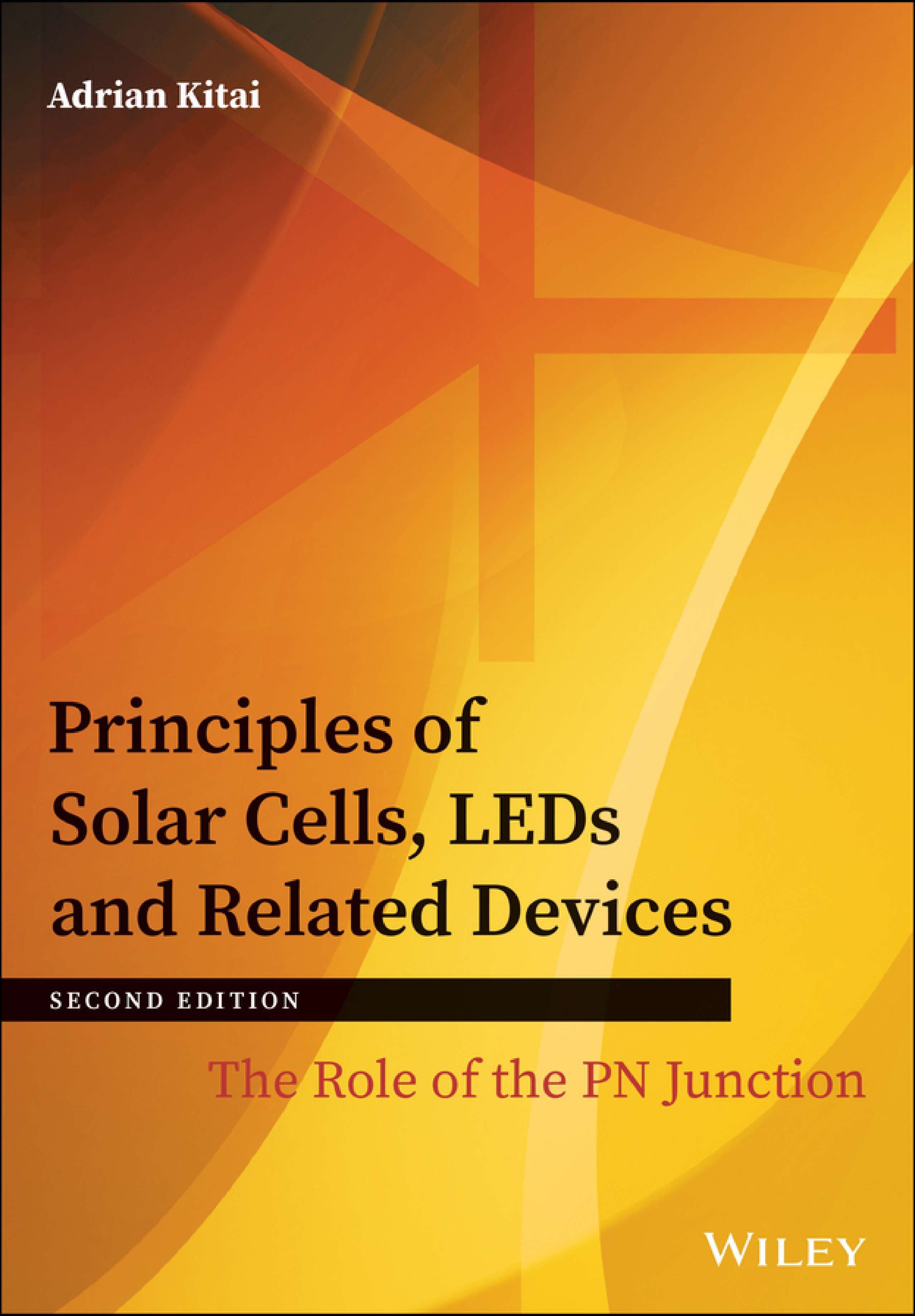

Adrian Kitai

Principles of Solar Cells, LEDs and Related Devices

SECOND EDITION

The Role of the PN Junction

WILEY

Principles of Solar Cells, LEDs and Related Devices

Principles of Solar Cells, LEDs and Related Devices

The Role of the PN Junction

ADRIAN KITAI
Departments of Materials Science and Engineering, and Engineering Physics, McMaster University, Canada

Second Edition

WILEY

This edition first published 2019
© 2019 John Wiley & Sons Ltd

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at http://www.wiley.com/go/permissions.

The right of Adrian Kitai to be identified as the author of this work has been asserted in accordance with law.

Registered Offices
John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Office
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty
In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data Applied for

ISBN 9781119451020

Cover design: Wiley
Cover image: © iStock.com/lovin-you

Set in 10/12pt TimesLTStd by SPi Global, Chennai, India

10 9 8 7 6 5 4 3 2 1

Dedicated to my wife Tomoko

Contents

Introduction

In the twenty-first century, p–n junction diode devices are revolutionising electronics, much as transistors did in the twentieth century. Diodes had been developed well before the transistor, and the properties of diodes were initially exploited in power supplies, radios, early logic circuits, and other more specialised applications. Diodes took a distant second place to transistors in the hierarchy of electronic devices after the transistor was developed. This paradigm has now changed decisively: Two semiconductor devices based directly on the p–n junction diode are currently enjoying unparalleled industrial growth. These two devices are the photovoltaic (PV) solar cell and the light-emitting diode (LED).

The consequences of this development constitute a revolution in two major industrial sectors:

1. Energy production has relied on hydrocarbons and nuclear power, and although these will continue to be important, the direct conversion of solar radiation into useful power is the key to a long-term, sustainable energy supply. Ninety-seven per cent of all renewable energy on earth is in the form of solar radiation. The twenty-first century has already seen the rapid growth of a global solar PV industry in conjunction with the involvement of governments worldwide. A scale of production and deployment of PVs that is unprecedented is now underway. The worldwide consumption of silicon semiconductor material for the entire microelectronics industry has been overtaken by its use for solar cells alone.
2. The twenty-first century has already witnessed the ongoing displacement of incandescent lamps, fluorescent lamps, and discharge lamps by LEDs. The world's major lighting companies are now dedicating their efforts to LED lighting products. Governments are recognising the benefits of LED lighting in their quest for sustainability.

More recently, both inorganic LEDs and organic light-emitting diodes (OLEDs) are enabling self-emissive displays in key display markets including handheld devices, televisions, and digital billboards. LEDs have also completely replaced fluorescent lamp backlighting in the well-established liquid crystal display (LCD) industry.

The purpose of this book is to present the physical concepts required for a thorough understanding of p–n junctions starting with introductory quantum mechanics, solid state physics, and semiconductor fundamentals. This leads to both inorganic and organic semiconductors and the associated p–n junction devices with a major emphasis on PV and LEDs. An introduction to transistors is also included since it builds readily on the p–n junction.

The book is aimed at senior undergraduate levels (years 3 and 4). The theory of the p–n junction can be quite dry in the absence of context. Students are inspired and motivated as they readily appreciate the relevance of both solar cells and LEDs. Chapter 1 motivates and presents introductory quantum mechanics for students who have not seen this elsewhere. As such, this book is designed to be accessible to all students with an interest in semiconductor devices. This is intentional since solar cells and LEDs involve a wide range of science and engineering concepts.

In Chapter 2, the physics of solid-state electronic materials is covered in detail starting from the basic behaviour of electrons in crystals. The quantitative treatment of electrons and holes in energy bands is presented along with the important concepts of excess carriers that become significant once semiconductor devices are connected to sources of power or illuminated by sunlight. A series of semiconductor materials and their important properties are reviewed. The behaviour of semiconductor surfaces and trapping concepts are introduced since they play an important role in solar cell and LED device performance.

In Chapter 3, the basic physics and important models of p–n junction devices are presented. The diode is presented as a semiconductor device that can be understood from band theory covered in Chapter 2. Diode device concepts are extended to include tunnelling, thermionic emission, metal–semiconductor contact phenomena, and the heterojunction.

Chapter 4 introduces the theory of radiation, a topic frequently overlooked in books on semiconductor devices. The deeper understanding of photon emission and absorption processes gained from this chapter is highly relevant to subsequent chapters on solar cells and LEDs. In this chapter, the physics of photon creation is explained with a minimum of mathematical complexity. Radiation theory of the oscillating electronic dipole is treated classically and then using simple quantum mechanics. The key role of the exciton in organic molecules is presented as preparation for OLEDs and organic solar cells in Chapter 7. In addition, line-shapes predicted for direct-gap semiconductors are derived. Finally, the subject of photometric units introduces the concepts of luminance and colour coordinates that are essential to a discussion of organic and inorganic LED devices.

Chapter 5 covers inorganic solar cells. The p–n junction fundamentals introduced in Chapter 3 are further developed to include illumination of the p–n junction. Readily understood modelling is used to explain the behaviour of a solar cell. Realistic solar cell structures and models are presented along with the attendant surface recombination and bulk absorption issues that must be understood in practical solar cells. A series of solar cell technologies are reviewed starting with bulk single and multicrystalline silicon solar cell technology. Amorphous silicon materials and device concepts are presented. Solar cells made using semiconductors such as CdTe are introduced followed by multijunction solar cells using layered, lattice-matched III–V semiconductor stacks.

Chapter 6 considers the basic LED structure and its operating principles. The measured lineshape of III–V LEDs is compared with the predictions of Chapter 4. LEDs are engineered to maximise radiative recombination, and key energy loss mechanisms are discussed. The series of developments that marked the evolution of today’s high-efficiency LED devices is presented starting from the semiconductors and growth techniques of the 1960s. This is followed by an in-depth presentation of wider band-gap semiconductors culminating in nitride materials and their synthesis methods for the LED industry. The double heterojunction is introduced and the resulting energy well is analyzed. Strategies to optimise optical outcoupling are discussed. Finally, the concept of spectral down-conversion using phosphor materials and the white LED are introduced along with topics of current importance including the ‘green gap’.

Chapter 7 introduces new concepts required for an understanding of organic semiconductors, in which conjugated molecular bonding gives rise to π bands and HOMO and LUMO levels. The organic LED is introduced by starting with the simplest single active layer polymer-based LED followed by successively more complex small-molecule LED structures. The roles of the various layers, including electrodes and carrier injection and transport layers, are discussed and the relevant candidate molecular materials are described. Concepts from Chapter 4, including the molecular exciton and singlet and triplet states, are used to explain efficiency limitations in the light generation layer of small-molecule OLEDs. In addition, the opportunity to use phosphorescent and delayed-fluorescence host–guest light-emitting layers to improve device efficiency is explained. The organic solar cell is introduced and the concepts of exciton generation and exciton dissociation are described in the context of the heterojunction and the bulk heterojunction. The interest in the use of fullerenes and other related nanostructured materials is explained for the bulk heterojunction. The most recent breakthrough in perovskites as a revolutionary hybrid organic/inorganic semiconductor material is presented.

Finally, Chapter 8 introduces, carefully explains, and models the two transistor types for which the p–n junction is most clearly relevant. Both the bipolar junction transistor (BJT) and the junction field effect transistor (JFET) permit the use of this book for introductory semiconductor device courses that are designed to include three-terminal devices and the concept of amplification. This lays the groundwork for subsequent courses on metal oxide field-effect transistors (MOSFETs) and other devices.

This Second Edition has been brought up to date throughout and colour has been added liberally throughout the book. A much improved and expanded set of homework problems has been developed. In addition to two new chapters, a more thorough treatment of solid-state physics to better develop band theory is included. Recent developments in telluride/selenide/sulfide solar cells, cadmium-free thin film solar cells, perovskite solar cells, triplet-harvesting strategies for OLEDs, phosphorescent, and thermally activated delayed fluorescence dopants, and LED optical outcoupling are included. A discussion of the LED colour-rendering index has been added, and a more in-depth analysis of carrier diffusion and recombination in solar cells is presented.

All the chapters are followed by problem sets that are designed to facilitate familiarity with the concepts and a better understanding of the topics introduced in the chapter. In many cases, the problems are quantitative and require calculations; however, conceptual problems are also presented. In Chapters 5 and 7, problems designed to give the reader experience in using Internet and library resources to look up information on on-going developments in solar cells and LEDs are included.

Adrian Kitai

Acknowledgements

I would like to acknowledge the people who have been invaluable to me during the preparation of this book. I would like to thank Oleg Rubel and Chris Cavalieri, who assisted in the preparation of the manuscript. I would also like to thank an excellent Wiley team including Mustaq Ahamed Noorullah, Lesley Jebaraj, Emma Strickland, and Jenny Cossham. Finally, I would like to offer special thanks to my wife, Tomoko Kitai, for her patience through this process.

1

Introduction to Quantum Mechanics

Objectives

1. Review the classical electron and motivate the need for a quantum mechanical model.
2. Present experimental evidence for the photon as a fundamental constituent of electromagnetic radiation.
3. Introduce quantum mechanical relationships based on experimental results and illustrate these using examples.

Principles of Solar Cells, LEDs and Related Devices: The Role of the PN Junction, Second Edition. Adrian Kitai.

4. Introduce expectation values for important measurable quantities based on the uncertainty principle.
5. Motivate and define the wavefunction as a means of describing particles.
6. Present Schrödinger's equation and its solutions for practical problems relevant to semiconductor materials and devices.
7. Introduce spin and the associated magnetic properties of electrons.
8. Introduce the Pauli exclusion principle and give an example of its application.

1.1 Introduction

The study of semiconductor devices relies on the electronic properties of solid-state materials and hence a fundamental understanding of the behaviour of electrons in solids.

Electrons are responsible for electrical properties and optical properties in metals, insulators, inorganic semiconductors, and organic semiconductors. These materials form the basis of an astonishing variety of electronic components and devices. Among these, devices based on the p–n junction are of key significance and they include solar cells and light-emitting diodes (LEDs) as well as other diode devices and transistors.

The electronics age in which we are immersed would not be possible without the ability to grow these materials, control their electronic properties, and finally fabricate structured devices using them, which yield specific electronic and optical functionality.

Electron behaviour in solids requires an understanding of the electron that includes the quantum mechanical description; however, we will start with the classical electron.

1.2 The Classical Electron

We describe the electron as a particle having mass

$$m = 9.11 \times 10^{-31}\ \text{kg}$$

and negative charge of magnitude

$$q = 1.602 \times 10^{-19}\ \text{C}$$

If an external electric field $\varepsilon(x, y, z)$ is present in three-dimensional space and an electron experiences this external electric field, the magnitude of the force on the electron is

$$F = q\varepsilon$$

The direction of the force is opposite to the direction of the external electric field due to the negative charge on the electron. If ε is expressed in volts per meter (V m^{-1}) then F will have units of newtons.

If an electron accelerates through a distance d from point A to point B in vacuum due to a uniform external electric field ε, it will gain kinetic energy ΔE in which

$$\Delta E = Fd = q\varepsilon d \tag{1.1}$$

This kinetic energy ΔE gained by the electron may be expressed in Joules within the Meter–Kilogram–Second (MKS) unit system. We can also say that the electron at point A has a potential energy U that is higher than its potential energy at point B. Since total energy is conserved,

$$|\Delta U| = |\Delta E|$$

There exists an electric potential $V(x, y, z)$ defined in units of the volt at any position in three-dimensional space associated with an external electric field. We obtain the spatially dependent potential energy $U(x, y, z)$ for an electron in terms of this electric potential from

$$U(x, y, z) = -qV(x, y, z)$$

We also define the *electron-volt*, another commonly used energy unit. By definition, one electron-volt in kinetic energy is gained by an electron if the electron accelerates in an electric field between two points in space whose difference in electric potential ΔV is 1 V.

Example 1.1

Find the relationship between two commonly used units of energy, namely the electron-volt and the Joule.

Consider a uniform external electric field in which $\varepsilon(x, y, z) = 1\ \mathrm{V\,m^{-1}}$. If an electron accelerates in vacuum in this uniform external electric field between two points separated by 1 m and therefore having a potential difference of 1 V, then from Eq. (1.1), it gains kinetic energy expressed in joules of

$$\Delta E = Fd = q\varepsilon d = 1.602 \times 10^{-19}\ \mathrm{C} \times \frac{1V}{\mathrm{m}} \times 1\ \mathrm{m} = 1.602 \times 10^{-19}\ \mathrm{J}$$

But, by definition, 1 eV in kinetic energy is gained by an electron if the electron accelerates in an electric field between two points in space whose difference in electric potential ΔV is 1 V, and we have therefore shown that the conversion between the joule and the electron volts is

$$1\ \mathrm{eV} = 1.602 \times 10^{-19}\ \mathrm{J}$$

If an external magnetic field $\boldsymbol{B}$ is present, the force on an electron depends on the charge q on the electron as well as the component of electron velocity $\boldsymbol{v}$ perpendicular to the magnetic field, which we shall denote as $\boldsymbol{v}_\perp$. This force, called the *Lorentz Force,* may be expressed as $\boldsymbol{F} = -q(\boldsymbol{v}_\perp \times \boldsymbol{B})$. The force is perpendicular to both the velocity component of the electron and to the magnetic field vector. The Lorentz force is the underlying mechanism for the electric motor and the electric generator.

This classical description of the electron generally served the needs of the vacuum tube electronics era and the electric motor/generator industry in the first half of the twentieth century.

In the second half of the twentieth century, the electronics industry migrated from vacuum tube devices to solid-state devices once the transistor was invented at Bell Laboratories in

1954. The understanding of the electrical properties of semiconductor materials from which transistors are made could not be achieved using a classical description of the electron. Fortunately, the field of *quantum mechanics*, which was developing over the span of about 50 years before the invention of the transistor, allowed physicists to model and understand electron behaviour in solids.

In this chapter we will motivate quantum mechanics by way of a few examples. The classical description of the electron is shown to be unable to explain some simple observed phenomena, and we will then introduce and apply the quantum-mechanical description that has proven to work very successfully.

1.3 Two Slit Electron Experiment

One of the most remarkable illustrations of how strangely electrons can behave is illustrated in Figure 1.1. Consider a beam of electrons arriving at a pair of narrow, closely spaced slits formed in a solid. Assume that the electrons arrive at the slits randomly in a beam having a width much wider than the slit dimensions. Most of the electrons hit the solid, but a few electrons pass through the slits and then hit a screen placed behind the slits as shown.

If the screen could detect where the electrons arrived by counting them, we would expect a result as shown in Figure 1.2.

In practice, a screen pattern as shown in Figure 1.3 is obtained. This result is impossible to derive using the classical description of an electron.

It does become readily explainable, however, if we assume the electrons have a wave-like nature. If light waves, rather than particles, are incident on the slits, then there are particular positions on the screen at which the waves from the two slits cancel out. This is because they are out of phase. At other positions on the screen the waves add together because they are in-phase. This pattern is the well-known interference pattern generated by light travelling

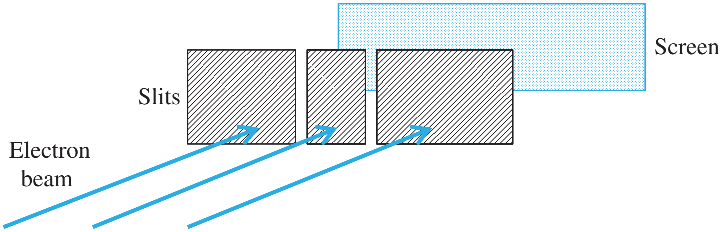

Figure 1.1 Electron beam emitted by an electron source is incident on narrow slits with a screen situated behind the slits

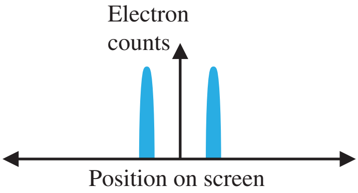

Figure 1.2 Classically expected result of two-slit experiment

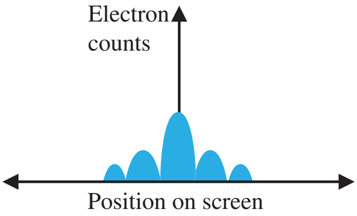

Figure 1.3 *Result of two-slit experiment. Notice that a wave-like electron is required to cause this pattern. If light waves rather than electrons were used, then a similar plot would result except the vertical axis would be a measure of the light intensity instead*

through a pair of slits. Interestingly we do not know which slit a particular electron passes through. If we attempt to experimentally determine which slit an electron is passing through we immediately disrupt the experiment and the interference pattern disappears. We could say that the electron somehow goes through both slits. Remarkably, the same interference pattern builds up slowly and is observed even if electrons are emitted from the electron source and arrive at the screen one at a time. We are forced to interpret these results as a very fundamental property of small particles such as electrons.

We will now look at how the two-slit experiment for electrons may actually be performed. It was done in the 1920s by Davisson and Germer. It turns out that very narrow slits are required to be able to observe the electron behaving as a wave due to the small wavelength of electrons. Fabricated slits having the required very small dimensions are not practical, but Davisson and Germer realised that the atomic planes of a crystal can replace slits. By a process of electron reflection, rows of atoms belonging to adjacent atomic planes on the surface of a crystal act like tiny reflectors that effectively form two beams of reflected electrons that then reach a screen and form an interference pattern similar to that shown in Figure 1.3.

Their method is shown in Figure 1.4. The angle between the incident electron beam and each reflected electron beam is θ. The spacing between surface atoms belonging to adjacent atomic planes is d. The path length difference between the two beam paths shown is $d\sin\theta$. A maximum on the screen is observed when

$$d\sin\theta = n\lambda \tag{1.2a}$$

or an integer number of wavelengths. Here, n is an integer and λ is the wavelength of the waves. A minimum occurs when

$$d\sin\theta = \left(\frac{2n+1}{2}\right)\lambda \tag{1.2b}$$

which is an odd number of half wavelengths causing wave cancellation.

In order to determine the wavelength of the apparent electron wave we can solve Eq. (1.2a) and (1.2b) for λ. We have the appropriate values of θ; however, we need to know d. Using X-ray diffraction and Bragg's law we can obtain d. Note that Bragg's law is also based on wave interference except that the waves are X-rays.

The results that Davisson and Germer obtained were quite startling. The calculated values of λ were on the order of angstroms, where 1 Å is one-tenth of a nanometre. This is much smaller than the wavelength of light, which is on the order of thousands of angstroms, and it

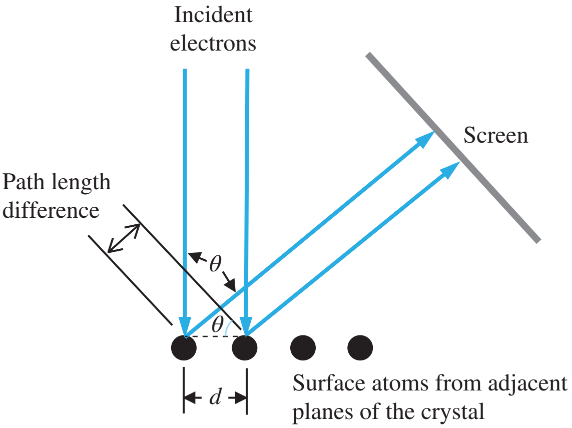

Figure 1.4 *Davisson–Germer experiment showing electrons reflected off adjacent crystalline planes. Path length difference is* $d\sin\theta$

explains why regular slits used in optical experiments are much too large to observe electron waves. But more importantly *the measured values of* λ *actually depended on the incident velocity* v *or momentum* mv *of the electrons* used in the experiment. Increasing the electron momentum by accelerating electrons through a higher potential difference before they reached the crystal caused λ to decrease, and decreasing the electron momentum caused λ to increase. By experimentally determining λ for a range of values of incident electron momentum, the following relationship was discovered:

$$\lambda = \frac{h}{p} \tag{1.3a}$$

This is known as the *de Broglie equation*, because de Boglie postulated this relationship before it was validated experimentally. Here p is the magnitude of electron momentum, and $h = 6.63 \times 10^{-34}$ Js is a constant known as *Planck's constant*. In an alternative form of the equation we define $\hbar$, pronounced *h-bar* to be $\hbar = \frac{h}{2\pi}$ and we define k, the *wavenumber* to be $k = \frac{2\pi}{\lambda}$. Now we can write the de Broglie equation as

$$p = \frac{h}{\lambda} = \hbar k \tag{1.3b}$$

Note that p is the magnitude of the momentum vector $\boldsymbol{p}$ and k is the magnitude of *wavevector* $\boldsymbol{k}$. The significance of wavevectors will be made clear in Chapter 2.

Example 1.2

An electron is accelerated through a potential difference $\Delta V = 10\,000$ V.

(a) Find the electron energy in both joules and electron-volts.
(b) Find the electron wavelength

Solution

(a) Assume that the initial kinetic energy of the electron was negligible before it was accelerated. The final energy is

$$E = q\Delta V = 1.6 \times 10^{-19}\ \mathrm{C} \times 10\,000\ \mathrm{V} = 1.6 \times 10^{-15}\ \mathrm{J}$$

To express this energy in electron-volts,

$$E = 1.6 \times 10^{-15}\ \mathrm{J} \times \frac{1\ \mathrm{eV}}{1.6 \times 10^{-19}\ \mathrm{J}} = 10^{4}\ \mathrm{eV}$$

(b) From Eq. (1.3)

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2mE}} = \frac{6.63 \times 10^{-34}\ \mathrm{Js}}{\sqrt{2 \times 9.11 \times 10^{-31}\ \mathrm{kg} \times 1.6 \times 10^{-15}\ \mathrm{J}}}$$

$$= 1.23 \times 10^{-11}\ \mathrm{m} = 0.123\ \mathring{\mathrm{A}}$$

1.4 The Photoelectric Effect

About 30 years before Davisson and Germer discovered and measured electron wavelengths, another important experiment had been undertaken by Heinrich Hertz. In 1887, Hertz was investigating what happens when light is incident on a metal. He found that electrons in the metal can be liberated by the light. It takes a certain amount of energy to release an electron from a metal into vacuum. This energy is called the *workfunction* Φ, and the magnitude of Φ depends on the metal.

If the metal is placed in a vacuum chamber, the liberated electrons are free to travel away from the metal and they can be collected by a collector electrode also located in the vacuum chamber shown in Figure 1.5. This is known as the *photoelectric effect*.

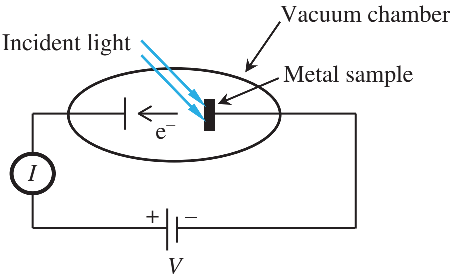

Figure 1.5 *The photoelectric experiment. The current I flowing through the external circuit is the same as the vacuum current*

By carefully measuring the resulting electric current I flowing through the vacuum, scientists in the last decades of the nineteenth century were able to make the following statements:

1. The electric current increases with increasing light intensity.
2. The colour or wavelength of the light is also important. If monochromatic light is used, there is a particular critical wavelength above which no electrons are released from the metal even if the light intensity is increased.
3. This critical wavelength depends on the type of metal employed. Metals composed of atoms having low ionisation energies such as cerium and calcium have larger critical wavelengths. Metals composed of atoms having higher ionisation energies such as gold or aluminium require smaller critical wavelengths.
4. If light having a wavelength equal to the critical wavelength is used, then the electrons leaving the metal surface have no initial kinetic energy and collecting these electrons requires that the voltage V must be positive to accelerate the electrons away from the metal and towards the second electrode.
5. If light having wavelengths smaller than the critical wavelength is used then the electrons do have some initial kinetic energy. Now, even if V is negative so as to retard the flow of electrons from the metal to the electrode, some electrons may be collected. There is, however, a maximum negative voltage of magnitude V_{max} for which electrons may be collected. As the wavelength of the light is further decreased, V_{max} increases (the maximum voltage becomes more negative).
6. If very low intensity monochromatic light with a wavelength smaller than the critical wavelength is used then individual electrons are measured rather than a continuous electron current. Suppose at time $t = 0$ there is no illumination, and then at time $t > 0$ very low intensity light is turned on. The first electron to be emitted may occur virtually as soon as the light is turned on, or it may take a finite amount of time to be emitted after the light is turned on. There is no way to predict this amount of time in advance.

Einstein won the Nobel Prize for his conclusions based on these observations. He concluded the following:

1. Light is composed of particle-like entities or *wave packets* commonly called *photons*.
2. The intensity of a source of light is determined by the photon flux density, or the number of photons being emitted per second per unit area.
3. Electrons are emitted only if the incident photons each have enough energy to overcome the metal's workfunction.
4. The photon energy of an individual photon of monochromatic light is determined by the colour (wavelength) of the light.
5. Photons of a known energy are randomly emitted from monochromatic light sources, and we can never precisely know when the next photon will be emitted.
6. Normally we do not notice these photons because there are a very large number of photons in a light beam. If, however, the intensity of the light is low enough, then the photons become noticeable and light becomes *granular*.
7. If light having photon energy larger than the energy needed to overcome the workfunction is used, then the excess photon energy causes a finite initial kinetic energy of the escaped electrons.

By carefully measuring the critical photon wavelength λ_c as a function of $V_{\max}$, the relationship between photon energy and photon wavelength can be determined: The initial kinetic energy E_k of the electron leaving the metal can be sufficient to overcome a retarding (negative) potential difference V between the metal and the electrode. Since the kinetic energy lost by the electron as it moves against this retarding potential difference V is $E_k = |qV|$, we can deduce the minimum required photon energy E using the energy equation $E = \Phi + qV_{\max}$. This experimentally observed relationship is

$$E = h\frac{c}{\lambda}$$

where c is the velocity of light. Since the frequency of the photon ν is given by

$$\lambda = \frac{c}{\nu}$$

we obtain the following relationship between photon energy and photon frequency:

$$E = h\nu = \hbar\omega \tag{1.4}$$

where $\omega = 2\pi\nu$. Note that Planck's constant h is found when looking at either the wave-like properties of electrons in Eq. (1.3) or the particle-like properties of light in Eq. (1.4). This *wave-particle duality* forms the basis for quantum mechanics.

Equation (1.4) that arose from the photoelectric effect defines the energy of a photon. In addition, Eq. (1.3) that arose from the Davisson and Germer experiment applies to both electrons *and* photons. This means that a photon having a known wavelength carries a specific momentum $p = \frac{h}{\lambda}$ even though a photon has no mass. The existence of photon momentum is experimentally proven since light-induced pressure can be measured on an illuminated surface.

In summary, electromagnetic waves exist as photons, which also have particle-like properties such as momentum and energy, and particles such as electrons also have wave-like properties such as wavelength.

Example 1.3

Ultraviolet light with wavelength 190 nm is incident on a metal sample inside a vacuum envelope containing an additional collector electrode. The collector electrode potential relative to the sample potential is defined by potential difference V as shown in Figure 1.5. Photoelectric current is observed if $V \geq -1.4$ V and ceases if $V \leq -1.4$ V. Identify the metal using the following table.

Metal	Aluminium	Nickel	Calcium	Caesium
Workfunction (eV)	4.1	5.1	2.9	2.1

Solution

190 nm photons have energy $E = h\nu = \frac{hc}{\lambda} = \frac{6.62 \times 10^{-34}\ \text{Js} \times 3 \times 10^{8}\ \text{ms}^{-1}}{190 \times 10^{-9}\ \text{m}} = 1.05 \times 10^{-18}\ \text{J}$

Expressed in electron-volts, this photon energy is $\frac{1.05 \times 10^{-18}\ \text{J}}{1.6 \times 10^{-19}\text{J eV}^{-1}} = 6.56\ \text{eV}$

The potential difference V will decelerate photoelectrons released from the metal surface if $V < 0$. This results in an energy loss of the photoelectrons and if $V = -1.4$ V, the energy loss will be 1.4 eV. The photons therefore supply enough energy to overcome both the metal workfunction Φ and the energy loss due to deceleration. Hence the predicted metal workfunction is $\Phi = 6.56 - 1.4\ \text{eV} = 5.16\ \text{eV}$. The metal having the closest workfunction match is nickel.

1.5 Wave Packets and Uncertainty

Uncertainty in the precise position of a particle is embedded in its quantum mechanical wave description. The concept of a wave packet introduced in Section 1.4 for light is important since it is applicable to both photons and particles such as electrons.

A wave packet is illustrated in Figure 1.6 showing that the wave packet has a finite size. A wave packet can be analysed into, or synthesised from, a set of component sinusoidal waves, each having a distinct wavelength, with phases and amplitudes such that they interfere constructively only over a small region of space to yield the wave packet, and destructively elsewhere. This set of component sinusoidal waves of distinct wavelengths added to yield an arbitrary function is a Fourier series.

The uncertainty in the position of a particle described using a wave packet may be approximated as Δx as indicated in Figure 1.6. The uncertainty depends on the number

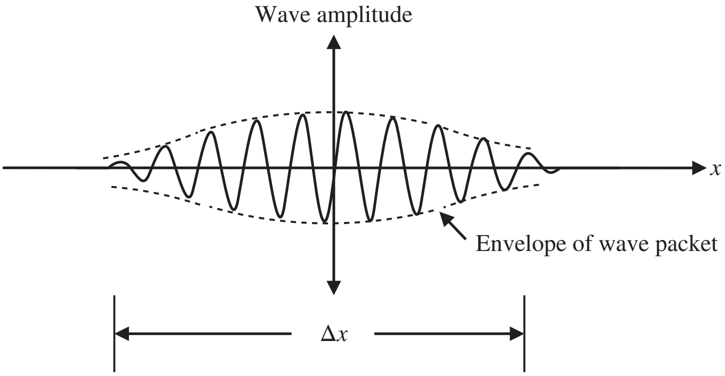

Figure 1.6 *Wave packet. The envelope of the wave packet is also shown. More insight about the meaning of the wave amplitude for a particle such as an electron will become apparent in Section 1.6 in which the concept of a probability amplitude is introduced*

of component sinusoidal waves being added together in a Fourier series: If only one component sinusoidal wave is present, the wave packet is infinitely long, and the uncertainty in position is infinite. In this case, the wavelength of the particle is precisely known, but its position is not defined. As the number of component sinusoidal wave components of the wave packet approaches infinity, the uncertainty Δx of the position of the wave packet may drop and we say that the wave packet becomes increasingly localised.

An interesting question now arises: If the wave packet is analysed into, or composed from, a number of component sinusoidal waves, can we define the precise wavelength of the wave packet? It is apparent that as more component sinusoidal waves, each having a distinct wavelength, are added together the uncertainty of the wavelength associated with the wave packet will become larger. From Eq. (1.3), the uncertainty in wavelength results in an uncertainty in momentum p and we write this momentum uncertainty as Δp.

By doing the appropriate Fourier series calculation, (see Appendix 2) the relationship between Δx and Δp can be shown to satisfy the following condition:

$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{1.5}$$

As Δx is reduced, there will be an inevitable increase in Δp. This is known as the *Uncertainty Principle*. We cannot precisely and simultaneously determine the position and the momentum of a particle. If the particle is an electron we know less and less about the electron's momentum as we determine its position more and more precisely.

Since a photon is also described in terms of a wave packet, the concept of uncertainty applies to photons as well. As the location of a photon becomes more precise, the wavelength or frequency of the photon becomes less well defined. Photons always travel with velocity of light c in vacuum. The exact arrival time t of a photon at a specific location is uncertain due to the uncertainty in position. For the photoelectric effect described in Section 1.4, the exact arrival time of a photon at a metal was observed not to be predictable for the case of monochromatic photons for which $E = h\nu = h\frac{c}{\lambda}$ is accurately known. If we allow some uncertainty in the photon frequency $\Delta\omega$ the energy uncertainty $\Delta E = \hbar\Delta\omega$ of the photon becomes finite, but then we can know more about the arrival time. The resulting relationship that may be calculated by the same approach as presented in Appendix 2 may be written as $\Delta E \Delta t \geq \frac{\hbar}{2}$. This type of uncertainty relationship is useful in time-dependent problems and, like the derivation of uncertainty for particles such as electrons, it results from a Fourier transform: The frequency spectrum $\Delta\omega$ of a pulse in the time domain becomes wider as the pulse width Δt becomes narrower.

A wave travels with velocity $v = f\lambda = \omega/k$. Note that we refer to this as a *phase velocity* because it refers to the velocity of a point on the wave that has a given phase, for example the crest of the wave. For a travelling wave packet, however, *the velocity of the particle described using the wave packet is not necessarily the same as the phase velocity of the individual waves making up the wave packet*. The velocity of the particle is actually determined using the velocity of the wave-packet envelope shown in Figure 1.6. The velocity of propagation of this envelope is called the *group velocity* v_g because the envelope is formed by the Fourier sum of a group of waves.

When photons travel through media other than vacuum, *dispersion* can exist. Consider the case of a photon having energy uncertainty $\Delta E = \hbar\Delta\omega$ due to its wave-packet description. In the case of this photon travelling through vacuum, the group velocity and the phase velocity are identical to each other and equal to the speed of light c. This is known as a *dispersion-free* photon for which the wave packet remains intact as it travels. But if a photon travels through a medium other than vacuum there is often finite dispersion in which some Fourier components of the photon wave packet travel slightly faster or slightly slower that other components of the wave packet, and the photon wave packet broadens spatially as it travels. For example, photons travelling through optical fibres typically suffer dispersion, which limits the ultimate temporal resolution of the fibre system.

It is very useful to plot ω versus k for the given medium in which the photon travels. If a straight line is obtained then $v = \omega/k$ is a constant and the velocity of each Fourier component of the photon's wave packet is identical. This is dispersion-free propagation. In general, however, a straight line will not be observed and dispersion exits.

In Appendix 3 we analyse the velocity of a wave packet composed of a series of waves. It is shown that the wave packet travels with velocity

$$v_\mathrm{g} = \frac{\mathrm{d}\omega}{\mathrm{d}k}$$

This is valid for both photons and particles such as electrons. For wave packets of particles, however, we can further state that

$$v_\mathrm{g} = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{1}{\hbar}\frac{\mathrm{d}E}{\mathrm{d}k}$$

This relationship will be important in Chapter 2 to determine the velocity of electrons in crystalline solids.

1.6 The Wavefunction

Based on what we have observed up to this point, the following four points more completely describe the properties of an electron in contrast to the description of the classical electron of Section 1.2:

1. The electron has mass m.
2. The electron has charge q.
3. The electron has wave properties with wavelength $\lambda = \frac{h}{p}$.
4. The exact position and momentum of an electron cannot be measured simultaneously.

Quantum mechanics provides an effective mathematical description of particles such as electrons that was motivated by the above observations. A *wavefunction* ψ is used to describe the particle and ψ may also be referred to as a *probability amplitude*. In general, ψ is a complex number, which is a function of space and time. Using Cartesian spatial coordinates, $\psi = \psi(x, y, z, t)$. We could also use other coordinates such as spherical polar coordinates in which case we would write $\psi = \psi(r, \theta, \phi, t)$.

The use of complex numbers is very important for wavefunctions because it allows them to represent waves as will be seen in Section 1.7.

Although ψ is a complex number and is therefore not a real, measurable or *observable* quantity, the quantity $\psi^*\psi = |\psi|^2$ where ψ^* is the complex conjugate of ψ, *is* an observable and must be a real number. $|\psi|^2$ is referred to as a *probability density*. At any time t, using