of the effect.

The experimental evidence for the effect may be obtained using apparatus of the type shown diagrammatically in Fig. 1.2. Light of frequency f illuminates

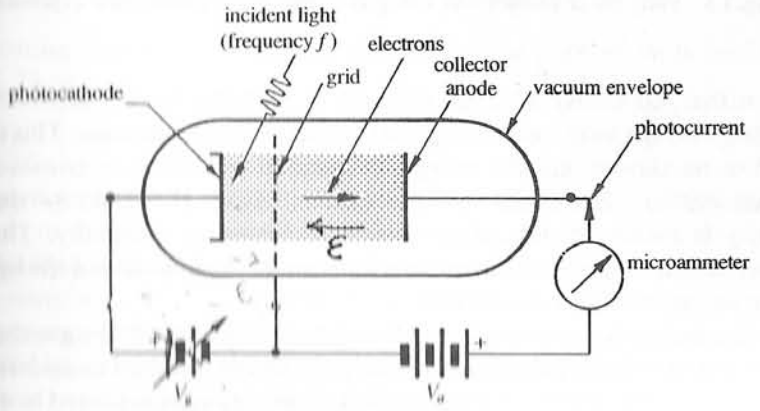


Fig. 1.2 A photoelectric experiment.

a cold cathode situated inside a vacuum envelope. If electrons are emitted, then, provided they have sufficient energy to overcome the retarding force field set up by the voltage V_g between grid and cathode, they will be swept to the positive collecting anode and a current will be registered on the microammeter in series with it.

The first thing that would be noticed when carrying out such an experiment is that, unless the frequency of the incident light is greater than some critical value, f_0, which is dependent on the material of the cathode, no emission is observed, *no matter how intense the light*. For constant light frequency, and provided f is greater than f_0, the photocurrent can be measured as a function of grid voltage V_g and light intensity, keeping the anode voltage constant, to give typical collector current data of the form shown in Fig. 1.3. The surprising

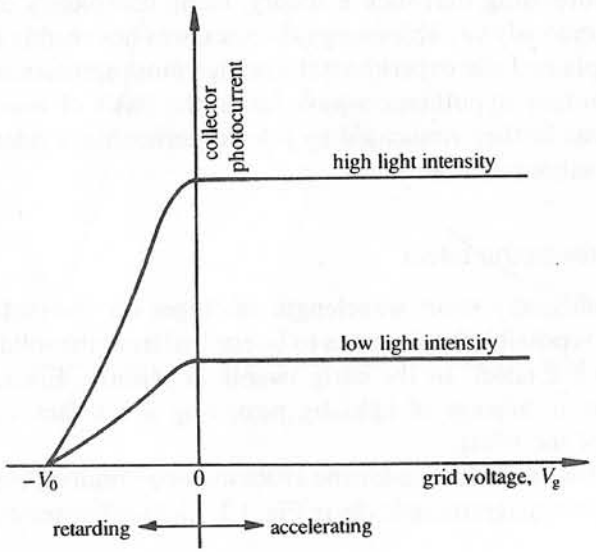


Fig. 1.3 Variation of photocurrent with grid voltage in the photoelectric experiment.

result is that, no matter what the intensity of the light, there is some constant retarding voltage, in this case $-V_0$, that entirely inhibits emission. This implies that the maximum kinetic energy of emitted electrons is constant and independent of the intensity of the incident light. However, as the light intensity is increased, the photocurrent increases in sympathy. Thus the number of emitted photoelectrons is a function of the intensity of the light but their maximum energy is constant.

Such experimental results cannot be explained by a classical wave theory of light and a satisfactory explanation can only be obtained by considering the light energy to be quantized. That is, the light energy is transported in discrete

Before discussing either theory we must digress a little to discuss briefly the reasons for the emission of electrons from a metal surface. We shall see later that a metal contains many highly mobile electrons, which can participate in the electrical conduction process, but these are confined to the interior of the metal by a binding energy. Thus, no conduction electrons can leave the surface of a metal unless they are in some way provided with additional energy to enable them to overcome this binding energy. The minimum energy required for an electron to be just emitted from a metal surface is called the *workfunction* of the particular metal and is usually designated $e\phi$, where ϕ is in volts.

In a classical theory of photoemission, conduction electrons in the cathode are accelerated by the electric field of the light wave and, if the light is bright enough, can gain sufficient energy to be emitted. Any surplus energy over and above the workfunction appears as kinetic energy of the emitted electron. Thus the brighter the light, the more energy is left over after overcoming the binding energy and the greater the kinetic energy of the emitted electron. This result is clearly at variance with the experimental evidence that the emitted electrons have a constant maximum energy. Further, a classical wave theory would not predict a threshold frequency f_0, which again is contrary to the experimental results.

If now we turn to a quantum theory based on Planck's photon hypothesis, we can give a simple explanation for the observed effects. We assume that the incident light is composed of discrete quanta or photons, each of energy hf. When the light impinges on the metal of the photocathode, each photon can transfer energy hf to a conduction electron. Some of the energy is used to overcome the binding forces and the remainder is converted to kinetic energy of the emitted electron. Thus

kinetic energy of emitted electron = photon energy − workfunction

or

$$\tfrac{1}{2}mv^2 = hf - e\phi \tag{1.4}$$

The limiting case occurs when an electron is just emitted with no kinetic energy. Then

$$f = f_0 = e\phi/h \tag{1.5}$$

At frequencies less than this critical value, the photon's energy, hf, is not even sufficient to overcome the workfunction and no emission occurs. Further, if the intensity is increased, the number of incident photons is increased but their energy remains constant, provided the frequency remains constant. Thus the kinetic energy of the emitted electrons, as given by Eq. (1.4), stays constant, which is again borne out by experiment. Incidentally, Eq. (1.4) gives the maximum kinetic energy at some particular frequency and in our experiment this is entirely converted to potential energy when the photocurrent is reduced

to zero by a retarding voltage $-V_0$. Hence

$$eV_0 = hf - e\phi \tag{1.6}$$

We see that V_0 does not vary with light intensity, a fact that we have already noted experimentally.

Thus, when considering the interaction of light with electrons, as in discussing the photoelectric effect, the quantum theory of light must be used in preference to the classical wave theory.

It will be useful to contrast this situation with the earlier experiments with light displaying such phenomena as diffraction, refraction, and so on. These effects could all be explained by a classical theory, which relied on a wave-like description of the light radiation. As an example, let us remind ourselves of the situation when light is diffracted by a mirror diffraction grating. A schematic diagram of the essential elements of the experiment is shown in Fig. 1.4. A light

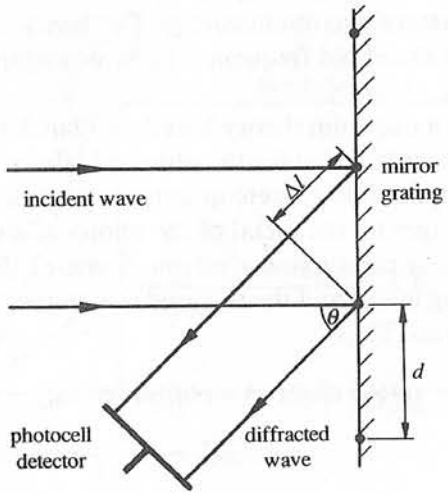


Fig. 1.4 Diffraction of light by a grating.

wave, wavelength λ, is incident normally on a reflecting diffraction grating of the type used in optical spectrographs, which has a grating spacing d. The light wave is diffracted by the grating and the diffracted wave is detected at some angle θ to the normal. The detector might, for example, be a photocell similar to the one just described. Experimentally, what is observed is that, as θ is varied, the intensity at the detector varies cyclically from maximum to minimum values. These are the well known diffraction fringes. A typical result might be as shown in Fig. 1.5. The appearance of the fringes can be explained quantitatively by invoking the wave description of light. The path length between light beams diffracted from adjacent rulings is indicated by Δl on

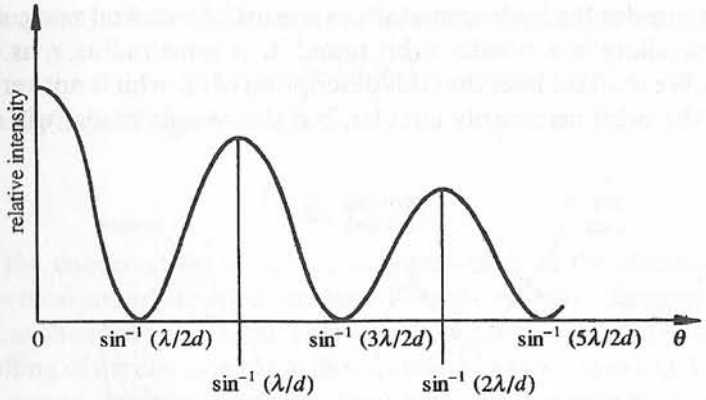


Fig. 1.5 Detected light intensity as a function of the angle of diffraction.

number of wavelengths, the diffracted waves interfere constructively and a bright fringe is observed. Thus, when

$$\Delta l = d \sin\theta = n\lambda$$

where n is an integer, there is constructive interference. This occurs at angles such that

$$\theta = \sin^{-1}(n\lambda/d) \tag{1.7}$$

Conversely, when the path difference Δl is equal to an odd number of half-wavelengths, which is equivalent to a 180° phase difference, there is destructive interference and a dark band is observed. This occurs at angles given by

$$\theta = \sin^{-1}\{[(2m+1)/2]\ \lambda/d\} \tag{1.8}$$

Both these results agree with the experimental evidence and a wave theory is entirely adequate to explain this phenomenon.

We see, then, that light can in some circumstances be considered to possess wave-like properties but on other occasions it must be treated in a quantized manner, its energy being transported by discrete photons, which are particle-like units each having energy hf. This wave–particle property of light radiation is sometimes referred to as the 'dual nature of light'.

If a light wave can behave as a particle, can a particle (say, for example, an electron) behave as a wave? The answer to this question will become apparent as we discuss further the historical development of quantum theory.

1.4 The Bohr atom

An excited hydrogen atom emits radiation at a discrete set of frequencies only. In 1913 Bohr produced a theoretical model that very accurately accounted for the observed sharp-line radiation spectrum from atomic hydrogen.

Let us consider the hydrogen atom to consist of a central nucleus with an electron travelling in a circular orbit round it, at some radius, r, as shown in Fig. 1.6(a). We shall see later that this description of an orbit is not very precise, neither is the orbit necessarily circular, but this simple model will suffice to

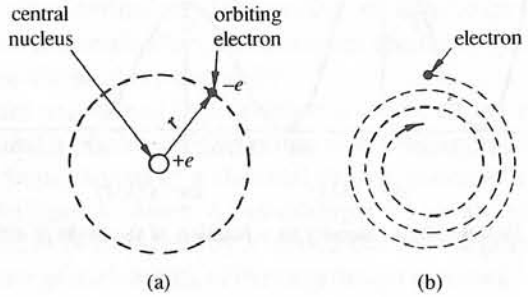


Fig. 1.6 (a) A possible model for a hydrogen atom and (b) spiral path due to radiated energy.

demonstrate the inadequacy of a classical theory. The Coulomb force on the electron due to the electric field of the positive nucleus is just sufficient to provide inward acceleration for circular motion at a constant radius, r.

It follows that

$$F=-\frac{e^2}{4\pi\epsilon_0 r^2}=-\frac{mv^2}{r} \tag{1.9}$$

Now, the total energy of the electron, E, is the sum of its potential energy, V, and its kinetic energy, T. Further

$$V=-e^2/(4\pi\epsilon_0 r)$$

and from Eq. (1.9)

$$T=\tfrac{1}{2}mv^2=e^2/(8\pi\epsilon_0 r)$$

Therefore, the total electron energy is

$$E=-e^2/(8\pi\epsilon_0 r) \tag{1.10}$$

Now, the electron in the circular orbit is constantly being accelerated and it can be shown by electromagnetic theory that such an accelerated charge radiates electromagnetic energy, with a corresponding loss of energy. Classical theory thus indicates that radiation can occur, its frequency corresponding to the periodic frequency of the circular motion. This frequency can be shown to be plausible by observing the electron's motion in the plane of the orbit. The electron will be seen to oscillate sinusoidally about a central position where the nucleus is located. The resultant sinusoidally varying current can then be likened to that occurring in an ordinary radio transmitting aerial, which radiates electromagnetic waves.

The frequency of the radiated wave from the classical atom is thus

$$f=v/(2\pi r)$$

which, using Eq. (1.9), gives

$$f=\frac{e}{4(\pi^3\epsilon_0 mr^2)^{1/2}} \tag{1.11}$$

Now, the conservation of energy indicates that, as the electron radiates energy, its total energy, E, must decrease. Thus the radius of the orbit must also decrease, as shown by Eq. (1.10). This would lead to a continual loss of energy and spiralling of the electron towards the nucleus, as shown in Fig. 1.6(b). This, in turn, would indicate that the frequency of the emitted radiation is continuously varying according to the dependence of f on the radius r as set out in Eq. (1.11). This is clearly in complete disagreement with the experimentally observed discrete frequency spectrum.

To overcome this difficulty, Bohr postulated that the electron could only exist in discrete energy levels, corresponding to certain allowed stable orbits, without radiating any energy. He further argued that radiation from the atom occurs only when the electron makes a transition from one allowed energy level to another, when the energy lost by the atom is converted into the energy of a single photon. Thus, if the electron is transferred from one stable orbit, corresponding to a total energy E_1, to another allowed orbit with lower energy E_2, a photon of radiation is emitted whose frequency, f_{12}, is given by

$$E_1-E_2=hf_{12} \tag{1.12}$$

Thus, since only a discrete set of energy levels is postulated, only a discrete set of characteristic frequencies is present in the output spectrum.

In order to calculate the value of the discrete allowed energy levels, Bohr was obliged to postulate, in a rather intuitive way, that the angular momentum, L, associated with a gyrating electron is quantized such that

$$L=mvr=n\hbar \qquad \text{where } n=1,2,3,\ldots \tag{1.13}$$

We can now eliminate v from Eqs (1.9) and (1.13) to obtain an expression for the radii of allowed orbits. From (1.9) and (1.13)

$$v^2=\frac{2e^2}{m8\pi\epsilon_0 r}=\frac{n^2\hbar^2}{m^2r^2}$$

or

$$r_n=\frac{4\pi n^2h^2\epsilon_0}{e^2m}=\frac{n^2h^2\epsilon_0}{\pi e^2m}\simeq 0.05n^2 \text{ nm} \tag{1.14}$$

Thus, the first possible orbit, when $n=1$, has a radius of about 0.05 nm ($1 \text{ nm}=10^{-9}$ m). Other possible orbits, corresponding to the various integer values of n, are 0.2 nm ($n=2$), 0.45 nm ($n=3$), and so on.

Each discrete orbit has a corresponding allowed energy level associated with it, which is evaluated by substituting the values of r_n in Eq. (1.10) to give

$$E_n = -\frac{e^2}{8\pi\epsilon_0}\frac{\pi e^2 m}{n^2 h^2 \epsilon_0} = -\frac{me^4}{8\epsilon_0^2 h^2 n^2} \simeq -\frac{13.6}{n^2}\text{eV} \tag{1.15}$$

Here the energy is expressed in electronvolts, a common practice in electronic engineering. One electronvolt corresponds to the energy acquired by an electron that has been accelerated through a potential difference of 1 V.

We see from Eq. (1.15) that the system energy is restricted to discrete levels corresponding to the various values of n. The lowest energy level or *ground state* for hydrogen is $E_1 = -13.6$ eV. Further allowed energy values are $E_2 = -3.4$ eV, $E_3 = -1.51$ eV, and so on. The allowed energy levels are usually represented in an energy level diagram, which for the hydrogen atom is shown in Fig. 1.7 (Note that in this representation the horizontal scale has no physical

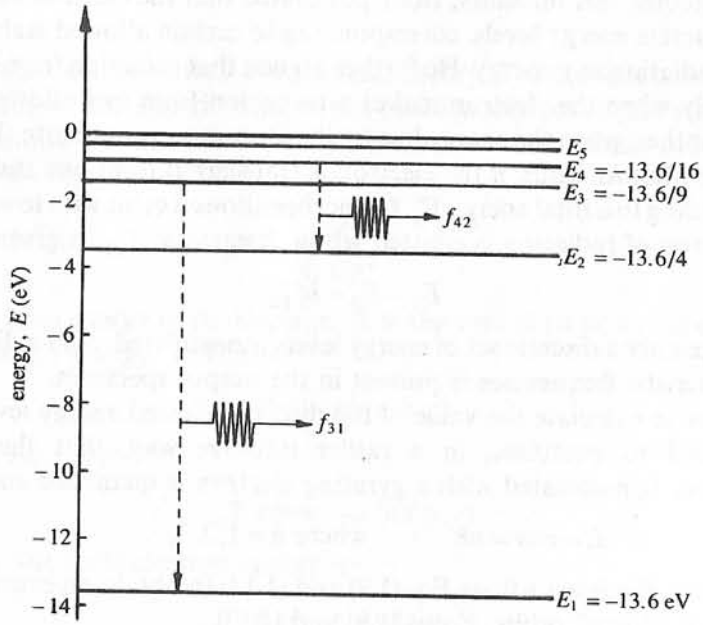


Fig. 1.7 Energy level diagram for the hydrogen atom.

significance.) As explained previously, it is possible for an electron to undergo a transition between any of the allowed energy levels as indicated by arrows on the diagram. The corresponding frequency of the emitted or absorbed radiation for such an occurrence is given by Eq. (1.12).

Although the characteristic emission frequencies predicted by the Bohr theory are very close to those observed for the hydrogen atom, the theory is only extendable to deal with hydrogen-like

more than one orbiting electron. A further limitation is the *ad hoc* manner in which the assumptions of quantized angular momentum and the relationship between energy change and frequency are introduced. We shall see later that there are further fundamental inadequacies in the theory. Meanwhile, it is well to remember that the Bohr theory, in a historical context, was a great step forward, in that it not only accounted for the predicted hydrogen frequency spectrum but, more important, it also clearly demonstrated a certain discreteness in some of the physical properties of matter, which is quite at variance with previous classical theory. A further advantage of the theory is that it indicates the importance of Planck's constant, h, for determining details of atomic structure.

1.5 Particle–wave duality

Let us return to our discussion of the possible dual nature of matter, in which waves can sometimes behave as particles and conversely particles can in some circumstances be considered to have wave-like properties.

We will first consider a photon of light, frequency f and energy hf. The photon travels at the velocity of light, c, which is related to frequency and wavelength, λ, by

$$c = f\lambda \tag{1.16}$$

Now, if energy E is transported with velocity c, the momentum of the photon is given by

$$p = E/c \tag{1.17}$$

This expression could be derived by finding the radiation pressure on a plate caused by an incident electromagnetic light wave and equating to the photon flux, but this is clearly too difficult, at this stage. Instead, we offer the following somewhat crude argument. Suppose that photons are subjected to an external force F, which acts over some distance $\mathrm{d}x$. The change in photon energy is then

$$\mathrm{d}E = F\,\mathrm{d}x$$

Also, if the photon momentum is p, then, by Newton's law,

$$F = \mathrm{d}p/\mathrm{d}t$$

and

$$\mathrm{d}E = \frac{\mathrm{d}p\,\mathrm{d}x}{\mathrm{d}t} = \frac{\mathrm{d}x}{\mathrm{d}t}\mathrm{d}p = c\,\mathrm{d}p$$

Integrating, we get

$$p = E/c$$

as before. Equations (1.16) and (1.17) can now be combined to give

$$p=\frac{E}{c}=\frac{E}{f\lambda}=\frac{hf}{f\lambda}$$

or

$$p=h/\lambda \tag{1.18}$$

Now, in 1924 de Broglie argued that if photons of light with wavelength λ have momentum $p=h/\lambda$, it might be possible for *particles* with momentum p to have some associated wavelength λ also and behave in a wave-like manner, under some circumstances. He further suggested that Eq. (1.18) might also be the correct relationship between p and λ for a particle.

Let us assume for the moment that this hypothesis, which as we shall see later can be substantiated by experimental evidence, is correct, and calculate the wavelength associated with various bodies. First, consider a classical Newtonian particle—an apple! If we let its mass be $m=0.2$ kg, and its velocity $v=10\,\mathrm{m\,s^{-1}}$ then its momentum

$$p=mv=0.2\times 10=2\,\mathrm{kg\,m\,s^{-1}}$$

and its associated wavelength is

$$\lambda=\frac{h}{p}=\frac{6.6\times 10^{-34}}{2}\approx 10^{-34}\,\mathrm{m}$$

Effects due to such a wavelength are much too small to be detected in ordinary laboratory experiments!

As a further example, consider an electron, mass m, charge $-e$, accelerated through a potential difference, V. Equating the gain in kinetic energy of the electron to its loss in potential energy,

$$\tfrac{1}{2}mv^2=eV$$

where v is its final velocity, gives

$$v=(2eV/m)^{1/2}$$

The momentum of the electron is then

$$p=mv=(2eVm)^{1/2}$$

and its associated wavelength is

$$\lambda=\frac{h}{p}=\frac{h}{(2eVm)^{1/2}}=\frac{1.225}{V^{1/2}}\,\mathrm{nm} \tag{1.19}$$

For example, if $V=50$ volts, $\lambda=1.7\times 10^{-7}$ mm, which again is small but, as we shall see, can produce measurable effects.

Soon after de Broglie suggested that particles could exhibit wave-like characteristics, Davisson and Germer provided experimental confirmation of

his ideas by diffracting electrons and producing interference between electron waves. Their experiment is similar to the grating experiment with light, which we have already discussed, a beam of electrons rather than light being diffracted by a 'grating'. We have seen that the wavelengths, λ, associated with electrons are small and, for a reasonable angular spread of the diffracted wave, the periodicity of the grating, d, should be of the same order as λ. The regular array of atoms within a single crystal of a metal satisfies this condition and can behave as a grating for electron diffraction. The apparatus for such an experiment is shown diagrammatically in Fig. 1.8(a). An electron gun accelerates a beam of electrons through a potential V and the beam then impinges

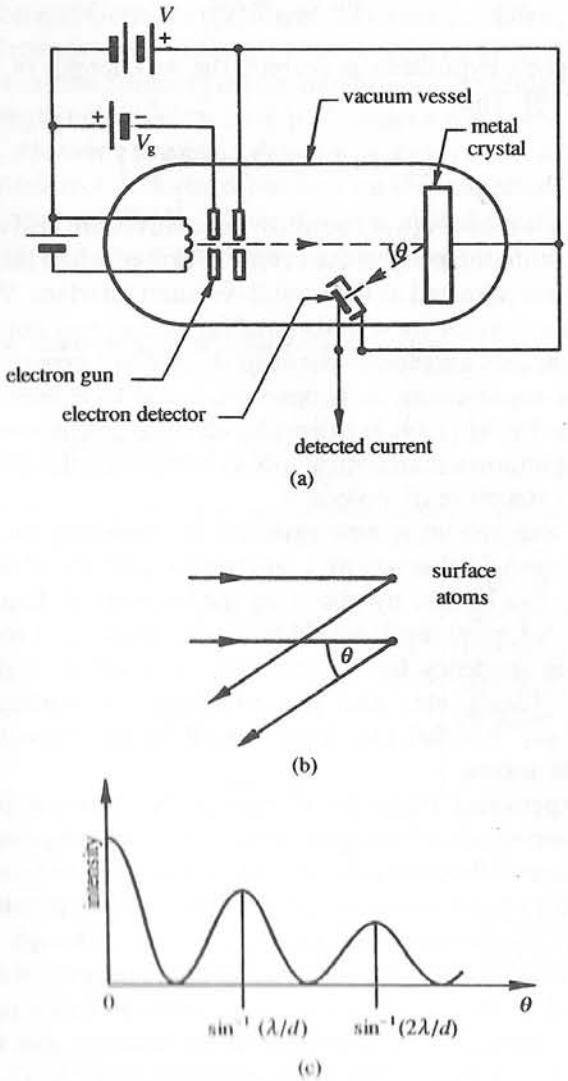


Fig. 1.8 An electron diffraction experiment.

normally on a plane single-crystal metal target. The detector is biased to provide a retarding field and only detects electrons that have been scattered with negligible loss of energy. It can be moved in angular direction, so as to measure the electron current diffracted by the target as a function of θ.

Let us now consider that electrons behave as waves, which are diffracted by the regular array of atoms on the surface of the target, spaced distance d apart, as illustrated in Fig. 1.8(b). By analogy with the mirror grating experiment, we would expect a maximum detected electron current at an angle θ such that the path difference between adjacent waves is an integral number of wavelengths and the first maximum detected signal to be given by

$$\sin \theta_{max} = \lambda/d$$

Now if de Broglie's hypothesis is correct, the wavelength of the electron is given by Eq. (1.19). Thus

$$\theta_{max} = \sin^{-1} [h/(2emVd^2)^{1/2}] \qquad (1.20)$$

Clearly, our simple model is not complete since waves are also diffracted from atomic planes within the body of the crystal. Further, when these waves leave the crystal they are refracted at the crystal–vacuum interface. When Eq. (1.20) is modified to account for these additional factors, as was done by Davisson and Germer, excellent agreement between theory and experiment is obtainable. In all their experiments, electrons were found to behave as waves with wavelength given by Eq. (1.19). Incidentally, electron diffraction apparatus has since become an important analytical tool to study such things as interatomic spacing and the structure of molecules.

Suppose the experiment is now extended by replacing the detector with a cathode-ray screen. In this way an average measurement of electron current over a long time can be take by observing the intensity of illumination at the screen versus θ. A typical result would be as shown in Fig. 1.8(c). We see that there is a strong tendency for electrons to come off at angles near $\theta = 0$, $\sin^{-1}(\lambda/d)$, $\sin^{-1}(2\lambda/d)$, etc., and the probability of finding electrons at $\theta \simeq \sin^{-1}(\lambda/2d)$, $\sin^{-1}(3\lambda/2d)$, etc., is very small. In this respect, the electrons behave like light waves.

A further experiment might be to replace the detector by an electron multiplier, connected via a high-gain amplifier to a loudspeaker. Each time a diffracted electron is detected a sharp 'click' is heard on the loudspeaker, each click being equally loud. Further, for a fixed value of θ, the number of electrons arriving when averaged over some time is constant, although the individual arrival times are erratic. Moving the detector alters the rate of clicks but their size, as measured by their loudness, remains constant. Lowering the cathode temperature or making the grid voltage in the electron gun more negative reduces the number of emitted electrons and hence alters the click rate but the loudness of each still stays the same. Thus, whatever is being detected arrives in

discrete amounts and in this respect the electrons are quantized and behave as particles.

We see, then, that the electrons are quantized but that they are in some way 'guided' by 'matter' waves (sometimes called probability waves, de Broglie waves, or ψ waves). Loosely speaking, wherever the matter waves have a large amplitude the probability of finding electrons is high, the converse being also true. Notice that we have had to abandon the absolute determinacy implicit in Newtonian mechanics since there is uncertainty as to where a particular electron will go after diffraction by the crystal lattice. We are forced to revert to discussing the probability of an electron being in a certain position at a given time.

The Davisson and Germer experiment, and a similar diffraction experiment carried out independently and almost concurrently by G. P. Thomson, clearly demonstrated the dual nature of matter and provide conclusive proof of the de Broglie wavelength relationship, $\lambda = h/p$. Whether a wave behaves as a particle or, conversely, whether a particle behaves as a wave is not only dependent on the type of experiment performed but also on the magnitude of the energies and momenta involved. Tables 1.1 and 1.2 are included to indicate under what conditions wave or particle properties become dominant.

Table 1.1 The electromagnetic spectrum of waves.

Frequency (Hz)	*Wavelength* (m)	*Typical wave type*	*Remarks*
	10^{-12}		
10^{20}		gamma rays X-rays	↑ photon energy ($\hbar\omega$) and momentum ($\hbar/\lambda$) increases—particle characteristics become important
	10^{-8}		
10^{16}		ultraviolet visible	
	10^{-4}		
10^{12}		infrared millimetre waves microwaves	wave characteristics dominant
	1		
10^{8}			
	10^{2}		
10^{4}		radio waves	↓

Table 1.2 Particle spectrum.

Mass (kg)	*Typical particle*	*Remarks*
	star	
10^{20}		↑ total energy increases—
	planet	particle behaviour
	car	predominates
	football	
	dust grain	
10^{-20}	molecule	
	atom	wavelength (h/p) increases
	electron	as p decreases—wave
10^{-40}		characteristics become
	neutrino	important ↓

1.6 Wavepackets: group and phase velocities of particles

It may be helpful to consider a geometric representation of how an object may simultaneously possess both wave and particle properties. We do this by studying the addition of waves of differing wavelength to produce a constructive interference pattern that has particle properties.

First, consider two waves of slightly different wavelength travelling in the same direction. These add together and produce regions of constructive interference that are periodically positioned in space, as in Fig. 1.9(a). This phenomenon is analogous to the beating of two sound waves to produce interference in the time domain. If, now, three waves, again slightly differing in frequency, are added, the interference maxima are not only larger but are spread at wider intervals, Fig. 1.9(b). The repetition in space of the regions of constructive interference is characteristic of the interference between finite numbers of waves. For an infinite number of waves, only one region of constructive interference exists; this is called a *wavepacket*, Fig. 1.9(c). The wavepacket geometrically represents an object with wave and particle properties. It obviously has wave properties since it is constructed from waves and has a wave-like form but it also behaves as a particle because of its localization in space. This is similar to the bow-wave of a ship, which has wave properties but travels with the ship and is always located relative to it. We have, so far, only considered waves travelling in the same direction but it is straightforward to extend the idea by describing a particle as consisting of an infinite number of waves travelling in all directions and constructively interfering at some point in space where the particle is positioned.

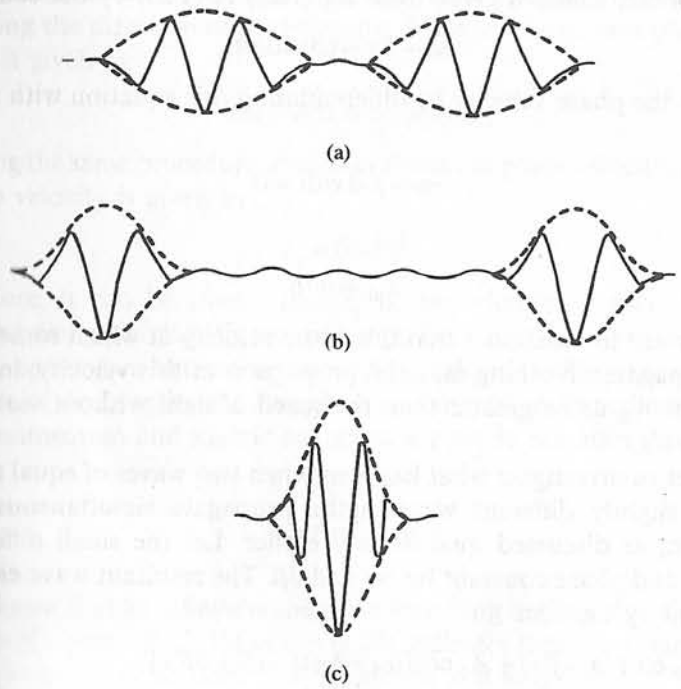


Fig. 1.9 Formation of a two-dimensional wavepacket.

What is the velocity of a wavepacket? In order to find this out, it will be necessary to digress slightly to summarize briefly the definitions and methods for calculating the various velocities associated with waves. A wave travelling in the positive x direction may be represented by an expression

$$A_0 \cos(\omega t - \beta x) \tag{1.21}$$

where A_0 is the amplitude of the wave, ω is its angular frequency and β is the phase constant, which is related to the wavelength by

$$\beta = 2\pi/\lambda \tag{1.22}$$

It is usually more convenient, mathematically, to represent the wave by an equivalent exponential function:

$$A_0 \,\mathrm{Re} \exp[\mathrm{j}(\omega t - \beta x)] \tag{1.23}$$

In most texts the Re is omitted but it is understood that only the real part of the result of any subsequent operation is valid.

The propagation of a wave is characterized by two velocities, the *phase velocity*, v_{ph} and the *group velocity*, v_g. Phase velocity is defined as the velocity of planes of constant phase along the propagation direction of the wave. To obtain an expression for this velocity we must examine the motion of a point of

constant phase, which is given from Eq. (1.21) or (1.23) by the condition

$$\omega t - \beta x = \text{constant}$$

We obtain the phase velocity by differentiating this equation with respect to time:

$$\omega - \beta \, dx/dt = 0$$

or

$$v_{ph} = \omega/\beta \tag{1.24}$$

It is important to remember that this is the velocity at which some arbitrary phase propagates. Nothing material propagates at this velocity; indeed, it is possible for v_{ph} to be greater than the speed of light without violating any physical laws.

Now, let us investigate what happens when two waves of equal amplitude but with slightly different wavelengths propagate simultaneously in the x direction, as discussed qualitatively earlier. Let the small differences of frequency and phase constant be $\delta\omega$ and $\delta\beta$. The resultant wave can then be represented by the sum

$$\begin{aligned} &A_0 \cos(\omega t - \beta x) + A_0 \cos[(\omega + \delta\omega)t - (\beta + \delta\beta)t] \\ &\quad = 2A_0 \cos\{\tfrac{1}{2}[(2\omega + \delta\omega)t - (2\beta + \delta\beta)t]\} \cos[\tfrac{1}{2}(\delta\omega t - \delta\beta x)] \\ &\quad \simeq 2A_0 \cos[\tfrac{1}{2}(\delta\omega t - \delta\beta x)] \cos(\omega t - \beta x) \end{aligned} \tag{1.25}$$

since $\delta\omega \ll 2\omega$ and $\delta\omega \ll 2\beta$ by our original assumption.

Thus, the resultant total wave consists of a high-frequency wave, varying as $\cos(\omega t - \beta x)$, whose amplitude varies at a much slower frequency rate represented by the other cosine term in (1.25); it is *modulated* by destructive and constructive interference effects, as shown in Fig. 1.10. The high-frequency wave has phase velocity ω/β as before. The variation of wave amplitude is called the envelope of the wavegroup; it varies sinusoidally with time and distance and is a travelling wave with the relatively long wavelength, $2\pi/\delta\beta$. Group velocity is defined as the velocity of propagation of a plane of constant

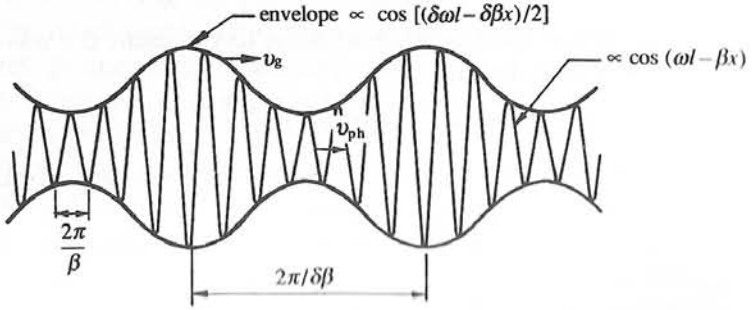


Fig. 1.10 Interaction of two waves of slightly different frequencies.

phase on the envelope. It corresponds to the velocity of the group or packet of waves along the direction of propagation. A plane of constant phase on the envelope is given by

$$\delta\omega t - \delta\beta x = \text{constant}$$

Thus, using the same procedure used to evaluate the phase velocity, we see that the group velocity is given by

$$v_g = \partial\omega/\partial\beta \tag{1.26}$$

Furthermore, it can be shown that v_g is the velocity at which energy is transmitted along the direction of propagation.

We can now return to the discussion of the velocities of particle waves. The hypotheses of de Broglie and Planck, corroborated by experiment, suggest that the momentum and kinetic energy of a particle are given by

$$\begin{aligned} p &= mv = h/\lambda \\ T &= \tfrac{1}{2}mv^2 = hf \end{aligned} \tag{1.27}$$

Now, we know that an infinite plane wave travelling in the x direction has the form $A_0 \exp[-j(\omega t - \beta x)]$. Equation (1.27) indicates that for a particle wave we might write the equivalent phase constant and frequency:

$$\begin{aligned} \beta &= \frac{2\pi}{\lambda} = \frac{2\pi p}{h} = \frac{p}{\hbar} \\ \omega &= \frac{2\pi T}{h} = \frac{T}{\hbar} \end{aligned} \tag{1.28}$$

This suggests that it might be possible to represent a particle by a function ψ, called a *wavefunction*, where

$$\psi = A_0 \exp[-j(Tt - px)/\hbar] \tag{1.29}$$

From (1.24) and (1.28) the phase velocity of such a wave is

$$v_{ph} = \frac{\omega}{\beta} = \frac{T}{p} = \frac{\tfrac{1}{2}mv^2}{mv} = \frac{v}{2}$$

Note that this result is not valid for a single particle, however, since we have seen that this must be represented by a discrete wavepacket and the concept of phase velocity is applicable only to infinite wavetrains.

A valid group velocity for the particle can, however, be found. From Eqs (1.27) and (1.28)

$$\partial\omega = (m/\hbar)\partial v \qquad \text{and} \qquad \partial\beta = (m/\hbar)\partial v$$

which, using Eq. (1.26), gives

$$v_g = \partial\omega/\partial\beta = v \tag{1.30}$$

Thus, a single electron or bunch of electrons can be represented by a wavepacket travelling with the same velocity as the electron. This seems physically reasonable when we remember the alternative definition of group velocity as being the rate at which energy is transported by waves.

Our discussion so far on wave–particle duality is summarized in Fig. 1.11, where the behaviour of light waves and electrons is compared diagrammatically.

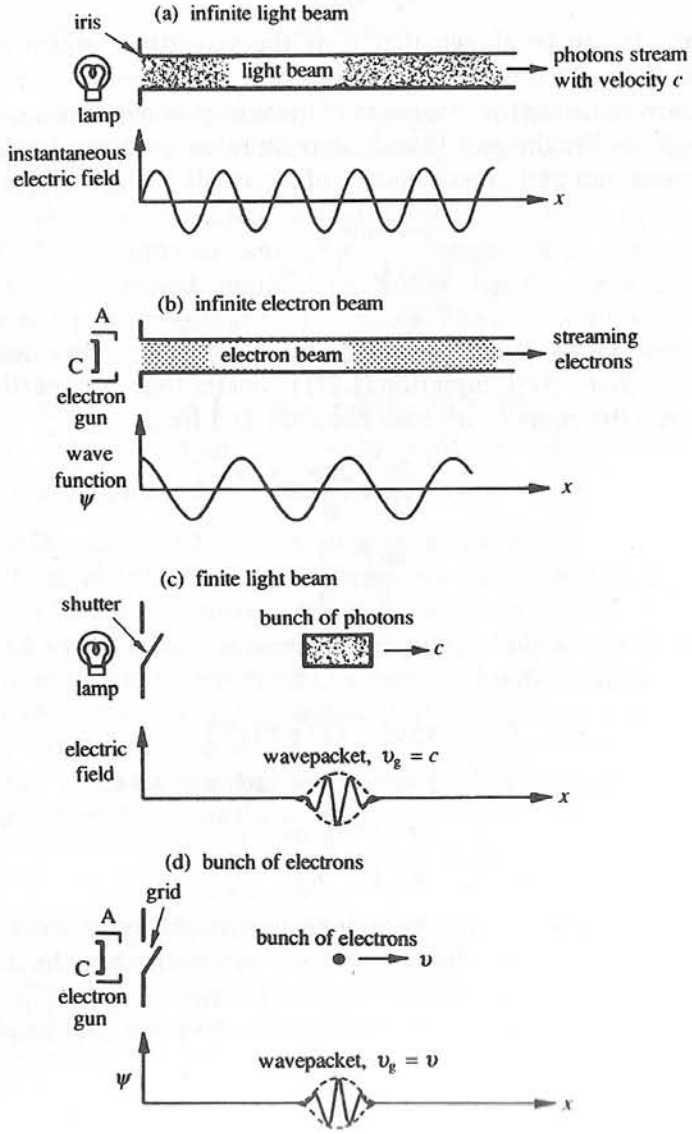


Fig. 1.11 Particle–wave duality.

1.7 The Schrödinger wave equation

So far, we have discussed experiments for which there exists no explanation based on classical concepts. However, if a series of assumptions is made, for instance the de Broglie hypothesis that $p = h/\lambda$, a quantitative description of the mechanism of each phenomenon can be provided. What is now required is some sort of unifying theory that will enable us to predict and explain other occurrences on the atomic scale. In 1926 Schrödinger provided a basis for such a theory by discovering an equation for predicting ψ, the wavefunction of a particle, in any particular circumstance. The equation he discovered is named after him; it replaces Newton's laws when atomic-sized particles are being considered. The theory based on Schrödinger's equation is called *wave* or *quantum mechanics*.

We have seen that particles can possess wave-like properties and that the probability waves associated with a beam of particles can be described in terms of a wavefunction, ψ, given by Eq. (1.29). This expression includes the kinetic energy, T, associated with a particle. In general, however, a particle can also possess potential energy. For instance, it might be an electron moving in a solid; the electron not only has kinetic energy but it also moves in the field due to the lattice and thus has a space-dependent potential energy. In general, the total energy, E, of a particle is therefore

$$E = \hbar\omega = T + V \tag{1.31}$$

where V denotes the potential energy. The one-dimensional wavefunction in the more general case then becomes

$$\psi = A_0 \exp[-\mathrm{j}(Et - px)/\hbar] \tag{1.32}$$

What equation does this generalized wavefunction satisfy? We would expect it to be some differential equation comparable to a one-dimensional wave equation, say

$$\frac{\partial^2 H}{\partial x^2} = \epsilon\mu \frac{\partial^2 H}{\partial t^2} \tag{1.33}$$

the solution of which

$$H = H_0 \exp[-\mathrm{j}(\omega t - \beta x)] \tag{1.34}$$

gives in this instance the magnetic field of a plane wave propagating in a medium with permittivity ϵ and permeability μ.

Let us try to find a wave equation similar to (1.33) but which has ψ given by Eq. (1.32) as its solution. First, differentiate (1.32) with respect to t

$$\frac{\partial \psi}{\partial t} = -\frac{\mathrm{j}}{\hbar} E\psi = -\frac{\mathrm{j}}{\hbar}(V + \tfrac{1}{2}mv^2)\psi \tag{1.35}$$

Also, let us differentiate ψ with respect to x, twice

$$\frac{\partial^2\psi}{\partial x^2}=-\frac{p^2}{\hbar^2}\psi=-\frac{m^2v^2}{\hbar^2}\psi \tag{1.36}$$

We may now compare the equations to give, from (1.35)

$$-\tfrac{1}{2}mv^2\psi=-j\hbar\frac{\partial\psi}{\partial t}+V\psi$$

and from (1.36)

$$-\tfrac{1}{2}mv^2\psi=\frac{1}{2m}\hbar^2\frac{\partial^2\psi}{\partial x^2}$$

Rearranging the equations gives

$$\frac{\partial^2\psi}{\partial x^2}-\frac{2m}{\hbar^2}V\psi+j\frac{2m}{\hbar}\frac{\partial\psi}{\partial t}=0 \tag{1.37}$$

This equation is called the one-dimensional time-dependent Schrödinger wave equation. It governs the behaviour in one dimension of all particles. Notice that we have not derived this equation rigorously, since for example the relationships $p=h/\lambda$ and $E=hf$ have been assumed in writing the wavefunctions, which was our starting point. The above steps are only an argument to demonstrate the plausibility of Schrödinger's equation. In fact, there is no proof of the equation. This situation is directly comparable to the lack of proof for Newton's laws. Agreement with experiment is the only check as to its validity; it has been found to be correct when applied to a wide number of circumstances concerning microscopic particles, particularly in its relativistic form. A further test is that, in the classical limit of laboratory-sized experiments, Schrödinger's equation must provide results that agree with those derivable from Newton's laws. It can be shown that this is so and that Eq. (1.37) is quite general and reduces to Newton's laws of motion for large-sized objects. Schrödinger's equation must always be used in preference to Newton's laws, however, when considering the interaction of atomic-sized particles.

If three-dimensional motion is allowed, the wavefunction becomes a function of three space coordinates and time, $\psi(x,y,z,t)$, and is a solution of the three-dimensional time-dependent Schrödinger equation

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}-\frac{2m}{\hbar^2}V\psi+j\frac{2m}{\hbar}\frac{\partial\psi}{\partial t}=0 \tag{1.38}$$

It should be noted that, although this equation or the simple one-dimensional version will be used predominantly to predict the wavefunction of electrons, it is applicable to *any* particle, provided the appropriate mass and potential

For a large class of problems in which the total particle energy is constant, the Schrödinger equation can be simplified by separating out the time- and position-dependent parts. This can be achieved by the standard separation-of-variables procedure. Accordingly we consider the one-dimensional equation for simplicity and assume solutions of the form

$$\psi=\Psi(x)\Gamma(t) \tag{1.39}$$

where Ψ and Γ are respectively functions of position and time only.

We substitute this in Eq. (1.37) to obtain

$$\frac{\hbar^2}{2m}\frac{1}{\Psi}\frac{d^2\Psi}{dx^2}-V=-j\frac{\hbar}{\Gamma}\frac{d\Gamma}{dt} \tag{1.40}$$

Now the left-hand side is a function of space coordinates only, provided V is time-independent, and the right-hand expression is a function of time only. Thus each equation must independently equal some constant, say C. Therefore

$$\frac{d\Gamma}{dt}=\frac{jC}{\hbar}\Gamma$$

or

$$\Gamma(t)=\exp(jCt/\hbar)$$

Comparison of this expression with the time-dependent part of the one-dimensional wavefunction, given in (1.32), indicates that the constant in the expression is equal to $-E$. Thus, if the energy is constant, the time-dependent part of the wavefunction is

$$\Gamma(t)=\exp(-jEt/\hbar) \tag{1.41}$$

and

$$\psi=\Psi(x)\exp(-jEt/\hbar) \tag{1.42}$$

The left-hand side of Eq. (1.40) is then equal to $-E$, to give

$$\frac{d^2\Psi}{dx^2}+\frac{2m}{\hbar^2}(E-V)\Psi=0 \tag{1.43}$$

This is the one-dimensional time-independent Schrödinger equation. It may be used to find the space-dependent part of the wavefunction whenever the system energy is constant, for example, when considering bound particles or a constant current of particles. In all other situations, for example, when discussing a varying current of particles, the more general time-dependent version of the equation must be employed. For most problems dealing with particles of constant energy, it is usually sufficient to solve the three-dimen-

position $\Psi(x,y,z)$, but if the complete wavefunction $\psi(x,y,z,t)$ is required then $\Psi(x, y, z)$ must be multiplied by $\Gamma(t)$, as indicated by Eq. (1.42).

1.8 Interpretation of the wavefunction ψ

We have shown that it is possible to describe the wave characteristics of a particle in terms of a wavefunction ψ, but we have not yet discussed precisely what property of the particle is behaving in a wave-like manner. There is little difficulty in this respect with other wave types; for example, it is the electric and magnetic field vectors that are oscillating in a radio wave and for sound waves the variable parameter is pressure. However, the physical significance of the wavefunction is not so readily apparent. Since $\psi(x,y,z,t)$ is a function of space and time coordinates, we might expect that it represents the position of a particle at some time t. However, we will see later that in general it is impossible to locate a particle exactly in space without there being any uncertainty as to its position. We can only consider the probability of a particle being at a particular point in space. A further complication is that, since ψ is a solution of Schrödinger's equation, it is usually a complex quantity.

Max Born, in 1926, overcame the difficulty of being unable to attach physical significance to ψ itself by showing that the square of its absolute magnitude, $|\psi|^2$, is proportional to the probability of a particle being in unit volume of space, centred at the point where ψ is evaluated, at time t. Thus, although the exact position of a particle at a particular time cannot be predicted, it is possible to find its most probable location. It follows that $|\psi|^2 \Delta V$ is proportional to the probability that a particle will be found in the volume element ΔV. For example, the probability of finding a particle in the range $x \to x + \mathrm{d}x$, $y \to y + \mathrm{d}y$ and $z \to z + \mathrm{d}z$ is proportional to

$$|\psi(x,y,z)|^2 \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \psi\psi^* \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \tag{1.44}$$

where ψ^* is the complex conjugate of the wavefunction. Although the direct physical significance of particle waves is not clear, if we solve Schrödinger's equation in particular circumstances and obtain a wavefunction ψ, the *probability density*, $|\psi|^2$, can be used to predict accurately what the spatial distribution of particles will be at some time, t, provided that a sufficiently large number of experiments has been performed.

If a particle exists at all, it is certainly located somewhere in space, and since the probability of its being in an elemental volume is proportional to $|\psi|^2 \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$, it is convenient to choose the constant of proportionality such that the integral of the probability density over all space equals unity or

$$\iiint_{-\infty}^{+\infty} \psi\psi^* \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = 1 \tag{1.45}$$

A wavefunction that satisfies this condition is said to be *normalized*. Whenever wavefunctions are normalized, $|\psi|^2 \Delta V$ *equals* the probability that a particle

Before solving Schrödinger's equation in any particular set of circumstances, it is first necessary to know what boundary conditions must be set on ψ. Since answers to physical problems are obtainable from wavefunctions, they must be well behaved in a mathematical sense. First, ψ must be a continuous and single-valued function of position. Supposing for the moment that this were not so, $\psi\psi^*$ would be discontinuous also, as indicated for the one-dimensional case shown in Fig. 1.12(a). This would imply that the

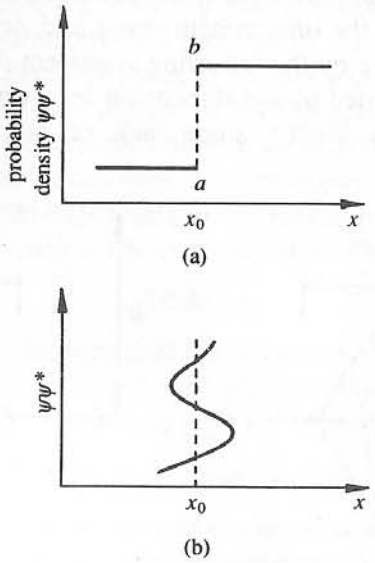


Fig. 1.12 (a) Discontinuous and (b) multivalued probability densities, which are not allowed.

probability of finding a particle is dependent on the direction from which the discontinuity is approached. Coming to x_0 from the left the probability density is a and from the right it is b, which in turn indicates that particles are instantaneously created or destroyed at x_0, which is clearly not allowable. Hence ψ must be continuous.

Consider now the possibility of ψ and hence $\psi\psi^*$ being multivalued, as shown in Fig. 1.12(b). This would imply that, at some position, x_0, there are several probabilities of finding a particle, which is obviously not physically possible. Hence ψ must be single-valued. By similar reasoning, it is possible to show that the spatial derivatives of ψ, $\partial\psi/\partial x$, $\partial\psi/\partial y$ and $\partial\psi/\partial z$ must be continuous and single-valued across any boundary.

1.9 The uncertainty principle

In 1927, Heisenberg published the uncertainty principle, which was subsequently named after him. He showed that if an attempt is made to measure

precision with which the two variables can be specified simultaneously becomes apparent. Such pairs of variables are momentum and position, and energy and time. Suppose, for example, in a particular experiment, that the energy of a particle can be measured to some accuracy ΔE and the time at which the measurement is taken is known to some accuracy Δt; then a classical theory would indicate that the precision to which these parameters can be measured is limited only by the experimental apparatus and technique. The uncertainty principle shows, however, that if the particle's energy is determined very accurately, so that ΔE is small, there is a proportional increase in the lack of precision in the time measurement and Δt increases. This can be made to sound plausible by the following argument. We know that a single particle can be represented in one dimension by a wavepacket, as shown in Fig. 1.13(a). There is evidently some lack of precision in locating the

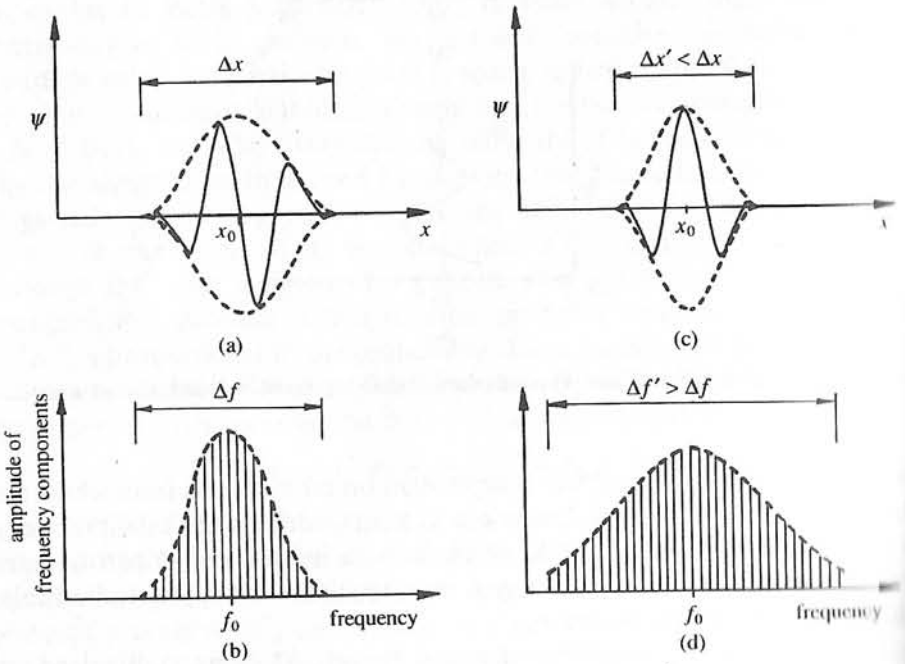


Fig. 1.13 Frequency spectra of wavepackets.

wavepacket and hence the particle, since it is spread over some distance Δx in space. Since the particle can be assumed to be travelling at some velocity v, the time of arrival of the wavepacket at some particular location can only be measured to an accuracy Δt, related to Δx by

$$\Delta x = v\Delta t$$

many waves, each of different frequency. The amplitudes of the various component waves can be obtained by Fourier analysis to give the frequency spectrum of the wavepacket as shown in Fig. 1.13(b). This indicates that the spread in frequency of the component waves is of order Δf, and since $E = hf$, or $\Delta E = h\Delta f$, there is necessarily a corresponding uncertainty in the particle energy.

If, now, the accuracy in the location of x is increased somewhat, Δx and hence Δt decrease and the wavepacket is shorter, as shown in Fig. 1.13(c). This shortening can be achieved only by the addition of further frequency components, with a corresponding increase in the width of the frequency spectrum, as in Fig. 1.13(d). This leads, as we have seen, to an increased ΔE and a greater uncertainty as to the precise value of the energy of the particle. Thus, as Δt is decreased, ΔE is increased and vice versa. If the frequency spectra are obtained in a more quantitative way, it can be shown that the best that can be done, no matter what the experimental apparatus, is that the product of the uncertainties to which each is known is equal to Planck's constant, or

$$\Delta E\, \Delta t \geqslant h \tag{1.46}$$

Notice that there is no fundamental restriction on the accuracy with which either quantity can be determined individually, only on their product.

Now, the energy of a free particle is $E = hf = \frac{1}{2}mv^2$, from which we obtain

$$\Delta E = h\Delta f = mv\Delta v = v\Delta p$$

Thus, the momentum spectrum for a wavepacket shows similar characteristics to the frequency spectrum; for a short wavepacket, Δx is small but the width of the momentum spectrum Δp is large, and if Δx is made bigger there is a corresponding decrease in Δp. Again, it can be shown that the product of the accuracies for a simultaneous measurement of momentum and position must be greater than $h/2$, or

$$\Delta p\, \Delta x \geqslant h/2 \tag{1.47}$$

We might, for example, try to locate the exact position of a particle, thus specifying $\Delta x = 0$, but this would be possible only if all knowledge of its momentum were sacrificed since to satisfy Eq. (1.47) Δp must become infinite.

Fortunately, the uncertainties implicit in Heisenberg's principle need not be too restricting for normal laboratory-scale experiments since h is very small, of order 10^{-34} J s. However, the limitations to accuracy, as given in expressions (1.46) and (1.47), become critical for atomic-sized particles, when the magnitudes of the experimental variables can become minute.

1.10 Beams of particles and potential barriers

We shall now consider the interaction of beams of particles with potential

right but also serve as an introduction to discussions of the solution of Schrödinger's equation for particles confined in space, which begin in the next chapter. Interactions in one dimension only will be considered, mostly for mathematical convenience, but the solutions obtained are directly applicable to the motion of particles in devices with large dimensions transverse to the current flow.

First, consider a beam of particles travelling in the x direction with energy E, impinging on a potential barrier at $x=0$, of height V_2-V_1, such that $V_1<E<V_2$, as shown in Fig. 1.14.

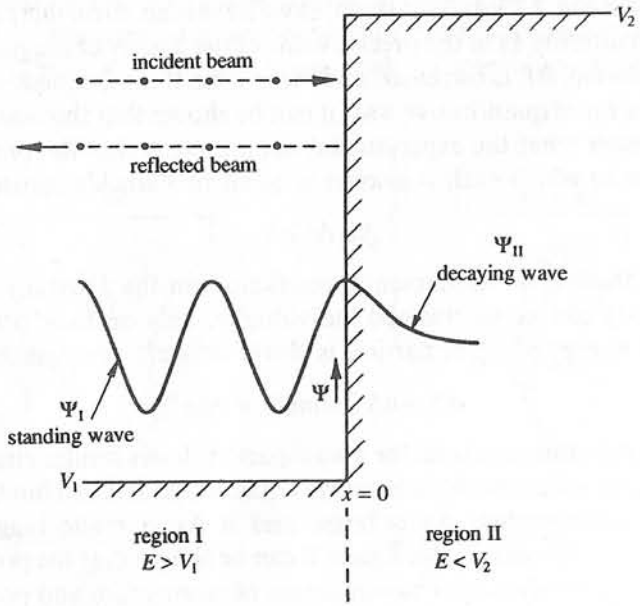


Fig. 1.14 An electron beam incident on a potential barrier.

Classically, we should expect particles in region I, where $E>V_1$, and none in region II, where $E<V_2$. We have deliberately chosen to investigate the motion of a beam of particles since, as we have seen, this can be represented by the simple wavefunction.

$$\psi = A_0 \exp[-j(Et-px)/\hbar]$$

Strictly speaking, each particle in the beam should be represented by a wavepacket. However, for a small spread in phase constant, $\Delta\beta$, the reflection of a packet is similar to the reflection of an infinite wave.

Let us apply Schrödinger's stationary-state equation to regions I and II, since E is constant. In region I

$$\frac{d^2\Psi}{dx^2}+\frac{2m}{\hbar^2}(E-V_1)\Psi=0 \qquad (1.48)$$

which, if we let

$$\frac{2m}{\hbar^2}(E-V_1)=\beta^2 \qquad (1.49)$$

becomes

$$\frac{d^2\Psi}{dx^2}+\beta^2\Psi=0 \qquad (1.50)$$

Solutions of this equation are of the form

$$\Psi_1 = A\exp(j\beta x)+B\exp(-j\beta x) \qquad (1.51)$$

The complete wavefunction for the region, including time dependence, is obtained by multiplying Ψ_1 by $\exp(-jEt/\hbar)$, giving

$$\psi_1 = A\exp\left[-j\left(\frac{Et}{\hbar}-\beta x\right)\right]+B\exp\left[-j\left(\frac{Et}{\hbar}+\beta x\right)\right] \qquad (1.52)$$

The first term represents an incident probability wave travelling in the positive x direction and the second term represents the wave reflected by the barrier, travelling in the negative x direction.

In region II, $(E-V_2)$ is negative and we let

$$\frac{2m}{\hbar^2}(E-V_2)=-\alpha^2 \qquad (1.53)$$

so that, in this region, the stationary-state equation becomes

$$\frac{d^2\Psi}{dx^2}-\alpha^2\Psi=0 \qquad (1.54)$$

This has a general solution

$$\Psi_{II}=Ce^{-\alpha x}+De^{\alpha x} \qquad (1.55)$$

Arguing on physical grounds, we would not expect Ψ_{II} to become infinite at large x, so D evidently must be zero, which gives

$$\Psi_{II}=Ce^{-\alpha x} \qquad (1.56)$$

We now find relationships between the magnitudes A,B and C by applying the boundary conditions on ψ at the barrier, $x=0$. That ψ and hence Ψ are continuous at the boundary gives

$$\Psi_1|_{x=0}=\Psi_{II}|_{x=0}$$

which, from Eqs (1.51) and (1.56), gives

Also, for continuity of the derivative of Ψ,

$$(\partial\Psi_{\mathrm{I}}/\partial x)|_{x=0}=(\partial\Psi_{\mathrm{II}}/\partial x)|_{x=0}$$

or

$$\mathrm{j}\beta A-\mathrm{j}\beta B=-\alpha C \tag{1.58}$$

Equations (1.57) and (1.58) give

$$A=\tfrac{1}{2}C(1-\alpha/\mathrm{j}\beta)$$
$$B=\tfrac{1}{2}C(1+\alpha/\mathrm{j}\beta) \tag{1.59}$$

Notice that the amplitude of incident and reflected waves are identical, i.e.

$$|A|=|B|$$

but, because there is a phase difference between the forward- and backward-travelling waves, the total Ψ wave or sum of incident and reflected waves is a standing wave. A further point of interest is that the solution Ψ_{II} given in Eq. (1.56), which applies to the right of the barrier, represents a wave whose amplitude decays exponentially with increasing x. Both of these situations are illustrated diagrammatically in Fig. 1.14. It will be seen that there is a finite probability that the incident beam penetrates some distance into the classically forbidden region, since $|\Psi_{\mathrm{II}}|^2$ is greater than zero there. However, if α is large, i.e. $V_2 \gg E$, few particles are found very far inside the boundary.

Now let us turn our attention to the situation where the barrier is not sufficiently high to cause complete reflection of the incident beam. In these circumstances, there is only partial reflection and part of the beam is transmitted through the barrier, as shown in Fig. 1.15. If, as before, we let

$$\beta_{1,2}=\frac{2m}{\hbar^2}(E-V_{1,2}) \tag{1.60}$$

then the solutions of Schrödinger's equation in the two regions are

$$\Psi_{\mathrm{I}}=A\exp(\mathrm{j}\beta_1 x)+B\exp(-\mathrm{j}\beta_1 x)$$
$$\Psi_{\mathrm{II}}=C\exp(\mathrm{j}\beta_2 x) \tag{1.61}$$

if it is assumed that there is no reflected wave in region II. Matching Ψ and $\partial\Psi/\partial x$ at the boundary as before we have, from (1.61)

$$A+B=C$$
$$\beta_1(A-B)=\beta_2 C \tag{1.62}$$

C can be eliminated from these equations to give

$$\frac{B}{A}=\frac{1-\beta_2/\beta_1}{1+\beta_2/\beta_1} \tag{1.63}$$

which is used to find a reflection coefficient:

$$\frac{\text{density of particles reflected}}{\text{density of particles incident}}=\frac{|\Psi_{\mathrm{ref}}|^2}{|\Psi_{\mathrm{inc}}|^2}=\frac{BB^*}{AA^*}=\frac{B^2}{A^2}$$
$$=\left(\frac{1-[(E-V_2)/(E-V_1)]^{1/2}}{1+[(E-V_2)/(E-V_1)]^{1/2}}\right)^2 \tag{1.64}$$

where we have substituted for the β's from Eq. (1.60).

Eliminating B from (1.62) gives the relative amplitude of the transmitted wave:

$$\frac{C}{A}=\frac{2}{1+\beta_2/\beta_1}=\frac{2}{1+[(E-V_2)/(E-V_1)]^{1/2}} \tag{1.65}$$

We thus find the surprising result that the amplitude of the transmitted wave is bigger than that of the incident wave, or the probable density of particles in the transmitted electron beam is greater than that in the incident beam! This may be explained by noting that the transmitted particles are moving more slowly than the incident particles. Thus, although $C>A$, the *rate* of flow of particles in the transmitted beam is less than that in the incident beam. Notice also that

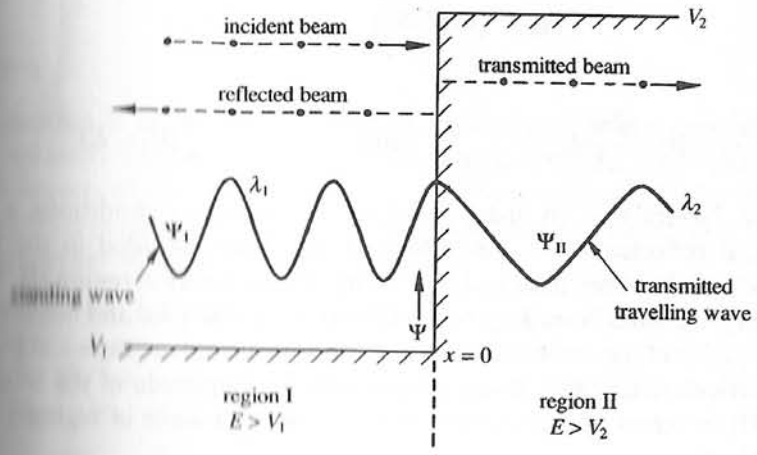


Fig. 1.15 An electron beam incident on a potential step.

Since $\beta_2<\beta_1$, the wavelength of the transmitted Ψ wave is greater than that for the incident wave, as shown in Fig. 1.15.

The final example in this series is the interaction of a constant-energy beam of particles incident on a classically impenetrable potential barrier of finite width d, as shown in Fig. 1.16. Applying Schrödinger's equation to the three

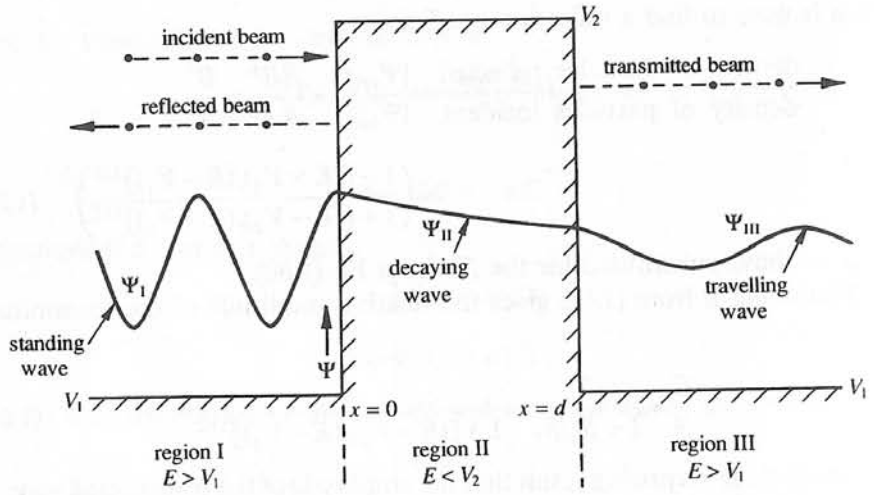


Fig. 1.16 Partial transmission of an electron beam through a narrow potential barrier.

regions gives wavefunctions of the form:

$$\Psi_I = A\exp(j\beta x) + B\exp(-j\beta x)$$
$$\Psi_{II} = C\exp(-\alpha x) + D\exp(\alpha x) \qquad (1.66)$$
$$\Psi_{III} = F\exp(j\beta x)$$

where

$$\beta^2 = \frac{2m}{\hbar^2}(E - V_1) \qquad \text{and} \qquad \alpha^2 = \frac{2m}{\hbar^2}(V_2 - E) \qquad (1.67)$$

as before. Notice that, in order to satisfy the boundary conditions, a highly attenuated reflected wave, magnitude D, has been included in the barrier region; also, it has been assumed that no reflection occurs in region III. We can now apply the usual boundary conditions at $x = 0$ and $x = d$ and hence find the relative values of the wavefunctions at various positions, as shown in Fig. 1.16. If, in particular, B,C and D are eliminated, the amplitude of the Ψ wave in region III in terms of the amplitude of the incident wave in region I can be obtained, thus

$$F = A\exp(-j\beta d)[\cosh(\alpha d) + \tfrac{1}{2}(\alpha/\beta - \beta/\alpha)\sinh(\alpha d)]^{-1} \qquad (1.68)$$

Now, the probability of a particle passing through the barrier, P_{I-III} is proportional to the ratio of the absolute square of ψ in region III to the absolute square of ψ in the incident region, or

$$P_{I-III} = |F|^2/|A|^2 \qquad (1.69)$$

which can be evaluated using Eq. (1.68).

In most practical cases, αd is large because of high attenuation of the

Ψ wave in the barrier region, and the hyperbolic functions in (1.68) can be simplified to give

$$P_{I-III} = \frac{|F|^2}{|A|^2} \simeq \frac{\exp(-2\alpha d)}{1 + \frac{1}{4}(\alpha/\beta - \beta/\alpha)^2}$$

We can now substitute values for α and β as defined in Eq. (1.67) to give

$$P_{I-III} \approx \exp\{-2[2m(V_2 - E)]^{1/2}\, d/\hbar\} \qquad (1.70)$$

Thus, if the barrier is sufficiently thin, i.e. d is small, there is a small but finite probability that particles can penetrate the barrier, even though classically this would not seem possible. The particle is said to have *tunnelled* through the barrier. The effect has important physical applications, as will be discovered later when, for example, tunnel diodes and electron emission from a cathode are discussed.

Notice that the probability of tunnelling falls off exponentially with increasing barrier height and thickness. As an example of the sort of figures involved, consider a 1 A electron beam approaching a barrier 1 V high ($V_2 - E = 1.6 \times 10^{-19}$ J) and 2 nm wide. The probability of tunnelling, using (1.70), is then of order e^{-20} and a current of about 10^{-9} A tunnels through the barrier. However, if the barrier thickness is reduced to 0.1 nm, the tunnelling current is increased to about 0.3 A!

Problems

1. An electron in an atom drops from an excited state to the ground state in a time normally lasting about 10^{-8} s. If the energy emitted is 1 eV, find (a) the relative uncertainty in the energy, (b) the relative uncertainty in the frequency of the emitted radiation and (c) the number of wavelengths in and the length of the wavepacket of emitted radiation.

Ans. 4×10^{-7}, 4×10^{-7}, 2×10^{6}, 3 m

2. What is the inherent uncertainty in the velocity of an electron confined in a crystalline solid of volume 10 mm cube?

Ans. 0.07 m s^{-1}

3. Show that the probability of reflection of a beam of electrons of energy E incident on an abrupt potential energy step of height V_0 is determined by a reflection coefficient r given by

$$r = \left(\frac{1 - (1 - V_0/E)^{1/2}}{1 + (1 - V_0/E)^{1/2}}\right)^2$$

Also show that the magnitude of the reflection coefficient is independent of

whether the electrons are incident from the high or low potential energy side. How is its phase affected?

4. Find the probability of transmission of a 1 MeV proton through a 4 MeV high, 10^{-14} m thick rectangular potential energy barrier ($m_p/m = 1836$).

Ans. 5×10^{-4}

2 The electronic structure of atoms

2.1 Introduction

In the previous chapter we showed that microscopic particles can be described in terms of waves and wavepackets and used such a description to discuss the interaction of particles with potential barriers of various shapes. In this section we confine our attention to particles that are constrained to remain localized in some part of space. We shall see that under these conditions the particle cannot have a continuous spectrum of possible energies and the system is such as to allow the particle to have only discrete values of energy; nor can the wavefunction of a confined particle take up any arbitrary value, and only a set of discrete wavefunctions is allowable.

The properties of particles confined to a finite region of space, or *bound*, can be studied by considering them to be trapped in a potential well with very steep sides. The shape of such containers with which we are able to deal is artificial in the extreme for reasons of mathematical simplicity and bears little resemblance to physical reality. However, the results obtained have general characteristics that are most relevant to the description of electrons bound to parent atoms and form a basis for the categorization of atoms in terms of their electronic structure.

2.2 A particle in a one-dimensional potential well

Consider the situation depicted in Fig. 2.1. A particle of mass m and total energy E moves only in the x direction and is constrained to remain in a region of length d by potential barriers at $x = 0$ and $x = d$. The barriers are made infinitely steep and high and are infinite in extent so that there is no possibility of the particle surmounting them and escaping. We assume that the potential energy of the particle inside the well is zero and apply Schrödinger's time-independent equation, Eq. (1.43), since the total energy E remains constant. Hence, for $0 \leq x \leq d$, $V = 0$ and

$$\frac{d^2\Psi}{dx^2} + \frac{2m}{\hbar^2} E\Psi = 0 \qquad (2.1)$$

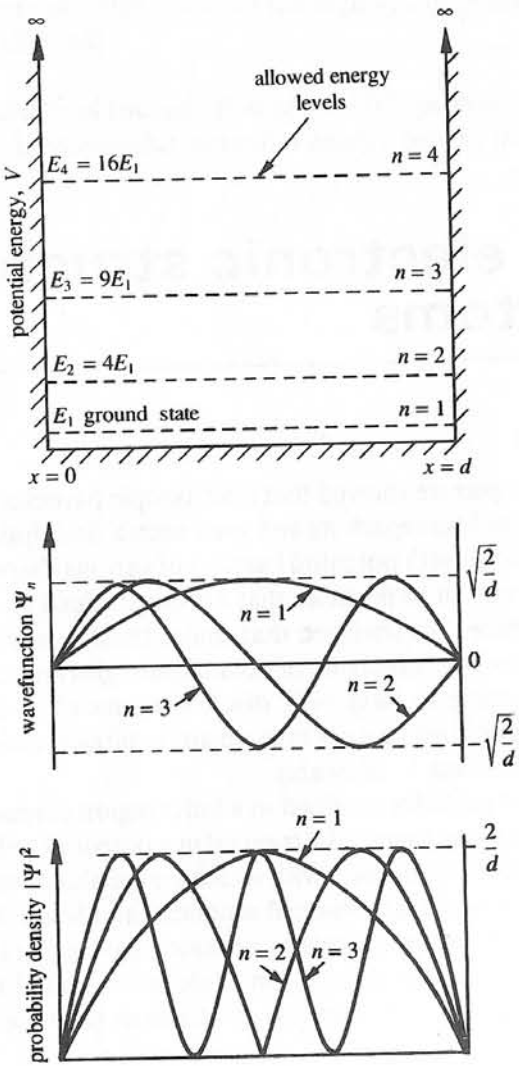


Fig. 2.1 Particle trapped in an infinitely deep, one-dimensional potential well.

We have seen that the general solution of this equation, which gives the wavefunction of the particle inside the well, is

$$\Psi = Ae^{j\beta x} + Be^{-j\beta x} \qquad (2.2)$$

where

$$\beta^2 = (2m/h^2)E$$

Now, since there is no possibility of the particle penetrating the containing walls and so gaining infinite potential energy, $\Psi\Psi^*$ and hence Ψ must be zero

value of Ψ inside the well, given by (2.2), must become zero at $x = 0$ and $x = d$. Applying these boundary conditions to Eq. (2.2) gives

$$B = -A$$

and

$$0 = A(e^{j\beta d} - e^{-j\beta d})$$

or

$$\sin(\beta d) = 0$$

Hence

$$\beta d = [(2mE)^{1/2}/\hbar]\, d = n\pi \qquad \text{where } n = 1,2,3,\ldots \qquad (2.3)$$

These conditions can now be substituted into Eq. (2.2) to give a general expression for the wavefunctions for the particle in the well:

$$\Psi = C \sin(n\pi x/d) \qquad (2.4)$$

where C is a newly defined constant. We see that the two travelling-wave components of Ψ, as represented by the two parts of Eq. (2.2), are of equal magnitude but move in opposite directions and combine to produce a standing wave in the usual way.

The constant C is a normalizing constant whose value can be obtained by stating that the probability that the particle is located *somewhere* in the well must be unity. Since the probability that it is located in length dx is $\Psi\Psi^* dx$, then

$$\int_0^d \Psi\Psi^* dx = 1$$

or

$$\int_0^d C^2 \sin^2(n\pi x/d)\, dx = 1$$

from which we find that $C = (2/d)^{1/2}$ and the normalized solution for Ψ becomes

$$\Psi = (2/d)^{1/2} \sin(n\pi x/d) \qquad n = 1,2,3,\ldots \qquad (2.5)$$

We see that the wavefunction for the bound particle is one of a set of discrete values each corresponding to a different value of the integer n. This general solution is not confined to this particular example but applies to all bound particles. Each allowed value of the wavefunction, for instance for each value of the integer in Eq. (2.5), is called an *eigenfunction*, which can be loosely translated from the German as 'particular function'.

Returning now to Eq. (2.3), it is evident that the total energy of the bound

particle also has only a discrete set of allowed values given by

$$E=\frac{\hbar^2 n^2 \pi^2}{2md^2}=\frac{n^2 h^2}{8md^2} \qquad n=1,2,3,\ldots \tag{2.6}$$

Thus, the total energy of the particle in the well has particular allowed values corresponding to the various integers; the energy is quantized and each particular energy level is called an *eigenvalue*. A set of possible energy levels is shown in Fig. 2.1. Notice that energy levels intermediate to those shown are forbidden and also that a particle in the lowest energy state or ground state, E_1, has a non-zero kinetic energy. Both of these general results are applicable to all bound particles and are at variance with classical mechanical ideas.

This result might have been obtained in a simpler, more intuitive way, using the de Broglie wavelength of the electron, $\lambda=h/p$, directly. Assuming that an electronic wave is reflected at either extremity of the well, since an electron is not allowed to penetrate outside the infinitely high potential walls, a standing wave will be established, analogous to the behaviour of electromagnetic waves in a cavity resonator or a vibrating string with fixed nodes. It follows that the well must contain an integral multiple of half electronic wavelengths of similar form to that shown in Fig. 2.1. Hence

$$d=n\lambda/2=(n/2)\ h/p$$

which can be substituted in the expression for electron energy

$$E=p^2/2m$$

to give

$$E=\left(\frac{nh}{2d}\right)^2 \frac{1}{2m}=\frac{n^2h^2}{8md^2}$$

as before.

The probability per unit length that the particle is located at some particular position in the well, x, can be found in the usual way by forming $|\Psi|^2$ from the value of Ψ given in Eq. (2.5). The eigenfunctions, Ψ_n, are plotted in Fig. 2.1 together with the corresponding probabilities of location per unit length, $|\Psi_n|^2$. Notice that the wavelength, λ, of the standing wave of Ψ is quantized, only having discrete values given by $\lambda=2\pi/\beta=2d/n$, and that n is the number of loops in the pattern. This result can be compared to the electrical resonances that occur in a transmission line short-circuited at each end; such resonances occur when the wavelength is equal to twice the line length divided by an integer. The analogy is valid and the results similar, since Schrödinger's equation is of the same form as the electromagnetic wave equation and the boundary conditions are similar in each case.

The diagrams of $|\Psi|^2$ indicate that a particle in the ground state is most probably located at the centre of the well. For higher energy states, for example

well and there is zero probability of the particle being located at certain other positions. These conclusions do not agree with the classical picture of the particle being reflected elastically from the walls of the well, which results in a constant probability amplitude since the particle may be located anywhere in the well with equal probability. However, at higher energy levels, for large values of n, the agreement is better since the quantum probability oscillates rapidly with position about a mean value, which is the classical probability, and its average value taken over a short length is equal to the classically anticipated one.

The arguments for the infinitely deep well can be extended either theoretically or on a more intuitive basis, as we shall do here, to consider the case of a particle trapped in a one-dimensional potential well of finite depth. Although this problem is slightly more realistic in that it more closely resembles the situation of an electron bound to an atom, it is still somewhat artificial. The problem is illustrated in Fig. 2.2. Consider first the classically

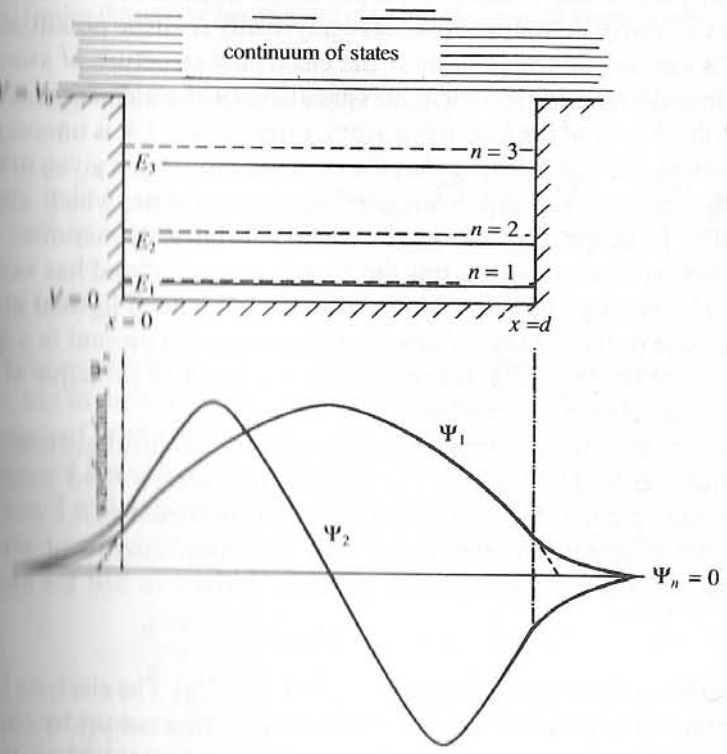


Fig. 2.2 Energy levels and wavefunctions for a particle in a finite one-dimensional potential well.

particle whose total energy, E, is less than V_0, the depth of the potential well. As we have seen earlier, there will be a small but finite probability that particles can penetrate some way through the boundaries of the well

have the characteristic exponentially decaying form, much the same as that depicted in Fig. 1.14. The wavefunction inside the well no longer fall to zero at the walls since there has to be smooth matching of wavefunctions across each boundary. Thus, although the wavefunctions have a shape generally similar to those for the infinite well, they are slightly modified in that they are no longer purely sinusoidal and their wavelength is increased. This leads to a set of eigenvalues slightly lower in value than the corresponding set for the infinite well, as shown in Fig. 2.2.

When the energy of the particle is greater than the depth of the well, $E > V_0$, then the particle is no longer classically bound and it can take up any level in a continuum of possible energy levels, as shown in the figure.

2.3 The hydrogen atom

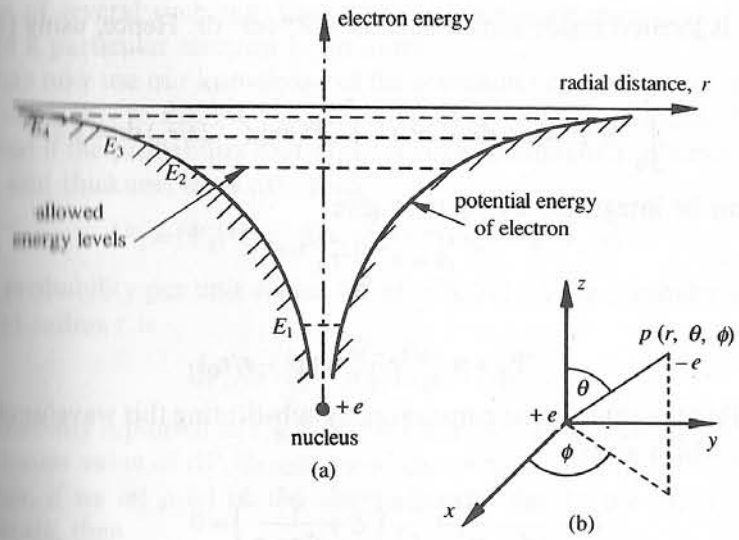
We now turn our attention to a quantum-mechanical solution of the hydrogen atom. Not only is this a useful example of the application of Schrödinger's equation to a particle trapped in a more physically realistic potential well, but the results also explain much about the electronic structure of atoms, which lays the foundations for a logical classification of the elements.

The Bohr theory of the hydrogen atom, given in Sec. 1.4, is unacceptable on several counts. First, it assumes that the electron moves in a given orbit, which implies that its position is known precisely at any time, which violates the uncertainty principle. Secondly, the theory arbitrarily assumes that the angular momentum of the orbiting electron is quantized and has values given by Eq. (1.13). Finally, the Bohr theory cannot be extended to treat atoms with more than one outer electron. These limitations are not present in a quantum-mechanical treatment and solutions can be obtained, in principle at least, for more complex atoms and molecules.

We assume that the hydrogen atom consists of a central nucleus of charge $+e$ surrounded by an electron. The nucleus is assumed fixed because of its relatively heavy mass, but to relax this condition would make only a slight quantitative difference to the result. The potential energy of an electron located at a distance r from the nucleus was derived in Sec. 1.4 and is

$$V = -e^2/(4\pi\epsilon_0 r) \tag{2.7}$$

This situation is shown diagrammatically in Fig. 2.3(a). The electron is trapped in the potential well created by the field of attraction set up by the positive nucleus; it is bound to the nucleus. The problem is more difficult than those encountered previously in that the contour of the well is not so simple mathematically and, also, the geometry is now three-dimensional. It is convenient to define the problem in spherical polar coordinates, as shown in Fig. 2.3(b), because of a possible spherical symmetry in the solution. The three-dimensional time-independent Schrödinger equation given in Eq. (1.38) is now written in the new coordinate system for the electron in the well and the



(a) Section through the potential well of a hydrogen atom; (b) spherical polar coordinates.

potential energy function given in Eq. (2.7) can be included at the same time to give

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Psi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Psi}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\Psi}{\partial\phi^2} + \frac{2m}{\hbar^2}\left(E+\frac{e^2}{4\pi\epsilon_0 r}\right)\Psi=0 \tag{2.8}$$

Finally, this equation is solved completely by the separation-of-variables technique, but in order to simplify the problem we shall look for solutions that are independent of the angular coordinates θ and ϕ. The wavefunctions that are solutions of Eq. (2.8) under these particular conditions will have spherical symmetry and will depend on the variable r only. Therefore, we let $\partial/\partial\theta$ and $\partial/\partial\phi$ be zero and carry out the differentiation implicit in the first term, to reduce Eq. (2.8) to

$$\frac{d^2\Psi}{dr^2}+\frac{2}{r}\frac{d\Psi}{dr}+\frac{2m}{\hbar^2}\left(E+\frac{e^2}{4\pi\epsilon_0 r}\right)\Psi=0 \tag{2.9}$$

There is a set of spherically symmetrical or *radial* wavefunctions that satisfy this equation, but the simplest is of the form

$$\Psi_1 = A\exp(-r/r_0) \tag{2.10}$$

where r_0 is a constant and A is the usual normalizing constant. Once again, A is evaluated by arguing that the probability of locating the electron somewhere in space must be unity. Since the probability density is $\Psi\Psi^*$ and the volume of the spherical shell of radius r and thickness dr is $4\pi r^2\,dr$, the probability that the

electron is located inside such a shell is $\Psi\Psi^* 4\pi r^2\,dr$. Hence, using (2.10), we have

$$\int_0^\infty \Psi\Psi^* 4\pi r^2\,dr = 4\pi A^2 \int_0^\infty r^2 \exp(-2r/r_0)\,dr = 1$$

which can be integrated by parts to give

$$A = \pi^{-1/2} r_0^{-3/2}$$

and

$$\Psi_1 = \pi^{-1/2} r^{-3/2} \exp(-r/r_0) \tag{2.11}$$

We can find the value of the constant r_0 by substituting this wavefunction into Eq. (2.9), which gives

$$\frac{1}{r_0^2} - \frac{2}{rr_0} + \frac{2m}{\hbar^2}\left(E + \frac{e^2}{4\pi\epsilon_0 r}\right) = 0 \tag{2.12}$$

For this equation to be true for all values of r, the terms containing r must equate to zero, or

$$\frac{2}{rr_0} = \frac{2me^2}{4\pi\hbar^2\epsilon_0 r}$$

which gives

$$r_0 = \frac{4\pi\hbar^2\epsilon_0}{me^2} = \frac{h^2\epsilon_0}{\pi e^2 m} \tag{2.13}$$

Comparing this result with Eq. (1.14) we see that the constant r_0 is numerically equal to the radius of the lowest-energy orbit of the Bohr atom; we will return to a discussion of the physical significance of this result later.

Meanwhile, if we now consider Eq. (2.12) with the sum of the r-dependent terms equal to zero, the total energy of the electron in the lowest energy state is

$$E_1 = -\frac{\hbar^2}{2mr_0^2} = -\frac{me^4}{8\epsilon_0^2 h^2} \tag{2.14}$$

where Eq. (2.13) has been used to evaluate the constant r_0.

A more complete solution of the wave equation so as to include higher-order radial wavefunctions Ψ_2, Ψ_3, etc., leads to the following general expression for the allowed energy levels:

$$E_n = \frac{E_1}{n^2} = -\frac{1}{n^2}\frac{me^4}{8\epsilon_0^2 h^2} \tag{2.15}$$

Again we see that the result of confining the electron to some localized portion of space is to quantize its possible energy levels, as shown in Fig. 2.3. The integer n is called the *principal quantum number*. We shall see later that this is

only one of several such numbers that are usually necessary to define the energy of a particular electron in an atom.

We may now use our knowledge of the wavefunction of the electron in the ground state, Ψ_1, to discuss the geometry of the hydrogen atom. We have shown that if the probability that an electron is located in a spherical shell of radius r and thickness dr is dP_r, then

$$dP_r = |\Psi_1|^2 4\pi r^2\,dr = (4r^2/r_0^3)\exp(-2r/r_0)\,dr \tag{2.16}$$

and the probability per unit radius for an electron in the ground state to be located at radius r is

$$dP_r/dr = (4r^2/r_0^3)\exp(-2r/r_0) \tag{2.17}$$

This probability is plotted in Fig. 2.4(a). It can be shown by differentiation that the maximum value of dP_r/dr occurs when $r = r_0$ as shown.

Further, if we let $\rho_1(r)$ be the charge density due to the electron in the ground state, then

$$\rho_1(r) = |\Psi_1|^2(-e) = -(e/\pi r_0^3)\exp(-2r/r_0) \tag{2.18}$$

which is shown in Fig. 2.4(b). The charge contained in a spherical shell of radius r and thickness dr is then q_r, where

$$q_r/dr = -(4er^2/r_0^3)\exp(-2r/r_0) \tag{2.19}$$

This expression is very similar to that in (2.17) and has a maximum value at r_0 again, as shown in Fig. 2.3(c). Hence a quantum-mechanical interpretation of the hydrogen atom indicates that the electron can no longer be considered as having a fixed orbit; it can be located at any distance from the nucleus, but when in the ground state its most probable location is at the Bohr radius. The charge associated with the electron is smeared out but there is a maximum in the probable charge distribution, again at the Bohr radius.

We have discussed so far only those solutions of the three-dimensional Schrödinger equation that depend on radial distance from the nucleus. A further series of wavefunctions is possible if this assumption of spherical symmetry is relaxed. When the wavefunction has angular dependence on θ and ϕ and neither $\partial/\partial\theta$ nor $\partial/\partial\phi$ is zero, Eq. (2.8) is usually solved by the method of separation of variables. This technique assumes a solution

$$\Psi(r, \theta, \phi) = f_r(r) f_\theta(\theta) f_\phi(\phi) \tag{2.20}$$

where the functions f_r, f_θ, f_ϕ are only dependent on r, θ and ϕ, respectively. When this trial wavefunction is substituted into Eq. (2.8), the three-dimensional equation in r, θ and ϕ is subdivided into three separate differential equations, one in r only, one in θ only, and one in ϕ only.

The solution for the differential equation in r gives only the radial wave functions and a *principal quantum number*, $n = 1,2,3,\ldots$, which defines the energy of an electron in a particular state, as we have already discussed.

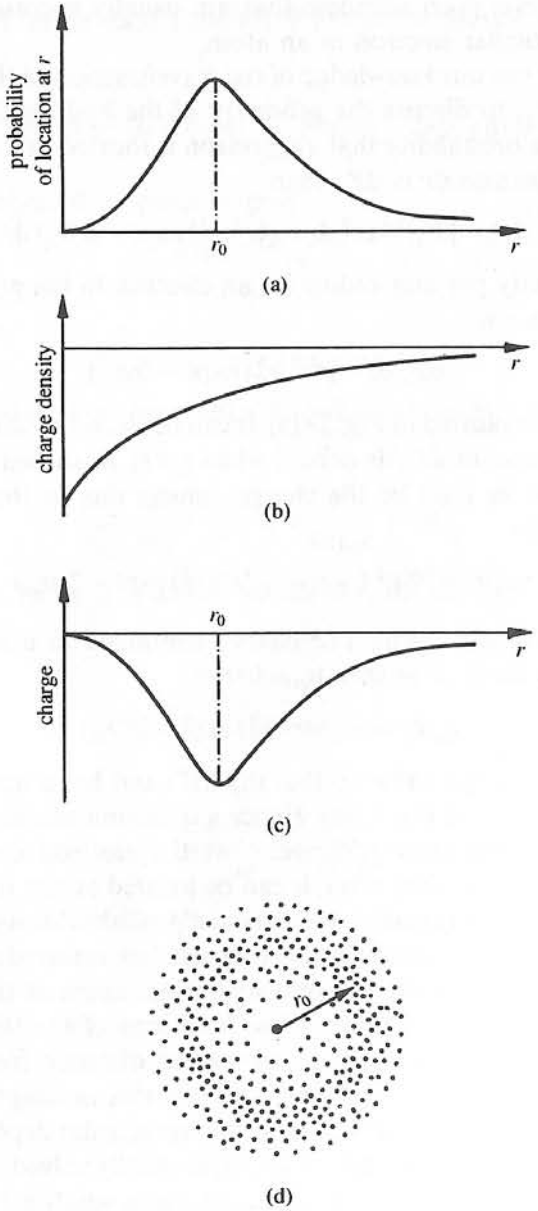


Fig. 2.4 (a)–(c) Geometry and charge configuration of the hydrogen atom in the ground state; (d) a two-dimensional representation of the smeared-out electronic charge.

A further set of *quantum numbers*, $l = 0, 1, 2, \ldots, (n-1)$, arises from a solution of the separated differential equation in θ. Such *azimuthal quantum numbers* are associated with the angular momentum of an electron, which is itself quantized.

Finally, that part of Schrödinger's equation which is separated to provide

an equation in ϕ can be solved to give a third set of possible quantum numbers, which vary from $m = -l$, including $m = 0$, to $m = +l$. These *magnetic orbital quantum numbers* are associated with the fact that an electron in an orbital constitutes a rotating charge and hence an electric current, which has associated magnetic field and magnetic moment. The orientation of this inherent magnetic moment with an externally applied magnetic field is quantized and the quantum number, m, arises because of the discrete number of possible orientations.

We see that in general there exists a set of three possible quantum numbers, l and m, and that a particular combination of these is necessary to specify an individual quantum state. In addition, a further quantum number is necessary to define completely a particular quantum state, which is not apparent from the solution to the three-dimensional Schrödinger equation. This was first introduced in an arbitrary manner to explain fine details of atomic spectra. The electron is assumed to spin about its axis in an either clockwise or anticlockwise direction. There are only two possible ways in which the inherent angular momentum vector due to the spin can be oriented with respect to an applied magnetic field. A *spin quantum number*, s, which can have only two values, accounts for this quantization. It has been shown more recently that the assumption of spin need not be introduced so arbitrarily since it arises directly as a solution of a more generalized form of the Schrödinger equation.

2.5 The exclusion principle and the periodic table of elements

We have seen that the particular quantum state of an electronic orbital can be specified by a set of quantum numbers, (n, l, m, s). Such numbers completely describe the wavefunctions for a given electron and are usually quoted instead of the wavefunction because they are less cumbersome.

The Pauli exclusion principle provides a method of classifying atoms according to their electronic structure. It states that in a multi-electronic system, which in general can be an atom, a molecule, or a complete crystal, no more than one electron can exist in any one quantum state. For the particular case of an atom, the principle implies that no two electrons can be described by an identical set of quantum numbers. Physically what this means is that no more than two electrons can have the same distribution in space and even then they must have opposite spins. There is no proof of the exclusion statement but its validity is supported by an abundance of experimental evidence.

Consider now the electronic structure of different atoms containing an increasing number of electrons. Electrons will tend to fill the lowest available energy levels first. A consequence of the exclusion principle is that each additional electron must have a different set of quantum numbers and possess a higher energy than the preceding electron. The energy levels are thus filled successively and additional electrons always have quantum numbers that

correspond to the lowest possible unoccupied energy state. A periodic table of elements based on their electronic structure can thus be constructed using this procedure in conjunction with the relationships between the various quantum numbers discussed in the previous section. Thus, for $n=1$, l and m must both be zero and two electronic states exist, corresponding to the two spin quantum numbers. But for $n=2$, l can equal 0 or 1, and for $l=1$, m may have values -1, 0, $+1$; the various combinations of (n, l, m) are thus $(2, 0, 0)$, $(2, 1, -1)$, $(2, 1, 0)$, $(2, 1, 1)$ and each of these is associated with two possible states because of spin, making a total of eight possible states with principle quantum number $n=2$. This process can be repeated for $n=3$, $l=0$, 1, 2, and so on, and the periodic table shown in Table 2.1 results.

Table 2.1 Electronic structure of the lighter elements.

Element	*Principal quantum number,* n	*Azimuthal quantum number,* $l=0, 1, \ldots, n-1$	*Magnetic quantum number,* $m=-l, \ldots, +l$	*Spectroscopic designation*
H	1	0	0	$1s$
He	1	0	0	$1s^2$
Li	2	0	0	$1s^2 2s$
Be	2	0	0	$1s^2 2s^2$
B	2	1	−1	$1s^2 2s^2 2p$
C	2	1	−1	$1s^2 2s^2 2p^2$
N	2	1	0	$1s^2 2s^2 2p^3$
O	2	1	0	$1s^2 2s^2 2p^4$
F	2	1	1	$1s^2 2s^2 2p^5$
Ne	2	1	1	$1s^2 2s^2 2p^6$
Na	3	0	0	$1s^2 2s^2 2p^6 3s$
Mg	3	0	0	$1s^2 2s^2 2p^6 3s^2$
Al	3	1	−1	$1s^2 2s^2 2p^6 3s^2 3p$
Si	3	1	−1	$1s^2 2s^2 2p^6 3s^2 3p^2$
P	3	1	0	$1s^2 2s^2 2p^6 3s^2 3p^3$
etc.				

Electrons that have the same principal quantum number, n, are said to be in the same shell. It is evident from the table that the maximum number of electrons per shell is $2n^2$. Within a shell, each state corresponding to a particular integer value of l, the azimuthal quantum number, is given a letter

designation:

l	*State or subshell*
0	s
1	p
2	d
3	f
etc.	

This peculiar nomenclature arises from the early spectroscopic identification of lines corresponding to various electronic transitions, namely, sharp, principal, diffuse, fundamental, etc.) Thus, in the electronic classification of the earlier elements, given in the last column of Table 2.1, the integers refer to the principal quantum number of each shell, the letters correspond to the value of the azimuthal quantum number, and the indices give the number of electrons in each subshell.

It should be noted that the periodic table does not progress continuously in such a numerically logical manner since for a group of the heavier elements the energy of some electrons in an outer shell is lower than that in an inner subshell and so these levels are filled before the inner subshells are fully occupied.

Problems

1. Compute the three lowest energy levels for electrons trapped in a one-dimensional well of length 0.1 nm.

Ans: 0.6×10^{-18}, 2.4×10^{-17}, 5.4×10^{-17} J

2. An electron is confined in a one-dimensional potential energy well of length 1 nm. Find (a) the kinetic energy of the electron when in the ground state and (b) the spectral frequency resulting from a transition from the next highest state to the ground level.

Ans: 18.8 eV, 3×10^{15} Hz