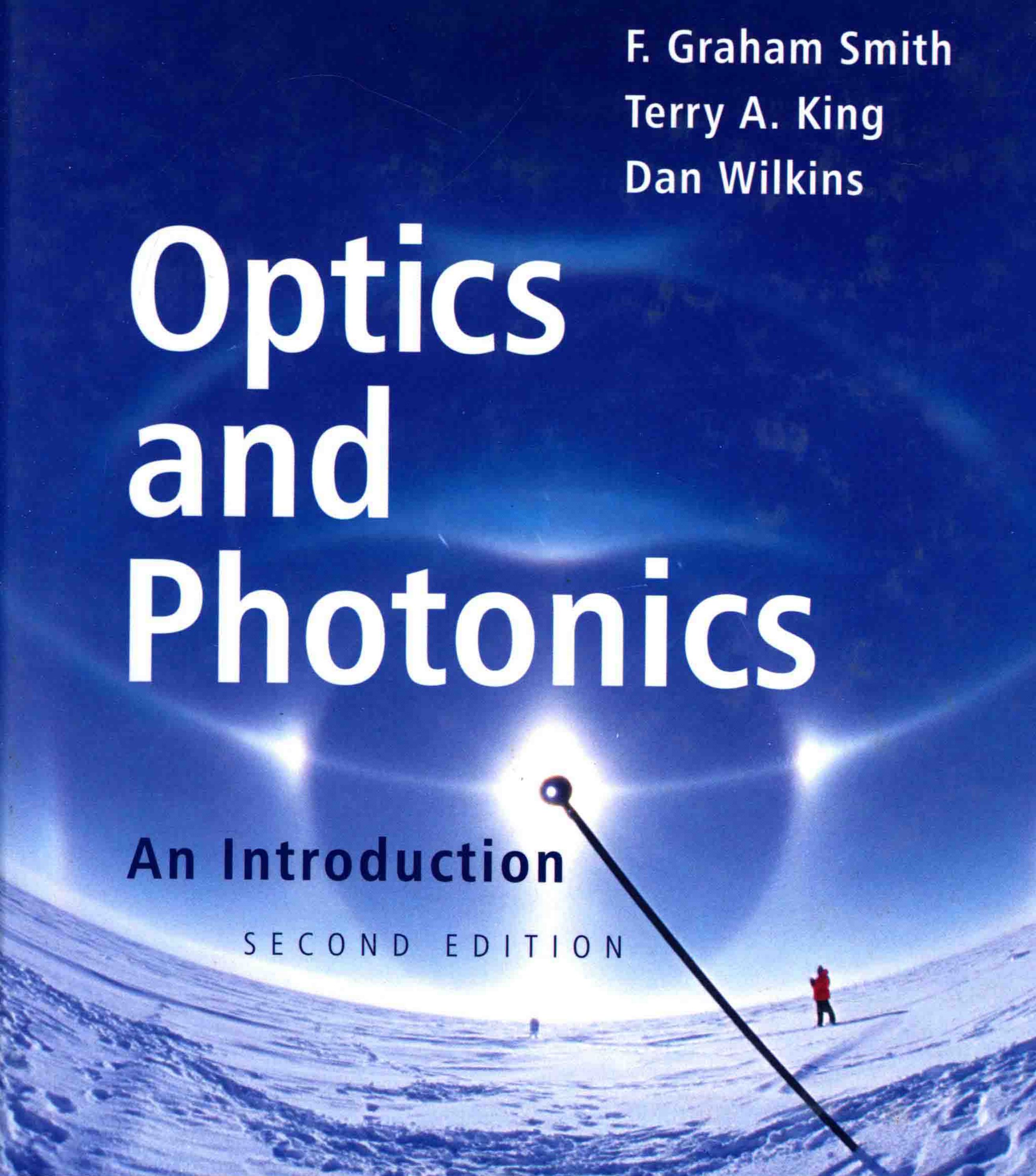
F. Graham Smith
Terry A. King
Dan Wilkins

Optics and Photonics

An Introduction

SECOND EDITION



WILEY

Optics and Photonics: An Introduction

Second Edition

F. Graham Smith

University of Manchester, UK

Terry A. King

University of Manchester, UK

Dan Wilkins

University of Nebraska at Omaha, USA

John Wiley & Sons, Ltd

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wileyeurope.com or www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 6045 Freemont Blvd., Mississauga, Ontario, Canada L5R 4J3

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Anniversary Logo Design: Richard J. Pacifico

Library of Congress Cataloging-in-Publication Data

Graham-Smith, Francis, Sir, 1923-
Optics and photonics : an introduction. – 2nd ed. / F. Graham Smith, Terry A. King, Dan Wilkins.
p. cm.
ISBN 978-0-470-01783-8 – ISBN 978-0-470-01784-5
1. Optics–Textbooks. 2. Photonics–Textbooks. I. King, Terry A. II. Wilkins, Dan, 1947- III. Title.
QC446.2.G73 2007
535–dc22 2006103070

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN: 9780470017838 (HB)
ISBN: 9780470017845 (PB)

Typeset in 10/12pt Times by Thomson Digital
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire.
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Optics and Photonics: An Introduction

SECOND EDITION

Preface

My Design in this Book is not to explain the Properties of Light by Hypothesis, but to propose and prove them by Reason and Experiments; In order to which I shall premise the following Definitions and Axioms.

The opening sentence of Newton's Opticks, 1717

Nature and Nature's laws lay hid in night: God said, Let Newton be! and all was light.

Alexander Pope, 1688–1744.

Teaching and research in modern optics must encompass the ray approach of geometric optics, the wave approach of diffraction and interferometry, and the quantum physics of the interaction of light and matter. *Optics and Photonics*, by Smith and King (2000), was designed to span this wide range, providing material for a two-year undergraduate course and some extension into postgraduate research. The text has been adopted for course teaching at the University of Omaha, Nebraska, by our third author, Dan Wilkins, and he has contributed many improvements that have proved to be essential for a rigorous undergraduate course. The material has been rearranged to give a more logical presentation and new subject matter has been added. The text has been completely revised, many of the figures have been redrawn, and new examples have been added.

The dominant factor in the recent development of optics has been the discovery and development of many forms of lasers. The remarkable properties of laser radiation have led to a wealth of new techniques such as non-linear optics, atom trapping and cooling, femtosecond dynamics and electro-optics. The laser has led to a deeper understanding of light involving coherence and quantum optics, and it has provided new optical coherence techniques which have made a major impact in atomic physics. Not only physics but also chemistry, biology, engineering and medicine have been enhanced by the use of laser-based methods, There is now a wonderful range of new applications such as holography, optical communications, picosecond and femtosecond probes, optoelectronics, medical imaging and optical coherence tomography. Myriad applications have become prominent in industry and everyday life.

A modern optics course must now place equal emphasis on the traditional optics, dealing with geometric and wave aspects of light, and on the physics of the recent developments, usually classified as photonics. The approach in this book is to emphasize the basic concepts with the objective of developing student understanding. Mathematical content is sufficient to aid the physics description but without undue complication. Extensive sets of problems are included, devised to develop understanding and provide experience in the use of the equations as well as being thought provoking. Some worked examples are in the text, and short solutions to selected problems are given at the end of the book. Notes and full solutions for all problems are posted on a website.

We now present the book as an introduction to the essential elements of optics and photonics, suitable for a one- or two-semester lecture course and including an exposition of key modern developments. We suggest that a first course, constituting minimal core material for the subject, might comprise:

- Chapter 1 Light as waves, rays, and photons.
- Chapter 2 Geometric optics, Sections 2.1–2.7.
- Chapter 4 Periodic and non-periodic waves.
- Chapter 5 Electromagnetic waves.
- Chapter 6 Fibre optics, Sections 6.1–6.8.
- Chapter 7 Polarization.
- Chapter 8 Interference by division of amplitude, Sections 8.1–8.2.
- Chapter 12 Spectra and spectrometers.
- Chapter 15 Lasers.

Selection of further material would then depend on the intended scope of the course and its duration; for example, if time permits, we recommend these additional chapters:

- Chapter 9 Interferometry.
- Chapter 10 Diffraction, Sections 10.1–10.3.
- Chapter 11 The diffraction grating.
- Chapter 14 Holography.

Communications engineers would want to include:

- Chapter 13 Coherence and correlation.
- Chapter 16 Laser light.
- Chapter 17 Semiconductors and semiconductor lasers.
- Chapter 20 The detection of light.

Those in the biosciences could well choose the following:

- Chapter 19 Interaction of light with matter.
- Chapter 20 The detection of light.
- Chapter 21 Optics and photonics in nature.

We welcome suggestions from lecturers on such course structures; we may be contacted c/o Celia Carden, Development Editor at John Wiley & Sons Ltd, email: ccarden@wiley.co.uk.

Contents

1 Light as Waves, Rays and Photons

Are not the rays of light very small bodies emitted from shining substances?

Isaac Newton, *Opticks*

All these 50 years of conscious brooding have brought me no nearer to the answer to the question 'What are light quanta?'. Nowadays every Tom, Dick and Harry thinks he knows it, but he is mistaken.

Albert Einstein, *A Centenary Volume*, 1951.

How wonderful that we have met with a paradox. Now we have some chance of making progress.

Niels Bohr (quoted by L.I. Ponomarev in *The Quantum Dice*).

Light is an *electromagnetic wave*: light is emitted and absorbed as a stream of discrete *photons*, carrying packets of energy and momentum. How can these two statements be reconciled? Similarly, while light is a wave, it nevertheless travels along straight lines or rays, allowing us to analyse lenses and mirrors in terms of geometric optics. Can we use these descriptions of waves, rays and photons interchangeably, and how should we choose between them? These problems, and their solutions, recur throughout this book, and it is useful to start by recalling how they have been approached as the theory of light has evolved over the last three centuries.

1.1 The Nature of Light

In his famous book *Opticks*, published in 1704, Isaac Newton described light as a stream of particles or corpuscles. This satisfactorily explained rectilinear propagation, and allowed him to develop theories of reflection and refraction, including his experimental demonstration of the splitting of sunlight into a spectrum of colours by using a prism. The particles in rays of different colours were supposed to have different qualities, possibly of mass, or size or velocity. White light was made up of a compound of coloured rays, and the colours of transparent materials were due to selective absorption. It was, however, more difficult for him to explain the coloured interference patterns in thin films, which we now call Newton's rings (see Chapter 9). For this, and for the partial reflection of light at a glass surface, he suggested a kind of periodic motion induced by his corpuscles, which reacted on the particles to give 'fits of easy reflection and transmission'. Newton also realized that double refraction in a calcite crystal (Iceland spar) was best explained by attributing a rectangular

Optics and Photonics: An Introduction, Second Edition F. Graham Smith, Terry A. King and Dan Wilkins

cross-section (or 'sides') to light rays, which we would now describe as *polarization* (Chapter 7). He nevertheless argued vehemently against an actual wave theory, on the grounds that waves would spread in angle rather than travel as rays, and that there was no medium to carry light waves from distant celestial bodies.

The idea that light was propagated as some sort of wave was published by René Descartes in *La Dioptrique* (1637); he thought of it as a pressure wave in an elastic medium. Christiaan Huygens, a Dutch contemporary of Newton, developed the wave theory; his explanation of rectilinear propagation is now known as 'Huygens' construction'. He correctly explained refraction in terms of a lower velocity in a denser medium. Huygens' construction is still a useful concept, and we use it later in this chapter.

It was not, however, until 100 years after Newton's *Opticks* that the wave theory was firmly established and the wavelength of light was found to be small enough to explain rectilinear propagation. In Thomas Young's double slit experiment (see Chapter 8), monochromatic light from a small source passed through two separate slits in an opaque screen, creating interference fringes where the two beams overlapped; this effect could only be explained in terms of waves. Augustin Fresnel, in 1821, then showed that the wave must be a *transverse* oscillation, as contrasted with the longitudinal oscillation of a sound wave; following Newton's ideas of rays with 'sides', this was required by the observed polarization of light as in double refraction. Fresnel also developed the theories of partial reflection and transmission (Chapter 5), and of diffraction at shadow edges (Chapter 10). The final vindication of the wave theory came with James Clerk Maxwell, who synthesized the basic physics of electricity and magnetism into the four Maxwell equations, and deduced that an electromagnetic wave would propagate at a speed which equalled that of light.

The end of the nineteenth century therefore saw the wave theory on an apparently unassailable foundation. Difficulties only remained with understanding the interaction of light with matter, and in particular the 'blackbody spectrum' of thermal radiation. This was, however, the point at which the corpuscular theory came back to life. In 1900 Max Planck showed that the form of the blackbody spectrum could be explained by postulating that the walls of the body containing the radiation consisted of harmonic oscillators with a range of frequencies, and that the energies of those with frequency ν were restricted to integral multiples of the quantity $h\nu$. Each oscillator therefore had a fundamental energy quantum

$$\boxed{E = h\nu} \tag{1.1}$$

where h became known as *Planck's constant*. In 1905 Albert Einstein explained the photoelectric effect by postulating that electromagnetic radiation was itself quantized, so that electrons are emitted from a metal surface when radiation is absorbed in discrete quanta. It seemed that Newton was right after all! Light was again to be understood as a stream of particles, later to become known as photons. What had actually been shown, however, was that light energy and the momentum carried by a light wave existed in discrete units, or quanta; photons should be thought of as events at which these quanta are emitted or absorbed.

If light is a wave that has properties usually associated with particles, could material particles correspondingly have wave-like properties? This was proposed by Louis de Broglie in 1924, and confirmed experimentally three years later in two classical experiments by George Thomson and by Clinton Davisson and Lester Germer. Both showed that a beam of particles, like a light ray encountering an obstacle, could be diffracted, behaving as a wave rather than a geometric ray. The diffraction pattern formed by the spreading of an electron beam passing through a hole in a metal sheet, for example, was the same as the diffraction pattern in light which we explore in Chapter 10. Furthermore, the wavelength λ involved was simply related to the momentum p of the electrons by

$$\boxed{\lambda = \frac{h}{p}.} \tag{1.2}$$

The constant h was again Planck's constant, as in the theory of quanta in electromagnetic radiation; for material waves λ is the *de Broglie* wavelength. A general wave theory of the behaviour of matter, *wave mechanics*, was developed in 1926 by Erwin Schrödinger following de Broglie's ideas. Wave mechanics revolutionized our understanding of how microscopic particles were described and placed limitations on the extent of information one could have about such systems – the famous Heisenberg uncertainty relationship.

The behaviour of both matter and light evidently has dual aspects: they are in some sense both particles and waves. Which aspect best describes their behaviour depends on the circumstances; light propagates, diffracts and interferes as a wave, but is emitted and absorbed discontinuously as photons, which are discrete packets of energy and momentum. Photons do not have a continuous existence, as does for example an electron in the beam of an accelerator machine; in contrast with a material particle it is not possible to say where an individual photon is located within a light beam. In some contexts we nevertheless think of the light within some experimental apparatus, such as a cavity or a laser, as consisting of photons, and we must then beware of following Newton and being misled by thinking of photons as particles with properties like those of material particles.

Although photons and electrons have very similar wave-like characteristics, there are several fundamental differences in their behaviour. Photons have zero mass; the momentum p of a photon in equation (1.1) is related to its kinetic energy E by $E = pc$, as compared with $E = p^2/2m$ for particles moving well below light speed. Unlike electrons, photons are not conserved and can be created or destroyed in encounters with material particles. Again, their statistical behaviour is different in situations where many photons or electrons can interact, as for example the photons in a laser or electrons in a metal. No two electrons in such a system can be in exactly the same state, while there is no such restriction for photons: this is the difference between Fermi–Dirac and Bose–Einstein statistics respectively for electrons and for photons.

In the first two-thirds of this book we shall be able to treat light mainly as a wave phenomenon, returning to the concept of photons when we consider the absorption and emission of electromagnetic waves.

1.2 Waves and Rays

We now return to the question: how can light be represented by a ray? Huygens' solution was to postulate that light is propagated as a wavefront, and that at any instant every point on the wavefront is the source of a wavelet, a secondary wave which propagates outward as a spherical wave (Figure 1.1)

Each wavelet has infinitesimal amplitude, but on the common envelope where countless wavelets intersect, they reinforce each other to form a new wavefront of finite amplitude. In this way, successive positions of the wavefront can be found by a step-by-step process. The envelope[1] of the

[1] To define the envelope evolved after a short time from a wavefront segment, take a finite number N of wavelets with evenly spaced centres, and note the intersection points between adjacent wavelets. In the limit that N goes to infinity, the intersection points crowd together and constitute the envelope, which is the new wavefront.

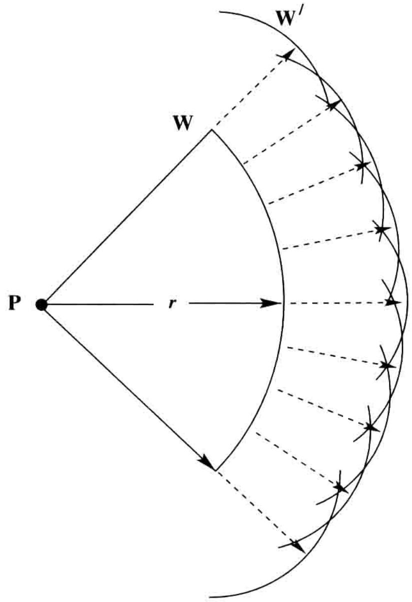


Figure 1.1 Huygens' secondary wavelets. A spherical wavefront W has originated at P and after a time t has a radius $R = ct$, where c is the speed of light. Huygens' secondary wavelets originating on W at time t combine to form a new wavefront W' at time t', when the radii of the wavelets are $c(t' - t)$

wavelets is perpendicular to the radius of each wavelet, so that the ray is the normal to a wavefront. This simple Huygens wavefront concept allows us to understand both the rectilinear propagation of light along ray paths and the basic geometric laws of reflection and refraction. There are obvious limitations: for example, what happens at the edge of a portion of the wavefront, as in Figure 1.1, and why is there no wave reradiated backwards? We return to these questions when we consider diffraction theory in Chapter 10.

Reflection of a plane wavefront W_1 reaching a totally reflecting surface is understood according to Huygens in terms of secondary wavelets set up successively along the surface as the wavefront reaches it (Figure 1.2(a)). These secondary wavelets propagate outwards and combine to form the reflected wavefront W_2. The rays are normal to the incident and reflected wavefronts. Light has travelled along each ray from W_1 to W_2 in the same time, so all path lengths from W_1 to W_2 via the mirror must be equal. The basic law of reflection follows: the incident and reflected rays lie in the same plane and the angles of incidence (i) and reflection (r) are equal.

Figure 1.2(b) shows the same reflection in terms of rays. Here we may find the same law of reflection as an example of *Fermat's principle of least time*, which states that the *time* of propagation is a minimum (or more strictly either a maximum or a minimum) along a ray path.[2] It is easy to see that the path of a light ray between the two points A and B (Figure 1.2 (b)) is a minimum if the angles i, r are equal. The proof is simple: construct the mirror image A′ of A in the reflecting surface, when the line A′B must be straight for a minimum distance. Any other path AP′B is longer.

[2]This explanation of the basic law of reflection was first given by Hero of Alexandria (First century AD).

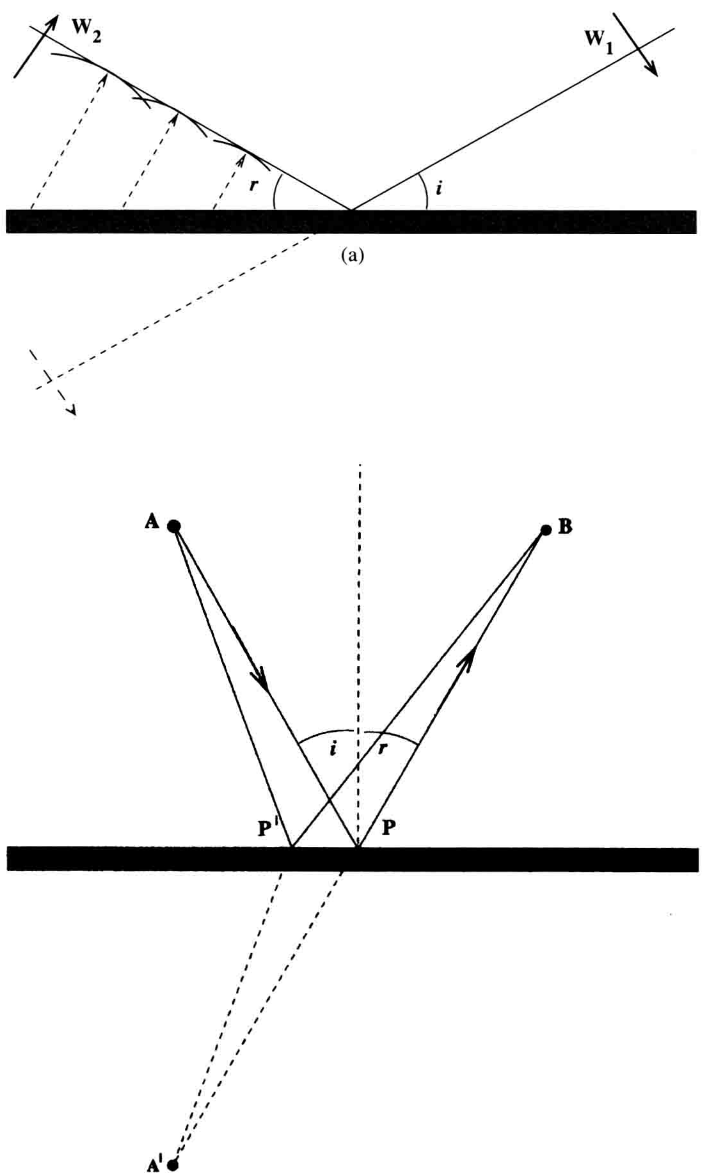


Figure 1.2 Reflection at a plane surface. (a) Huygens wave construction. The reflected wave W_2 is made up of wavelets generated as successive points on the incident plane wave W_1 reach the surface. (b) Fermat's principle. The law of reflection is found by making the path of a reflected light ray between the points A and B a minimum

Why are these two approaches essentially the same? Fermat tells us that the time of travel is the same along all paths close to an actual ray. In terms of waves this means that waves along these paths all arrive together, and reinforce one another as in Huygens' construction. When we consider periodic waves, we will express this by saying that they are *in phase*.

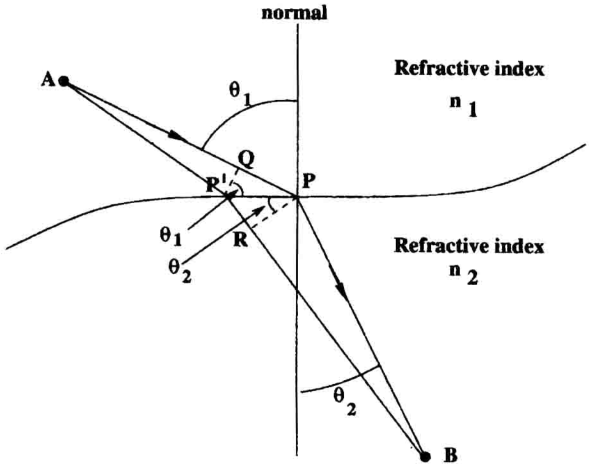


Figure 1.3 Refraction at a surface between transparent media with refractive indices n_1 and n_2. We assume the light rays and the surface normal all lie in the plane of the paper. Snell's law corresponds to a stationary value of the optical path $n_1\mathrm{AP} + n_2\mathrm{PB}$ between the fixed endpoints A, B; for small virtual variations such as shifting the point P to P′, the optical path changes negligibly

The basic law of refraction (Snell's law) may be found by applying either Huygens' or Fermat's principles to a boundary between two media in which the velocities of propagation v_1, v_2 are different; as Huygens realized, his secondary waves must travel more slowly in an optically denser medium. The *refractive indices* are defined as $n_1 = c/v_1, n_2 = c/v_2$ where c is the velocity of light in free space. As we now show, the Fermat approach shown in Figure 1.3 leads to Snell's law via some simple trigonometry.

The Fermat condition is that the travel time $(n_1\mathrm{AP} + n_2\mathrm{PB})c$ is stationary (minimum, maximum, or point of inflection); this means that for any small change in the light path of order ϵ, the change in travel time vanishes as ϵ^2 (or even faster). The distance $n_1\mathrm{AP}+n_2\mathrm{PB}$ is called the optical path. We consider a small virtual displacement of the light rays from APB to AP′B. Denote the length PP′ as ϵ. By dropping perpendiculars from P and P′, we create two thin triangles AP′Q and BPR that become perfect isosceles triangles in the limit of zero displacement. Fermat requires then that the change of the optical path satisfies[3]

$$n_1\mathrm{QP} - n_2\mathrm{P'R} = n_1\epsilon\sin\theta_1 - n_2\epsilon\sin\theta_2 = \mathrm{O}(\epsilon^2). \tag{1.3}$$

Dividing by ϵ, and going to the limit $\epsilon = 0$, this leads directly to Snell's law of refraction:

$$n_1\sin\theta_1 = n_2\sin\theta_2. \tag{1.4}$$

Notice that this derivation works for a smoothly curving surface of any shape.

In Chapter 5 we show how the laws of reflection and refraction may be derived from electromagnetic wave theory.

[3]The notation $\mathrm{O}(\epsilon^2)$ designates a quantity that varies as ϵ^2 in the limit of vanishing epsilon.

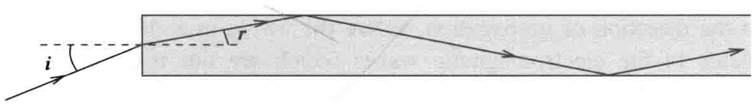


Figure 1.4 The light pipe. Rays entering at one end are totally internally reflected, and can be conducted along long paths which may include gentle curves

1.3 Total Internal Reflection

Referring again to Figure 1.3, and noting that the geometry is the same if the ray direction is reversed, we consider what happens if a ray inside the refracting medium meets the surface at a large angle of incidence θ_2, so that $\sin\theta_2$ is greater than n_1/n_2 and equation (1.4) would give $\sin\theta_1 > 1$. There can then be no ray above the surface, and there is *total internal reflection*. The internally reflected ray is at the same angle of incidence to the normal as the incident ray.

The phenomenon of total internal reflection is put to good use in the light pipe (Figure 1.4), in which light entering the end of a glass cylinder is reflected repeatedly and eventually emerges at the far end. The same principle is applicable to the transmission of light down thin optical fibres, but here the relation of the wavelength of light to the fibre diameter must be taken into account (Chapter 6).

1.4 The Light Wave

We now consider in more detail the description of the light wave, starting with a simple expression for a plane wave of any quantity ψ, travelling in the positive direction z with velocity v:

$$\boxed{\psi = f(z - vt).} \tag{1.5}$$

The function $f(z)$ describes the shape of ψ at the moment $t = 0$, and the equation states that the shape of ψ is unchanged at any later time t, with only a movement of the origin by a distance vt along the z axis (Figure 1.5). The minus sign in $(z - vt)$ indicates motion in the $+z$ direction; a plus sign would correspond to motion in the $-z$ direction. The variable quantity ψ may be a scalar, e.g. the pressure in a sound wave, or it may be a vector. If it is a vector, it may be *transverse*, i.e.

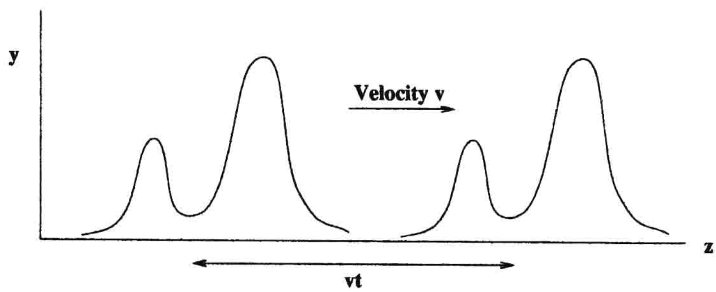


Figure 1.5 A wave travelling in the z direction with unchanging shape and with velocity v. At time $t = 0$ the waveform is $\psi = f(z)$, and at time t it is $\psi = f(z - vt)$