14

Optical aspects of solids

- What are lasers?
- Why are thin films often brightly coloured?
- What produces the colour in opal?

When light falls onto a solid, it is absorbed and/or scattered. Scattering is a broad term that has been subdivided into several categories, owing to the historical development of the subject. Thus, the topic 'scattering' tends to refer to the interaction of light with small particles. Other aspects of scattering include reflection and refraction, which is the scattering of light by a surface, and diffraction, the scattering of light by objects that are of similar dimensions to that of a light wave.

In this chapter, sources of light are considered before absorption and scattering are described, under the traditional headings. In addition, the appearance of objects, and fibre optics, both of interest in the context of optical properties of solids, are considered.

14.1 The electromagnetic spectrum

14.1.1 Light waves

Light is the form of energy detected by the eye. When light interacts with atoms and molecules it behaves as a stream of small particles, called photons. At larger scales, photons can be ignored and light can be treated as a wave. Light waves are part of the electromagnetic spectrum, with wavelengths ranging continuously from very long radiowaves, with wavelengths of gigametres (Gm), to high-energy 'cosmic rays', with wavelengths of the order of femtometres (fm); (see Figure 14.1). The wave and particle descriptions are linked by the fact that the wave equations describe the statistical behaviour of large numbers of photons.

A light wave has an electric and magnetic component, consisting of an oscillating electric and magnetic field, each described by a vector. As far as the topics in this chapter are concerned, the magnetic component need not be considered and only the electric component, specified by an electric field vector, needs to concern us. Light waves can

Understanding solids: the science of materials. Richard J. D. Tilley

ISBNs: 0 470 85275 5 (Hbk) 0 470 85276 3 (Pbk)

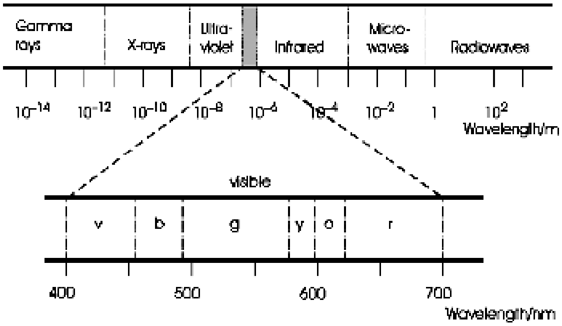

Figure 14.1 The electromagnetic spectrum. The visible region occupies only a small part of the whole, from approximately 400–700 nm. Note: v, violet; b, blue; g, green; y, yellow; o, orange; r, red

then be represented by the equation:

$$y = a_0 \sin\left[\left(\frac{2\pi}{\lambda}\right)(x + \mathrm{v}t)\right]$$

where y is the magnitude of the electric field vector at position x and time t, a_0 is the amplitude of the wave, a constant; the wavelength of the light is λ, and the velocity of the wave is v. The peaks in the wave are referred to as crests and the valleys as troughs (Figure 14.2). Any point on the wave, a crest say, is moving in the x direction with a velocity v.

The velocity of a light wave, v, is related to the frequency, ν, of the wave by the equation:

$$\mathrm{v} = \lambda\nu$$

The frequency, ν, has units of Hz (Hertz) or s^{-1}. The velocity of light in a vacuum, 2.99792×10^8 m s^{-1}, which has a special significance in physics, is given the symbol c. The velocity of light in anything other than a vacuum is less than c. The units commonly used in discussions of light, optics and spectra are given in Section S4.9.

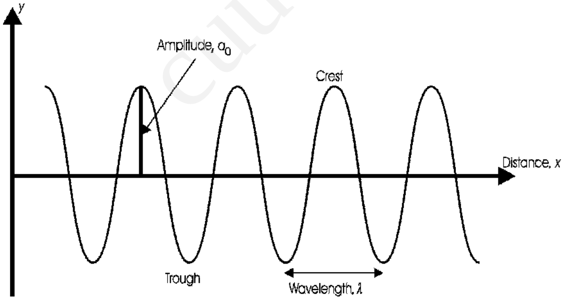

Figure 14.2 A light wave is characterised by an amplitude, a_0, and a wavelength, λ. For optical purposes, only the electric component, with an amplitude in the x–y plane, is needed

The part of the electromagnetic spectrum detected by human eyes is called the visible spectrum. Perception of the different wavelengths is called colour. The shortest wavelength of light that an average observer can perceive corresponds to the colour violet, $\lambda = 400$ nm, and the longest wavelength of light perceived by an average observer corresponds to the colour red, $\lambda = 700$ nm. Between these two limits, the other colours of the spectrum occur in the sequence from red, orange, green, blue, indigo and violet (Table 14.1).

Ultraviolet light has wavelengths shorter than violet. 'Ultraviolet A' is closest to the violet region and has the longest wavelengths. Although invisible to humans, many animals can detect ultraviolet A radiation. 'Ultraviolet B' and 'ultraviolet C' are at shorter wavelengths, and are energetic enough to damage biological cells. Radiation with wavelengths longer than red are referred to as infrared radiation. Although not visible, the longer wavelengths of infrared radiation, called thermal infrared, are detectable as the feeling of warmth on the skin.

A beam of light is said to be monochromatic when it is composed of only a very narrow range of wavelengths, and is said to be coherent when all of the waves that make up the beam are completely in phase, that is, when the crests and troughs of the waves are 'in step'. Normal light is incoherent, and laser light is coherent.

Many of the effects of the interaction of light with solids can be explained in terms of interference between light waves. If two light waves occupy the same region of space at the same time, they can add together, or interfere, to form a product wave (Figure 14.3). If two identical waves are exactly in step (Figure 14.3a.i) then they will add to produce a resultant wave with twice the amplitude by the process of constructive interference (Figure 14.3a.ii). If the two waves are out of step, the resultant amplitude will be less, as a result of

Table 14.1 The visible spectrum

Colour	λ/nm	ν/Hz	Energy J	Energy eV
Deep red	700	4.29×10^{14}	2.84×10^{-19}	1.77
Orange-red	650	4.62×10^{14}	3.06×10^{-19}	1.91
Orange	600	5.00×10^{14}	3.31×10^{-19}	2.06
Yellow	580	5.17×10^{14}	3.43×10^{-19}	2.14
Yellow-green	550	5.45×10^{14}	3.61×10^{-19}	2.25
Green	525	5.71×10^{14}	3.78×10^{-19}	2.36
Blue-green	500	6.00×10^{14}	3.98×10^{-19}	2.48
Blue	450	6.66×10^{14}	4.42×10^{-19}	2.75
Violet	400	7.50×10^{14}	4.97×10^{-19}	3.10

destructive interference. If the waves are sufficiently out of step that the crests of one correspond to the troughs of the other (Figure 14.3b.i), the resulting amplitude will be zero (Figure 14.3b.ii).

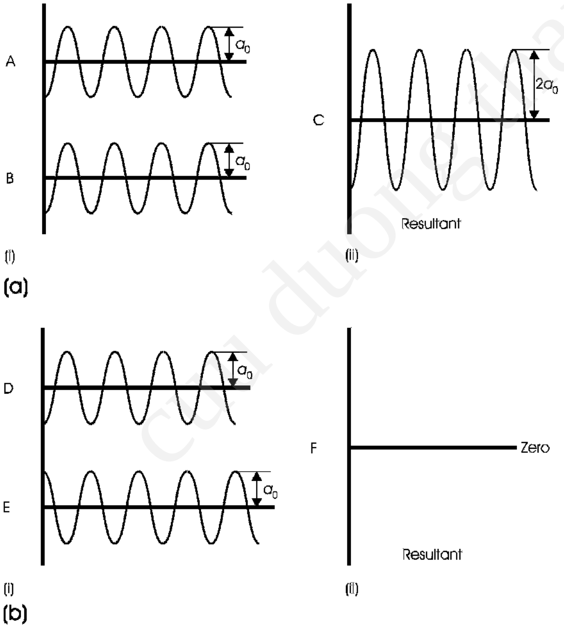

Figure 14.3 The interference of light waves: (a) two waves, A and B, in step (part i) add to give a resultant wave, C, with twice the amplitude of the original waves (part ii); (b) two waves, D and E, out of step (part i) add to zero, F (part ii)

14.1.2 Photons

When dealing with events at an atomic scale it is necessary to regard light as composed of particles, called photons. The energy of a photon is given by:

$$E = h\nu = \frac{hc}{\lambda}$$

where h is Planck's constant, and c is the velocity of light in a vacuum. The relationship between the wavelength and the frequency is:

$$\nu\lambda = c$$

Lattice vibrations in solids, which are at infrared frequencies, are often treated as particles rather than waves, and are called phonons.

14.1.3 The interaction of light with matter

When light interacts with individual atoms or molecules it is best regarded as a stream of photons. Isolated atoms and molecules have sharp energy levels. When a stream of photons encounters an atom or a molecule, the light will be absorbed only if the photon energy precisely matches the energy between the occupied energy level and one of the higher energy levels:

$$E(\text{photon})_{\text{abs}} = h\nu = E_2 - E_1$$

where E_1 represents the energy of the initial state, and E_2 represents the energy of the final (higher-energy) state. If the photon energy does not correspond exactly to the energy level separation between the occupied state and an upper energy state, it will not interact. Under normal circumstances, atoms or molecules are usually to be found in the lowest energy level, called the ground state, E_0, and the photon energy must be equal to the energy between the ground state and a higher energy state.

When an atom or molecule is in a high energy state, it may lose energy by dropping from the higher energy state to the lower state. At the same time, a photon is emitted, which carries off the excess energy. The energy of the photon is given by

$$E(\text{photon})_{\text{emit}} = h\nu = E_2 - E_1$$

where E_2 represents the energy of the initial high-energy state, and E_1 represents the energy of the final, lower-energy state. Ultimately, by emitting one or more photons, an atom or molecule will return to the ground state. Because the energy levels are sharp, the photons have precise energies, and the emission spectrum will consist of one or more sharp lines.

Although the photon energy may match the energy-level separation, photon absorption or emission need not necessarily occur. If the energy requirement is met, the actual probability of the transition occurring is governed by quantum mechanical selection rules. When the transition is allowed, which is equivalent to being highly probable, many photons will be absorbed or emitted. Colours caused by the absorption or emission of photons in allowed transitions are strong. The colours of dyes originate in allowed transitions. In cases where the transitions are forbidden – those that have a low probability of occurring – very few photons will be absorbed or emitted. Colours arising from forbidden transitions are weak and are typified by the colours displayed by 3d transition metal ions in solution.

As atoms or molecules are crowded together densely, the sharp energy levels give way to broader energy bands (see Section 2.3.1). In these cases, the sharp lines in the emission spectrum change to broader lines. In solids and liquids the interactions between the atoms is so great that the sharp outer energy levels are transformed into broad energy bands. The emission and absorption of light then becomes more complex. Nevertheless, in favourable circumstances, the interaction of light with the solid is still understood in simple terms. For example, the energy gap between the valence and conduction band in semiconductors can be measured by observing the energy at which photons are just absorbed. When photons with an energy less than the semiconductor band gap, E_g, fall on a semiconductor, they are not absorbed, as electrons in the valance band do not gain enough energy to reach the valence band. When the photon energy is greater than E_g, the transition can occur, and the photons are absorbed. The energy at which absorption just starts is a measure of the optical band gap of the semiconductor.

14.2 Sources of light

14.2.1 Luminescence

The emission of radiation by solids at relatively low temperatures is called luminescence. Luminescence is subdivided into a number of categories, some of which are listed in Table 14.2. Fluorescence is

Table 14.2 Types of luminescence

Type	Source of energy
Fluorescence	Electronic excitation between allowed states
Phosphorescence	Electronic excitation between forbidden states
Triboluminescence	Mechanical bond breaking, fracture or friction
Chemiluminescence	Chemical reactions
Photoluminescence	Visible or ultraviolet light
Cathodoluminescence	Electron bombardment
Thermoluminescence	Increase of temperature
Electroluminescence	Applied electric field
Bioluminescence	Luminescence in a living organism

characterised by the immediate re-release of the exciting energy as light. Laser action is a form of fluorescence. Phosphorescence is typified by the slow conversion of the exciting energy into light, so that light emission is delayed, often by considerable lengths of time. Phosphors are materials that absorb radiation and re-emit a portion at a lower wavelength. They are widely used in fluorescent lamps, which convert ultraviolet radiation to visible light.

Fireflies and many other animals emit radiation by chemiluminescence. In this process, chemicals mix and produce light as one of the products of reaction. Many solids emit light when crushed, a manifestation of triboluminescence. There are no general mechanisms for luminescence, and each type needs to be treated independently.

14.2.2 Incandescence

Incandescence is the emission of light by a hot body. When light from an incandescent object is spread out according to wavelength by a prism the result is the continuous fan of colours listed in Table 14.1, called a continuous spectrum. However, the radiation emitted extends over a continuous range of wavelengths much broader than the visible spectrum and is both incoherent and unpolarised. For a solid body a little above room temperature all the wavelengths of the emitted energy lie in the infrared and are discernable as a sensation of warmth. As the temperature increases the overall energy of the radiation increases and the peak of the wavelength range moves towards shorter wavelengths. At a temperature of about 700 °C, the shortest wavelengths emitted creep into the red end of the visible spectrum. The colour of the emitter is seen as red and the object is said to become red hot. At higher temperatures, the wavelengths of the radiation given out extend increasingly into the visible region and the colour observed changes from red to orange and thence to yellow. When the temperature of the emitting object reaches about 2500 °C, all visible wavelengths are present and the body is said to be white hot.

The most important incandescent object for us is the Sun, which is the ultimate source of energy on Earth. The solar spectrum has a maximum near 560 nm. Light is perceived as white if it has a make-up like that of the solar spectrum. The human eye is most sensitive to the maximum in the solar spectrum, which corresponds to yellow-green, and is noticeably less sensitive to violet and red light.

The intensity of the radiation emitted by incandescent solids can be understood by considering a black body, which is an object that absorbs and emits all wavelengths perfectly. A graph of the intensity of the radiation issuing from a black body as a function of wavelength is called a black-body emission spectrum (Figure 14.4). The shape of the curve is found to be dependent only on the temperature of the body. As the temperature increases the peak in the curve moves to shorter wavelengths (higher energies). The solar spectrum has a form quite similar to the emission spectrum of a black body with temperature of about 5700 °C (about 6000 K).

The form of the black-body emission spectrum cannot be explained by classical physics, and its

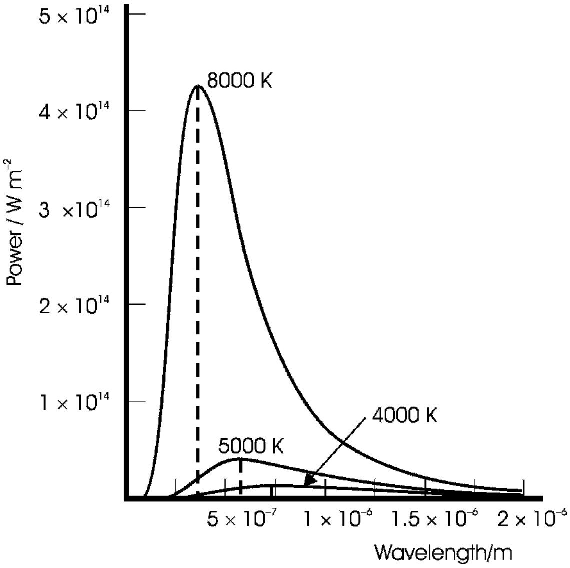

Figure 14.4 Black-body emission spectra: the maxima move to shorter wavelengths as the temperature of the emitter increases

successful theoretical explanation, by Planck in 1901, initiated quantum theory. In order to reproduce the form of the curve, Planck postulated that the energy absorbed or given out by the atoms and molecules in the black body had to be delivered in packets, or quanta. Energy emission was quantised. The relationship between the energy of a single quantum, E, and the frequency of the radiation, ν, was given by:

$$E = h\nu$$

The energy of the radiation emitted by a black body, E (in W m^{-2}), into a hemispherical region of space, in the wavelength interval, $\delta\lambda$, is given by:

$$E\mathrm{d}\lambda = \frac{2\pi hc^2 \mathrm{d}\lambda}{\lambda^5 \left[\exp\left(\frac{hc}{\lambda kT}\right) - 1\right]}$$

where T is the absolute temperature of the black body.

14.2.3 Fluorescence and solid-state lasers

The word laser is an acronym for the expression *l*ight *a*mplification by *s*timulated *e*mission of *r*adiation. The process of stimulated emission, described below, produces light that is coherent, giving it quite different properties from those of light from an incandescent source, which is incoherent.

When a photon of energy $h\nu$ is absorbed by an ion, atom or molecule it passes from a lower-energy state, the ground state, to an upper or excited state. The transition will take place if the frequency of the photon, ν, is given exactly by:

$$\nu = \frac{E_2 - E_0}{h}$$

where E_0 is the energy of the ground state, E_2 is the energy of the excited state, and h is Planck's constant. If an atom in the excited state, E_2, makes a transition to the ground state, E_0, energy will be emitted with the same frequency, given by the same equation. In 1917, Einstein suggested that there should be *two* possible types of emission process rather than just one. The most obvious is that an atom in an excited state can randomly change to the ground state: a process called spontaneous emission. Alternatively, a photon having an energy equal to the energy difference between the two levels, that is, $E_2 - E_0$, can interact with the atom in the excited state causing it to fall to the lower state and emit a photon at the same time: a process called stimulated emission.

Spontaneous emission occurs in incandescence. The light photons all have the same frequency but the waves possess random phases and the light is incoherent. Stimulated emission occurs in lasers. The photons produced have the same energy and frequency as the one that caused the emission, and the light waves of all photons are coherent.

The key to laser action is to obtain atoms or molecules in an excited state and keep them there long enough for photons to pass and trigger stimulated emission. There are two theoretical difficulties to overcome. Under normal circumstances the number of atoms in the excited energy level, N_2, relative to those in the ground state, N_0 – the relative population – will be extremely small for energy levels that are sufficiently separated to give rise to visible light. This may be confirmed by the Boltzmann law, which gives the relative populations under conditions of thermal equilibrium by:

$$\frac{N_2}{N_0} = \exp\left[-\frac{(E_2 - E_0)}{kT}\right]$$

where k is the Boltzmann constant, T is the absolute temperature, and E_2 and E_0 are the energies of the excited state and the ground state, respectively. Under equilibrium conditions it is not possible to increase the population N_2 over N_0, a situation called a population inversion.

A second problem compounds the difficulty. The ratio R of the rate of spontaneous emission to stimulated emission under conditions of thermal equilibrium is given by:

$$R = \exp\left(\frac{h\nu}{kT}\right) - 1$$

At 300 K, at visible wavelengths, R is much greater than 1. Thus, stimulated emission will be negligible compared with spontaneous emission (see Section S4.10).

The solution to the difficulties lies in recognising that the transition from one energy level to another is associated with a transition probability. If atoms can be excited into energy levels from which the probability of a transition is small, the atoms will remain in this state for long enough to produce a population inversion and so be available for stimulated emission.

14.2.4 The ruby laser: three-level lasers

The ruby laser, invented in 1960, was the first device to put the ideas just described into practice. Rubies are crystals of alumina (aluminium oxide, corundum, Al_2O_3), containing about 0.5 % chromium ions, Cr^{3+}, in place of aluminium ions, Al^{3+}. Ruby is a dilute solid solution, and the Cr^{3+} ions form substitutional defects. The laser action involves only the Cr^{3+} ions and is due to the transition of electrons from the ground state to higher energy levels among the 3d orbitals.

The energy levels of the Cr^{3+} ions are derived from the term scheme of an isolated Cr^{3+} ion on which crystal field interactions have been imposed (see Sections S1.3.1 and S4.5). The three electrons on Cr^{3+} ions give rise to two sets of energy levels, one set associated with all electron spins parallel (↑↑↑) and one set with one spin-paired couple (↑↓↑) (Figure 14.5). The ground state, E_0, has all spins parallel. Electron transitions involved in colour (optical transitions) are only allowed between levels in which the total amount of electron spin does not change. As the ground state has electrons with parallel spins, (↑↑↑), the allowed transitions are also to states with parallel spins. Such transitions are called spin-allowed transitions.

In the case of ruby there are two important spin-allowed transitions:

- $E_0 \rightarrow E_2$, at 556 nm, absorbs yellow-green;
- $E_0 \rightarrow E_3$, at 407 nm, absorbs violet.

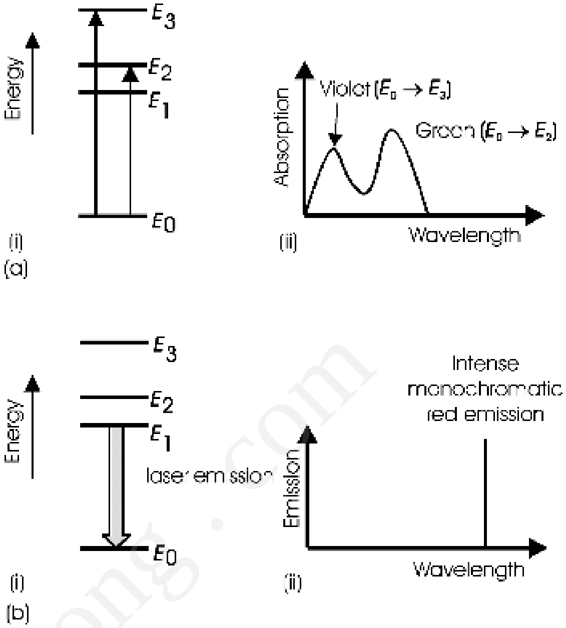

Figure 14.5 The energy levels involved in the ruby laser: (a) transitions from the ground state, E_0, to energy levels E_2 and E_3 (part i) produce the normal colour of ruby, revealed in the absorption spectrum (part ii); (b) the transition from the level E_1 to the ground state, E_0 (part i) is responsible for the laser emission (part ii)

Lying at a lower energy than E_2 is an energy level E_1. This energy level is due to an electron configuration in which two electrons are spin-paired, (↑↓↑). A transition from E_0, E_2 or E_3 to E_1 is not allowed under the total electron spin rule. None of these transitions would normally be involved in transitions that produce colour. However, in ruby, excited Cr^{3+} ions in states E_2 or E_3 can lose energy to the crystal structure and drop down to level E_1. This process operates under different conditions from the optical transitions and is independent of spin. The energy is taken up in lattice vibrations and the ruby crystal warms up. This is called a radiationless or phonon-assisted transition. Typical rates of the transitions are:

- $E_2 \rightarrow E_0$, $3 \times 10^5\ s^{-1}$;
- $E_2 \rightarrow E_1$, $2 \times 10^7\ s^{-1}$.

The second of these two transitions is about 100 times faster than the first. The rates of the transitions from the E_3 energy level to E_1 and E_0 are of a similar magnitude. This means that on irradiating the ruby with white light, Cr^{3+} ions will be excited to energy levels E_2 and E_3, and then a significant number end up in the E_1 state rather than returning to the ground state. The transition from E_1 to the ground state is not allowed because of the spin rule and so atoms in the E_1 state have a long lifetime. (The spontaneous emission rate is $2 \times 10^2\,s^{-1}$.) Thus, it is possible to build a population inversion between the E_1 and E_0 levels.

Laser operation takes place in the following way. An intense flash of white light is directed onto the crystal. This process is called optical pumping. This excites the Cr^{3+} ions into the E_2 and E_3 states. These then lose energy by radiationless transitions and 'flow over' into state E_1. An intense initial flash will cause a population inversion to form between E_1 and E_0. About 0.5 ms after the start of the pumping flash, some spontaneous emission will occur from E_1. In order to prevent these first photons from escaping from the crystal without causing stimulated emission from the other excited ions one end is coated with a mirror and the other with a partly reflecting mirror. In this arrangement, the photons are reflected back and forth, causing stimulated emission from the other populated E_1 levels. Once started, the stimulated emission rapidly depopulates these levels in an avalanche. There will be a burst of red laser light of wavelength 694.3 nm, which emerges from the partly reflecting surface.

Following the light burst, the upper levels will be empty and the process can be repeated. The ruby laser generally operates by emitting energy in short bursts, each of which lasts about 1 ms, a process referred to as pulsed operation. The ruby laser is called a three-level laser, because three energy levels are involved in the operation. These are the ground state (E_0), an excited state reached by optical absorption or pumping (E_2 or E_3), and an intermediate state of long lifetime (E_1) reached by radiationless transfer and from which stimulated emission (laser emission) occurs to the ground state.

It is energetically costly to obtain a population inversion in a three-level laser because one must pump more than half the population of the ground state to the middle level. Very little of the electrical energy supplied to the flash lamp ends up pumping photons, and carefully designed reflectors are essential. The energy lost in the transitions from E_3 and E_2 to E_1 ends up as lattice vibrations, which cause the crystal to heat up considerably. To make sure that the ruby does not overheat and shatter it is necessary to cool the crystal and to space the pulses to allow the heat to dissipate. Although the ruby laser was the first laser made, the three-level mode of operation makes it inefficient.

14.2.5 The neodymium (Nd^{3+}) solid-state laser: four-level lasers

A more energy-efficient device can be made employing four-level lasers. Laser operation takes place in the following sequence of steps (Figure 14.6). Atoms in the ground state, E_0, are excited to a rather high energy level, E_2, by optical pumping. This process needs to be fast and efficient. Subsequently, the atoms in E_2 lose energy without radiating light, to an intermediate state I_1. This step should also be fast and efficient. However, once in

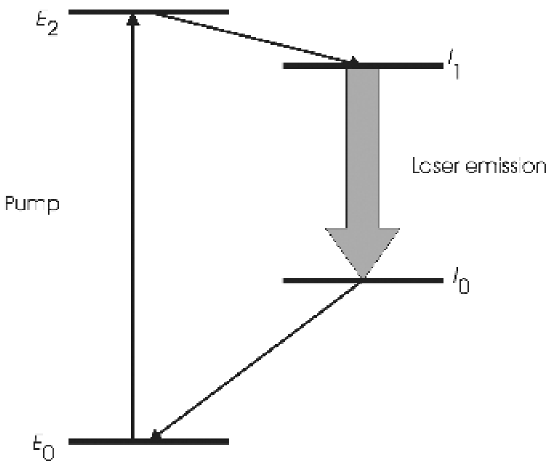

Figure 14.6 The principle transitions in a four-level laser: the laser transition occurs between two intermediate energy levels, I_1 and I_0; the transition that drives the laser, the pump transition, is between the ground state, E_0, and the energy level E_2

I_1, atoms should have a long lifetime and not lose energy quickly. When another intermediate state, I_0, is present and sufficiently high above the ground state to be effectively empty, a small population in I_1 gives a population inversion between I_1 and I_0. Ultimately, a few photons will be released as some atoms drop from I_1 to I_0. These can promote stimulated emission between I_1 and I_0, allowing laser action to take place. Atoms return from I_0 to E_0 by a step that needs to be rapid. If the energy corresponding to the transitions from E_2 to I_1 and from I_0 to E_0 can be easily dissipated, continuous operation rather than pulsed operation is possible.

The most important four-level solid-state laser uses neodymiumions ions (Nd^{3+}) as the active centres. The important transitions taking place in Nd^{3+} ion lasers are due to the transitions of f electrons (Section S1.2.3). The f-electron levels are rather sharp and can be approximated to free ion energy levels, because these levels are shielded from the effects of the surrounding crystal lattice by outer electron orbitals. Above the f-electron energy levels lie energy bands of considerable width derived from the interaction of the 5d and 6s orbitals (Figure 14.7). Optical pumping excites the ions from the ground state to these wide bands. This process is very efficient because broad energy bands allow a wide range of wavelengths to pump the laser and because the transitions are allowed. In addition, loss of energy from the excited state down to the f-electron energy levels is fast. The energy loss halts at the ^4F pair of levels. The principal laser transition is from these ^4F levels to a set of levels labelled $^4I_{11/2}$. The emission is at approximately 1060 nm, in the infrared. This is a useful wavelength as it coincides with a reasonably low-loss region of silica-based optical fibres.

Practical lasers contain about 1 % Nd^{3+} and can have quite high power outputs. The most common host materials are glass, yttrium aluminium garnet (YAG) and calcium tungstate, $CaWO_4$. They can be operated continuously or pulsed. At higher Nd^{3+} concentrations the lifetime of the ^4F upper state drops from about 200 μs in a typically 1 % doped material to about 5 μs at higher dopant concentrations. This is due to Nd–Nd interactions and associated changes in lattice vibration characteristics. Under these conditions, laser operation is no longer possible.

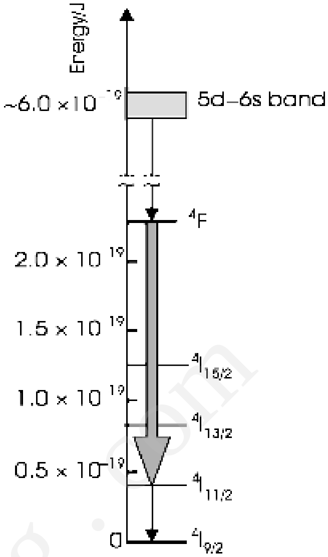

Figure 14.7 The energy levels of most importance in the neodymium laser. The pump transitions are from the ground state to a broad 5d – 6s band. The main laser transition occurs between the ^4F and $^4I_{11/2}$ levels

14.2.6 *Light-emitting diodes*

The structure of a p–n junction diode was described in Section 13.2.6. A key feature of this device is the potential barrier that builds up in the junction region. Under equilibrium conditions, this serves to separate the electrons in the n-type region from the holes in the p-type region. When a positive voltage, called a forward bias, is applied to the p-type side of the junction, the equilibrium barrier height falls and the junction region broadens (Figure 14.8). Electrons and holes now enter the junction and recombine. The energy released is approximately equal to the band gap. If this appears as light, the diode acts as a lamp and is the familiar light-emitting diode, or LED.

There are a number of aspects of importance for a working device. Clearly, the band gap must be such as to give out visible wavelengths. The most widely

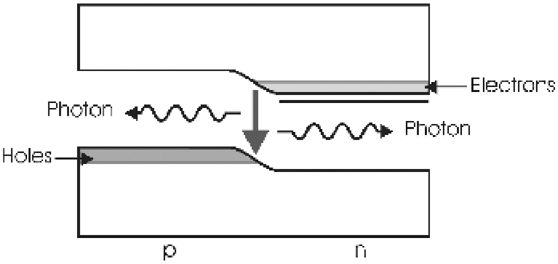

Figure 14.8 The principle of LED (light-emitting diode) operation: under a forward bias, electrons and holes recombine in the junction region and emit radiation

used LED materials are solid solutions based on the semiconductors gallium arsenide (GaAs), gallium phosphide (GaP), gallium nitride (GaN), aluminium arsenide (AlAs) and aluminium phosphide (AlP). By varying the amounts of each of these elements in the solid solution, light emission across the visible spectrum can be achieved.

A second consideration is vitally important. Ideally, LEDs must be highly efficient and give an adequate light output under a small voltage. To achieve this, the nature of the excitation of an electron from the valence band to the conduction band and the reverse process of annihilation must be efficient. This efficiency depends on the detailed band structure of the semiconductors, and the flat-band model used in Chapter 13 is no longer adequate.

There are two possibilities of importance. The first is that the lowest point of the valance band corresponds to the highest point of the conduction band; that is, they are at the same value of the wave vector, ***k*** (Figure 14.9a). In this case, an optical transition between the bands can take place without a change of ***k*** or electron momentum. Such a transition is called a direct transition. In terms of quantum mechanics, the transition has a high probability of occurring when a photon of the correct energy hits the semiconductor, and the efficiency of the process is high.

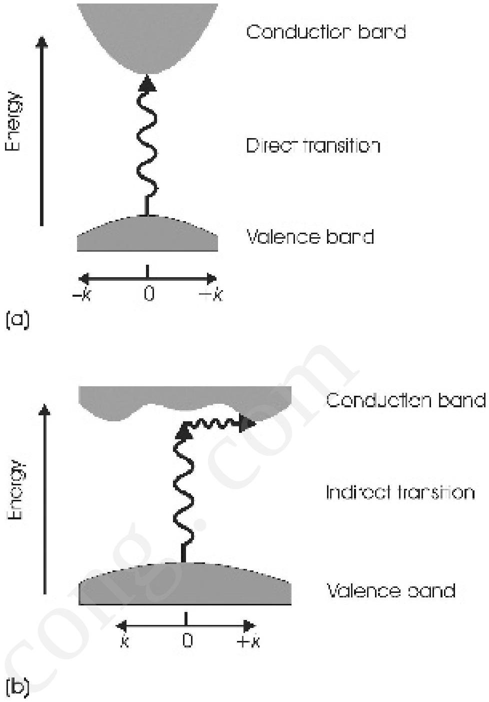

Figure 14.9 (a) In a direct transition between the valence band and the conduction band, the wave vector, ***k***, of the excited electron does not change; (b) in an indirect transition, the wave vector, ***k***, changes, making the transition much less probable than a direct transition

In the second case, the lowest point of the valence band is not at the same value of the wave vector ***k*** as the highest point in the conduction band (Figure 14.9b). In this case, the photon that is to promote the electron from the top of the valance band to the conduction band must also interact with the lattice to pick up (or lose) sufficient momentum to make the transition possible. That is, the photon must interact with a phonon to make the transition occur. Such a transition is called an indirect transition. The probability of an indirect transition occurring is quite low.

The indirect process is represented by the equation

$$h\nu = E_g \pm h\nu_{\text{phonon}}$$

where $h\nu$ is the energy of the photon needed to transfer the electron, E_g is the optical band gap, and $h\nu_{\text{phonon}}$ is the energy of the phonon involved. The ± term depends on whether the phonon is absorbed or emitted. In general, the phonon energy is small, from 0.01–0.03 eV (1.6×10^{-21}–4.8×10^{-21} J), so that this results in only a small error when optical absorption is used to measure the band gap.

The energy band structure of crystalline silicon gives rise to an indirect transition. The same is true of crystalline germanium, Ge, and for this reason neither of these materials is used in LEDs. Note, though, that amorphous silicon seems to show a direct transition. The energy band structure of gallium arsenide, GaAs, favours a direct transition, and for this reason GaAs is the preferred material for LEDs. The reason why gallium arsenide solid solutions are limited to the red and yellow colours is because of the fact that the transitions become indirect when the band gap becomes appropriate to green or blue emission. To overcome this problem gallium nitride (GaN) and related materials are being explored as blue and violet emitting LEDs.

14.2.7 Semiconductor lasers

Semiconductor lasers are, in essence, identical to LEDs, although the physical structure of lasers tends to be more complex. A semiconductor laser consists of a p–n junction in which one component has been heavily doped. The junction is placed under forward bias and a high current is passed across the device. Initially, the junction acts as an LED, and electrons recombine at random with holes to give out light by way of spontaneous emission. If the current is high enough, at some point in the junction region the number of electrons in the conduction band (from the n-type region) exceeds the number in the valence band (from the p-type region). When this population inversion is achieved, stimulated emission occurs and light emission from the junction region is coherent laser light (Figure 14.10). To increase the chance of stimulated emission occurring and to make the beam directional, two ends of the device are polished, and the whole is constructed of a number of carefully engineered layers with varying electronic characteristics.

An advantage of semiconductor lasers is that they are very efficient. Moreover, because the emission comes from the host material itself (not a small quantity of dopant as in ruby or neodymium lasers) these lasers are very powerful for their size.

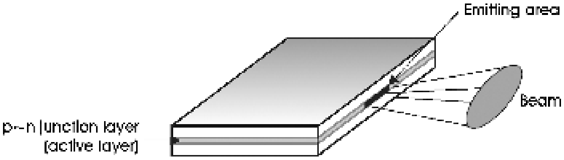

Figure 14.10 Schematic diagram of a semiconductor laser; the beam is emitted from a thin p–n junction active layer

14.3 Colour and appearance

14.3.1 Luminous solids

The colour of a solid depends on the wavelengths of light that travel from the solid to the observer's eye. Luminous objects emit radiation directly and the colour of the object will be the overall perception of the wavelength range recorded in the eye. When all wavelengths are present at a reasonable intensity the object is regarded as white. Individual wavelengths are not perceived separately, and a mixture of wavelengths corresponding to, say, red and yellow, is seen as orange and not as a combination of colours. This combination of different wavelengths of light is called additive coloration. Mixing just three different wavelengths of light in various proportions can reproduce the perceived colours of all light sources. The three wavelengths are called additive primary colours and the process of mixing lights to obtain other colours is called additive mixing. There is no fixed set of primary colours, and any three colours loosely designated as red, blue and yellow will suffice for the purpose. Mixing the three additive primary colours in equal proportions will produce white light.

14.3.2 Nonluminous solids

Nonluminous solids interact with light passively. When light of a particular wavelength falls onto such a solid it might be absorbed, by absorption centres in the material, in which case the energy of the light may end up in the solid structure as heat, and the solid is described as opaque to the absorbed wavelength. Other absorption centres may re-emit light as luminescence. Alternatively, the light may pass through the solid, no energy transfer is made and the solid will be transparent to that wavelength. The light that leaves the material is the transmitted light. Finally, for the material to be visible, at least some of the incident light has to be scattered towards the observer. The commonest scattering process is reflection, which takes place at surfaces. Smooth surfaces reflect light uniformly, known as specular reflection. Rough surfaces will reflect incident light in all directions, called diffuse reflection. Scattering centres, for example crystallites, act to reflect the light internally, giving an opalescent appearance. The appearance of a solid will depend on which of these processes occur (Figure 14.11).

For example, a solid that has no internal surfaces that reflect the light and hinder its passage, and no internal absorption centres that subtract energy from the incident beam, will appear to be transparent. The same transparent material in powder form will appear white. This is because no light is absorbed but each granule of the powder scatters light of all wavelengths. A proportion of the incident light is soon deflected in the direction of the observer by multiple reflections. Crystallites of transparent material deposited inside a glass also scatter light. If there are sufficient numbers, the glass will appear milky white. Opal glass is deliberately made to produce large concentrations of internal surfaces and appear uniformly white. Many thermoplastic materials consist of crystalline regions embedded in amorphous material. This behaviour is typified by polyethylene. The crystalline regions are of high density and of high refractive index (see Section 14.4.1) and so scatter light, which is why polyethylene appears milky.

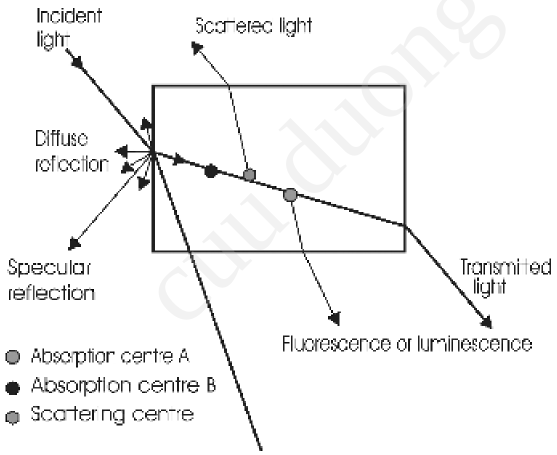

Figure 14.11 The interaction of light with a solid. The light can be transmitted, reflected, scattered or absorbed. Some absorption centres are able to re-emit light as fluorescence or luminescence. The appearance of the solid depends on which of these interactions occur

Nontransparent objects become coloured by the selective absorption of radiation. If white light falls onto a material that absorbs blue and yellow light, the object will transmit and reflect red light. It will appear red. This process is called subtractive coloration. There are three subtractive primary colours which when blended produce the subtractive colour spectrum. These are cyan, which is red-absorbing, magenta, which is green-absorbing, and yellow, which is blue-absorbing. An object containing the appropriate amounts of the three subtractive primary colours will absorb all of the light falling onto its surface and appear black to the eye. Printing inks function by way of subtractive absorption.

Metals are a particular category of opaque solids. Metals are characterised by free electrons, and those at the Fermi surface have empty energy levels readily available. Thus, any light falling on the surface of a metal will be absorbed in the surface layers. However, the excited electrons rapidly fall back to lower energy levels. Most of the light is re-emitted, but a proportion is absorbed and converted into heat. This strong absorption prevents the light from penetrating much below the surface of the metal. If the surface is smooth, the light appears to be reflected. Most metals reflect all wavelengths equally, giving them a silver appearance. Copper and gold are notable exceptions, in that they do not reflect evenly across the visible spectrum. Powdered metals repeatedly reflect light but, as some is absorbed at each interaction, much less escapes from the mass, and most finely divided metals look black.

The appearance of a surface is bound up not only with the colour leaving it but also with the texture of the surface. A smooth surface reflects light and looks shiny, even if coloured, because a certain amount of specular reflection takes place. A rough surface exhibits diffuse reflection. Skin has a different texture and appearance from a plastic film. The rendition of objects showing different amounts of specular and diffuse reflection is a difficult task for artists and especially for those wishing to create computer graphics, and much effort has been devoted to this objective.

14.3.3 The Beer–Lambert law

The cause of light absorption in a solid is due to the presence of absorption centres. These are atoms or groups of atoms that have energy levels with a spacing equivalent to the energy of light. Ignoring processes such as fluorescence and phosphorescence, in which some light is re-radiated, the interactions of light with a solid can be expressed thus:

$$\begin{aligned}&\text{Incident intensity } (I_0) = \text{amount of light reflected } (I_r)\\&\quad + \text{amount scattered } (I_s) + \text{amount absorbed } (I_a)\\&\quad + \text{amount transmitted } (I_t),\\&\qquad I_0 = I_r + I_s + I_a + I_t\end{aligned}$$

or

$$1 = R + S + A + T$$

where R is the fraction of light reflected (I_r/I_0), S is the fraction of light scattered (I_s/I_0), A is the fraction of light absorbed (I_a/I_0), and T is the fraction of light transmitted (I_t/I_0). In good-quality optical materials, the amount of light scattered and absorbed is small and it is often adequate to write:

$$I_0 = I_r + I_t$$

or

$$1 = R + T$$

When absorption centres are distributed uniformly throughout the bulk of the sample, the amount of light absorbed in a transparent plate is given by Lambert's law:

$$I = I_0 \exp(-\alpha_a l) \tag{14.1}$$

where I is the intensity leaving the plate, I_0 is the incident intensity, l is the thickness of the plate (in m), and α_a is the linear absorption coefficient (in m^{-1}). The amount of absorption will be a function of the concentration of the absorbing centres throughout the bulk of the material. This is taken into account in the Beer–Lambert law:

$$\log\left(\frac{I}{I_0}\right) = -\varepsilon[J]\,l$$

where I is the intensity after passage through a length of sample l, I_0 is the incident intensity, $[J]$ is the molar concentration of absorption centres or absorbing species, and ε is the molar absorption coefficient or extinction coefficient, which has units of $\mathrm{dm}^3\,\mathrm{mol}^{-1}\,\mathrm{m}^{-1}$, or $\mathrm{l\,mol}^{-1}\,\mathrm{m}^{-1}$. The dimensionless product A, equal to $\varepsilon[J]l$, is called the absorbance or optical density. The ratio I/I_0 is the transmittance, T, hence:

$$\log T = -A$$

The Beer–Lambert law finds use in the measurement of concentrations. For example the clarity or otherwise of polluted air is often measured by comparing the intensity of light at a certain time (I) with the intensity on a fine day (I_0).

14.4 Refraction and dispersion

14.4.1 Refraction

When light enters a transparent and insulating material, it is refracted. Refraction is the cause of the apparent bending of a ray of light when it enters water or glass and is the physical effect used in lenses (Figure 14.12). The magnitude of the effect is given by the index of refraction, or refractive index, n, where:

$$n = \frac{\sin\theta_1}{\sin\theta_2}$$

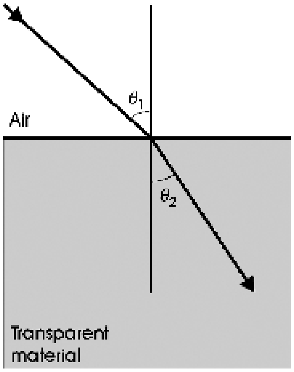

Figure 14.12 Refraction of a light beam on entering a transparent solid

θ_1 being called the angle of incidence, and θ_2 the angle of refraction. This equation is known as Snell's law (even though the originator was named Snel). The above equation is a special case of the more general relation that applies to light passing from a medium of refractive index n_1 to one of refractive index n_2:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1}$$

Some refractive indices are listed in Table 14.3 and Table 11.1 (page 342). In many crystals, the index of refraction varies with direction. These are called optically anisotropic materials. In amorphous materials such as glass, or in crystals with a cubic structure, the index of refraction is the same in all directions. These are called optically isotropic solids.

In effect, the refractive index is a manifestation of the fact that the light is slowed down on entering a transparent material. This is due to the interaction of the light with the electrons around the atoms that make up the solid. It is found that the refractive index, n, of a transparent substance is given by:

$$n = \frac{\text{velocity of light in a vacuum (c)}}{\text{velocity of light in the medium (v)}}$$

The frequency of the light does not alter when it enters a transparent medium and, because of the relationship between the velocity of a light wave and its frequency,

$$\nu\lambda = \text{velocity}$$

$$n = \frac{c}{\mathrm{v}} = \frac{\lambda_{\mathrm{vac}}}{\lambda_{\mathrm{subs}}}$$

where λ_{vac} is the wavelength of the light wave in a vacuum, and λ_{subs} is the wavelength in the transparent substance. Light has a smaller wavelength in a transparent material than in vacuum (Figure 14.13).

Table 14.3 Some refractive indices

Substance	Refractive index[a]	Substance	Refractive index[a]
Vacuum	1.0[b]	Dry air, 1 atm, 15 °C	1.00027
Water	1.3324	Na_3AlF_6 (cryolite)	1.338[c]
MgF_2	1.382[c]	Fused silica (SiO_2)	1.4601
KCl (sylvite)	1.490	Crown glass	1.522
Extra-light flint glass[d]	1.543	NaCl (halite)	1.544
Flint glass[d]	1.607	MgO (periclase)	1.735
Dense flint glass[d]	1.746	Al_2O_3 (corundum)	1.765[c]
ZrO_2 (zirconia)	2.160[c]	C (diamond)	2.418
$CaTiO_3$ (perovskite)	2.740	TiO_2 (rutile)	2.755[c]

[a] A value appropriate to the yellow light emitted by sodium atoms, the sodium D-lines, with an average wavelength 589.3 nm, is given.
[b] Definition.
[c] The refractive index varies with direction; the average value is given.
[d] The flint glasses contain significant amounts of lead oxide, PbO, as follows: extra-light flint, 24 mass% PbO; flint, 44 mass% PbO; dense flint, 62 mass % PbO.

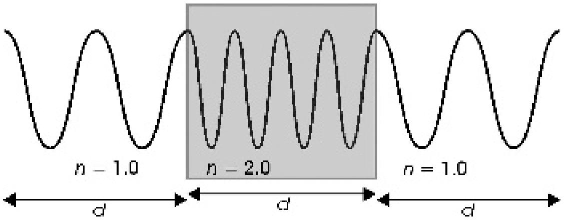

Figure 14.13 The effect of refractive index n on the wavelength of light. The wavelength is compressed in materials with a high refractive index

This can introduce confusion when a light ray traverses several different materials. To overcome this it is useful to define the optical path or optical thickness $[d]$, and distinguish it from the real or physical thickness of a material, d. The relationship is given by:

$$[d] = nd$$

where d is the physical thickness (in m) and n is the refractive index. The optical thickness of a material is frequently quoted as a number of wavelengths. Thus, a thin film with an optical thickness of λ has a real thickness given by the wavelength (in m) of the light involved divided by the refractive index. For several transparent materials traversed in sequence,

$$[d] = n_1 d_1 + n_2 d_2 + n_3 d_3 + \cdots$$

When light passes from a higher refractive index material such as glass to one of lower refractive index such as air, the refraction causes the emerging ray to bend towards the interface. As the angle, θ, at which the ray approaches the surface increases the angle of the emerging ray becomes closer to the surface, until, at the critical angle, θ_c the emerging ray actually travels exactly along the surface (Figure 14.14). If θ_c is exceeded then *no light escapes* and the light behaves as if it were reflected from the under surface. This effect is called total internal reflection. The critical angle follows from the general relation given above when θ_i is equal to 90°:

$$\sin \theta_c = \frac{n(\text{low})}{n(\text{high})}$$

where $n(\text{low}) < n(\text{high})$.

14.4.2 Refractive index and structure

The relationship between refractive index and the atomic or molecular structure of a material was considered in Sections 11.1.4 and 11.1.5. Recall that light can be treated as a varying electric field and this interacts with the internal charges on the solid. At the frequency of a light wave, only the electrons respond, and so contribute to the refractive

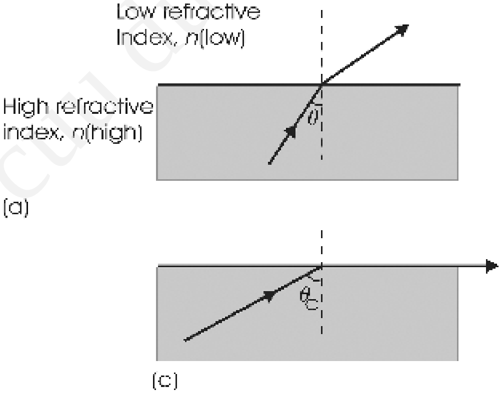

Figure 14.14 Total internal reflection of a beam of light in a solid with a higher refractive index than the surrounding medium. As the angle of the beam, θ, approaches the critical angle, θ_c, (parts a and b), the emerging beam approaches the external surface ever more closely. When the critical angle, θ_c, is reached the emerging beam lies along the surface (part c). If the critical angle is exceeded (part d) no light escapes, and it suffers total internal reflection

index. The refractive index of a material, n, is a reflection of the electronic polarisability of the solid and is related to the relative permittivity by the equation:

$$n \approx \sqrt{\varepsilon_r}$$

where ε_r is the relative permittivity of the material.

In general, strongly bound electrons, trapped at atomic nuclei or in strong chemical bonds, have a low polarisability, and this leads to a low refractive index. Loosely bound electrons, outer electrons on large atoms, or lone-pair electrons, are highly polarisable and so will yield materials with a larger refractive index.

The refractive index of a solid can be estimated via the Gladstone–Dale formula. It is especially useful for complex oxides, for which the Gladstone–Dale formula can be written:

$$n = 1 + \rho(p_1k_1 + p_2k_2 + p_3k_3 \cdots)$$

or

$$n = 1 + \rho \sum p_i k_i$$

where ρ is the density of the complex oxide. The factors k_i, called the refractive coefficients, are empirically determined constants; some examples are given in Table 14.4. The amount of each oxide is taken into account by multiplying the refractive coefficient by its weight fraction in the compound, p.

The assumption underlying the formula is that the refractive index of a complex oxide is made up by adding together the contributions from a collection of simple oxides, oxide 1, oxide 2 and so on, for which optical data are known. The rule works well and usually gives answers within about 5 %. Note, however, that the value obtained is an average refractive index. The Gladstone–Dale relationship ignores the fact that many oxides have refractive indices that vary according to crystallographic direction.

14.4.3 The refractive index of metals and semiconductors

The refractive index of transparent materials is mainly a function of electronic polarisation, arising from strongly bound electrons. In metals and many semiconductors there are considerable numbers of free electrons present. The refractive index of these materials is written as a complex number:

$$N = n + \mathrm{i}k$$

where N is the complex refractive index of the solid, i is the square root of -1, n is the real part of the refractive index, and k is called the absorption index, absorption coefficient or attenuation coefficient. At optical frequencies, the complex refractive index of a metal manifests itself in the reflectivity (see Section 14.5.1).

14.4.4 Dispersion

The refractive index of a solid varies with wavelength (Figure 14.15). This is called dispersion. In general, the index of refraction of transparent materials increases as the wavelength decreases so that the refractive index of red light in a material is less than that of violet light. The dispersion can be formally defined as $\mathrm{d}n/\mathrm{d}\lambda$, which is the slope of the curve of refractive index, n, against wavelength, λ. Although the dispersion of many materials is

Table 14.4 Refractive coefficients, k, for some oxides

Oxide	k	Oxide	k	Oxide	k	Oxide	k	Oxide	k	Oxide	k	Oxide	k	Oxide	k
H_2O	0.34														
Li_2O	0.31	BeO	0.24							B_2O_3	0.22	CO_2	0.22	N_2O_5	0.24
Na_2O	0.18	MgO	0.20							Al_2O_3	0.20	SiO_2	0.21	P_2O_5	0.19
K_2O	0.19	CaO	0.23			TiO_2	0.40								
		SrO	0.14	Y_2O_3	0.14	ZrO_2	0.20	Nb_2O_5	0.30			SnO_2	0.15		
		BaO	0.13	La_2O_3	0.15							PbO	0.15	Bi_2O_3	0.16

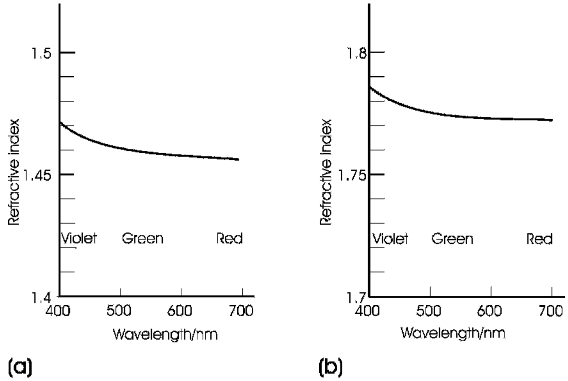

Figure 14.15 The dispersion of (a) fused silica glass and (b) corundum

rather small, it is important to include it when calculating the optical properties of optical components.

14.5 Reflection

14.5.1 Reflection from a surface

Light is reflected from smooth surfaces; following the law of reflection, the angle of incidence, θ_i, is equal to the angle of reflection, θ_r (Figure 14.16). The amount of light reflected from a surface at normal incidence (i.e. perpendicular to the surface) is given by the coefficient of reflection, r:

$$r = \frac{n_0 - n_1}{n_0 + n_1}$$

where n_0 is the refractive index of the entrance medium and n_1 is the refractive index of the material making up the reflecting surface (Figure 14.16). The coefficient of reflection is defined such that if a wave of amplitude a_0 falls on the surface, then the amplitude of the reflected wave is ra_0. For reflection at a surface between a substance of low refractive index and a substance of high refractive index, r is *negative*. This signifies a phase change of π radians on reflection, which means, in terms of a light wave, that a peak turns into a trough on reflection (Figure 14.17).

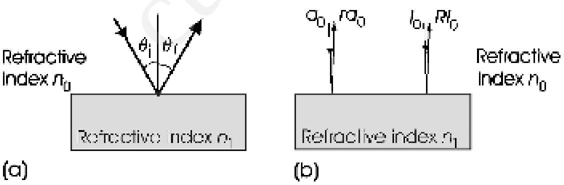

Figure 14.16 The reflection of a beam of light: (a) the angle of incidence, θ_i is equal to the angel of reflection, θ_r; (b) the coefficient of reflection at normal incidence, r, relates the amplitude reflected to the incident amplitude, a_0, whereas the reflectivity, R, relates the intensity reflected to the incident intensity, I_0

The eye detects intensity changes rather than amplitude changes, and so it is more convenient to work with the reflectivity or reflectance, R:

$$R = r^2 = \left(\frac{n_0 - n_1}{n_0 + n_1}\right)^2$$

This is because the intensity, I_0, is proportional to the square of the amplitude, $(a_0)^2$. The reflected intensity, $R(I_0)$, is proportional to $r^2(a_0)^2$. The reflectivity, R, for a plate of a transparent material of refractive index n in air is:

$$\begin{aligned} R &= \left(\frac{1-n}{1+n}\right)^2 \\ &= \left(\frac{n-1}{n+1}\right)^2 \end{aligned} \tag{14.2}$$

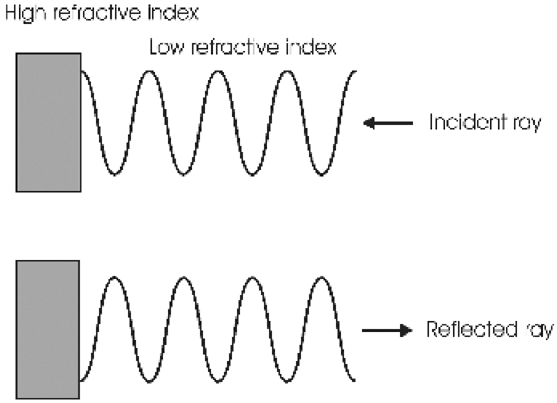

Figure 14.17 A light beam incident on a plate with a higher refractive index than the surrounding medium suffers a phase change on reflection so the incident peak is reflected as a trough

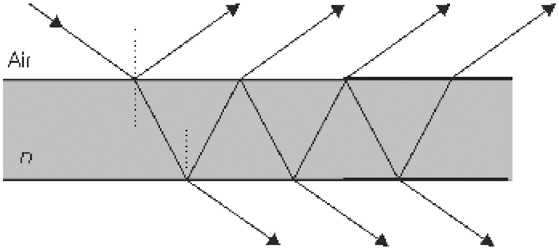

Figure 14.18 Multiple reflections at the upper and lower surfaces of a thin film of refractive index *n*

As *n* depends on wavelength, the reflectivity will vary across the spectrum. When the reflecting surface is a metal, it is necessary to use the complex refractive index, $N = N_1 + ik$. In this case, the reflectivity of a metallic surface at normal incidence is

$$R = \frac{(n_1 - n_0)^2 + k^2}{(n_1 + n_0)^2 + k^2}$$

For the case of a metallic film in air this becomes:

$$R = \frac{(n_1 - 1)^2 + k^2}{(n_1 + 1)^2 + k^2}$$

14.5.2 Reflection from a single thin film

Monochromatic light travelling through air, falling on a homogeneous thin film of refractive index *n*, will be reflected from the top surface to give a reflected ray. The light transmitted into the film will be repeatedly reflected from the bottom surface and the underside of the top surface (Figure 14.18). At each reflection, some of the light will escape to produce additional reflected and transmitted rays. As the reflectivity is rather small, the first reflected ray and the first transmitted ray are of most importance. Because of the difference in the paths taken by the repeatedly reflected rays, the waves will interfere with each other.

The appearance of the film will depend on the extent of this interference. In the case of an observer looking down on a film in a perpendicular direction, some of the light incident on the surface and seen by the observer will have been reflected at the top surface (ray 1, Figure 14.19). In addition, some light travels through the film and is reflected from the bottom surface before reaching the observer (ray 2, Figure 14.19). In addition, because ray 1 is reflected at a surface of higher refractive index, a wave peak will turn into a trough. Interference between the two waves will occur, which will cause the film to look either dark or bright. The optical path difference between rays 1 and 2 will be [*p*]:

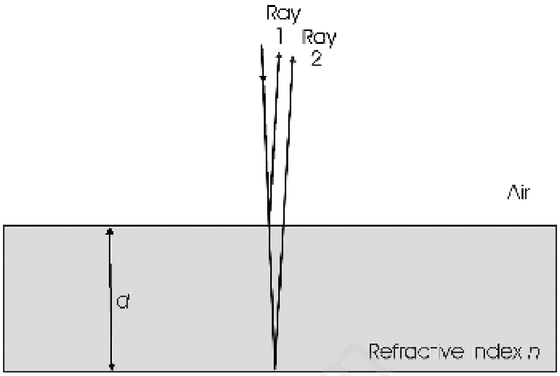

Figure 14.19 Reflection at a thin film in air

$$[p] = 2nd$$

where *d* is the physical thickness, and *n* is the refractive index of the film. If the path difference, [*p*], is equal to an integral number of wavelengths the film will appear dark as a result of destructive interference:

$$[p] = m\lambda, \quad m = 1, 2, 3, \ldots, \quad \text{minimum (dark)}$$

If the path difference, [*p*], is equal to a half-integral number of wavelengths the film will then appear bright, because constructive interference will occur:

$$[p] = \left(m + \tfrac{1}{2}\right)\lambda$$
$$m = 1, 2, 3, \ldots, \quad \text{maximum (bright)}$$

At other path differences, the film will appear to have an intermediate tone, depending on the exact phase difference between the rays.

Should the light beam fall on the surface at an angle of incidence θ, producing an angle of refraction θ', the path difference, $[p]$, between rays 1 and 2 is:

$$[p] = 2nd \cos \theta'$$

Thus

$$[p] = 2nd \cos \theta' = \begin{cases} m\lambda \text{ gives a minimum (dark);} \\ \left(m+\frac{1}{2}\right)\lambda \text{ gives a maximum (bright).} \end{cases}$$

14.5.3 The reflectivity of a single thin film in air

The reflectivity of a thin film in air will be different from that for a thick plate, given in Section 14.5.1, as interference effects from the bottom surface also need to be considered.

For light at normal incidence on a transparent solid, the reflectivity is given by:

$$R = \frac{2r_1^2 - 2r_1^2 \cos 2\delta}{1 - 2r_1^2 \cos 2\delta + r_1^4}$$

where

$$r_1 = \frac{n_0 - n_f}{n_0 + n_f}$$

where n_0 is the refractive index of the surrounding medium, usually air ($n_0 = 1.0$) n_f is the refractive index of the film; and

$$\delta = \frac{2\pi[d]}{\lambda}$$

where $[d]$ is the optical thickness of the film, given by

$$[d] = n_f d$$

with d being the physical thickness of the film.

The reflectivity is found to vary in a cyclic fashion, with zero for values of $[d]$ equal to 0, $\lambda/2$, λ, etc. and a maximum for values of $[d]$ given by $\lambda/4$, $3\lambda/4$, etc. (Figure 14.20). Because the value of n depends on wavelength, the reflectivity will also vary across the spectrum.

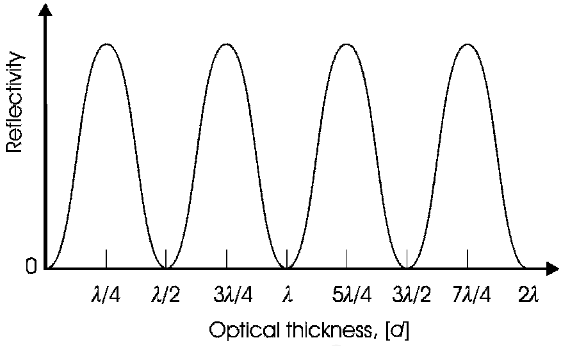

Figure 14.20 The reflectivity of a thin film varies sinusoidally with the film thickness, being zero when the optical thickness is a multiple of a half wavelength, and maximum when the thickness is an odd multiple of a quarter wavelength

14.5.4 The colour of a single thin film in air

When a thin transparent film is viewed in white light, the same reflection and interference discussed in Section 14.5.2 will occur, except that we have to take into account the effects of all of the different wavelengths present. In order to determine the reflected colour of a thin film when viewed in white light it is necessary to add the contributions of all of the wavelengths. The intensity pattern and perceived colour generated by adding the contributions as a function of the optical path difference of the film, [p], is given in Section S4.11. The sequence of colours seen will repeat in a cyclical fashion as the film thickness increases or decreases, as certain colours are either reinforced or cancelled. Each sequence of spectral colours is called an order, which starts with the first order for the thinnest of films. A new order begins every 550 nm of retardation.

Since the fraction of incident white light that is reflected is coloured, it follows that the transmitted light will be depleted in this colour. The transmitted colour seen will therefore be the complementary colour to that strongly reflected (Section S4.11).

If the angle of viewing is not perpendicular to the film, the optical path difference is given by:

$$[p] = 2[d] \cos \theta'$$

where θ' is the angle of refraction. This formula indicates that as the viewing angle moves away from perpendicular to the film the colour observed will move towards that appropriate to lower optical path length. For example, second-order orange-red will change towards green and blue.

14.5.5 *The colour of a single thin film on a substrate*

The behaviour of a single thin film on a substrate is similar to that discussed for the case of a single thin film in air. Now, however, it is necessary to take into account any change of phase that might occur on reflection at the back surface of the film. If the substrate has a lower refractive index than the film on the surface then the treatment will be identical to that for a thin film in air. In this case, the reflected colours observed when the film is viewed at normal incidence in white light will be the same as those listed in the 'colour reflected' column of Table S4.1, Section S4.11.

If the refractive index of the substrate is greater than that of the film then a phase change will be introduced both at the air–film interface and at the film–substrate interface. In this case, the reflected colour seen at normal incidence when viewed in white light will be the complementary colour to that just described. These are listed in Section S4.11 in the column labelled 'colour transmitted'.

14.5.6 *Low-reflectivity (antireflection) and high-reflectivity coatings*

The reflectivity, R, of a homogeneous nonabsorbing thin film on a substrate, illuminated by light of one wavelength perpendicular to the surface, is given by:

$$R = \frac{r_1^2 + 2r_1r_2 \cos 2\delta + r_2^2}{1 + 2r_1r_2 \cos 2\delta + r_1^2r_2^2}$$

where

$$r_1 = \frac{n_0 - n_\mathrm{f}}{n_0 + n_\mathrm{f}}$$
$$r_2 = \frac{n_\mathrm{f} - n_\mathrm{s}}{n_\mathrm{f} + n_\mathrm{s}}$$

and n_0 is the refractive index of the surrounding medium, n_f is the refractive index of the film and n_s is the refractive index of the substrate;

$$\delta = \frac{2\pi[d]}{\lambda}$$

where $[d]$ is the optical thickness of the film, given by

$$[d] = n_\mathrm{f}d$$

where d is the physical thickness of the film.

For values of $[d]$ given by $[d] = \frac{\lambda}{2}, \lambda, \frac{3\lambda}{2}$, etc.,

$$R = \frac{(n_0 - n_\mathrm{s})^2}{(n_0 + n_\mathrm{s})^2}$$

This equation is interesting, because if n_0 is set as 1.0, for air, and n_s is set as the refractive index of the material, n, it is identical to the equation for as uncoated surface, Equation (14.2). Thus a layer of optical thickness $\lambda/2$, etc. can be considered to be optically absent and the surface has normal uncoated reflectivity.

When the optical thickness of the film is $[d] = \lambda/4, 3\lambda/4$, etc.,

$$R = \left(\frac{n_\mathrm{f}^2 - n_0n_\mathrm{s}}{n_\mathrm{f}^2 + n_0n_\mathrm{s}}\right)^2$$

The reflectance will be either a maximum or a minimum. When the film has a higher refractive index than the substrate the reflectivity will be a maximum, $n_0 < n_\mathrm{f} > n_\mathrm{s}$. When the film has a lower refractive index than the substrate the reflectivity will be a minimum, $n_0 < n_f < n_s$.

To make a nonreflective coating [antireflection (AR) coating] on a glass surface in air the value of n_f must lie between that of air and the glass. The reflectivity will be a minimum for a $\lambda/4$ film. Putting R equal to zero yields a value of the refractive index of a film that will give no reflection at all:

$$n_f = \sqrt{n_s}$$

For glass, n_S is about 1.5, so the antireflecting film must have a refractive index:

$$n_f = \sqrt{1.5} = 1.225$$

Very few solids have such a low index of refraction, and a compromise material often used is magnesium fluoride, MgF_2, for which n in the middle of the visible is 1.384.

A similar strategy can be used to optimise the reflectivity of a surface by coating it with a high-reflection coating. The object is to make the reflectivity, R, as close to 1 as possible. A film of thickness $\lambda/4$ will increase the reflectivity in the case when the refractive index of the film, n_f, is greater than the refractive index of the substrate, n_S. Two materials frequently used are SiO ($n = 2.0$) and TiO_2 ($n = 2.90$). A TiO_2 film of thickness $\lambda/4$ on glass will have a reflectivity of about 0.48 (48 %). As R for a single glass surface in air is about 0.04 (4 %), almost 48 % represents a great improvement. The effect is used in costume jewellery. Rhinestones are made of glass coated with an approximately $\lambda/4$ thickness film of TiO_2.

14.5.7 Multiple thin films and dielectric mirrors

Multiple thin films of transparent materials can be laid down one on top of the other in such a way as to form perfect mirrors. These are often called dielectric mirrors. The simplest formulae for the reflectance of such a mirror refers to the specific case in which all layers are $\lambda/4$ thick and of alternating high (H) and low (L) refractive indices, n_H and n_L, illuminated by light falling *perpendicular* to the surface (Figure 14.21). The arrangement is called a quarter-wave stack. For a quarter-wave stack deposited on a substrate in the sequence:

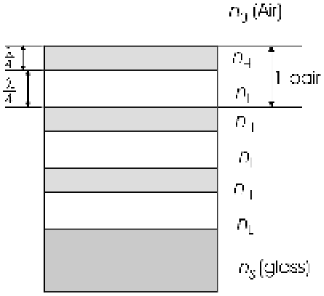

Figure 14.21 A quarter-wave stack of alternating-layers of high and low refractive index solids, n_H, and n_L, each of optical thickness one quarter of a wavelength, $\lambda/4$

substrate; L; H; L; H; L; H; . . . , L; H; air

maximum reflectance is given by the formula:

$$R = \left(\frac{n_s f - n_0}{n_s f + n_0}\right)^2$$

where f is equal to $(n_H/n_L)^{2N}$, n_0 is the refractive index of the surrounding medium, usually air ($n_0 = 1.0$), n_s is the refractive index of the substrate, usually glass ($n_s \approx 1.5$), and N is the number of LH pairs of layers in the stack. For a stack in air:

$$R = \left[\frac{n_s - \left(\frac{n_L}{n_H}\right)^{2N}}{n_s + \left(\frac{n_L}{n_H}\right)^{2N}}\right]^2$$

The general approach used to make a dielectric mirror is to lay down a stack of thin films that have alternately higher and lower refractive indices. Manipulation of the thickness and the refractive index of each layer in the stack allows the optical properties to be modified at will to produce virtually perfect mirrors and virtually perfect antireflection coatings – both of which can be tuned to respond to very specific wavelengths – as well as a variety of optical filters. The fabrication of such devices falls into the area of photonic or thin-film engineering. Filters utilising multilayers, sometimes referred to as interference filters, fall into three different categories (Figure 14.22). Shortpass filters transmit visible wavelengths and cut out infrared radiation (Figure 14.22a). They are often used in surveillance cameras to eliminate heat radiation. Longpass filters block ultraviolet radiation and transmit the visible (Figure 14.22b). Bandpass filters pass only a limited section (or band) of the electromagnetic spectrum, (Figure 14.22c).

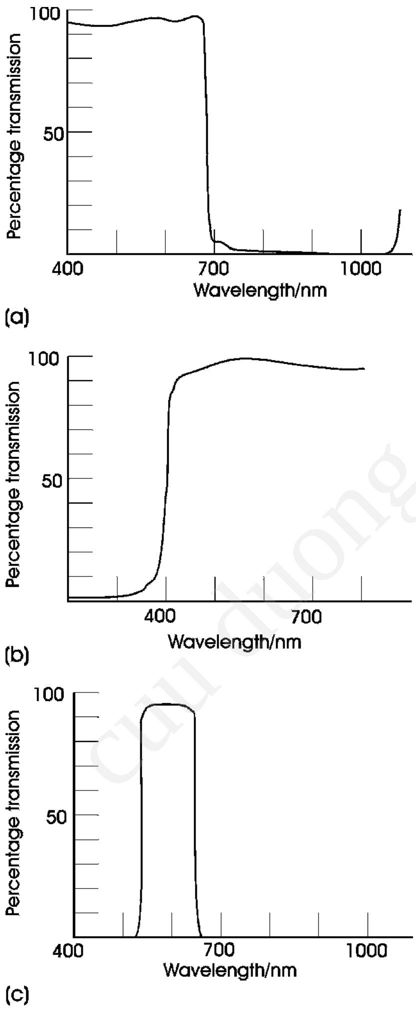

Figure 14.22 The transmission characteristics of interference filters: (a) a shortpass filter, (b) a longpass filter, and (c) a bandpass filter

The reflectivity of a stack of transparent thin films that are not carefully engineered in terms of thickness or refractive index will give a reflection that appears to be metallic silver. This can be seen in a less-than-perfect fashion with a stack of microscope slides or a roll of thin transparent plastic film, such as 'cling film'.

14.6 Scattering

14.6.1 Rayleigh scattering

If a transparent medium contains scattering centres, the intensity of light traversing the medium in the incident direction will gradually fall as the light is scattered into other directions. The reduction in the intensity of such a beam can be written as:

$$I = I_0 \exp\left(-\alpha_s l\right)$$

where I_0 is the incident beam intensity, I is the intensity after travelling a distance l in the turbid medium, and α_s is an experimentally determined linear scattering coefficient. The form of this equation is identical to that of Lambert's law for absorption.

Rayleigh scattering applies to spherical insulating particles with a diameter less than about a tenth of the wavelength of the incident light. When a beam of unpolarised light of intensity I_0 is scattered *once only* by the scattering centre, the intensity of the scattered light, I_s at a distance r from the scattering centre is:

$$I_s = I_0\left(\frac{9\pi^2 V^2}{2r^2\lambda^4}\right)\left(\frac{\mathrm{m}^2-1}{\mathrm{m}^2+2}\right)^2\left(1+\cos^2\theta\right)$$

where V is the volume of the scattering particle, λ is the wavelength of the light, θ is the angle between the incident beam and the direction of the scattered beam, and m is the relative refractive index of the particle:

$$m = \frac{n_{\text{particle}}}{n_{\text{medium}}}$$

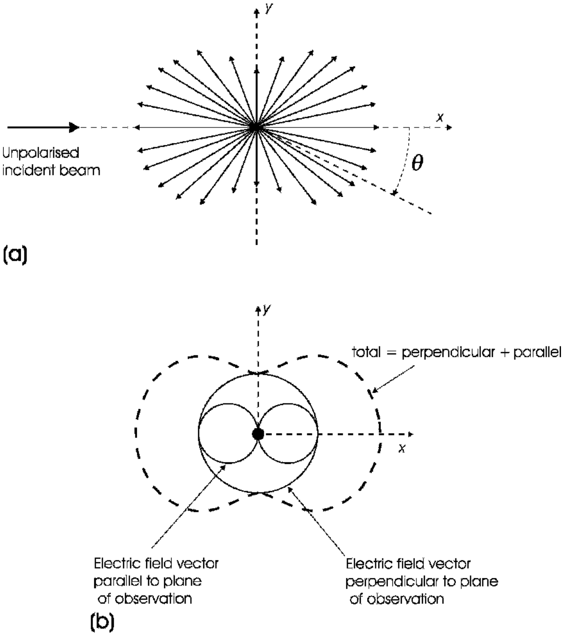

Figure 14.23 (a) The Rayleigh scattering pattern from a small spherical particle; (b) the scattering pattern is the sum of radiation scattered with its electric field vector parallel and perpendicular to the plane of observation

In this case, n_{particle} is the refractive index of the particle, and n_{medium} is the refractive index of the surrounding medium. For air, n_{medium} is 1.0.

As much light is scattered backwards as forwards, and only half as much intensity is scattered normal to the beam direction (Figure 14.23). All wavelengths scatter in this pattern, but the shorter wavelengths are more strongly scattered than are the longer wavelengths. Because the scattering of light is proportional to $1/\lambda^4$, violet light is scattered far more than is red light. The blue appearance of the sky on a sunny day is the result of the preferential scattering of violet light combined with the maximum sensitivity of the eye, which lies in the green-yellow region. Similarly, the red skies visible at dawn and dusk, and the rare blue moon, are caused by scattering of light from small particles in the upper atmosphere.

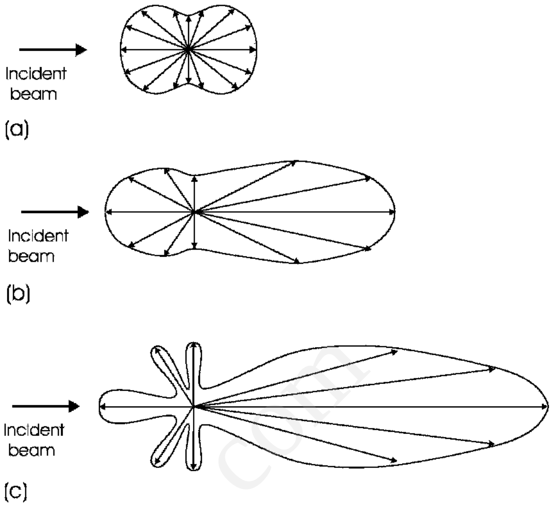

Figure 14.24 (a) The Rayleigh scattering pattern from a small particle. As the particle size increases, the scattering pattern becomes asymmetrical parts (b) and develops side lobes (part c). The scattering in parts (b) and (c), by particles larger than that in (part a), is called Mie scattering

14.6.2 Mie scattering

The term Mie scattering is generally reserved for scattering by particles that are somewhat larger than those for which Rayleigh scattering is valid, about a third the wavelength of light or more. As the particle size increases from that appropriate to Rayleigh scattering, forward-scattering begins to dominate over backward scattering (Figure 14.24). As particle size passes the wavelength of light, the forward-scattering lobes increase further, and side bands develop, due to maxima and minima of scattering at definite angles. The position of these lobes depends on the wavelength of the scattered light and so they are strongly coloured. These coloured bands, referred to as higher-order Tyndall spectra, are dependent on the particle size. Mie scattering is responsible for the colours produced in ruby glass, which contains a dispersion of gold particles as the scattering centres.

With even larger particles, white light becomes reflected (rather than scattered, as we are discussing here) evenly in all directions. This situation holds in fogs and mists.

14.7 Diffraction

Diffraction effects occur when waves interact with objects having a size similar to the wavelength of the radiation. In general, two regimes have been explored in most detail: (1) diffraction quite close to the object which interacts with the light, called Fresnel diffraction; and (2) the effects of diffraction far from the object which interacts with the light, called Fraunhofer diffraction. The result of diffraction is a set of bright and dark fringes, due to constructive and destructive interference, called a diffraction pattern.

14.7.1 Diffraction by an aperture

If a long narrow slit is illuminated by monochromatic light the intensity pattern observed far from the slit (the Fraunhofer diffraction pattern) is given by the expression:

$$I_x = I_0 \left[\frac{\sin x}{x}\right]^2$$

where

$$x = \frac{\pi w \sin\theta}{\lambda}$$

and w is the width of the slit, θ is the angular deviation from the 'straight-through' position, and λ is the wavelength of the light. This produces a set of bright and dark fringes with minima given by:

$$\sin\theta_{\min} = \frac{m\lambda}{w}$$

where m takes values 1, 2, 3, etc. For $\theta_{\min}$ to be appreciable, w must be close to λ. In fact, the formula shows that the spacing between the minima will be proportional to the reciprocal of the slit width, so that the narrower the opening the wider the fringe spacing (Figure 14.25). The positions of the maxima between these dark bands are not given by a simple formula, but are approximately midway between the minima.

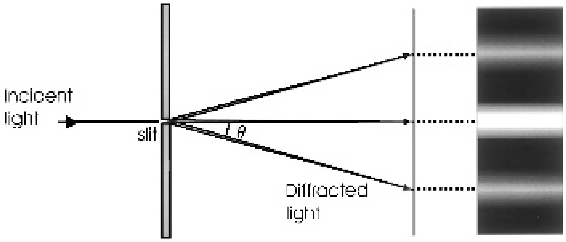

Figure 14.25 The diffraction pattern from a thin slit shows regions of higher and lower intensity

The sine of the angle through which a ray is diffracted is related to its wavelength. This indicates that each wavelength in white light will be diffracted through a slightly different angle and that red light will be diffracted through a greater angle than will violet light. In this way, white light will produce a set of diffraction patterns, each belonging to a different wavelength. These patterns look like, and are called, spectra. They are referred to as first-order, second-order and so on as they are recorded further and further from the undeviated beam.

When the slit is shortened to form a rectangular aperture the diffraction maxima will take the form of small rectangular spots running in two perpendicular directions. White light will produce coloured spots via the same mechanism as described above.

The form of the diffraction pattern produced by a circular aperture consists of a series of bright and dark circles concentric with the original aperture. The spacing of the maxima and minima is given by:

$$\sin\theta = \frac{n\lambda}{d}$$

where θ is the angle between the directly transmitted ray and the diffraction ring, λ is the wavelength of the light, and d is the diameter of the aperture. The computation of n requires rather sophisticated mathematics, the results of which show that n takes the values 0 (central bright spot), 1.220 (first dark ring), 1.635 (first bright ring), 2.333

(second dark ring), 2.679 (second bright ring) and 3.238 (third dark ring).

Just as with the slit, the dependence of the diffraction angle on wavelength means that a circular aperture illuminated with white light will produce a set of coloured rings, rather like miniature circular rainbows. The formula indicates that each ring will have a violet inner edge and a red outer edge.

Diffraction by circular apertures plays an important part in the overall performance of many optical instruments such as telescopes and microscopes. The resolution of such instruments, which is, roughly speaking, equivalent to the separation of two points which can just be distinguished as separate objects, is controlled by diffraction. It is of the order of the wavelength of the observing radiation. Because of this limitation, optical microscopes are unable to image atoms. Electron microscopes, using radiation with a wavelength of the order of 0.002 nm, are able to do so.

14.7.2 Diffraction gratings

Planar diffraction gratings consist of an object inscribed with a set of parallel lines with spacing similar to that of the wavelength of light. A transmission grating has alternating clear and opaque lines, and diffraction effects are observed in light transmitted by the clear strips (Figure 14.26a). A reflection grating consists of a set of grooves or blazes, and diffraction effects are observed in the light reflected from the patterned surface (Figure 14.26b). The effectiveness of a grating is the same whether light is transmitted through it or reflected from it.

The positions of the diffraction maxima from a transmission grating illuminated by monochromatic light normal to the surface is given by:

$$\sin\theta = \frac{n\lambda}{d}$$

where d is the repeat spacing of the grating, λ is the wavelength of the radiation, and θ is the angle through which the beam has been diffracted (Figure 14.26). The positions of the maxima for light at grazing incidence to a reflection grating are given by:

$$1 - \cos\theta = \frac{n\lambda}{d}$$

where d is the repeat spacing of the grating, λ is the wavelength of the radiation, and θ is the angle through which the beam in question has been diffracted.

The term n in both formulae can take integer values of 0, ±1, ±2 and so on. Each of these corresponds to a different diffraction maximum, called an order. When illuminated by white light, each wavelength will be diffracted through a slightly different angle so that each order will consist of a spectrum, similar to those produced by a long narrow slit but, because each line on the grating acts as a contributing slit, they are of much greater intensity.

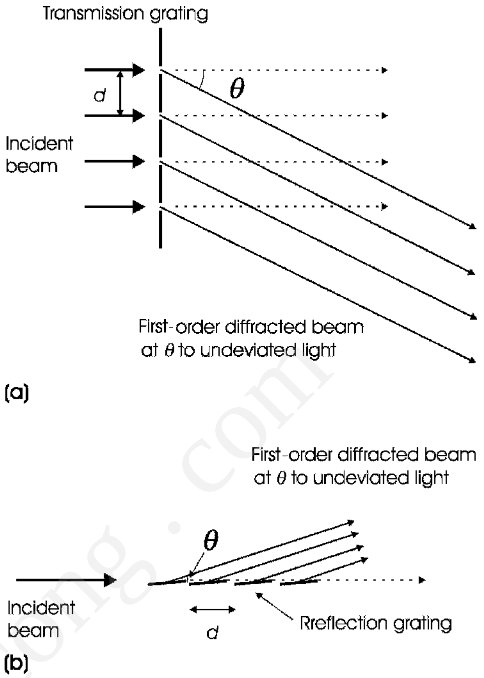

Figure 14.26 (a) Diffraction by a transmission grating and (b) diffraction by a reflection grating

14.7.3 Diffraction from crystal-like structures

Crystal structures can be determined by X-ray diffraction (Section 5.2.2). The position of the strongly diffracted beams is given by Bragg's law:

$$n\lambda = 2d \sin\theta$$

where d is the separation of the planes of atoms that are responsible for the diffraction, λ is the X-ray wavelength, and θ is the angle between the X-ray beam and the atom planes (Section 5.2.2). The theory holds for any three-dimensional array no matter the size of the 'atoms'. Thus, any arrangement of particles, or even voids, which are spaced by distances similar to the wavelength of light, will diffract light according to Bragg's law. When white light is used, each wavelength will diffract at a slightly different angle, and colours will be produced.

The colour of precious opal is due to the diffraction of white light. The regions producing the colours are made up of an ordered packing of spheres of silica (SiO_2), which are embedded in amorphous silica or a matrix of disordered spheres (Figure 14.27). These small volumes of ordered spheres resemble small crystallites. They interact with light because the spacing of the ordered spheres is similar to that of the wavelength of light.

The conditions under which diffraction takes place are the same as those specified by the Bragg equation. However, because the diffraction takes place within a silica matrix, it is necessary to use the optical path instead of the vacuum path. The layer spacing, d, must be replaced by the optical thickness, $[d]$, equal to $n_s d$, where n_s is the refractive index of the silica in opal, about 1.45. The correct equation to use for opal is thus:

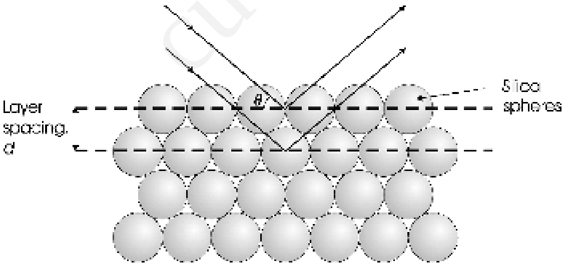

Figure 14.27 The diffraction of light by ordered arrays of silica spheres gives colour to precious opals. The geometry of the diffraction is identical to that of the diffraction of X-rays

$$\begin{aligned} n\lambda &= 2n_s d \sin\theta \\ &\approx 2.9\, d \sin\theta \end{aligned}$$

with

$$\begin{aligned} \lambda_{\max} &= 2n_s d \\ &\approx 2.9d \end{aligned}$$

The relationship between the radius of the spheres, r, and the distance between the layers, d, will depend on the exact geometry of the packing. If each layer of spheres is arranged in hexagonal closest packing, the relationship between the sphere radius and the layer spacing is:

$$d = \frac{2\sqrt{2}r}{\sqrt{3}} = 1.633\, r$$

A useful general relationship is that the radius of the spheres is given, to a reasonable approximation, by one fifth of the wavelength of the colour observed at normal incidence.

14.7.4 Photonic crystals

Photonic crystals are artificial structures that diffract light in specified ways. The dimensions of the diffracting centres in the 'crystals' are approximately the same as the wavelength of light, and the diffraction can generally be understood in terms of the Bragg equation. For example, artificial opals are photonic crystals, and Section 14.7.3 has explained how these diffract light of certain wavelengths. However, the terminology employed to describe diffraction in photonic crystals is that of semiconductor physics. The transition from a diffraction description to a physical description can be illustrated with respect to a one-dimensional photonic crystal.

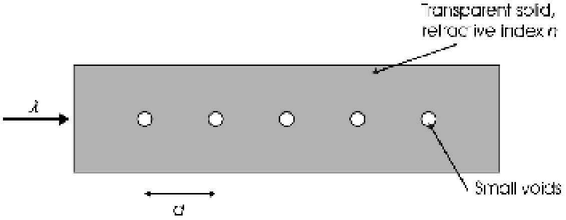

Figure 14.28 A one-dimensional photonic crystal can be thought of as a regularly spaced linear array of diffracting particles, such as small voids in a transparent solid

A one-dimensional photonic crystal is simply a stack of transparent layers of differing refractive indices. They are also called Bragg stacks or, when built into an optical fibre, fibre Bragg gratings. The simplest model is that of a transparent material containing ‘atoms’ consisting of regularly spaced air voids (Figure 14.28). When a beam of light is incident on such a grating, a wavelength, λ, will be diffracted when the Bragg law is obeyed:

$$\lambda = 2\,[d]\,\sin\theta$$

where n is the refractive index of the material, and $[d]$ is the repeat spacing. (As for opal, it is necessary to use the optical thickness, $[d] = nd$, and not the physical thickness, d, for the repeat spacing.) For a beam normal to the ‘atoms’, $\sin\theta = 1$, hence

$$\lambda = 2\,n[d]$$

and the light will be diffracted back on itself and not be transmitted (Figure 14.29).

This same idea was used in the description of Brillouin zones in crystals, in Section 2.3.9. The result of which was the creation of an energy band gap for the electrons. In terms of semiconductor physics, the array of voids has opened a photonic band gap (PBG) in the material. A photonic band gap blocks transmission of the light wave with an energy equal to the band gap. In real materials, the ‘atoms’ have thickness, and a small range of wavelengths is blocked rather than just one. The result is similar to that described for the multilayer interference filters, and the range of wavelengths blocked increases as the difference between the refractive indices of the atoms and the surrounding medium increases.

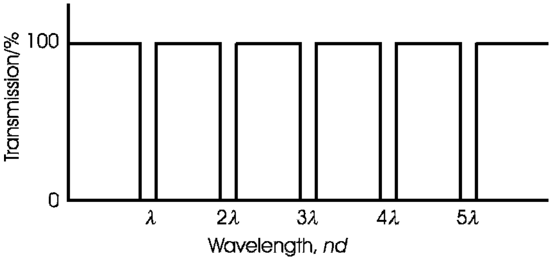

Figure 14.29 The idealised transmission profile of a one-dimensional photonic crystal. Wavelengths that are a multiple of the optical thickness between the scattering centres (voids) are completely diffracted back on themselves and not transmitted. Other wavelengths pass unhindered

Two-dimensional PBG crystals can be thought of as a two-dimensional array of ‘atoms’ in a transparent medium, and opal is an example of a three-dimensional PBG crystal. The light reflected by an opal gives a measure of the photonic band gap of the gemstone. Many insects also use ‘photonic crystal’ structures for the production of vivid colours, and a study of insect colours, especially of iridescent butterfly wings, has led to advances in understanding structures similar to those in PBG materials.

14.8 Fibre optics

14.8.1 *Optical communications*

The transmission of light along thin fibres of glass, plastic or other transparent materials is referred to as fibre optics. Data are carried by a series of pulses of light encoded so that information can be stored and retrieved. In this brief survey, the properties of the materials used in the fibre will be outlined. The engineering of an optical communications system can be explored by reference to some of the sources listed in the Further Reading section.

The transparent optical wave carrier used for communications is silica (SiO_2) glass. The light pulses launched into the fibre are constrained to stay within the fibre by total internal reflection.

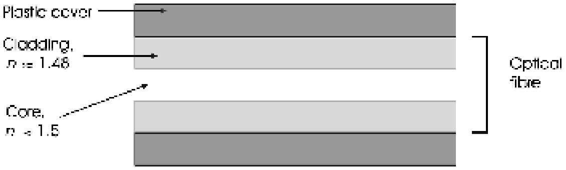

Figure 14.30 The structure of an optical fibre. The fibre is made of glass with a higher refractive index – the core – and glass of a slightly lower refractive index – the cladding. The fibre is covered in a plastic coating to protect the glass from damage

Thus, the core of the fibre, along which light travels, must possess a higher refractive index than the outer surface of the fibre. Moreover, a glass surface at which the total internal reflection is to occur is easily damaged, and needs protection. Both of these objectives are met by providing a surface cladding of lower refractive index glass compared with the core of the fibre. The core and the cladding make up a single glass fibre (Figure 14.30). The cladding should not be confused with a plastic protective covering, which has no optical role to play.

14.8.2 Attenuation in glass fibres

Attenuation describes the loss of light intensity as the signal is transmitted along the fibre. This is of major concern, as any degradation of the signal must be minimised.

The unit of loss is the decibel, dB – the base unit, the bel, is almost never used. The loss is defined as:

$$\text{loss (dB)} = -10\log_{10}\left(\frac{\text{power in}}{\text{power out}}\right)$$
$$= -10\log_{10}\left[\frac{P(x)}{P(0)}\right]$$

where $P(0)$ is the power input at $x = 0$, $P(x)$ is the power at a remote point, x.

The attenuation is defined as the loss per kilometre, thus

$$\text{attenuation} = \frac{-10}{x}\log_{10}\left[\frac{P(x)}{P(0)}\right]$$

where $P(x)$ is the power at a point x kilometres along the cable. The units of attenuation are dB km^{-1}. For a material showing an attenuation of 1 dB km^{-1} an input power of 10 W would give an output power of 7.9 W after 1 km. Ordinary window glass has an attenuation of about 10 000 dB km^{-1}. Attenuation, like dispersion, varies with wavelength. The spectral response of a fibre defines the way in which the fibre attenuation changes with the frequency of the radiation being transmitted.

Attenuation is caused by a combination of absorption and scattering within the glass. Extrinsic attenuation is due to poor processing or fabrication techniques, and may be due to artefacts such as bubbles, particles, impurities and variable fibre dimensions. These problems have been eliminated in modern optical fibre manufacture. Intrinsic attenuation is a property of the pure material itself and cannot be removed by processing. It is the ultimate limit on the performance of the fibre and mainly arises from two factors, Rayleigh scattering and lattice vibrations.

Rayleigh scattering arises from small inhomogeneities in the glass, which cause changes in refractive index. This variation is an inevitable feature of the noncrystalline state and cannot be removed by processing. As Rayleigh scattering is proportional to λ^{-4}, where λ is the wavelength of the optical pulse, the effect is more important for short-wavelength radiation. For any particular glass, most of the factors affecting Rayleigh scattering are constant and cannot be easily changed. However, materials with a low refractive index and glass transition temperature tend to exhibit low Rayleigh scattering.

Absorption due to lattice vibrations, referred to as phonon absorption, occurs when the lattice vibrations of the solid match the energy of the radiation. This occurs for infrared wavelengths and converts the signal energy into heat. It is a function of the mass of the atoms in the glass and the strength of the chemical bonds between them and results in a decrease in the transparency of the glass at long wavelengths.

Absorption due to electronic transitions, mostly at high energies and associated with ultraviolet wavelengths, do not figure significantly in present-day applications but may become important if shorter signal wavelengths are to be used in the future. The dependence on wavelength of absorption due to electronic transitions can often be expressed by a formula of the type:

$$\text{Electronic absorption} = B_1 \exp\left(\frac{B_2}{\lambda}\right)$$

where B_1 and B_2 are constants relating to the glass used, and λ is the wavelength of the radiation.

Attenuation in early fibres was mainly due to metallic impurities. The gravest problem was iron, present as Fe^{2+} – the ion that also imparts a greenish tint to window glass. Even a concentration as low as 1 part per million (ppm) of iron can result in an attenuation of 15 dB km^{-1}. The presence of transition metal cations was avoided by the preparation of silica from very high-purity chemicals made available by the semiconductor industry.

The most important impurity in silica fibres today is hydroxyl (—OH). Hydroxyl arises from water or hydrogen incorporation into the glass during fabrication. An impurity level of 1 ppm can give an attenuation of 10^4 dB km^{-1} at a 1.4 μm signal wavelength. Despite careful processing, fibres currently in production still contain significant amounts of this impurity, and hydroxyl remains an important source of attenuation.

By 1979, the best silica fibres showed only intrinsic attenuation and had a loss of about 0.2 dB km^{-1} at 1.5 μm wavelength. This is currently the industry standard.

14.8.3 Dispersion and optical fibre design

A short pulse of light launched into a fibre will tend to spread out, as a result of dispersion. In optical fibres, the dispersion is defined as the delay between the arrival time of the start of a light pulse and its finish time relative to that of the initial pulse. It is measured at half peak amplitude. If the initial pulse has a spread of t_i seconds at 50 % amplitude and the final pulse a spread of t_f seconds at 50 % amplitude after having travelled d kilometres, the dispersion is given by:

$$\text{dispersion} = \frac{t_f - t_i}{d}$$

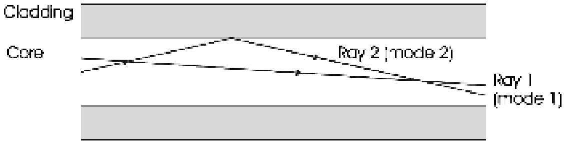

Figure 14.31 The allowed paths that light beams can take in the core region of an optical fibre are called modes; although drawn here as ray paths, in reality they are alternative light wave patterns in the core

The units of dispersion in optical fibres are ns km^{-1}.

Dispersion will result if the light source is not strictly monochromatic. An initially sharp pulse consisting of a group of wavelengths will spread out as it travels down the fibre, because the refractive index depends on wavelength. Thus, different wavelengths will travel at different speeds. This effect is known as wavelength dispersion.

Even with completely monochromatic light, pulse spreading can still occur, because the radiation can take various paths, or modes, through the fibre, as sketched in Figure 14.31. It is apparent that a ray that travels along the axis of a fibre will travel less than one that is continually reflected on its journey. [In fact, the dispersion that results cannot be properly understood in terms of the transmission of light rays, and the various modes are better described in terms of the allowed wave patterns that can travel down the fibre.] The resultant pulse broadening, due to the various modes present, is called modal (or intermodal) dispersion. In order to overcome modal dispersion a number of different fibre types have evolved.

The earliest fibres were called stepped index multimode fibres. These fibres have a large core region, allowing many modes to propagate (Figure 14.32a). The ray labelled H in Figure 14.32(a) is known as a high-order mode, whereas the ray L is a low-order mode. Stepped index multimode fibres are easy to make and join but have a lower performance compared with those described below.

The first advance on stepped index fibres was the graded index fibre. In this design, the refractive index of the fibre varies smoothly, from high at the centre to low at the periphery of the core region (Figure 14.32b). The refractive index gradient

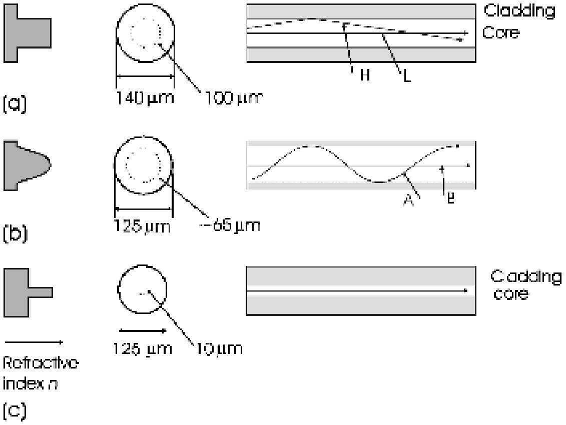

Figure 14.32 Types of optical fibre: (a) stepped index fibre, (b) graded index fibre and (c) monomode fibre

means that light travels faster as it approaches the edge regions of the fibre. The velocity of mode A will vary smoothly from lowest at the fibre centre to greatest near to the fibre edge. The velocity of mode B will be fairly constant and lower on average than mode A. The differences in path length between high-order and low-order modes is thus minimised by this velocity variation.

For best results, monomode fibres (Figure 14.32c) are now used. The number of possible modes is reduced by decreasing the diameter of the core. When the core diameter reaches 10 µm or less only one mode can propagate and, in principle, modal dispersion is zero for these fibres. Monomode fibres have a high performance but are harder to make and join.

The material used for optical communications fibre is highly purified silica (SiO_2) glass. The cladding and core regions are created by doping with carefully chosen impurities. A commonly used production method utilises a tube of pure silica glass. Layers of germanium dioxide, GeO_2, are laid down in the centre of the tube. Germinium dioxide is chemically and physically very similar to silica and readily forms a solid solution with the silica glass. As the germanium atoms are heavier than silicon atoms, they increase the refractive index of the doped inner region relative to the undoped outer region of the tube. When a sufficient amount of GeO_2 has been laid down, the tube is heated until it collapses into a solid rod called a preform. The preform has a germanium-rich higher refractive index core and a lower refractive index periphery. The preform is drawn into a fibre and, because of the nature of the way in which glass flows at elevated temperatures, the refractive index profile of the preform is maintained in the fibre.

Although fibres in commercial use are made of silica glass, it is not perfect. The dispersion is lowest at 1.3 µm, but the minimum attenuation occurs at 1.5 µm, leading to some sacrifice of performance irrespective of the signal wavelength chosen. The search for new materials to resolve this conflict continues in many research laboratories.

14.8.4 Optical amplification

The amplification of signals in fibre optic transmission systems is of great importance. Amplification uses a section of optical fibre doped with erbium (Er^{3+}) as the activator. The amplifying section consists of about 30 m of monomode fibre core containing just a few hundred parts per million of Er^{3+} (Figure 14.33a). This section of the fibre is illuminated by a semiconductor diode laser at the frequency related to that of the carrier signal. The commonest wavelengths used are 980 nm and 1480 nm. The erbium ions have the remarkable ability to transfer energy from the laser to the signal pulses as they traverse this section of fibre.

The energy transfer comes about in the following way. With reference to the schematic energy-level diagram of the Er^{3+} ion (Figure 14.33b), illumination of the erbium-containing section of fibre with energy of wavelength 980 nm excites the ions from the ground state (${}^4I_{15/2}$) to the upper state (${}^4I_{11/2}$) from whence they rapidly decay to the ${}^4I_{13/2}$ level shown by the dashed diagonal line. The use of radiation of 1480 nm wavelength excites the Er^{3+} ions directly from the ground state to the ${}^4I_{13/2}$ level. This process is referred to as pumping, and the laser involved is described as the pump. The excited state has quite a long lifetime, and so a passing light pulse, with a wavelength close to 1480 nm, can empty it via stimulated emission. This achieves

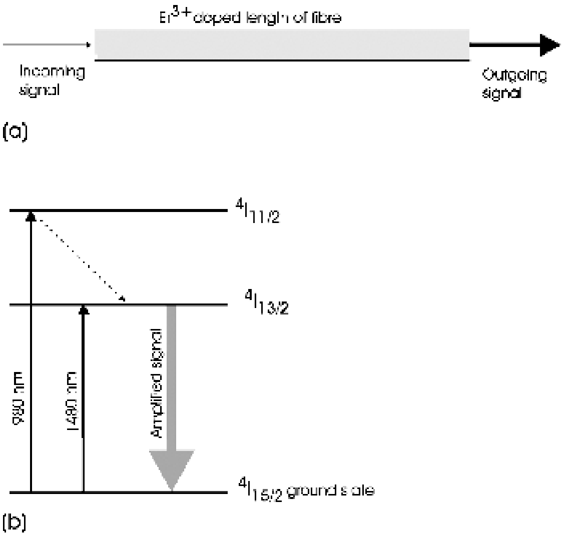

Figure 14.33 Signal amplification in an optical fibre using erbium doping: (a) an incoming weak signal is amplified on passage through a length of fibre in which the core has been doped with Er^{3+} ions. (b) The principle energy levels used by the Er^{3+} amplifier. Pump wavelengths of 980 nm and 1480 nm (narrow, upward-pointing arrows) excite the Er^{3+} ions. This energy is transferred to the signal as the Er^{3+} ions lose energy (wide, downward-pointing arrow)

signal amplification while retaining the coherence of the pulse constituting the signal.

14.9 Nonlinear optical materials

14.9.1 Nonlinear optics

The electric field of a light beam induces a polarisation in a solid. For light beams of ordinary intensity the polarisation, $\boldsymbol{P}$, is a linear function of the electric field, $\boldsymbol{E}$:

$$\boldsymbol{P} = \varepsilon_0 \chi \boldsymbol{E}$$

where ε_0 is the permittivity of free space, and χ is the dielectric susceptibility of the material, which is proportional to the relative permittivity and refractive index of the substance. This approximation is perfectly adequate for normal optics, but laser beams in particular can be associated with very high electric fields and, in this case, it is necessary to write the polarisation as a series:

$$\boldsymbol{P} = \varepsilon_0 \chi^{(1)} \boldsymbol{E} + \varepsilon_0 \chi^{(2)} \boldsymbol{E}^2 + \varepsilon_0 \chi^{(3)} \boldsymbol{E}^3 + \cdots \qquad (14.3)$$

where $\chi^{(1)}$ is the linear dielectric susceptibility, $\chi^{(2)}$ is the second-order dielectric susceptibility, $\chi^{(3)}$ is the third-order dielectric susceptibility and so on. The polarisation is no longer a simple linear function of the electric field.

The extra 'nonlinear' terms are only high enough to be of importance in relatively few materials. In general, the magnitudes of the values of the dielectric susceptibilities decrease rapidly as the order increases, so that the second-order, $\chi^{(2)}$, term is the most important nonlinear coefficient. Moreover, all even-order terms, including the second order, $\chi^{(2)}$, term, are zero in centrosymmetric crystals. The second-order term, $\chi^{(2)}$, has a nonzero value in noncentrosymmetric crystals, and it is these that are generally known as nonlinear optical materials.

The nonlinear terms in the polarisation equation allow photons to be added and subtracted in certain specific ways. For example, if the crystal is irradiated with laser light characterised by angular frequencies ω_1 and ω_2 a collection of frequencies $2\omega_1$, $2\omega_2$, $\omega_1 + \omega_2$ and $\omega_1 - \omega_2$ can all be produced. The production of a frequency $2\omega_1$, the second harmonic, from a single input frequency ω_1, is known as frequency doubling or second harmonic generation (Figure 14.34a). The production of the other frequencies, which can also occur, is known as frequency mixing (Figure 14.34b).

Second harmonic generation comes about in the following way. The sinusoidally varying electric field associated with a light beam can be written:

$$\boldsymbol{E} = \boldsymbol{E}_0 \cos \omega t$$

where ω is the angular frequency of the light. If this is substituted into equation (14.3):

$$\boldsymbol{P} = \varepsilon_0 \chi^{(1)} \boldsymbol{E}_0 \cos \omega t + \varepsilon_0 \chi^{(2)} (\boldsymbol{E}_0 \cos \omega t)^2 + \varepsilon_0 \chi^{(3)} (\boldsymbol{E}_0 \cos \omega t)^3 + \cdots$$

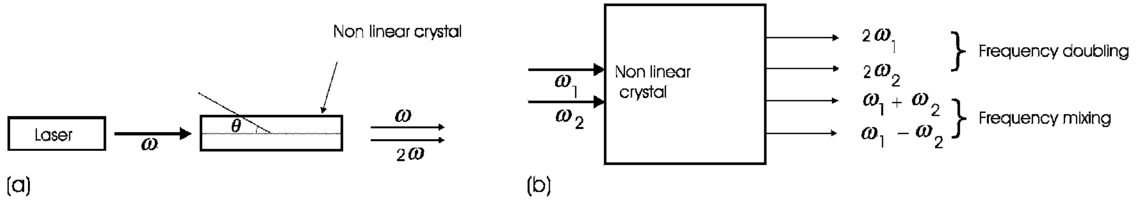

Figure 14.34 (a) Frequency doubling by a nonlinear crystal: an input signal of angular frequency ω, from a laser, is partly converted into a signal with angular frequency 2ω on passing through the crystal. (b) Frequency mixing by a nonlinear crystal: signals of angular frequencies ω_1 and ω_2 are partly mixed to produce signals $(\omega_1 + \omega_2)$ and $(\omega_1 - \omega_2)$ on passing through a nonlinear crystal as well as producing frequency-doubled signals

This series can be written more simply as:

$$\boldsymbol{P} = A + B\cos\omega t + C\cos 2\omega t + D\cos 3\omega t + \cdots$$

where A, B, C and D are constants. At field strengths of the order of those found in laser light the wave that emerges from the crystal can have both the ω and the higher 2ω, 3ω and even greater angular frequencies present if the symmetry is suitable. [Because $\chi^{(2)}$ is zero in centrosymmetric crystals, a second harmonic is not produced, and the generation of a second harmonic from a crystal when illuminated by a laser is usually taken as a good test for the lack of a centre of symmetry.]

If two input waves are used, frequency mixing can occur. Suppose the crystal is irradiated with two beams simultaneously:

$$\boldsymbol{E}_1 = \boldsymbol{E}_{01}\cos\omega_1 t$$
$$\boldsymbol{E}_2 = \boldsymbol{E}_{02}\cos\omega_2 t$$

The electric field in the sample is then:

$$\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2 = \boldsymbol{E}_{01}\cos\omega_1 t + \boldsymbol{E}_{02}\cos\omega_2 t$$

Substituting this into Equation (14.3) will yield a second-order polarisation, $\boldsymbol{P}$:

$$\begin{aligned}\boldsymbol{P} &= \varepsilon_0\chi^{(2)}[(\boldsymbol{E}_{01}\cos\omega_1 t)(\boldsymbol{E}_{02}\cos\omega_2 t)] \\ &= \varepsilon_0\chi^{(2)}(\boldsymbol{E}_{01}\boldsymbol{E}_{02}\tfrac{1}{2}[\cos(\omega_1+\omega_2)t + \cos(\omega_1-\omega_2)t]\end{aligned}$$

The term

$$\tfrac{1}{2}[\cos(\omega_1+\omega_2)t + \cos(\omega_1-\omega_2)t]$$

corresponds to the production of two output waves, one of frequency $(\omega_1 + \omega_2)$, the sum wave, and one of frequency $(\omega_1 - \omega_2)$, the difference wave.

14.10 Energy conversion

14.10.1 *Photoconductivity and photovoltaic solar cells*

If radiation of a suitable wavelength falls on a semiconductor it will excite electrons across the band gap. The most obvious effect of this is that the conductivity of the material increases. The magnitude of the effect is roughly proportional to the light intensity. This effect, called the photoconductive effect, has been used in light meters, exposure meters and automatic shutters in cameras and many other devices. In practice, a dc voltage is applied to the ends of a semiconductor. On illumination the resistance of the semiconductor falls, which provides a means of measurement.

A p–n junction can act in a similar way to a single piece of semiconductor. However, the control afforded by the junction makes the device, called a photodiode, far more flexible and, as a result, photodiodes are used in a number of devices, including solar cells. A solar cell is a large-area p–n junction. It is fabricated so that the depletion region is approximately 500 nm thick, and the electric field across the junction is high (Figure 14.35a). The junction is not connected to any

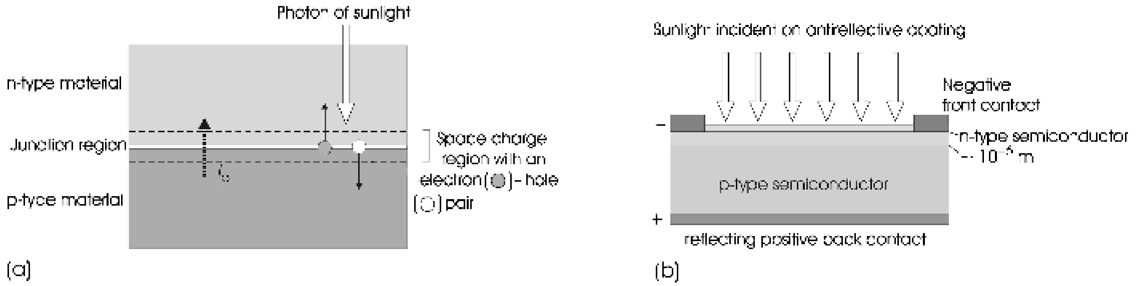

Figure 14.35 (a) Sunlight falling on a p–n junction creates an electron – hole pair. These are swept into the external circuit by the field in the junction region, generating a photocurrent, I_p. (b) A solar cell needs a thin antireflection coating on the front surface, a thin n-type layer and a junction region near to the front surface. A reflecting layer below the cell helps to increase efficiency by reflecting back photons that have not been absorbed in the semiconductor

external power source. Holes and electrons produced in the junction region by sunlight are swept across the depletion region by the high field present, the electrons going from p to n and the holes from n to p. This charges the p region more positive and the n region more negative, and produces a current, I_p, across the junction which generates a photovoltage. Materials that allow a voltage to be produced on illumination are called photovoltaic materials. The photovoltage corresponds to a forward bias, and so will cause a current I to flow. At equilibrium $I = I_p$. Should an external load, R, be connected, some current can flow through it, and so do useful work.

A number of practical considerations influence the design of solar cells (Figure 14.35b). Obviously, the band gap of the semiconductor must be such that as much as possible of the solar spectrum is absorbed. At sea level, the energy available in sunlight amounts to approximately 1000 W m^{-2}, and has a wavelength spread of approximately 400 nm to 2500 nm, with a peak in the yellow-green at 550 nm. Indirect band gap materials have a lower efficiency than direct band gap materials, which, from this perspective, are preferred. Materials used include copper indium selenide (CuInSe), cadmium telluride (CdTe), gallium arsenide (GaAs) and amorphous silicon. Because impurities and defects trap mobile electron and holes, which greatly reduces the efficiency of the cell, high-purity materials are mandatory.

Solar cells must have a large area, to collect as much sunlight as possible. Some systems use reflectors or lenses to 'concentrate' the sunlight on the cell surface. Moreover, all incoming photons must be utilised, and it is necessary to coat the front surface with an antireflection layer and to make the lower surface reflecting so that transmitted photons cross the junction again. Similarly, carriers must be produced in the junction region, so that they can separate and not recombine. The junction region must therefore be placed close to the upper surface, and is generally within 1000 nm of the upper surface.

Recently, much effort has been put into the construction of solar cells with use of polymers. These have the great advantages of low weight and flexibility. However, efficiencies are not yet adequate for commercial purposes.

14.10.2 *Photoelectrochemical cells*

In a conventional solar cell the conversion of the light to charge carriers is carried out by the solid semiconductor, which then has to move these away from the junction in order to obtain energy. The method of conversion of sunlight to energy of most importance on the Earth, photosynthesis, uses slightly different methods of achieving the same objective. The central reactions are oxidation and reduction. Photoelectrochemical cells aim to mimic this process. A number of cells have been devised to utilise redox reactions in a liquid electrolyte for energy conversion.

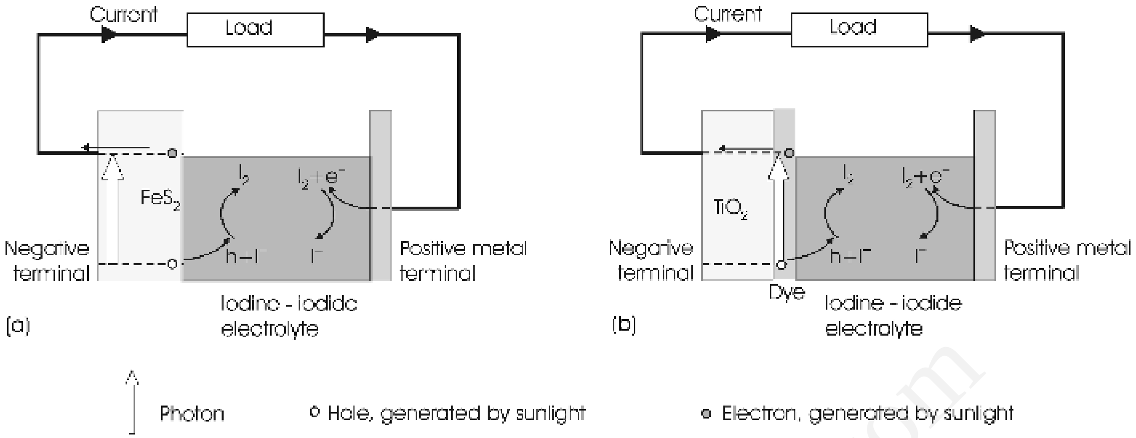

Figure 14.36 Schematic construction of (a) a pyrite–iodine (FeS_2–I_2) photoelectrochemical cell and (b) a dye sensitised titanium dioxide (TiO_2) photoelectrochemical cell. In both cells the generation of electron and hole pairs by the interaction of photons from the sun, with the semiconducting pyrite crystal in part (a) and with the dye coated onto the TiO_2 in part (b), leads to power output. The cell is completed by an internal oxidation–reduction cycle involving iodide ions. The thin arrows show the electron and hole paths in the cell

The iron pyrites cell uses FeS_2, iron pyrites, as the light absorber, and the interconversion of iodine – iodide as the redox reaction (Figure 14.36a). The pyrites crystal has a broad absorption spectrum and is able to absorb photons from across the visible spectrum to generate a supply of electrons and holes. The electrons are pulled through the crystal by the internal field, and the holes are pushed into the electrolyte. At the pyrite surface the holes (h) react with iodide ions in the following way:

$$2I^-(aq) + 2h \rightarrow I_2(aq)$$

Electrons (e^-) travel the external circuit, doing work, before entering the electrolyte to regenerate iodide ions:

$$I_2(aq) + 2e^- \rightarrow 2I^-(aq)$$

A cell that is rather closer in design to a natural photosynthesis system uses a dye to absorb the radiation. This type of cell uses titanium dioxide as the semiconductor that separates the charges. The dye is used as a more efficient photon absorber than pure titanium dioxide. The cell is shown schematically in Figure 14.36b. Because the charge separation takes place in the dye, the purity and defect structure of the solid is not crucial to satisfactory operation. The holes are pushed into the electrolyte, and the electrons travel the external circuit and power any external devices. The titanium dioxide is prepared in nanocrystalline form in order to achieve a large surface area in a compact cell. The circuit is completed with an electrolyte that can undergo oxidation and reduction. The iodine – iodide electrolyte is often used; the reactions are as above.

14.11 Nanostructures

Nanostructures are structures at a scale that gives properties to the solid that are noticeably different from those of bulk material. In optical terms, this implies dimensions of the order of the wavelength of light. Thus, these effects have been already been discussed with respect to diffraction by photonic crystals and reflection by multiple thin films. In this section the optical consequences of quantum wells, quantum wires and quantum dots, will be outlined.

14.11.1 The optical properties of quantum wells

The physical and electronic structure of quantum wells has been outlined in Sections 3.2.5 and 13.3. In a quantum well, the electrons and holes occupy energy levels that are approximately given by Equation (13.7) in Section 13.3.1, as shown in Figure 14.37. The electrons in the upper energy levels can drop to the lower hole levels and emit a photon. The energy separation of these levels is greater than that of the bulk conduction band – valence band energy gap, E_g, and hence the photons will be of higher energy, or shorter wavelength, than the bulk. The emission is said to be blue-shifted compared with the bulk, and the transitions are called interband transitions.

The photon energy derived from an interband transition is:

$$E(\text{photon}) = E_g + E_{\text{electron}} + E_{\text{hole}}$$

$$h\nu = E_g + \left(\frac{h^2}{8a^2}\right)\left(\frac{n^2}{m_e^*} + \frac{n^2}{m_h^*}\right)$$

where h is the Planck constant, ν is the frequency of the radiation, E_g is the band gap of the bulk well material, a is the dimension of the quantum well, m_e^* is the electron effective mass, and m_h^* is the hole effective mass. In the approximation that the effective mass of the electron and the hole are identical:

$$h\nu = E_g + \frac{n^2 h^2}{4a^2 m^*}$$

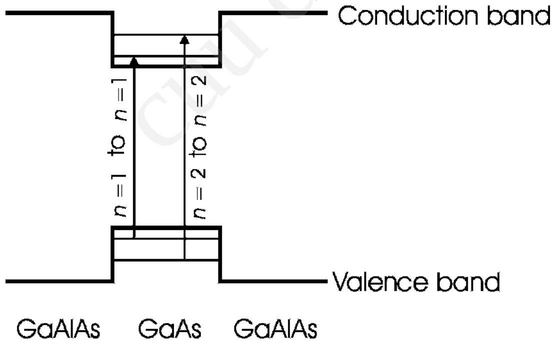

Figure 14.37 The energy levels in a quantum well. Electron transitions between the sharp energy levels resemble atomic transitions rather than the broad transitions normally seen in solids

The selection rule for the transition is $\Delta n = 0$, that is, transitions can take place only between levels with the same quantum number. (As with all selection rules, these are never perfectly obeyed, and transitions between levels with differing n values do occur infrequently, giving rise to weak lines in the emission spectrum.)

Electrons can also be excited from one electron level, say $n = 1$, to another electron level, say $n = 2$, both levels lying in the electron subband. Holes can make similar transitions between levels in the hole subband. These transitions, which give rise to extra peaks in the emission spectrum, are known as intersubband transitions.

Because the dimensions of the quantum well can be varied, the emission spectrum can be varied or tuned. This feature, in quantum wells and in quantum wires and dots, discussed below, is called band-gap engineering.

14.11.2 Quantum wires and quantum dots

Carbon nanotubes behave as quantum wires, and emission spectra from isolated carbon nanotubes have confirmed a blue shift in the radiation as expected.

The optical properties of quantum dots have been extensively investigated. They have discrete atomic-like energy levels that alter predictably as the size of the particle changes, so that the light emitted can be tuned. This finds application in many areas, including displays and lasers.

Quantum dots can be prepared in a number of ways. Conventional semiconductor techniques can be used to grow small islands of the dot on the surface of a semiconductor crystal. Isolated quantum dots are synonymous with nanoparticles, and small particles precipitated in glasses, colloids or solutions have been obtained. The most studied of these are the compounds cadmium sulphide (CdS), zinc sulphide (ZnS), cadmium selenide (CdSe) and zinc selenide (ZnSe), all of which find applications in display technology. Layers of nanocrystals, laid down by size, will give a rainbow-like emission, as each layer will emit light at a frequency appropriate to the size of the particles in the layer.

Answers to introductory questions

What are lasers?

The word laser is an acronym for the expression light amplification by stimulated emission of radiation. Lasers are devices that emit light by the process of stimulated emission. This produces light that is coherent and often polarised. The key to laser action is to obtain atoms or molecules in an excited state and to keep them there long enough for photons to pass and trigger stimulated emission. This situation is called a population inversion. A population inversion can be achieved in several ways. For solid-state lasers the high energy state that is populated is often protected by having a low transition probability to the ground state. This happens in ruby lasers. In semiconductor lasers, doping and current levels are chosen to ensure that electron numbers in the conduction band outnumber the population of holes in the valence band.

Why are thin films often brightly coloured?

Thin films are often coloured because of interference between light that is reflected from the upper surface of the film and the lower surface of the film. The extent of the interference is dependent on the film thickness. When a film of gradually increasing thickness is illuminated with monochromatic light, alternating strips of dark and light contrast are seen as a result of alternating regions of constructive and destructive interference. When white light is used, each wavelength undergoes constructive and destructive interference at slightly different film thickness. The addition of these effects results in bands of rainbow-like colours, called orders, running parallel to the thickness contours of the film.

What produces the colour in opal?

The colour of precious opal is due to the diffraction of white light by ordered crystallite-like regions in the gemstone. The regions producing the colours are made up of an ordered packing of spheres of silica (SiO_2), which are embedded in amorphous silica or a matrix of disordered spheres. They interact with light because the spacing of the ordered spheres is similar to that of the wavelength of light. The conditions under which diffraction takes place are the same as those specified by the Bragg equation. However, because the diffraction takes place within a silica matrix, it is necessary to use the optical path instead of the vacuum path. A useful general relationship is that the radius of the spheres is given, to a reasonable approximation, by one fifth of the wavelength of the colour observed at normal incidence.

Further reading

The properties of light in general are covered in:

M. Fox, 2001, *Optical Properties of Solids*, Oxford University Press, Oxford.

O.S. Heavens, R.W. Ditchburn, 1993, *Insight into Optics*, John Wiley & Sons, Chichester, Sussex.

The properties of light with respect to colour are found in:

K. Nassau, 2001, *The Physics and Chemistry of Colour*, 2nd edn, John Wiley & Sons, Chichester, Sussex.

R.J.D. Tilley, 2000, *Colour and the Optical Properties of Materials*, John Wiley & Sons, Chichester, Sussex.

A series of articles on photovoltaics is to be found in *Materials Research Society Bulletin* **XVIII** (October 1993).

A series of articles on photonic materials for optical communications is to be found in *Materials Research Society Bulletin* **28** (May 3002).

Problems and exercises

Quick quiz

1 The long wavelength part of the electromagnetic spectrum is associated with:
(a) Radiowaves
(b) X-rays
(c) Infrared radiation

2 The short wavelength region of the visible spectrum is associated with the colour:
(a) Red
(b) Violet
(c) Green

3 A beam of light is said to be coherent when:
(a) All of the waves have the same wavelength
(b) All of the waves are in step
(c) All of the waves travel at the same speed

4 The emission of light by a solid at ordinary temperatures is called:
(a) Incandescence
(b) Phosphorescence
(c) Luminescence

5 When light is emitted slowly from an excited solid it is called:
(a) Fluorescence
(b) Phosphorescence
(c) Electroluminescence

6 Some sugar is placed in a bowl in a dark room and is ground. Light is clearly emitted during this process, which is called:
(a) Triboluminescence
(b) Chemoluminescence
(c) Bioluminescence

7 The emission of light by a solid at high temperatures, for example the Sun, is called:
(a) Luminescence
(b) Incandescence
(c) Phosphorescence

8 The word laser is an acronym derived from:
(a) Light amplification by spontaneous emission of radiation
(b) Light amplification by stimulated emission of radiation
(c) Light amplification by simultaneous emission of radiation

9 The light produced by spontaneous emission is:
(a) Coherent
(b) Incoherent
(c) Partly coherent

10 The centres that produce the laser action in a ruby crystal are:
(a) Cr^{3+} ions
(b) Al^{3+} ions
(c) O^{2-} ions

11 Electron transitions involved in colour (optical transitions) are:
(a) Spin-allowed transitions
(b) Phonon-assisted transitions
(c) Radiationless transitions

12 The four-level neodymium laser utilises:
(a) p-electron transitions
(b) d-electron transitions
(c) f-electron transitions

13 The acronym LED stands for:
(a) Light emitting device
(b) Light emitting diode
(c) Laser emitting diode

14 Mixing the three additive primary colours gives:
(a) Black
(b) White
(c) No colour (colourless)

15 Mixing the three subtractive primary colours gives:
(a) Black
(b) White
(c) No colour (colourless)

16 Which of the following subtractive primary colours absorbs *green*:
  (a) Cyan
  (b) Yellow
  (c) Magenta

17 A ray of light passes from a material A, with a refractive index 1.33, to material B, with refractive index 1.5. The wavelength of the light is:
  (a) Longer in A than in B
  (b) Longer in B than in A
  (c) The same in both materials

18 The optical thickness of a film of transparent material is:
  (a) Less than the physical thickness
  (b) Greater than the physical thickness
  (c) The same as the physical thickness

19 The refractive index of a transparent material is mainly attributable to:
  (a) Ions in the material
  (b) Dipoles in the material
  (c) Electrons in the material

20 The refractive index of a transparent material:
  (a) Increases as the wavelength of the light increases
  (b) Decreases as the wavelength of the light decreases
  (c) Does not change as the wavelength of the light changes

21 A simple glass lens is used to form an image of a white circular object. The edges of the image appear coloured. Which colour is outermost:
  (a) White?
  (b) Violet?
  (c) Red?

22 Rayleigh scattering applies to particles of:
  (a) Diameter less than half the wavelength of the radiation
  (b) Diameter less than a third of the wavelength of the radiation
  (c) Diameter less than a tenth of the wavelength of the radiation

23 Mie scattering obscures a distant object:
  (a) It will be clearer if imaged in ultraviolet light
  (b) It will be clearer if imaged in infrared light
  (c) There will be no difference between imaging in ultraviolet or infrared light

24 Diffraction is a form of scattering of light that occurs when the light has:
  (a) A wavelength comparable in size to the object
  (b) A wavelength much larger than the object
  (c) A wavelength much smaller than the object

25 A beam of white light is diffracted during passage through a small circular aperture. The colour of the diffracted rings will have:
  (a) Red on the inside and violet on the outside
  (b) Violet on the inside and red on the outside
  (c) White on the outside and red on the inside

26 An opal strongly diffracts green light of wavelength 550 nm. The approximate diameter of the silica spheres in the opal is:
  (a) 110 nm
  (b) 220 nm
  (c) 550 nm

27 The cladding of an optical fibre consists of:
  (a) Air
  (b) Plastic
  (c) Glass

28 Intrinsic attenuation in an optical fibre can be caused by:
  (a) Minute air bubbles in the glass
  (b) Impurities in the glass
  (c) Density fluctuations in the glass

29 The ‘spreading out’ of a pulse in an optical communications fibre is called:
  (a) Pulse dispersion
  (b) Modal dispersion
  (c) Wavelength dispersion

30 The addition of germanium dioxide to a silica fibre:

(a) Increases the refractive index of the glass

(b) Lowers the refractive index of the glass

(c) Lowers the dispersion of the glass

31 Amplification of signals in optical communications fibres uses:
(a) Yttrium ions doped into the glass
(b) Ytterbium ions doped into the glass
(c) Erbium ions doped into the glass

32 Nonlinear optical materials are mostly:
(a) Centrosymmetric crystals
(b) Noncentrosymmetric crystals
(c) Noncrystalline

33 Photovoltaic materials:
(a) Produce a voltage when illuminated
(b) Produce light when a voltage is applied
(c) Produce chemical (oxidation–reduction) energy when illuminated

Calculations and questions

14.1 Calculate the frequency and energy of photons associated with wavelengths 425 nm, 575 nm and 630 nm. What colour is attributed to these wavelengths?

14.2 Light of wavelength 400 nm is shone through a gas of absorbing molecules. Calculate the energy absorbed by one mole of gas if each molecule absorbs one photon.

14.3 Carbon dioxide, CO_2, is a greenhouse gas that absorbs infrared radiation escaping from the Earth. What is the energy per mole absorbed by the gas if each CO_2 molecule absorbs one photon of wavelength 15 μm?

14.4 The energy required to break the bond linking the two oxygen atoms in a molecule of O_2 is 495 kJ mol^{-1}. What is the longest wavelength light that could cause this decomposition to occur?

14.5 The energy required to dissociate ozone, O_3, into O_2 plus O is 142.7 kJ mol^{-1}. What is the longest wavelength light that will dissociate ozone in the upper atmosphere?

14.6 The light absorbed by the complex ion $[FeF_6]^{3-}$ peaks at 719 nm. This absorption is due to the promotion of an electron from the lower (t_{2g}) to the upper (e_g), state in the Fe^{3+} ion.
(a) Calculate the magnitude of the crystal field splitting of the Fe^{3+} d orbitals due to F^- (see Section S4.5 for information on crystal field splitting in magnetic ions).
(b) What is the relative population of the two levels at 300 K?

14.7 The light absorbed by the complex ion $[Fe(CN)_6]^{3-}$ peaks at 333 nm. This absorption is due to the promotion of an electron from the lower (t_{2g}) to the upper (e_g) state in the Fe^{3+} ion.
(a) Calculate the magnitude of the crystal field splitting of the Fe^{3+} d orbitals due to CN^- (see Section S4.5 for information on crystal field splitting in magnetic ions).
(b) What is the relative population of the two levels at 300 K?

14.8 Calculate the wavelength at which the rates of spontaneous and stimulated emission become equal at 300 K.

14.9 The optical transitions in ruby are:
- ground state $E_0 \rightarrow E_2$, 556 nm;
- ground state $E_0 \rightarrow E_3$, 407 nm.

What are the energies of these states above the ground state? What are the relative populations N_2/N_0 and N_3/N_0 at 300 K?

14.10 The laser light from a ruby laser is at 694.3 nm.
(a) What is the energy of the lasing state, E_1, above the ground state, E_0?
(b) Estimate the fraction of Cr^{3+} ions in this upper state due to thermal equilibrium alone at 300 K.

14.11 What is the separation of the energy levels in neodymium ions that give rise to laser lines at (a) 0.914 μm and (b) 1.06 μm?

14.12 Gallium arsenide, GaAs, has a band gap of 1.35 eV, and aluminium arsenide, AlAs, has a band gap of 2.16 eV. What is the wavelength and colour of photons emitted by these solids? In order to make an orange LED with an emission at a wavelength of 600 nm, it is proposed to make a solid solution $Ga_xAl_{1-x}As$. Taking the variation in band gap with composition as linear, what value of x is required?

14.13 Gallium nitride, GaN, has a band gap of 3.34 eV, and indium nitride, InN, has a band gap of 2.0 eV. What is the wavelength and colour of photons emitted by these solids? In order to make a green LED with an emission at a wavelength of 525 nm, it is proposed to make a solid solution $Ga_xIn_{1-x}N$. Taking the variation in band gap with composition as linear, what value of x is required?

14.14 A solid solution of indium phosphide, InP, and aluminium phosphide, AlP, $In_xAl_{1-x}P$ is made up. At what value of x will the light emitted by an LED made from this compound just be visible? What colour will it be? The band gap of InP is 1.27 eV and that of AlP is 2.45 eV.

14.15 A solution is quoted as having a 22 % transmittance. What is the absorbance?

14.16 The linear absorption coefficient of zinc metal for X-rays from a nickel target is $5.187\ \mathrm{m^{-1}}$ and for cadmium metal is $18.418\ \mathrm{m^{-1}}$. What thickness of plates of these metals is needed to reduce the intensity of the radiation passing through a plate to 0.1 of the incident radiation? What will be the transmittance and absorbance of the plates?

14.17 A plate of a cadmium – zinc alloy 21.7 cm thick is used to reduce the X-radiation from a nickel target to 0.05 of its incident value. What is the transmittance and absorbance of the plate? Assuming that the absorption coefficients, given in Question 14.16, can be added, what is the composition of the alloy, in atom%?

14.18 A lead glass fibre has a refractive index of 1.682. What is the critical angle for total internal reflection at the interface with an acrylic coating with refractive index 1.498?

14.19 A ray of light passing through water (refractive index, $n = 1.33$) in a glass tank (refractive index 1.58) hits the water/glass surface at an angle of 23°.

(a) What angle does it make with the surface as it continues through the glass?

(b) What is the critical angle for light passing through the water striking the water/glass interface?

(c) What is the critical angle for light passing through the glass striking the glass/water interface?

14.20 Estimate the refractive indices of the ceramics barium titanate, $BaTiO_3$, and lead titanate, $PbTiO_3$, using the Gladstone – Dale formula. Densities are as follows: $BaTiO_3$, $6017\ \mathrm{kg\ m^{-3}}$; $PbTiO_3$, $8230\ \mathrm{kg\ m^{-3}}$.

14.21 Estimate the refractive index of the minerals spinel, $MgAl_2O_4$, and akermanite, $CaMgSl_2O_7$, using the Gladstone – Dale formula. Densities are as follows: spinel, $3600\ \mathrm{kg\ m^{-3}}$; akermanite, $2940\ \mathrm{kg\ m^{-3}}$.

14.22 Estimate the refractive index of the minerals beryl, $Be_3Al_2(SiO_3)_6$, and garnet, $Mg_3Al_2Si_3O_{12}$, using the Gladstone – Dale formula. Densities are as follows: beryl, $2640\ \mathrm{kg\ m^{-3}}$; garnet, $3560\ \mathrm{kg\ m^{-3}}$.

14.23 Estimate the refractive coefficient of Al_2O_3 using the information that the mineral andalusite, Al_2SiO_5, has a density of $3150\ \mathrm{kg\ m^{-3}}$ and a refractive index of 1.639. The refractive coefficient of SiO_2 is 0.21.

14.24 Calculate the reflectivity of the surfaces of the following transparent materials in air: cryolite, Na_3AlF_6, $n = 1.35$; glass, $n = 1.537$; corundum, Al_2O_3, n (average) = 1.63; 1.63; cubic zirconia, ZrO_2, $n = 2.05$; tantala, Ta_2O_5, $n = 2.15$ (average).

14.25 Calculate the reflectivity of the surfaces of the following metals in air (the refractive indices are for a wavelength of 550 nm): aluminium, $n = 0.82$, $k = 5.99$; silver, $n = 0.255$, $k = 3.32$; gold, $n = 0.33$, $k = 2.32$; chromium, $n = 2.51$, $k = 2.66$; nickel, $n = 1.85$, $k = 3.27$.

14.26 A thin film on a substrate is viewed in air and has an optical thickness of $\lambda/4$. Will the film be reflecting or not if (a) the substrate has a lower refractive index than the film and (b) the substrate has a higher refractive index than the film?

14.27 A thin film on a substrate is viewed in air and has an optical thickness of $\lambda/2$. Will the film be reflecting or not if (a) the substrate has a lower refractive index than the film and (b) the substrate has a higher refractive index than the film?

14.28 Describe the differences between the reflectivity of a soap film, thickness $\lambda/4$, in air, compared with the same film on an oil surface. Assume that the refractive index of the soap film is very slightly more than that of pure water ($n = 1.33$), the refractive index of the oil is 1.44.

14.29 (a) What is the minimum real (physical) thickness of a film of titanium dioxide, TiO_2, in air for constructive interference of green light, $\lambda = 550$ nm? The refractive index of TiO_2 is 2.875 (average). Estimate the colour that the film would appear when viewed in white light by (b) reflection and (c) transmission. Use the information in Section S4.11.

14.30 Derive the relationship

$$n_f = \sqrt{n_s}$$

for an antireflection coating, refractive index n_f, on a substrate, refractive index n_s, in air. [Note: derivation is not provided in the answers at the end of this book.]

14.31 Determine the reflectivity of a $\lambda/4$ film of (a) silicon oxide, SiO, $n = 2.0$, and (b) titanium dioxide, TiO_2, $n = 2.504$, on glass, $n = 1.504$, in air.

14.32 Determine the reflectivity of a $\lambda/4$ and a $\lambda/2$ film of magnesium difluoride, MgF_2 ($n = 1.384$), on glass ($n = 1.504$), in air. What changes would occur if the substrate was titanium dioxide, TiO_2, $n = 2.875$ (average)?

14.33 Plot a graph of reflectivity versus the number of pairs of layers for a quarter wave stack on a glass substrate ($n = 1.495$), in air, using alternating layers of magnesium fluoride, MgF_2 ($n = 1.384$) and titanium dioxide ($n = 2.875$, average). How many pairs are needed to achieve a reflectivity of at least 99.9 %? [Note: graph is not shown in the answers at the end of this book.]

14.34 A quarter wave stack on a glass substrate ($n = 1.545$), in air, is required to reflect light of 650 nm from a laser. The materials chosen have refractive indices of 2.15 (tantalum pentoxide, Ta_2O_5) and 1.35 (cryolite, Na_3AlF_6). What are the real (physical) thicknesses of the layers?

14.35 The intensity of a light beam traversing a solution placed into a cell of 10 cm path length drops to 80.3 % of the incident intensity. Calculate the linear absorption coefficient of the solution.

14.36 The visibility of the atmosphere is reported as being 10 km. Assuming that at this distance an object can just be perceived, with 1 % of the initial light falling on the object reaching the observer, determine the linear absorption coefficient of the atmosphere.

14.37 Plot the relative scattering of light by the Rayleigh model across the visible spectrum, 400–700 nm. [Note: graph is not shown in the answers at the end of this book.]

14.38 Determine the amount of light scattered by dust particles with a refractive index of 1.45 relative to that of water droplets with a refractive index 1.33 if scattering takes place in accordance with the Rayleigh formula.

14.39 Will air visibility be improved if water droplets responsible for scattering are replaced by similar sized limestone dust particles? The refractive index of limestone, $CaCO_3$, is 1.53.

14.40 The linear scattering coefficient of limestone dust in the air is $0.0002194\ \mathrm{m}^{-1}$. Over what distance will the intensity of a light beam diminish to 10 % of its initial value?

14.41 What is the relative amount of light scattered (a) in the incident direction, (b) at 45° to the incident direction, (c) perpendicular to the incident direction and (d) in the reverse direction, for a monochromatic beam of light, according to the Rayleigh model of scattering?

14.42 (a) What slit width is needed in a transmission diffraction grating to cause the red light, $\lambda = 700\ \mathrm{nm}$, to be deviated by 7.5° to the normal? (b) What will be the deviation of the violet light, $\lambda = 400\ \mathrm{nm}$?

14.43 The first photonic crystal made was formed by drilling an array of holes in a block of material with a refractive index of 3.6 so as to form a face-centred cubic array of holes throughout the block, with a unit cell parameter of 2 mm. What wavelength radiation will not pass in the [1 0 0] direction?

14.44 A photonic crystal was made by laying down close-packed layers of polystyrene spheres of refractive index 1.595 to form a hexagonal closest packed array. The sphere diameter was 250 nm, and the layer separation normal to the layers can be taken as 0.8 of the sphere diameter.

(a) What is the wavelength that will not be transmitted in light normal to the close-packed layers?

(b) What colour does this correspond to?

(c) If the ordering is not perfect and the close-packed layers are sometimes in hexagonal (AB) packing and sometimes in cubic (ABC) packing, how will the result change?

14.45 Estimate the attenuation of an optical fibre in which 0.5 % of the initial power is lost in a distance of 1 km.

14.46 The attenuation of ordinary window glass is of the order of $10^6\ \mathrm{dB\ km}^{-1}$. What thickness of glass would cause the incident light intensity to fall to 50 %?

14.47 An instantaneous pulse of wavelength $1000 \pm 60\ \mu\mathrm{m}$ is introduced into a silica optical fibre. What will be the pulse spread in km after 1 second? The refractive indices of silica are as follows: $n(400\ \mathrm{nm}) = 1.47000$; $n(\mathrm{average}) = 1.46265$; $n(700\ \mathrm{nm}) = 1.45530$.

14.48 A thin film of silver oxide, Ag_2O, forms on a silver surface. The band gap for silver oxide is 2.25 eV. What colour would be emitted by bulk silver oxide? Treating the thin film as a quantum well, what film thickness is needed to obtain an emission in the blue region of the spectrum, at a wavelength of 413 nm? Assume that the effective mass of electrons and holes in Ag_2O is equal to $0.3m_e$.

14.49 A film of zinc sulphide, ZnS, 4 nm in thickness, forms on metallic zinc. Treating the film as a quantum well, what is the wavelength of the transition $\Delta n = 2$? The bandgap of ZnS is 3.54 eV, and the effective mass of electrons and holes is $0.4m_e$.