Collins

PHYSICS

Medical physics
AQA A-level
Year 2

Chris Bishop

William Collins' dream of knowledge for all began with the publication of his first book in 1819.

A self-educated mill worker, he not only enriched millions of lives, but also founded a flourishing publishing house. Today, staying true to this spirit, Collins books are packed with inspiration, innovation and practical expertise. They place you at the centre of a world of possibility and give you exactly what you need to explore it.

Collins. Freedom to teach

HarperCollins*Publishers*

The News Building,

1 London Bridge Street

London

SE1 9GF

Browse the complete Collins catalogue at
www.collins.co.uk

This optional topic is part of the Collins AQA A-Level Physics Year 2 Student Book.

10 9 8 7 6 5 4 3 2 1

ISBN 978-0-00-759764-2

Collins® is a registered trademark of HarperCollins Publishers Limited
www.collins.co.uk

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

Authored by Chris Bishop
Commissioned by Emily Pither
Development by Jane Roth
Editorial management by Mike Appleton and Kate Ellis
Edited by Geoff Amor
Proofread by Mitch Fitton and Sue Glover
Artwork and typesetting by Jouve
Cover design by We are Laura

The publisher would like to thank Sue Glover and Peter Robinson.

Approval Message from AQA

This textbook has been approved by AQA for use with our qualification. This means that we have checked that it broadly covers the specification and we are satisfied with the overall quality. Full details for our approval process can be found on our website.

We approve textbooks because we know how important it is for teachers and students to have the right resources to support their teaching and learning. However, the publisher is ultimately responsible for the editorial control and quality of this book.

Please note that when teaching the A-level Physics course, you must refer to AQA's specification as your definitive source of information. While this book has been written to match the specification, it cannot provide complete coverage of every aspect of the course.

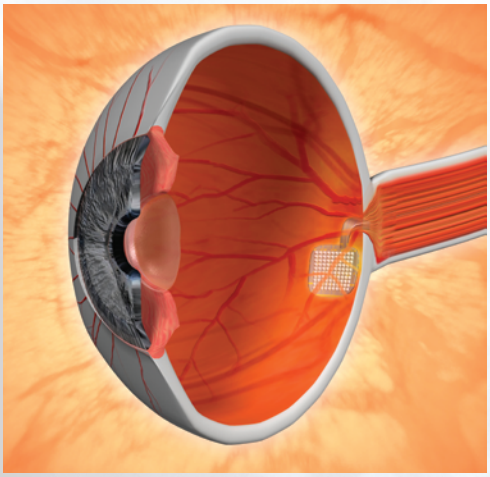
A wide range of other useful resources can be found on the relevant subject pages of our website: www.aqa.org.uk

CONTENTS

MEDICAL PHYSICS

Some of the greatest advances in modern medicine have only been possible because of the input of physicists in developing technologies such as X-ray cameras, magnetic resonance imaging, ultrasound scanners, particle accelerators, radioisotopes and prosthetics to help medical specialists to diagnose and treat disease and trauma.

The loss of sight is a major and distressing event that impacts on a person's life in a multitude of ways. The eye's retina is a collection of photoreceptor cells at the back of the eye that respond to visible light and transmit electrical signals to the brain, which interprets them as images. There are many diseases, such as age-related macular degeneration and retinitis pigmentosa, that can damage these cells, impairing or destroying vision.



A retinal implant

Medical physicists are currently working on artificial retina technology, to restore some sense of vision to those who are blind. An electronic chip is implanted in the eye in the part of the retina that is damaged – this is called a retinal implant (see figure). A miniature camera, mounted on eyeglasses, captures images and sends the information to a body-mounted video processing unit (VPU) that converts it to an electronic signal. The signal is sent to the implant, which consists of an array of microelectrodes, and these stimulate ganglion nerve cells under the retina to fire electrical impulses. The impulses are sent along the optic nerve to the brain, which perceives them as patterns of light and dark, corresponding to which microelectrodes in the array have been stimulated. Patients who have been fitted with the artificial retinal implant (see figure) report that they now have the ability to perceive colours, recognise large letters, locate objects and some report even being able to read short sentences.

A retinal implant user. The camera is at the centre of the glasses. The box lower left is an antenna that receives signals from the VPU (worn on the body) and transmits them to the implant.

The development of the artificial retina has been made possible because of advances in microelectronics, charge-coupled devices and the fabrication of miniature electrodes, all at the forefront of modern physics. In this unit you will see how physics enables us to push back the frontiers in many areas of modern medicine. You will learn about the physics of the eye and the ear, how we can measure the electrical responses of the human body, how diagnostic imaging using X-rays, magnetic fields and radionuclides can give us clear pictures of its interior, and how particle accelerators are used in the fight against cancer.

1 PHYSICS OF THE EYE AND THE EAR

PRIOR KNOWLEDGE

You should have a good understanding of light as a wave and the wave property of refraction from *Chapter 7 in Year 1 Student Book*, and you will be familiar with the use of ray diagrams. You will know that a lens changes the path of light by refraction, and can form an image. From GCSE Biology you will have learnt about the basic structure of the eye and the ear, and their use as biological sensors that enable us to see and hear. You will also know about the properties of sound as a longitudinal wave motion, including its description in terms of amplitude, frequency and period. You need to have an understanding of logarithms and their manipulation.

LEARNING OBJECTIVES

In this chapter you will learn about the physics that governs the ways in which the human eye and ear work. You will learn how we can measure their performance as biological sensory organs, about common defects in their working, and how these may be identified and in many cases corrected.

(Specification 3.10.1, 3.10.2)

1.1 PHYSICS OF VISION

The human eye is a remarkable optical system, a living video camera. It contains 130 million light-detecting cells. These are sensitive enough to form an image in the darkness of night and adaptable enough to cope with summer sunshine, when the light may be 10 million times brighter. With perfect eyesight, you can see details as small as 1 mm up to 6 m away, and switch your focus from a 20 cm close-up to infinity in less than one-tenth of a second.

This section deals with the eye as an optical refracting system, explains how it forms an image and discusses its sensitivity and ability to resolve detail.

The human eye and a video camera (Figure 1) have similar demands. They need to:

- produce an image in 'real time', which means updating the image around 20 times per second
- automatically adjust the focus from close-up to infinity
- automatically adapt to changing light conditions.

There are similarities about the way that the eye and the camera meet these demands. They both receive light through a variable aperture, and they both

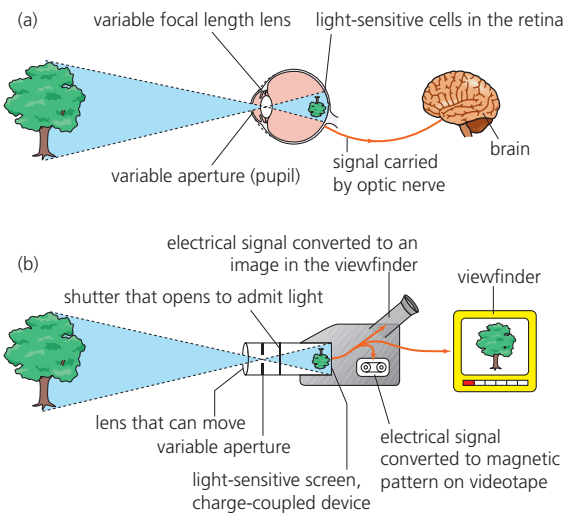


Figure 1 *(a) The eye–brain system compared with (b) the video camera–TV system*

use a convex lens to **focus** light as an image onto a light-sensitive surface. A focused image is produced when light coming from one particular point on an object arrives at one point after passing through the optical system, so that, when considering all points on an object, the image is a replica of the object.

Controlling the amount of light

The amount of light entering a video camera is controlled by a variable aperture between the lens and the charge-coupled device (CCD) that detects the light. In the eye (Figure 2), the aperture is the **pupil** in front of the lens. The diameter of the pupil is varied by the **iris**, which is a ring of smooth muscle. Some of the muscle fibres in the iris run radially, like bicycle spokes. Other muscle fibres encircle the pupil. When the circular fibres contract, the pupil becomes smaller (constricts). When the radial fibres contract, the pupil opens wider (dilates).

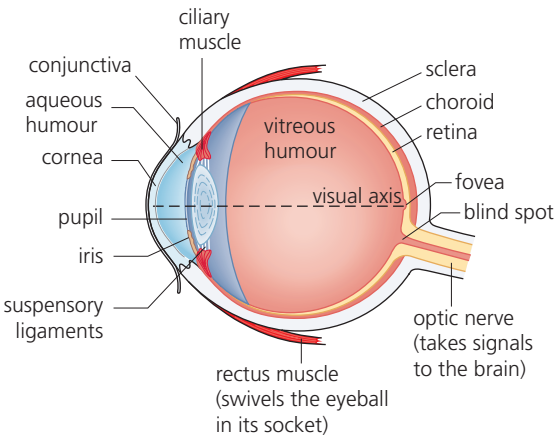


Figure 2 *Vertical section through the eye*

Forming a focused image

Light rays enter the eye through the **cornea**, the transparent membrane that covers the front of the eye (see Figure 2). Refraction occurs as light passes from air into the cornea (Figure 3). The amount of refraction depends on the ratio of the speed of light in air to its speed in the cornea – the **refractive index** of the cornea – which is similar to that in water (*see sections 7.2 and 7.3 in Chapter 7 in Year 1 Student Book*)

After the cornea, light passes through a clear watery fluid known as the aqueous humour before reaching the **lens**. At each boundary between media, light is refracted, but about 60% of the refraction that occurs in the eye happens when light strikes the cornea. The cornea bulges slightly from the eye to present a curved surface to the light. Variations and defects in this shape lead to problems in focusing the light.

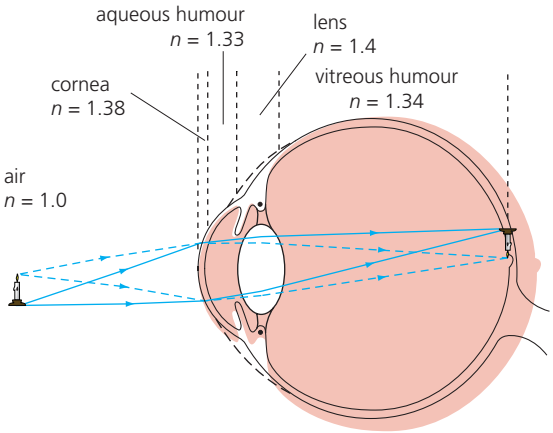


Figure 3 *Refraction in the eye. As a ray of light passes through each part of the eye, it is refracted at each boundary between media. A large change in the refractive index (n) causes more refraction. In the eye, the largest deviation occurs as the light passes from air into the cornea.*

In an eye without defects, light from an object that enters the eye is focused as an image on the **retina** at the back of the eyeball (see Figures 2 and 3).

QUESTIONS

1. **a.** Refer to Figure 3. At which boundary in the eye does the second largest deviation of light occur?
   b. Does this add to or subtract from the deviation caused at the front surface of the cornea?

Sensitivity of the eye

The human eye can perceive objects in dark conditions and in very bright conditions, and its response can vary by a factor of about one million. In low levels of light, the diameter of the pupil increases (dilates) up to about 8 mm; and in bright light levels, it reduces (constricts) to 1.5 mm. But is the eye's different response to different light levels entirely due to the change in area of the pupil?

Area of pupil when constricted $= \pi r^2$
$= \pi(0.75 \times 10^{-3})^2$

Area of pupil when dilated $= \pi r^2 = \pi(4.0 \times 10^{-3})^2$

The ratio of these is

$$\frac{\pi(4.0 \times 10^{-3})}{\pi(0.75 \times 10^{-3})^2} = 0\left(\frac{4.0}{0.75}\right) = 28$$

So there must be another, much more important, mechanism for increasing the ability of the eye to see at low levels of light.

The retina is the eye's **photodetector**. It contains about 130 million light-sensitive cells. Each cell reacts to light from one tiny point of the image. The light-sensitive cells contain a chemical called a **photopigment**. The molecules of the photopigment absorb light and thereby stimulate the cell to generate an electrical impulse, which carries a signal through a network of nerves to the **optic nerve** and then to the brain.

There are two types of light-sensitive cells in the retina, known as **rods** and **cones**.

- Rods are very sensitive to light – they can respond to low levels of illumination, even individual photons. However, they give no information about colour, although they are most sensitive at blue–green wavelengths. Rods contain a photopigment called **Rhodopsin** which is a protein sensitive to light.
- Cones are less sensitive, but their photopigments respond to a narrower range of wavelengths. Their response pattern according to wavelength – or **spectral response** – gives us colour vision (see below).

Under normal daylight conditions, our vision relies on cone cells. In the middle of the retina is a small dimple called the **fovea** (see Figure 2). The density of cones in this region is the highest, and this is the centre of the eye's sharpest vision and the location of most colour perception. Figure 4 shows the distribution of rod and cones on the retina. The eye has a blind spot, which is an area on the retina without any receptors that respond to light. An image

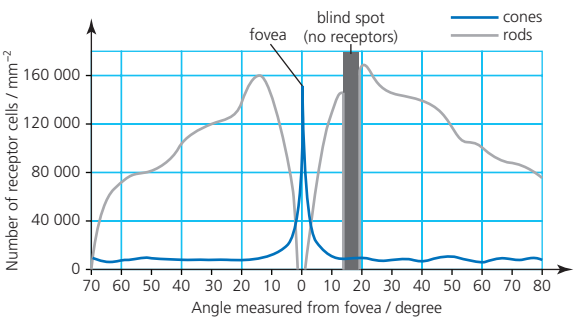


Figure 4 *The distribution of rod and cones on the retina*

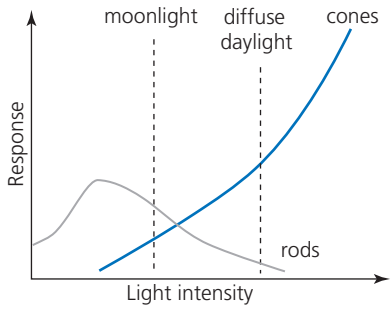


Figure 5 *The response of rod and cone cells in different light conditions*

that falls on this region will not be seen. It is in this region where the optic nerve exits the eye on its way to the brain.

Cone cells only work well in bright light or 'photopic' conditions. It is the rod cells that are important in low light or 'scotopic' conditions. Figure 5 shows these different responses.

In rods, when the photopigment absorbs a photon, it leads to a tiny pulse of current in the cell that lasts for approximately 300 ms. During this time, the rod cannot detect any other photons that arrive. Rhodopsin is destroyed, or 'bleached', by light, and it takes time for the levels to recover. In normal daylight, much of the rhodopsin is bleached, so the rods only function at a low level and we rely on the cones. When you step from bright sunlight into a darkened room, you cannot see clearly at first. There is not enough light to stimulate the cones, and the rods are still inactive. However, after some time has passed, more and more rhodopsin becomes active again, and your eyes become sensitive to low levels of light (Figure 6). You are said to be 'dark-adapted'. It takes about 30 minutes for the retina to become fully dark-adapted, when rod cells are up to 100 000 times more sensitive than in bright light.

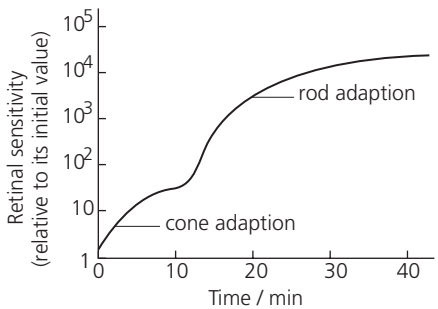


Figure 6 *Retinal sensitivity versus time*

Colour vision

The eye can respond to electromagnetic waves over a range of wavelengths from about 380 nm to about 750 nm. Rod cells respond to almost this whole range, with a peak response around 500 nm. Cone cells, however, contain one of three different photopigments, each of which responds to a different wavelength range. These responses are shown in Figure 7. When a photon of light strikes the retina, it may be absorbed by one of these photopigments in a cone cell. The probability of absorption depends on the wavelength of the light. A given wavelength will stimulate a certain ratio of the three different types of cone cell. This enables the brain to identify wavelengths and produce the sensation of colour. For example, the sensation of 'yellow' is created by the 'red' and 'green' cones being triggered in the right proportion. (In fact, the yellow light from a TV screen is actually a combination of red and green light.)

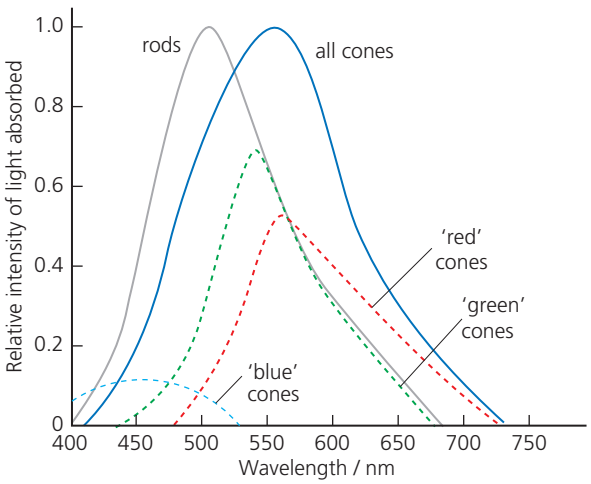


Figure 7 *The spectral sensitivity of the different cells that have different photopigments*

QUESTIONS

2. **a.** How does the range of wavelengths that the eye can detect change as the intensity of light decreases to a very low value?
   b. What effect does this have on our ability to differentiate colours in very dim light?

Spatial resolution

The eye's ability to see detail depends crucially on the spacing of the rod and cone cells in the retina. We can only tell that two objects are separate if there is at least one unstimulated cell lying between stimulated cells (Figure 8).

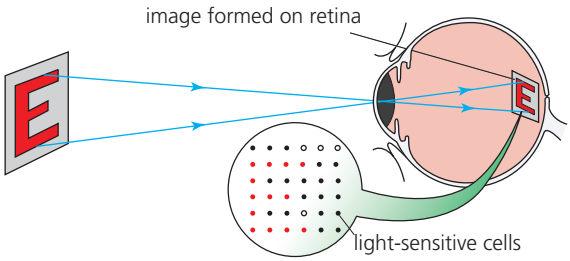


Figure 8 *To be able to read the letter E, there must be at least one unstimulated retinal cell between those that are stimulated*

The average separation of light-sensitive cells on the retina is 0.003 mm. The distance from the retina to the centre of the eye lens along the principal axis is about 15 mm, which using the small-angle approximation $\theta \approx \tan\theta$ gives a minimum angle that can be resolved of

$$2 \times \frac{0.003}{15} = 0.0004 \text{ rad}$$

The eye's ability to see detail is termed its **visual acuity**, defined as

$$\text{visual acuity} = \frac{1}{\text{minimum angle resolved in arcminutes}}$$

The **arcminute** (**arcmin**) is a unit of angular measurement equal to 1/60th of one degree:

$$1 \text{ rad} = \left(\frac{180}{\pi}\right)^\circ$$

so the angular resolution of 0.0004 rad is

$$0.0004 \times \left(\frac{180}{\pi}\right)^\circ = 0.0229^\circ = 0.0229 \times 60$$
$$= 1.375 \text{ arcmin}$$

Therefore the eye's visual acuity is $\frac{1}{1.375} = 0.73 \text{ arcmin}^{-1}$.

Our eyes can actually do better than this because rods and cones are not uniformly distributed across the retina. At the edge of our field of view there are more rods than cones, but towards the centre of our field of view cones dominate. The fovea has no rod cells at all and the cones are packed to a density of in excess of 150 000 per square millimetre. This area, which is directly in the centre of our field of view, has no large blood vessels near it and the nerve fibres run radially from it, giving an unobstructed passage for light. This part of the retina gives the maximum visual acuity, about 2 arcmin^{-1}, so a resolution of $\frac{1}{2} = 0.5$ arcmin. Image resolution in bright light is better because cone cells are smaller than rod cells.

QUESTIONS

3. Why is the eye able to see more detail in the centre of the field of view than at the edges?

KEY IDEAS

- The eye forms an image by refracting light through the cornea, the fluids in the eye and the eye lens, on to the back of the retina.
- The retina contains light-sensitive cells called rods and cones. These contain photopigments that have different spectral responses.
- Rods do not differentiate between colours. They are sensitive at low levels of illumination.
- Cones respond differently to different ranges of colour and are sensitive at higher levels of illumination.
- The spatial resolution of the eye is its ability to see fine detail and depends on the spacing between the rods and cones.

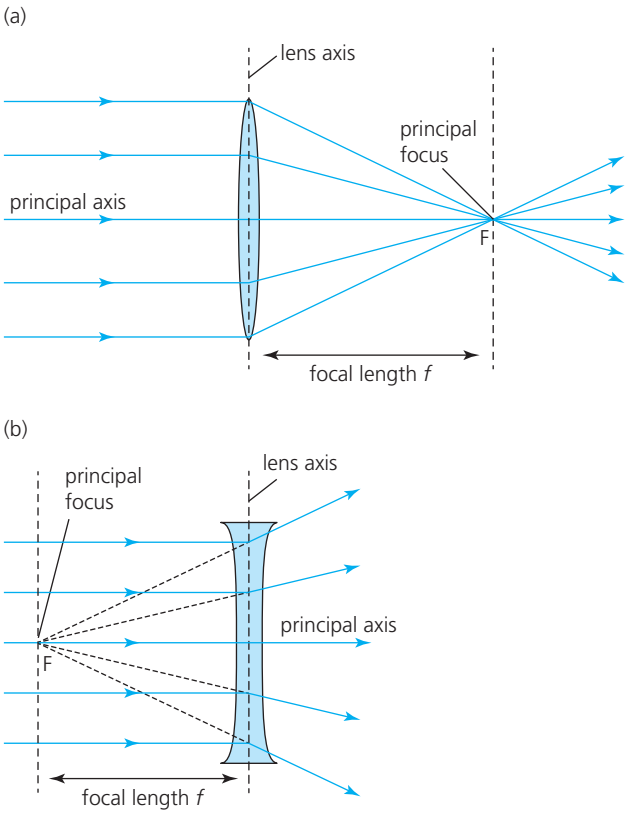


Figure 9 *The action of (a) a converging lens and (b) a diverging lens*

1.2 THE ACTION OF LENSES

Converging and diverging lenses

A lens is a transparent block of glass, or, in the case of the eye lens, a transparent organic material, that causes light passing through it to refract.

A **convex lens** is curved outwards on both sides. When parallel rays are incident on it, the emerging rays come together or converge at a point (Figure 9a), so it is alternatively called a *converging lens*. A converging lens can form a **real image**, for example on a screen. The eye lens is convex.

A **concave lens** is curved inwards on both sides. When parallel rays are incident on it, the emerging rays spread out or diverge (Figure 9b), hence it is also called a *diverging lens*. A diverging lens always, by itself, forms a **virtual image** – an image that cannot be projected onto a screen or focused onto a photodetector (*see sections 7.1 in Chapter 7 in Year 1 Student Book*).

As well as in the eye, lenses are used in optical instruments, such as binoculars, telescopes, projectors, cameras, spectacles, magnifying glasses and microscopes, to produce an image.

A horizontal line through the centre of a lens and at right angles to the lens is called the **principal axis**. The **principal focus** F is a point at which light rays parallel to the principal axis are focused – or, in the case of a diverging lens, appear to diverge from. The **focal length** *f* is the distance from the centre of the lens to the principal focus. The focal length of a lens is determined partly by its shape: a lens that is highly curved (small radius of curvature) will have a short focal length.

The lens of the eye

The lens in the eye converges rays of incident light from an object so that they come to a focus at the fovea on the retina. In order to be able to focus light from both distant and near objects, the focal length of the lens needs to change. This is done by the muscles changing the shape of the lens.

The lens is held in tension by ligaments, which are connected to the **ciliary muscle** (see Figure 2 in Medical Physics section 1.1). As this muscle contracts or relaxes, the lens changes shape and therefore its focal length changes (Figure 10).

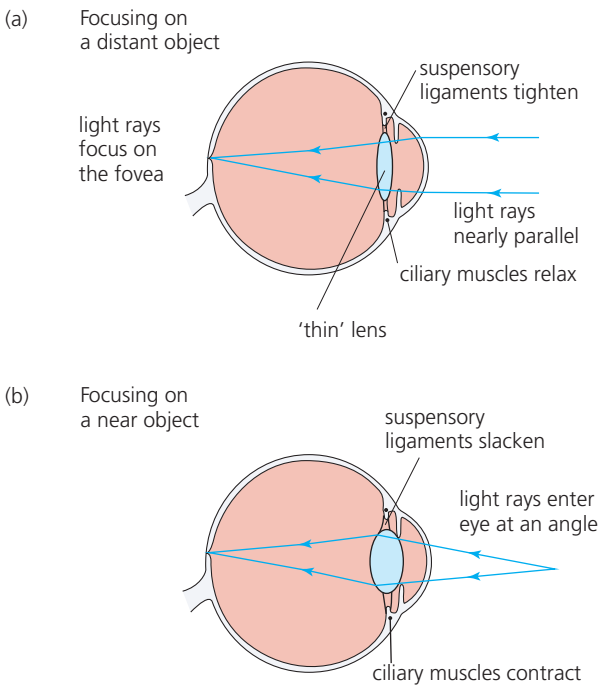


Figure 10 *Focusing in the eye for an object that is (a) distant and (b) near*

This ability of the lens in the eye to change its focal length, so that objects at different distances are brought into sharp focus on the retina, is called **accommodation**. When the lens adopts a highly curved surface, it increases its focusing power. The **power** of a refracting surface is defined as

$$\text{power} = \frac{1}{f}$$

where f is the focal length in metres. The unit of power is called the **dioptre** (D), and is equivalent to m^{-1}. By convention, a converging lens has a positive power, and a diverging lens has a negative power.

The total refracting power of a combination of lenses, or surfaces, is the sum of their powers.

The human cornea has a refracting power of about +43 D. The power of the lens varies from +17 D at its flattest to about +31 D at its most curved.

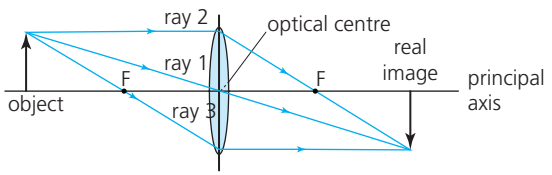
The total power of the eye therefore varies from about +60 D to +74 D.

There is a limit on the range of object distances at which the eye can focus.

- The **near point** is the closest distance at which an object can be brought into focus. For a young, healthy eye, the near point is at about 25 cm. For older people, a stiffer eye lens can make the near point further away.
- The **far point** is the furthest distance at which an object can be brought into focus. For a normal eye, the far point is infinity.

Ray diagrams

For a lens of known power, it is possible to find out where an image will be formed by constructing a ray diagram. Figures 11 and 12 show ray diagrams for a convex lens and a concave lens, respectively. The axis of the lens is shown as a straight vertical line, the principal axis as a horizontal line, the principal foci F are dots on the principal axis either side of the lens, and an extended object is drawn as an arrow standing on the principal axis. The diagram is drawn to scale. Three specific rays are drawn leaving the object. The passage of these through the lens can be predicted according to the rules stated below each figure, and the focused image can be located. The distance of the image from the lens and its size will be to scale. Note that the rays are shown to change direction only at the lens axis in ray diagrams.

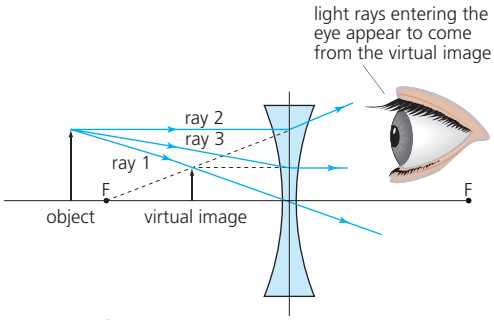


Ray 1 Any ray of light that passes through the optical centre will be undeviated

Ray 2 A ray of light that travels parallel to the principal axis will be refracted through the principal focus

Ray 3 Any ray of light that passes through the principal focus will be refracted so as to travel parallel to the principal axis (this is rule 2 in reverse)

Figure 11 *Ray diagram of light passing through a convex lens. The focused image forms where the rays of light from the tip of the object meet again. This is a real image, because rays of light actually pass through it – an image would be formed on a screen placed at this point.*



Ray 1 Any ray of light that passes through the optical centre will be undeviated
Ray 2 A ray of light that travels parallel to the principal axis will be refracted away from the axis, so that it appears to have come from the principal focus
Ray 3 A ray of light that would have passed through the principal focus will be refracted so as to travel parallel to the principal axis (this is rule 2 in reverse)

Figure 12 *Ray diagram of light passing through a concave lens. The focused image forms where the rays of light from the tip of the object appear to have come from. This is a virtual image, because no rays of light actually pass through it – it is not possible to form an image on a screen placed at this point.*

Worked example 1

A converging lens has a power of 3 D. Construct a ray diagram to find out where this lens will produce a focused image of a candle that is 1 m from the lens.

The focal length of the lens can be found by rearranging the equation for power:

$$\text{focal length} = \frac{1}{\text{power}} = \frac{1}{3} = 0.33\,\text{m} = 33\,\text{cm}$$

Then the ray diagram shown in Figure 13 can be constructed to scale. For the greatest accuracy, a scale should be chosen that gives the largest diagram that will fit on the piece of paper. The vertical scale, showing the object and image size, need not be the same as the horizontal scale.

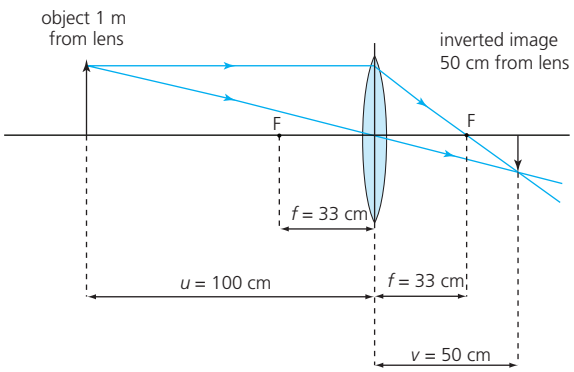


Figure 13 *Using a ray diagram to locate an image*

The diagram shows that the image is 50 cm from the lens, on the other side of the lens from the object, and inverted.

The lens formula

The position of a focused image can also be calculated, using the **lens formula**. The formula connects the focal length, f, of a lens with the object distance, u, and the image distance, v:

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$$

Figure 13 in Worked example 1 shows f, u and v for a convex lens.

In the use of the lens formula, there is a sign convention. It is possible for an image distance, v, to be negative. A negative value of v means that the image is **virtual**, on the same side of the lens as the object, and upright. Figure 14 shows the situation that occurs for convex lenses whenever the object distance is less than the focal length.

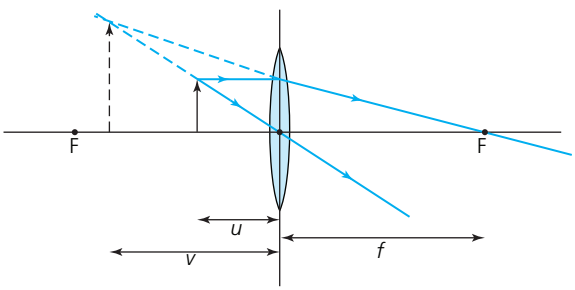


Figure 14 *Formation of a virtual image by a convex lens*

When using the lens formula for a *concave* lens, the focal length f used *must be negative*, since the power is negative. The resulting image distance v is always negative, as ray diagrams show – see Figure 12.

Worked example 2

a. Use the lens formula to find where a convex lens with a 10.0 cm focal length will form an image of an object that is 30.0 cm away.

b. The object in part **a** is now placed 5.0 cm from the lens. Where will the image be?

a. We have $f = 10\,\text{cm}$ and $u = 30\,\text{cm}$. Rearranging the lens formula gives

$$\frac{1}{v} = \frac{1}{f} - \frac{1}{u} = \frac{1}{0.100} - \frac{1}{0.300} = 10.0 - 3.33$$
$$= 6.67\,\text{m}^{-1}$$

So
$$v = \frac{1}{6.67} = 0.15\,\text{m}$$

The image will be formed 15 cm away from the lens and will be real and inverted.

b. The calculation now becomes:

$$\frac{1}{v} = \frac{1}{f} - \frac{1}{u} = \frac{1}{0.100} - \frac{1}{0.050} = 10.0 - 20.0$$
$$= -10.0\,\text{m}^{-1}$$

So
$$v = \frac{1}{-10} = -0.10\,\text{m}$$

If a ray diagram were constructed, the light rays from the object would be seen to diverge after passing through the lens. The image would be on the same side of the lens as the object. If we looked through this lens towards the object, the light rays would appear to come from a point 0.10 m behind the lens. A virtual image has been formed.

QUESTIONS

4. Draw a ray diagram to scale for the situation in Worked example 2, for an object of height 6.0 cm. What height is the image?

Magnification

The image formed by a lens may be smaller or larger than the object. The **magnification** of an image produced by a lens is the ratio of image height to object height. So, in Figure 15, we have

$$\text{magnification} = \frac{XY}{AB}$$

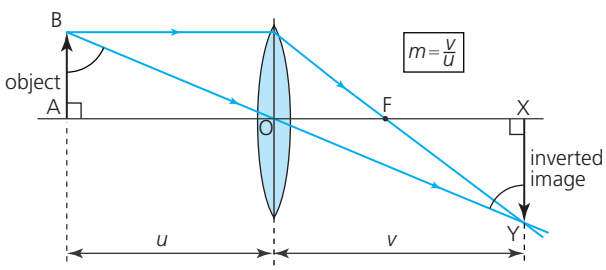


Figure 15 *The definition of magnification*

But

$$\tan \text{AOB} = \frac{AB}{AO} \quad \text{and} \quad \tan \text{XOY} = \frac{XY}{XO}$$

Angles AOB and XOY are equal, so

$$\frac{AB}{AO} = \frac{XY}{XO}$$

Rearranging gives
$$\frac{XY}{AB} = \frac{XO}{AO}$$

and this gives an alternative definition for magnification m:
$$m = \frac{v}{u}$$

If v is negative, m is negative, which indicates that the image is virtual and upright, rather than real and inverted.

Worked example 3

A convex lens with a focal length of 20 cm is used to produce an image of an object that is 30 cm high and 1 m away from the lens. How big will the image be?

Use the lens formula to find the image distance, v. Rearranging the lens formula gives

$$\frac{1}{v} = \frac{1}{f} - \frac{1}{u} = \frac{1}{0.20} - \frac{1}{1.0} = 5.0 - 1.0 = 4\,\text{m}^{-1}$$

so

$$v = \frac{1}{4} = 0.25\,\text{m} = 25\,\text{cm}$$

The magnification is

$$m = \frac{v}{u} = \frac{25}{100} = 0.25$$

so the image will be $0.25 \times 30 = 7.5$ cm high.

QUESTIONS

5. An object of height 2.0 cm is placed 75 cm from a converging lens of focal length 10 cm. Find the position, magnification and size of the image.

6. An object of height 4.0 cm is placed at a distance of 35 cm from a concave lens of focal length 12 cm. What are the image distance and the magnification of the image so formed?

7. Draw ray diagrams for an object producing an image with a convex lens of focal length f for object distances as follows:

a. $u = 2f$

b. $2f > u > f$

State the nature and size of the image in each case.

8. For a concave (diverging) lens, draw the ray diagrams and state the nature and size of the image formed for an object placed:

a. at an object distance greater than twice the focal length f

b. at an object distance between f and $2f$.

9. What is the power of the lens in

a. question 5

b. question 6?

KEY IDEAS

- A convex lens is a converging lens. It forms a real image if the object distance is greater than the focal length. The lens of the eye is a converging lens.
- A concave lens is a diverging lens. It forms a virtual image.
- The power of a lens is the reciprocal of its focal length:

$$\text{power} = \frac{1}{f}$$

It is measured in the unit dioptre (D), which is equivalent to m^{-1}.

- Converging lenses have positive powers; diverging lenses have negative powers.
- Accommodation is the process by which the eye changes the power of its lens to maintain a focused image for objects at different distances.
- Ray diagrams may be drawn to scale to trace the passage of light rays through a lens and locate the position of the image.
- The lens formula is

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$$

where u is object distance, v is image distance, f is negative for a concave lens, and v is negative for a virtual image.

- Magnification is given by

$$m = \frac{\text{image distance}}{\text{object distance}} = \frac{v}{u}$$

1.3 DEFECTS OF VISION AND THEIR CORRECTION

Short sight

Short-sighted people can focus on nearby objects, but cannot focus on distant ones (Figure 16a). The far point is nearer to the eye than infinity. They are said to have **myopia**. Someone suffering from myopia has a cornea and lens combination that is too powerful, or an eyeball that is too long. Distant objects are brought to a focus before the light reaches the retina. A diverging spectacle or contact lens is needed to correct this (Figure 16b).

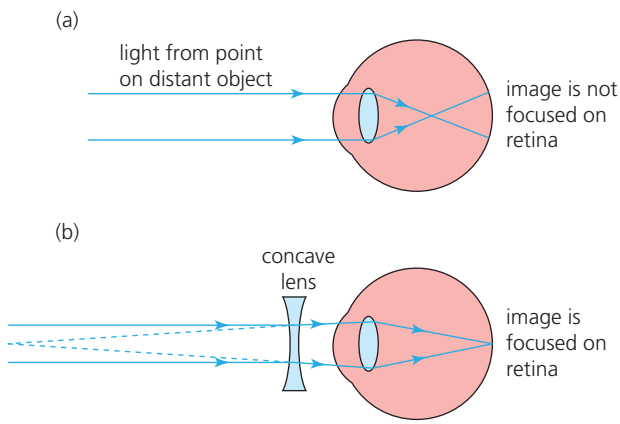


Figure 16 *The myopic (short-sighted) eye and the use of a concave lens to correct the vision for distant objects*

Worked example 1

Sarah is short-sighted. Without glasses, her far point is 50 cm. If the distance between her eye lens and her retina is 20 mm, find the power of the lens needed to correct her vision.

Using the lens formula, we can work out the power of Sarah's eye when she tries to focus on a distant object. We take the image distance to be the lens–retina distance, so $v = 0.02\,\text{m}$. For an object at her far point, $u = 0.50\,\text{m}$, so

$$\text{power} = \frac{1}{f} = \frac{1}{0.50} + \frac{1}{0.02} = 52\,\text{D}$$

If Sarah is to focus on an object at infinity, then $u = \infty$, and the power needs to be

$$\text{power} = \frac{1}{f} = \frac{1}{\infty} + \frac{1}{0.02} = 50\,\text{D}$$

(since division by infinity gives zero). We need to reduce the total refracting power of her eyes by 2 D. Since the total refracting power of a combination of lenses is the sum of their powers, Sarah needs lenses of power -2 D.

Long sight

When the lens–cornea combination is not powerful enough, the light from nearby objects cannot be brought to a focus on the retina. The near point is further away than the normal 25 cm. This is long sight, or **hypermetropia**. Someone with hypermetropia can focus on objects at infinity, but not on nearby things (Figure 17a). A convex correcting lens is needed to converge the light more (Figure 17b).

A second cause of long sight, common in older people, is reduced flexibility of the eye lens, which means that accommodation to focus on nearby objects is not so good.

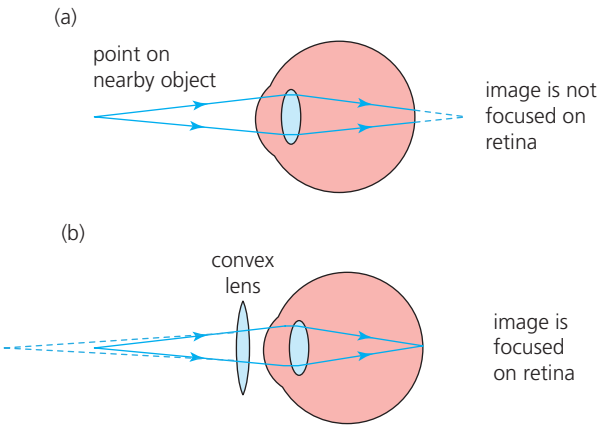


Figure 17 *The hypermetropic (long-sighted) eye and the use of a convex lens to correct the vision for near objects*

Worked example 2

James is long-sighted. Without glasses, his near point is 50 cm. If the distance between his eye lens and his retina is 20 mm, find the power of the lens needed to bring his near point to 25 cm.

Using the lens formula, we can work out the power of James's eye when he tries to focus on a near object. Taking the image distance as the lens–retina distance, $v = 0.02$ m. For an object at his near point, $u = 0.50$ m, so

$$\text{power} = \frac{1}{f} = \frac{1}{0.50} + \frac{1}{0.02} = 52\,\text{D}$$

If James is to focus on an object at 25 cm, then $u = 0.25$ m, and the power needs to be

$$\text{power} = \frac{1}{f} = \frac{1}{0.25} + \frac{1}{0.02} = 54\,\text{D}$$

We need to increase the total refracting power of his eyes by 2 D, so he needs lenses of power $+2$ D.

Astigmatism

Corneas are not spherical in shape. They have different curvatures, and so different refracting powers, in different directions. When this difference in curvature is large, or there are irregularities in the cornea, the image formed on the retina can be unevenly focused, a condition known as **astigmatism** (Figure 18). The lenses discussed so far in this chapter have been spherical lenses, that is, they have spherical surface curvature and focus parallel light rays to a point. Astigmatism is corrected by using a *cylindrical* lens, as shown in Figure 19 (or by laser surgery). A cylindrical lens has curved faces that are sections of a geometric cylinder. It thus has differing radii along perpendicular axes and focuses parallel light rays to a line.

Figure 18 *A person suffering from astigmatism sees some lines in this optician's chart clearer than others.*

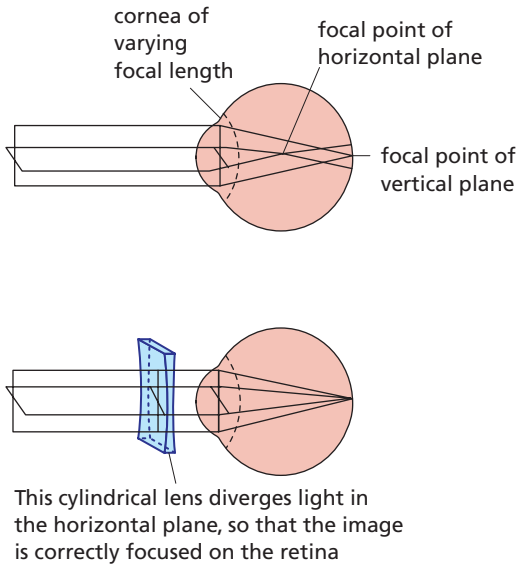


Figure 19 *Astigmatism and its correction with a cylindrical lens*

An optician's prescription for a lens to rectify astigmatism needs to give the power of the lens, in dioptre, *and* the required direction of the cylindrical axis. Figure 20 shows an example.

	Sphere (SPH)	Cylinder (CYL)	Axis (AXI)
Right Eye	+0.25	–1.25	75
Left Eye	+0.25	–1.25	100

Figure 20 *An example prescription for a person who has astigmatism*

In Figure 20, the 'Sphere (SPH)' refers to the spherical portion of the prescription, which is the correction in dioptre needed for short-sightedness (a negative correction) or long-sightedness (a positive correction).

The 'Cylinder (CYL)' refers to the cylindrical correction, which depends on the degree of astigmatism, and can be a negative or a positive number in dioptre. If nothing appears in this column on a prescription, there is no astigmatism that needs to be corrected.

The 'Axis (AXI)' is a number of degrees of angle between 1° and 180°. This specifies in what direction the difference in curvature is, and so where in the spectacle lens the cylindrical axis should be. The angle is measured using an imaginary semicircle with a horizontal base line that starts with 0° in the 3 o'clock direction (as viewed by the optician who is taking the measurement) and increases to 180° in an anticlockwise direction (Figure 21). It represents the angle at which the spectacle lens is set into the frame.

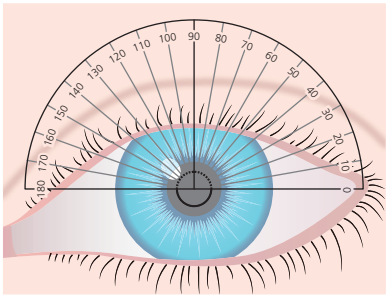


Figure 21 *Measuring the axis for a cylindrical lens prescription*

QUESTIONS

10. A long-sighted person has a near point of 70 cm. What is the power of a corrective lens that is needed to bring the near point to a distance of 20 cm from the eye? Assume that the distance from the cornea to the retina is 17 mm and the eye acts as a single thin lens situated at the front surface of the cornea.

11. The furthest distance a short-sighted person can see an object clearly is 50 cm. Assuming that the eye acts as a single thin lens with a cornea–retina distance of 17 mm, calculate:

- **a.** the focal length of the eye when viewing at object at 60 cm
- **b.** the lens power of the eye when viewing the object at 60 cm
- **c.** the lens power needed by the eye to see an object clearly at infinity
- **d.** the power of a corrective lens that when placed over the unaided eye will allow the person to see an object clearly at infinity.

12. A prescription for Harry is shown in Figure 22.

	SPH	CYL	Axis
Right Eye	–2.00	+1.50	180
Left Eye	+3.50	+3.00	45

Figure 22

- **a.** Explain what the prescription tells the optician about Harry's right eye.
- **b.** Explain what the prescription tells the optician about Harry's left eye.

KEY IDEAS

- Myopia is short-sightedness. The far point is nearer to the eye than infinity. It is corrected with a spherical lens of negative power (diverging).
- Hypermetropia is long-sightedness. The near point is further than 25 cm from the eye. It is corrected with a spherical lens of positive power (converging).
- Astigmatism is a defect of the eye in which the cornea has different degrees of curvature in different directions. It is corrected with a cylindrical lens.
- The prescription for correcting astigmatism needs to give the power of the cylindrical lens and the orientation of its axis.

1.4 PHYSICS OF THE EAR AS A SOUND DETECTION SYSTEM

Sound waves are longitudinal vibrations that travel through a medium, such as air, as a pressure wave. The energy transferred by the wave is proportional to the square of its amplitude, where the amplitude is a pressure difference (*see section 5.1 in Chapter 5 in Year 1 Student Book*).

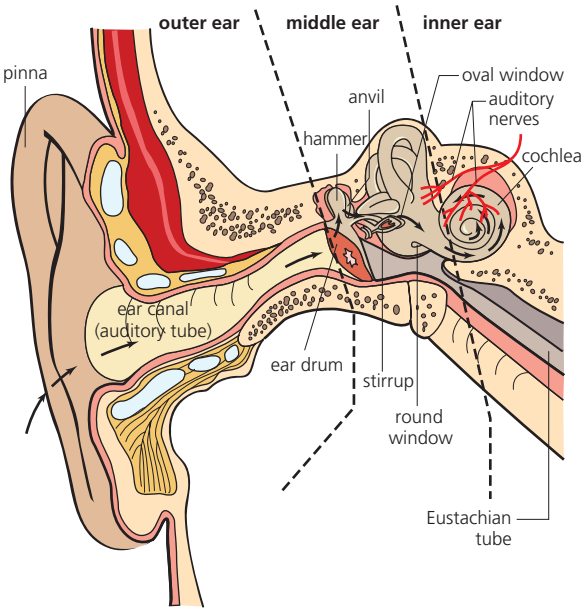


Figure 23 *A cross-section through the ear. A sound wave passes through the three sections of the ear in about 20 ms.*

The human ear is an incredibly sensitive device for collecting and detecting sound waves. The visible part of the human ear is known as the pinna (Figure 23). Its function is to collect sound waves and funnel them down the ear canal, or external **auditory tube**. The pinna is shaped so that sound sources in front of the head are detected more easily than those behind. This helps us to determine the direction of a sound source.

- The *outer ear* collects sound waves and relays them to the **eardrum**.
- The *middle ear* amplifies the vibrations of the eardrum and transmits them to the inner ear.
- The *inner ear* or **cochlea** converts the vibrations to electrical signals, which are transmitted to the brain via the **auditory nerves**.

The external auditory tube is about 2.5 cm long and has a diameter of 7 mm. It modifies the sound we hear through its resonant properties, which are like those of a pipe closed at one end (*see section 5.6 in Chapter 5 in Year 1 Student Book*).

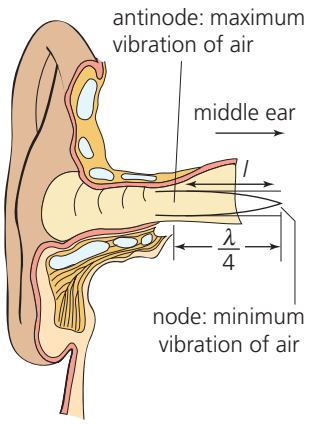


Figure 24 *Resonance in the external auditory tube*

The ear has its maximum response at a sound frequency of 3300 Hz. To see why, consider the air in the external auditory tube vibrating in its fundamental mode (resonating). There is a node at the closed end and an antinode at the open end (Figure 24), so the length l of the tube is one-quarter of the wavelength of the sound:

$$\frac{\lambda}{4} = l = 2.5\,\text{cm}$$
$$\lambda = 10\,\text{cm}$$

The speed of sound in air, v, is about $330\,\text{m}\,\text{s}^{-1}$. Since $v = f\lambda$,

$$f = \frac{v}{\lambda} = \frac{330}{10 \times 10^{-2}} = 3300\,\text{Hz}$$

The sound waves in the ear canal cause the eardrum to vibrate. These vibrations have a remarkably small amplitude. For very quiet sounds, your eardrum may vibrate with an amplitude of only 10^{-11} m (less than the diameter of an atom).

The bones of the middle ear – the malleus or hammer, incus or anvil, and stapes or stirrup – collectively termed the **ossicles**, can be seen in Figure 23, and are shown enlarged in Figure 25. They amplify the *pressure changes* of the sound wave at the eardrum and transfer them to the oval window, which is the entrance to the cochlea.

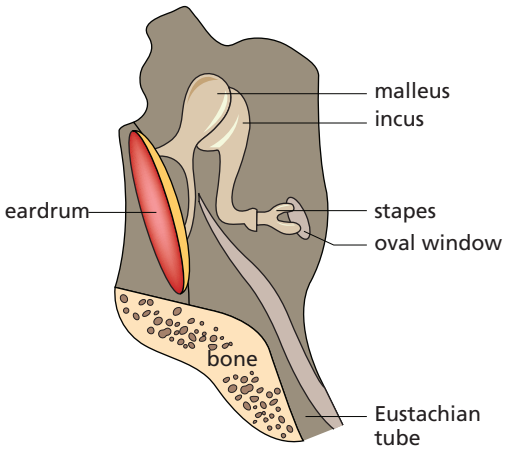


Figure 25 *The middle ear. The ossicles transmit the vibrations of the eardrum to the oval window.*

The ossicles together act as a lever, which magnifies a force on it by a factor of 1.3. Also, the area of the oval window is only 1/20 of that of the eardrum, so the overall effect in the middle ear is to increase the pressure (force per unit area) acting on the fluid in the inner ear by a factor of 26 (= 1.3 × 20). Without this amplification, the attenuation as the sound wave passed between different media in the ear would be significant.

The vibrations of the oval window are transferred to the inner ear as pressure waves in the fluid of the **cochlea** (Figure 26). The cochlea is a coiled structure, about the size of a pea, in a cavity in the thickest part of the skull (see Figure 23). It contains hair cells that convert vibrations to electrical impulses. There are about 30 000 sensory hair cells in each ear. They are amazingly sensitive: a distortion of 0.0003° can be detected.

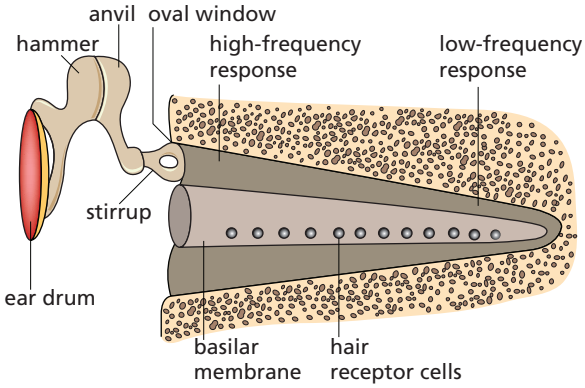


Figure 26 *The cochlea, unrolled and magnified. The basilar membrane holds a row of hair cells. When the stirrup vibrates against the oval window, the basilar membrane vibrates and the hair cells are distorted.*

Hair cells can fire quickly enough to directly detect sound waves and transmit impulses up to a frequency of 1 kHz. However, another mechanism is needed to detect higher-frequency sounds. As the stirrup bangs against the oval window, a wave is transmitted through the basilar membrane. The distance this wave travels (and, subsequently, the hair cells that are stimulated) are dictated by the frequency of the sound wave. Different parts of the basilar membrane therefore respond to different frequency ranges. Loudness can be perceived by the number and duration of hair cell stimulation at that point.

QUESTIONS

13. How is the ear able to distinguish sounds of different frequencies?

14. Explain how the components of the ear amplify the pressure changes of the sound wave.

KEY IDEAS

- The outer ear collects sound waves and relays them to the eardrum.
- The ossicles in the middle ear amplify the vibrations of the eardrum and transmit them to the inner ear.
- The cochlea in the inner ear converts the vibrations to electrical signals, which are transmitted to the brain via the auditory nerves.

1.5 LOUDNESS, SENSITIVITY AND FREQUENCY RESPONSE

The perceived loudness of a sound is a difficult thing to measure, because it depends on the ears of the listener as well as on the sound. What is loud for one person may be barely audible for someone else. The sensation of loudness also depends on the frequency of the sound, and we will consider this later in the section. However, for a given person, the loudness of a sound at a certain frequency depends on the rate at which energy is transferred through a unit area by the sound wave. This is known as the sound **intensity** and it is an objectively measurable quantity.

The sound intensity, I, is the power (energy per second) through unit area perpendicular to the direction of propagation (travel) of the wave. It is measured in watts per square metre ($W m^{-2}$).

The intensity of sound with distance from a point source follows an inverse square law (Figure 27). Sound spreads out from a point source as a spherical wave. The intensity of the wave is its power per unit area, and, as the wave travels further from the source, the area of the wave front increases and so the intensity decreases.

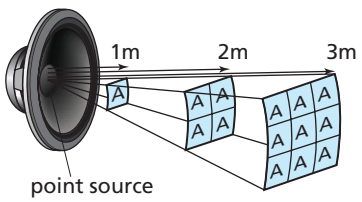


Figure 27 *Sound and the inverse square law. If the distance doubles from a sound source, the intensity drops to one-quarter $(\frac{1}{2})^2$ of its initial value. If the distance trebles, the intensity drops to one-ninth $(\frac{1}{3})^2$.*

Worked example 1

A firework explodes in the air at a distance of 60 m above the ground. If the sound power is 5.0 W and spreads out equally in all directions as a spherical wave, what would be the intensity of the sound experienced by person A 60 m from the firework and by person B 100 m from the firework?

At location of person A,

$$\text{intensity} = \frac{\text{power}}{4\pi r^2} = \frac{5.0}{4\pi \times (60)^2} = 1.1 \times 10^{-4}\ \text{W m}^{-2}$$

At location of person B,

$$\text{intensity} = \frac{\text{power}}{4\pi r^2} = \frac{5.0}{4\pi \times (100)^2} = 4.0 \times 10^{-5}\ \text{W m}^{-2}$$

In Worked example 1, which person perceives the firework as louder depends on the sensitivity of their ears. The ear's sensitivity is its ability to detect small changes in intensity. Our ears respond differently to different frequencies of sound, so the sensitivity depends on frequency. As mentioned above, the human ear is most sensitive around 3 kHz. At this frequency, a healthy ear can detect changes of intensity of around 12%.

The human ear can detect and respond to an enormous range of intensities. At a frequency of 1 kHz, the lowest intensity that a normal human ear can detect is about $1 \times 10^{-12} W m^{-2}$ (perhaps a pin dropping). The **threshold of hearing**, I_0, is defined as

$$I_0 = 1.0 \times 10^{-12} W m^{-2}$$

The loudest sound that we can detect without severe discomfort corresponds to an intensity of about $1 W m^{-2}$ (perhaps a jet aircraft taking off nearby). This is known as the 'threshold of pain' (see Figure 28).

Between these very quiet and very loud sounds, the intensity changes by a factor of 10^{12}. However, the sensation of loudness that we experience does not vary by this much. Our ears have an 'automatic volume control' that turns down the amplification when we are exposed to high-intensity sounds. The ear does not have a linear response, so turning up the intensity of a sound wave in equal steps does not increase the loudness by equal amounts. When it is quiet, a small increase in the sound intensity, such as someone coughing in a library, would seem loud. The same increase in sound intensity at a rock concert would go unnoticed.

In fact, the change in the loudness that we hear is proportional to the fractional change in intensity:

$$\text{perceived change in loudness} \propto \frac{\text{change in intensity}}{\text{initial intensity}}$$

We hear the same change in loudness when the sound intensity increases from $1 \times 10^{-12} W m^{-2}$ to $2 \times 10^{-12} W m^{-2}$ as we do when it increases from $1 W m^{-2}$ to $2 W m^{-2}$. This means that the loudness that we experience is proportional to the logarithm of the intensity, rather than to the intensity itself:

$$\text{loudness} \propto \log_{10}\left(\frac{I}{I_0}\right)$$

This logarithmic response gives the ear its enormous range.

Relative intensity level: the decibel

It is useful to use a logarithmic scale to compare sound intensities, because of the ear's logarithmic response. Relative intensity level, or simply **intensity level**, is measured in **decibels** (dB) on the **dB scale**:

$$\text{intensity level (dB)} = 10\log_{10}\left(\frac{I}{I_0}\right)$$

where I_0 is the threshold of hearing. All sound levels are compared to the threshold of hearing, which is 0 dB. A sound that has twice the intensity of the threshold of hearing will have a sound intensity level given by

$$\text{intensity level} = \log_{10}\left(\frac{2I_0}{I_0}\right) = 10\log_{10} 2 = 3\,\text{dB}$$

Doubling the sound intensity always corresponds to a change of 3 dB in the sound intensity level. By a similar calculation (see question 20):

Adding 10 dB to the sound intensity level means a factor of 10× the sound intensity.

Note that the dB scale measures sound intensity level, *not* loudness. A sound intensity level of 20 dB would seem louder to some people than to others, because of their varying sensitivity (*see Assignment 3 in Chapter 7 in Year 1 Student Book*). Table 1 relates the decibel scale to some everyday sounds.

dB	Typical example
120	Jet aircraft taking off close by
100	Disco/rock gig
80	Vacuum cleaner close by
60	General classroom noise
40	Low-level noise in a quiet room
20	Natural sounds in isolated countryside
0	Threshold of hearing

Table 1 *The dB scale quantifies everyday sounds*

Worked example 2

At a rock concert, the sound intensity level is 120 dB at a distance of 1.0 m from the speakers.

a. Calculate the sound intensity at this distance.
b. Estimate how far away you would have to be for the sound to drop to the 'safe' level of 90 dB.

a. $\text{Intensity level (dB)} = 10\log_{10}\left(\frac{I}{I_0}\right)$

Taking inverse logs, and remembering that inverse $\log_{10} x = 10^x$:

$$\text{inverse}\log_{10}\left(\frac{\text{intensity level}}{10}\right) = \frac{I}{I_0}$$

$$10^{\left(\frac{120}{10}\right)} = \frac{I}{I_0}$$

so

$$I = I_0 \times 10^{12}$$
$$= 1.0 \times 10^{-12} \times 10^{12}$$
$$= 1.0\,\text{W m}^{-2}$$

b. The sound intensity that corresponds to 90 dB is $0.001\,\text{W m}^{-2}$, because it is 30 dB less than 120 dB, and so the intensity is 10^{-3} that corresponding to 120 dB.

The ratio of intensities is therefore

$$\frac{I_2}{I_1} = \frac{0.001}{1} = 0.001.$$

Applying the inverse square law, this is equal to the ratio of the distances $\left(\frac{r_1}{r_2}\right)^2$, and since $r_1 = 1.0$ m, we have

$$\left(\frac{1}{r_2}\right)^2 = 0.001$$
$$r_2^2 = \frac{1}{0.001} = 1000$$
$$r_2 = 31\,\text{m}$$

This is an estimate, because it assumes that the speakers behave as point sources of sound from which the sound waves spread out equally in all directions. In practice, this will not be the case, because loudspeakers are designed to focus the sound in front, rather than behind.

QUESTIONS

15. a. The Saturn V moon rocket when launched generated sound of power 1×10^8 W. Assuming that the sound from the launch radiated out as a spherical sound wave, what was the sound intensity at the ear for spectators watching the launch 8 km away?

b. At what maximum distance would you have to be from the rocket before experiencing sound at $1\,W\,m^{-2}$, the threshold of pain?

16. By how many times does the dB value change when the sound intensity increases by a factor of

a. 100

b. 1000?

17. What increase in sound intensity does a change of $+6$ dB represent?

18. a. An alarm clock produces a sound of $50 \times 10^{-5}\,W\,m^{-2}$ at the ear. What is the sound intensity level in dB of the alarm clock at the ear?

b. For many people, the sound intensity at the threshold of pain is $1.0\,W\,m^{-2}$. What intensity level is this in dB?

19. a. In Table 1 the general classroom noise is 60 dB. What is the intensity in $W\,m^{-2}$?

b. A man is walking along a very quiet lane. From Table 1 the intensity level in the very quiet countryside has a value of 20 dB. All of a sudden, a motorcyclist roars past and the intensity level rises to 70 dB. Calculate the fractional change in intensity corresponding to this change in intensity level.

Stretch and challenge

20. Show that a 10 dB increase in intensity level corresponds to an increase in sound intensity of 10×.

Equal-loudness curves

The range of sound frequencies that can be heard varies from person to person, but a young adult with healthy ears can typically detect sounds from 20 Hz to 20 kHz (Figure 28). Sounds above 20 kHz are referred to as ultrasonic or ultrasound. The threshold of hearing is defined as 0 dB at 1 kHz, but our ears can detect 3 kHz sounds at even lower intensities than this. Very high-frequency sounds can be detected if the intensity is high enough. Figure 28 shows how the minimum intensity of sound that can be heard varies with frequency. A sound at around 2–3 kHz will seem louder than one at 10 kHz, even though the intensities are the same.

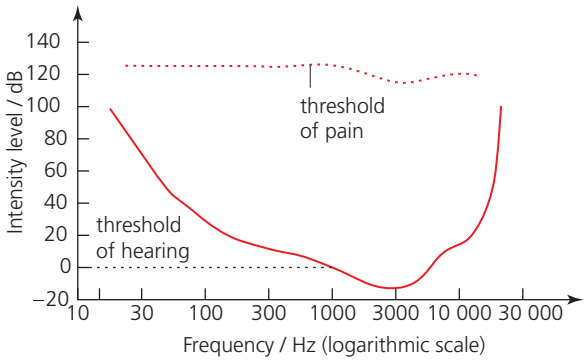


Figure 28 *The frequency range of the normal ear. The solid line represents the intensity level necessary for a sound at that frequency to be heard by a normal ear.*

While intensity, in $W\,m^{-2}$, and intensity level, in dB, are objectively measurable, the loudness of a sound is a subjective experience that varies from person to person. Also, sounds with *equal intensity* but *different frequency* are perceived by the same person to be unequal in loudness. For example, a 60 dB sound with a frequency of 1 kHz sounds louder than a 60 dB sound with a frequency of 500 Hz.

The loudness of a sound, for a given person, is measured by comparison with the loudness of a standard source of sound. A 1 kHz standard sound source is placed next to the source of unknown loudness. The intensity of the 1 kHz sound is adjusted until it sounds just as loud as the unknown source. If the 1 kHz source is then measured to be 70 dB, the unknown source is said to have a loudness of 70 phon. The **phon** is thus a unit of apparent loudness and, by definition:

1 phon is equivalent to 1 dB at 1 kHz.

The 'equal-loudness' curves for the normal ear in Figure 29 show the sound intensity level that is required to produce the same perception of loudness at different frequencies. The frequency scale is logarithmic. The curves show that the human ear is much more sensitive to sounds in the frequency range of about 1 kHz to 4 kHz than to very low- or high-frequency sounds.

It is possible to take into account the frequency dependence of our hearing when we are trying to specify sound intensity levels. For this reason, sound level meters are usually fitted with a filter called an 'A weighting filter' whose response to frequency is similar to that of the human ear. The sound intensity

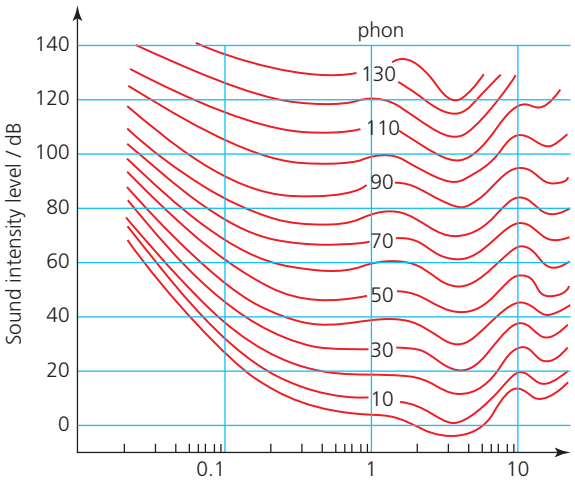


Figure 29 *Equal-loudness curves*

level is then given in units of **dBA** (or dB_A or dB(A)) and this A-weighted intensity level reading is a measure of the effect on the ear (Figure 30). The sound level meter is therefore less sensitive to low and very high frequencies. The dBA scale is used for environmental monitoring (Figure 31) and in hearing tests.

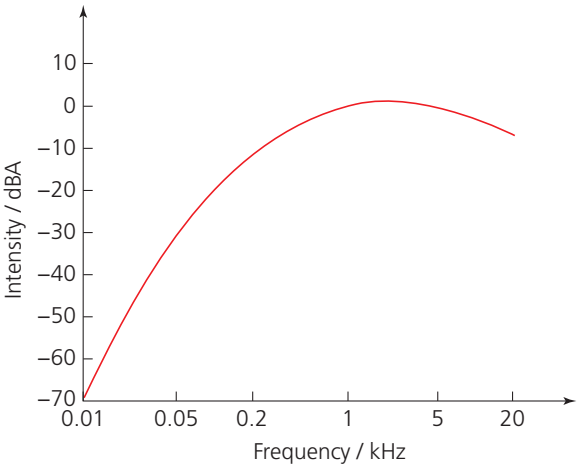


Figure 30 *The dBA scale. These weightings are used to simulate the response of the human ear at an intensity level of 40 dB . The negative dB scale indicates how the A-weighted sound intensity level is reduced to simulate the way the ear responds at different frequencies. At 1 kHz there is zero reduction, whereas at frequencies below 1 kHz there is a steady decrease to simulate the ear's lesser sensitivity to lower frequencies.*

Figure 31 *A sound level meter with a readout in dBA. This type of meter is used for environmental and industrial noise monitoring.*

QUESTIONS

21. Figure 29 shows equal-loudness curves that have been determined experimentally by subjecting volunteers to sounds of different frequencies, compared with a 1 kHz sound.

a. What is the apparent loudness, in phon, of a 1 kHz sound at 60 dB?

b. What would be the phon rating of a sound of frequency 100 Hz that has a decibel rating of 60 dB?

KEY IDEAS

- The ear's ability to detect small changes in sound intensity is called its sensitivity and is measured in decibels (dB).
- A change in the loudness that we sense is proportional to the fractional change in intensity.
- The ear thus has a logarithmic response to sound intensity. This is measured by the intensity level measured in decibels (dB):

$$\text{intensity level (dB)} = 10\log_{10}\left(\frac{I}{I_0}\right)$$

- The perceived loudness of sound is a subjective measurement and dependent on frequency. It is measured in the unit phon:

1 phon is equivalent to 1 dB at 1 kHz.

- Equal-loudness curves for a normal ear show the sound intensity level required to produce the same sensation of loudness at different frequencies.
- Sound level meters are weighted so they simulate the response of the human ear, using the dBA scale.

1.6 DEFECTS OF HEARING

The sensitivity of our hearing deteriorates as we get older. Hearing may also be damaged by exposure to loud noises, genetic factors and disease. Hearing damage can be assessed by testing a person's response to sounds at different frequencies (Figure 32). An **audiometer** is used to produce sounds at various frequencies and intensities. These are played to the person under test through headphones, or using an electronic vibrator held on the bone behind the ear. The audiometer is adjusted to read 0 dB at the intensity level at which the sound can just be detected by a person with normal hearing. If a patient can only detect sounds at this frequency that are 50 dB

Figure 32 *An audiometry test*

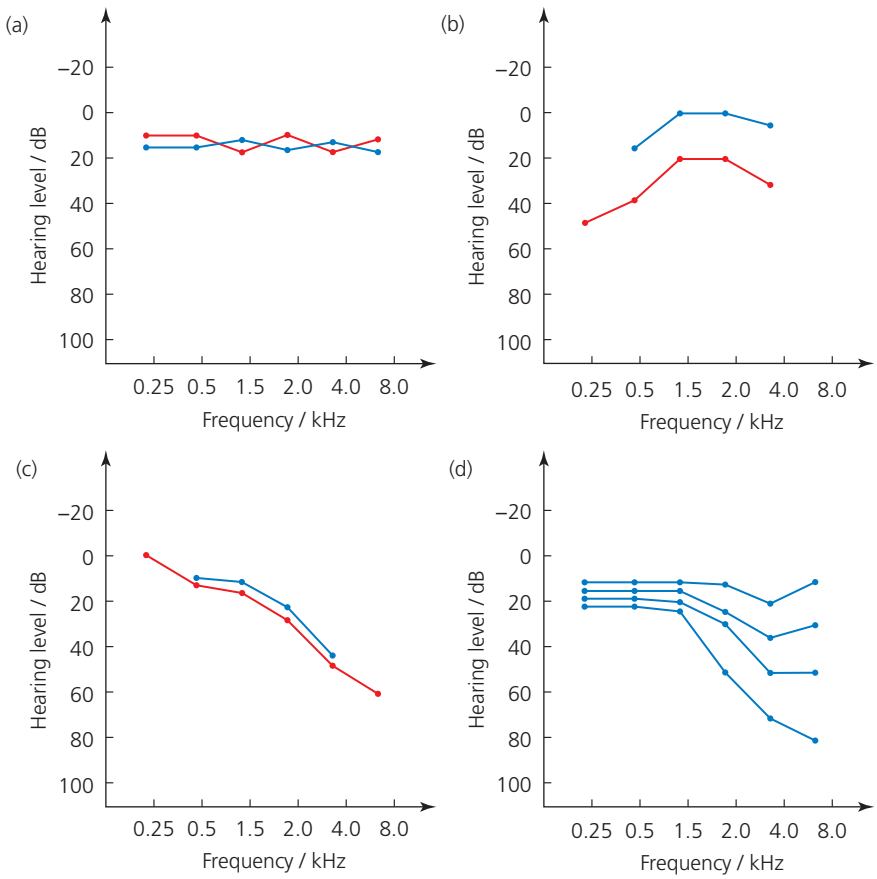


Figure 33 *Typical audiograms for: (a) normal ear; (b) conductive hearing loss due to damage to structures in the outer or middle ear; (c) sensorineural deafness due to damage in the inner ear; and (d) progressive deafness due to exposure to high levels of noise. In (a) to (c), the tests are conducted via headphones (red lines) or using an electronic vibrator held on the bone behind the ear (blue lines). In (d), the curves are for various exposure times (increasing downwards from top to bottom).*

louder than this, they are said to have a hearing loss of 50 dBA. The test is carried out at a number of frequencies, and the resulting plot of hearing loss against frequency is known as an **audiogram**. Some typical audiograms are shown in Figure 33.

The audiogram of a person who has suffered years of exposure to loud noise is shown in Figure 33d. Damage may have occurred to the hair cells in the cochlea. Exposure to high-intensity, high-frequency sound is particularly dangerous, but audiograms always show maximum loss at 4 kHz, no matter what the frequency of noise that caused the damage.

Age-related hearing loss tends to affect high-frequency sounds the most (Figure 34). Frequencies below 500 Hz are barely affected. Although human speech is mostly covered by this range, there are higher-frequency harmonics in sounds such as 's' and 't', which older people may find difficult to detect.

The red curve in Figure 35 is an equal-loudness curve at the threshold of hearing for a person with hearing loss due to age, compared with a person with normal hearing (black line). Again, it is clear that the loss is worse at higher frequencies than at lower. The purple curve is for a person suffering hearing loss due to excessive exposure to noise – there is a marked 'notch' in threshold hearing over a narrow range of high frequencies.

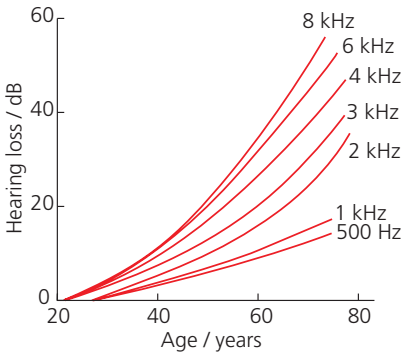


Figure 34 *Age-related hearing loss. The curves show progressive hearing loss for sounds of different frequencies.*

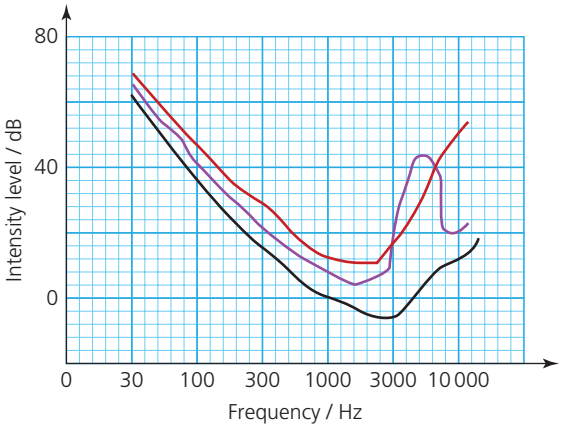


Figure 35 *The effect on the threshold of hearing equal-loudness curve (black) due to age (red) and exposure to excessive noise (purple)*

QUESTIONS

22. In Figure 36 graphs A, B and C show three different audiograms for three people who have had their hearing tested.

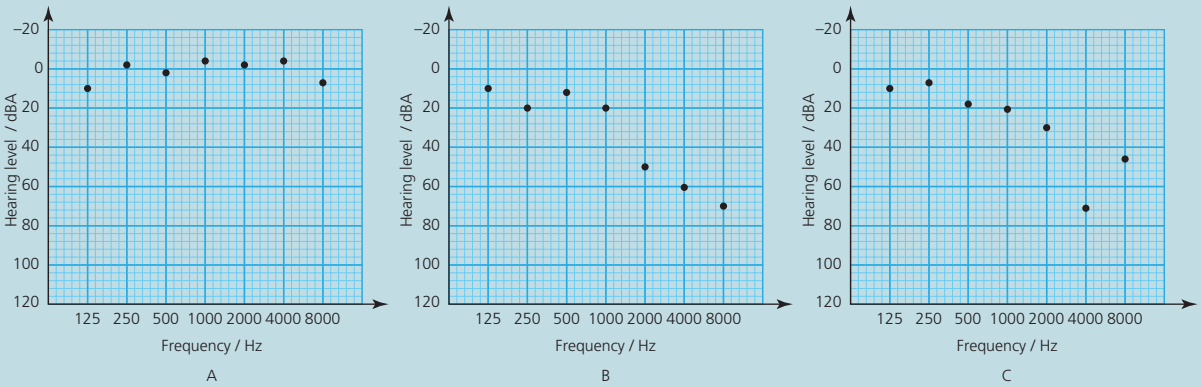


Figure 36

a. Explain why the dBA scale is used for sound intensity for audiograms.

b. Describe the significant difference between graph A and graph B, and suggest a possible cause.

c. Describe the significant difference between graph A and graph C, and suggest a possible cause.

KEY IDEAS

- Defects in hearing can be due to exposure to excessive noise, to genetic factors, to ageing or to disease.
- An audiometer is used to test the frequency response of the human ear.
- An audiogram is a graph of sound intensity level in dBA against frequency. The shape of the graph can show if hearing level has changed due to hearing loss and what type of defect might be causing it.

ASSIGNMENT 1: UNDERSTANDING ELECTRONIC HEARING

(PS 1.2)

Read the text here and then answer the questions that follow.

Our hearing deteriorates with age and is an example of sensorineural hearing loss. We start to lose sensitivity to higher-frequency sound, and this can be detected on an audiogram (see Figure 33c). Sensorineural hearing loss is defined as damage to the hair cells in the cochlea and/or damage to the nerve cells that transmit electrical impulses along the auditory nerve to the brain. This type of hearing loss can also be caused by infections or be inherited, and can lead to severe or even total deafness at any range of frequencies. The hair cells in the cochlea do not repair themselves once they are damaged, so the hearing loss is permanent.

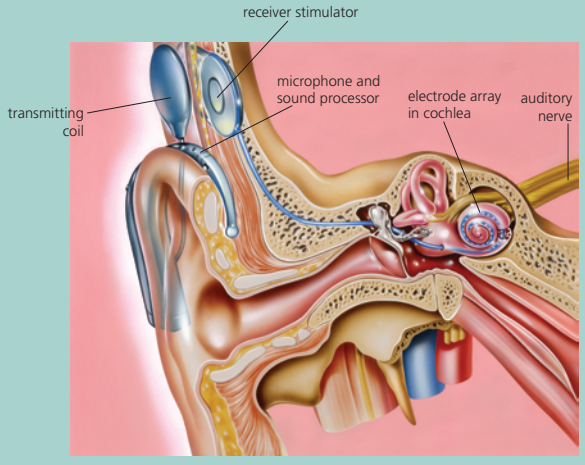


Figure A1 *The components of a cochlear implant*

However, help is available in the form of a cochlear implant (Figure A1), which can provide a sense of hearing to people who have severe to total deafness caused by sensorineural hearing loss. Unlike a hearing aid, which simply amplifies sound, the implant provides electrical stimulation directly to the nerve fibres, bypassing non-functional hair cells in the cochlea.

Sound is picked up by a microphone and sound processor in a unit positioned behind the ear (Figure A1), which sends an electrical signal to a transmitting coil attached to the head. Implanted inside the skull is a receiver stimulator, which sends signals to an electrode array inserted in the cochlea. These electrodes send signals along the auditory nerve system to the brain.

In the functioning human ear, the analysis of speech sounds relies on the separation of frequency information in the electrical impulses that come from different parts of the cochlea (see Medical Physics section 1.4). Groups of hair cells on the basilar membrane in the cochlea respond selectively to different narrow bands of frequency as sound travels through the membrane. Cells at the base of the cochlea (near the oval window, see Figure 26) respond to the highest frequencies and cells at the apex respond to the lowest. The nerve endings are thus stimulated in a highly organised manner.

The electrode array in the cochlear implant stimulates cochlear nerve cells where the hair cells have been damaged or are non-functional. Modern implants contain up to 22 separate electrodes that can stimulate different parts of the cochlea, corresponding to different frequency ranges, along

its length. Incoming sound in the approximate range (or 'bandwidth') of 200 to 7500 Hz is separated by the sound processor into as many as 20 frequency bands (or 'channels'), generates matching electrical impulses and distributes these among the electrodes. The bandwidth assigned to each individual electrode depends on the patient's needs. The result is that the cochlea is stimulated in a way that bears some resemblance to the stimulation of groups of hair cells along the basilar membrane in a normal ear, producing the perception of sound at different frequencies.

A cochlear implant does not restore or create normal hearing. However, it can give a deaf person a useful auditory understanding of the environment and help him or her to understand speech and even enjoy music.

QUESTIONS

A1 What is meant by *sensorineural* hearing loss?

A2 On axes like those in Figure A2, sketch an audiogram (the frequency response) for someone with age-related deafness.

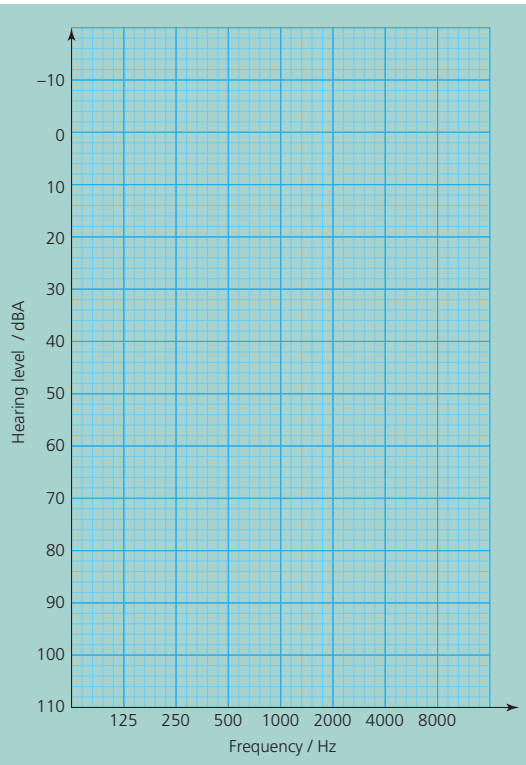


Figure A2

A3 Explain why the sound processor in a cochlear implant divides the incoming sound into separate channels of a particular frequency bandwidth.

A4 Normal human hearing has a bandwidth of about 20 Hz to 20 kHz. Suggest why the processor needs to provide channels only over a bandwidth of 200 Hz to 7500 Hz to enable speech to be recognised.

PRACTICE QUESTIONS

1. **a.** With the aid of a ray diagram, explain what is meant by *hypermetropia*.

   b. A long-sighted person has a near point of 65 cm. Calculate the power of a corrective lens needed to bring the near point to a distance of 25 cm from the lens.
   You may assume that the refractive system of the eye acts as a single thins lens at the cornea and the distance from the cornea to the retina is 17 mm.

2. **a.** Copy and complete the ray diagram in Figure Q1 to show the formation of the image of a real object by a diverging lens.

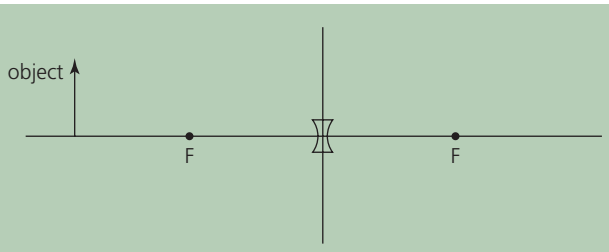


Figure Q1

b. Define the power of a lens.

c. A lens of focal length – 0.56 m is used to correct a defect of vision of an eye.

i. Name this defect of vision .

ii. The defective eye has an unaided near point at 0.15 m from the eye. Calculate the aided near-point distance, giving your answer to an appropriate number of significant figures.

iii. Another person was found to suffer from astigmatism. State the format of the prescription to correct this defect.

AQA June 2010 Unit 5B Q1

3. In the eye, rods and cones are used to detect light incident on the retina.

a. Describe how the rods and cones are distributed over the surface of the retina.

b. On axes like those in Figure Q2 sketch **three** curves to show how the response of each of the three types of cone found in the retina varies with the wavelength of light. Label each of the three curves with the cone colour to which it refers.

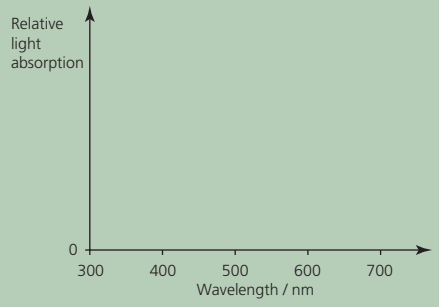


Figure Q2

c. State the condition that must be satisfied for two objects to be resolved as individual images on the retina.

d. Explain how the resolution of the image of an object seen in very dim white light compares to that of the image of the same object seen in bright white light.

AQA June 2013 Unit 5B Q1

4. a. i. What is meant by the *intensity of sound* and the *intensity level*?

ii. Define the threshold of hearing, I_0.

b. If the threshold of hearing of a human ear is 55 dB at a certain frequency, what is the intensity of the sound incident on the ear at the same frequency?

c. In a firework display, a rocket explodes with a bang of 5.0 W. Assuming that a spectator is standing 40 m from the explosion and the sound radiates outwards over a spherical area, calculate:

i. the intensity of the sound incident on the spectator's ear

ii. the intensity level.

5. a. Explain what the *ossicle*s in the human ear are and how they act to amplify the pressure changes of a sound wave in the human ear.

b. Why is an A-weighting filter used in sound intensity level measurement?

c. An environmental health officer is called in to monitor complaints of loud music coming from a garden in a house. She records an intensity level reading of 110.2 dB on her sound level meter.

i. Calculate the sound intensity at the meter.

ii. The sound source in the house is emitting sound equally all directions. If she is standing at a distance of 5.0 m, what is the power emitted by the music source in the garden?

2 BIOLOGICAL MEASUREMENT

PRIOR KNOWLEDGE

You will have a good understanding of electric potential difference and resistance. From your GCSE Science (Biology) you should have a knowledge of the structure of the human heart and how blood is pumped around the body. You should also know about the actions of nerves and how they conduct electrical impulses to control muscle action.

LEARNING OBJECTIVES

In this chapter you will learn how it is possible to make electrical measurements of the heart, how the resultant electrocardiograph (ECG) trace arises from the electric potentials in the heart muscle, and how the interpretation of the ECG trace can indicate how the heart is functioning and provide important diagnostic information.

(Specification 3.10.3)

2.1 A BRIEF HISTORY OF ELECTROCARDIOGRAPHY

Since the 18th century, scientists have been aware that there is a relationship between electric potential and the movement of muscles in the body. The Italian scientist Luigi Galvani was the first to discover this. He performed experiments at the University of Bologna involving frogs. He found that a frog's leg (from a dead frog) could be made to twitch by applying a potential difference across it. This led him to believe that animals and humans produce electricity, which he surmised was stored in the muscles. Another Italian scientist contemporary with Galvani, Alessandro Volta, showed that, in fact, the electricity that produced the muscular contractions originated from the combination of the metals used in Galvani's experiment together with the fluid in the frog's leg. These were behaving like a battery. Although Galvani's conclusions regarding 'animal electricity' had been wrong, his experiments showed that biological tissue could conduct electricity and were important in laying the foundations of what we now understand as the electrical basis of nerve impulses and bioelectricity, and ultimately electrocardiography. 'Electrocardiography' means monitoring the electrical activity of the heart (*kardia* is the Greek for 'heart').

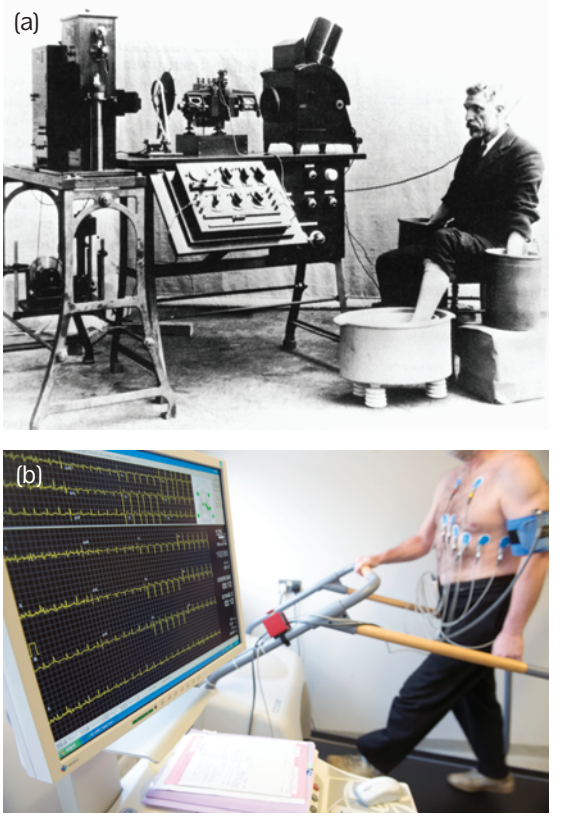


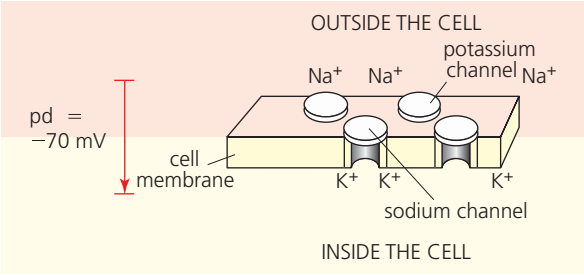
Figure 1 *ECG machines in use: (a) in 1908, and (b) today*

In 1903, the Dutch physiologist Willem Einthoven invented an instrument that was called a string galvanometer and with it he developed an improved method for measuring the electrical changes that take place in the body upon the contraction of the heart. Einthoven was able to detect and identify a number of different kinds of electrical waves associated with a beating heart. The equipment was very large and bulky and required a team of five people to operate it, but it was the forerunner of today's **electrocardiograph** or **ECG** machine, an important aid in medical diagnosis (Figure 1). Although the modern machine is much smaller and more accurate, the underlying principles of measuring electric potentials produced by the action of heart muscle are fundamentally the same.

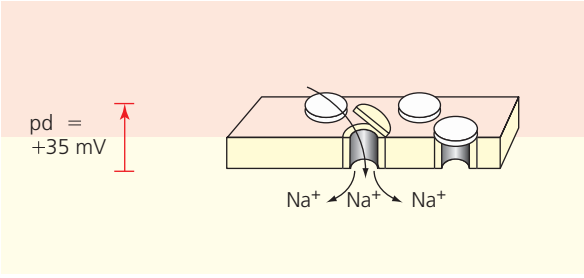
2.2 ELECTRICAL SIGNALS IN THE BODY

In order to understand the pattern of electric potentials in the heart muscle, it helps first to consider nerve cells. Electrical signals are conducted around the body along these specialised cells. Nerves are made of fibres that can be up to a metre long, but are usually only a few micrometres in diameter. Signals travel along these fibres in the form of a changing potential difference, which is generated by the movement of ions across cell membranes.

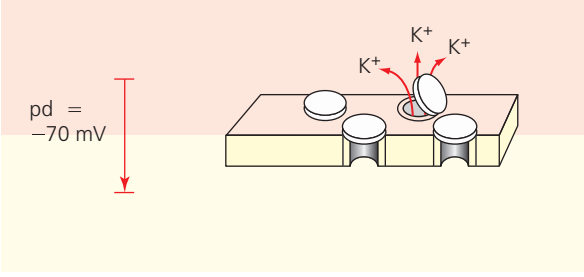
The membrane of nerve cells allows water to diffuse freely into or out of the cell. The membrane is much less permeable to the passage of sodium Na^+ and potassium K^+ ions, but there are 'channels' through which ions can be transferred into and out of the cell. In all cells in their resting state, there is a high potassium concentration and a low sodium concentration, whereas the fluid surrounding the cell has a low potassium concentration and a high sodium concentration. It is the imbalance in the numbers of sodium and potassium ions that causes a potential difference across the cell membrane. The concentration gradient tends to cause potassium ions to leave the cell, carrying their single positive charges with them. Potassium ions continue to leave the cell until the excess positive charge outside the cell is large enough to stop them. This happens when the inside of the cell is at a potential of –70 mV compared to its surroundings. At this **resting potential**, an equilibrium exists between the concentration gradient and the potential gradient, and the cell is said to be **polarised** (Figure 2a).



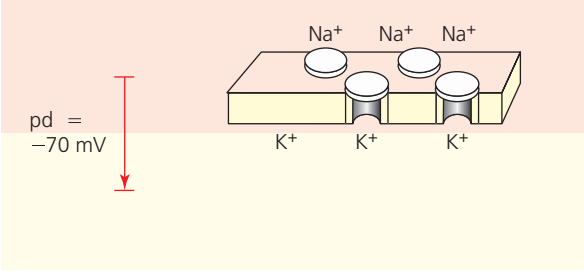
(a) The cell resting and polarised



(b) An electrical signal is initiated: depolarisation



(c) Repolarisation



(d) Original ion concentrations are restored

Figure 2 *(a) In this resting state, there is an imbalance of ions – more potassium ions inside the cell and more sodium ions outside. This causes a potential across the membrane of –70mV. The cell is* **polarised**. *(b) Sodium channels open and sodium ions rush into the cell. The inside of the membrane becomes positive with respect to the outside – the membrane is* **depolarised**. *(c) The increase in potential in the cell closes the sodium channels and opens the potassium channels. Potassium ions leave the cell. The potential drops again to –70mV; it is* **repolarised**. *(d) The original sodium and potassium ion concentrations are restored.*

When a nerve impulse is initiated, channels in the cell membrane suddenly become much more permeable to sodium ions. For a period of about 1 ms, sodium ions move into the cell (Figure 2b).

This causes the potential difference across the cell membrane to rise to +35 mV before the sodium channels close. The cell is **depolarised**. Reaching this positive potential triggers another set of channels, which permit potassium ions to leave the cell. The resting potential of –70 mV is then restored and the cell is **repolarised** (Figure 2c). The original ion concentrations are restored (Figure 2d) (by means of a 'sodium–potassium pump', which need not concern us here).

This pattern of changing potentials is called an **action potential** (Figure 3). A changing potential in one part of the nerve initiates a new action potential at an adjacent site. In this way, the action potential propagates along the fibre.

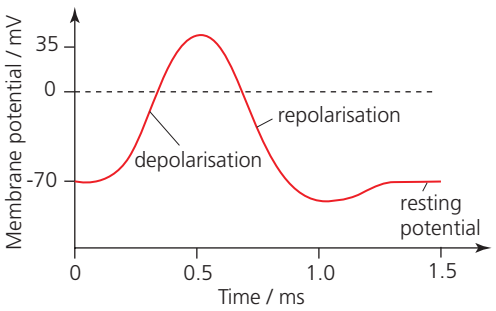


Figure 3 *The action potential of a nerve cell*

Electrical activity in the heart

During a heart beat, the four chambers of the heart contract in a sequence that is controlled by electrical signals from a 'node' in the heart that acts as a natural pacemaker. This node, rather than nerve cells, generates the electrical impulses that spread rapidly across the upper chambers of the heart – the atria – causing them to contract. The signal is then transferred to the lower chambers – the ventricles. The walls of the ventricles contract uniformly, exerting pressure on the blood from all sides. The action potential in the heart (Figure 4) is similar to that in nerve fibres, but it travels more slowly and lasts longer. In particular, the depolarisation stage lasts longer, about 0.2 s compared to 5 ms in the action potential generated by a nerve cell. The longer cycle overall allows the heart muscle to relax completely before the next contraction.

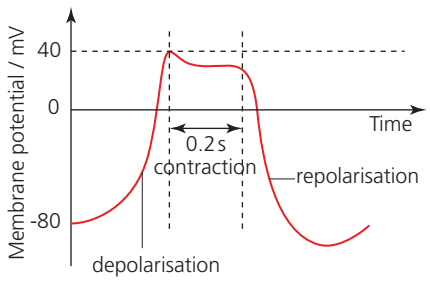


Figure 4 *The action potential of heart muscle*

QUESTIONS

1. **a.** Explain what is meant by a *resting potential* and an *action potential* in nerve activity.
   b. How does an action potential in heart muscle differ from that in nerve cells?
   c. What happens in the cycle of the heart beat at
      i. depolarisation
      ii. repolarisation?

KEY IDEAS

- The heart muscle is stimulated by pulses of electric potential in the muscle cells.
- The pulse of electrical potential is called the action potential and causes the heart to beat (see Figure 4):

  relaxed state – cells polarised

  contraction of muscle – depolarisation of cells

  relaxation of muscle – repolarisation of cells.

2.3 THE NORMAL ECG WAVEFORM

The electrical signal from the heart spreads through the surrounding tissues and fluid, and it can be detected by electrodes taped to the skin. It can be difficult to make a good electrical contact between the electrodes and the skin, so gentle abrasion is used to remove dead skin cells and an electrically conducting gel is used. Silver electrodes coated with silver chloride are used, because these do not react with chemicals produced by the skin. The adhesive tape prevents voltage spikes caused by electrostatic effects due to motion between the electrodes and the skin.

The measured voltages are so small that they need to be amplified with a high-gain amplifier. Unfortunately, unwanted signals (noise) are also amplified. A particular problem comes from induced voltages due to nearby

mains-operated equipment. Mains voltage alternates at 50 Hz, so small induced electromotive forces (emfs) occur in the ECG equipment. To combat this effect, all electrical leads are screened and the ECG equipment is carefully shielded. In addition, the patient should remain relaxed and still during the measurement.

The resulting voltage trace, or electrocardiograph (ECG), is viewed on a screen in real time, as seen in Figure 1b, and can be printed out as a record.

The shape of a normal healthy ECG waveform, with its characteristic PQRST curve, is shown in Figure 5. While the action potentials are tens of millivolts (see Figure 4), the actual electrical signals measured on the surface of the skin are much less than this – about 1 mV in amplitude.

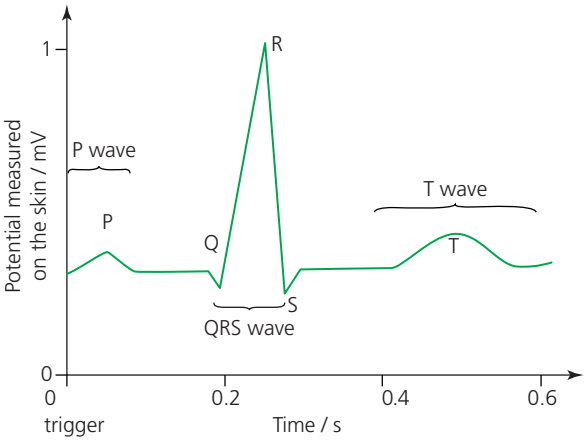


Figure 5 *One cycle of a normal ECG trace*

The features of the waveform arise from the nature of the action potential:

- The small change in potential at P, or 'P wave', in Figure 5 corresponds to the depolarisation of the muscle cells of the atria.
- The 'QRS wave' is due to the depolarisation of the muscle cells of the ventricles and corresponds to the contraction of the ventricles.
- The repolarisation of the atria is masked by the much larger QRS potential from the ventricles.
- The 'T wave' is due to the repolarisation of the muscle cells of the ventricles and corresponds to the relaxation of the ventricles.

The shape of the ECG trace gives important diagnostic information about the state of the patient's heart. The repeating pattern of the trace (Figure 6) enables the **pulse rate** to be easily determined. The pulse rate is the number of ventricular contractions per minute.

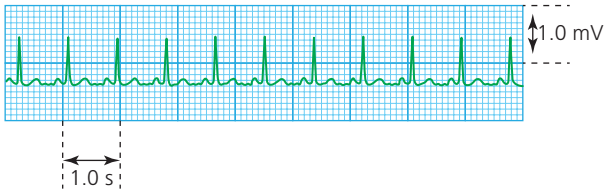


Figure 6 *An ECG trace from a patient with normal heart function*

QUESTIONS

2. a. Estimate the pulse rate of the patient whose ECG is shown in Figure 6.

b. How would the ECG trace differ for a patient with

i. high pulse rate

ii. low pulse rate?

3. A patient is diagnosed with 'ventricular standstill'. There is no contraction of the ventricles. Sketch the likely ECG trace from this patient. Include a scale on the potential axis.

KEY IDEAS

- An electrocardiograph (ECG) is a graphical display of the heart's electrical activity as a function of time.
- Precautions need to be taken when making ECG measurements, which include ensuring good electrical contact, shielding leads from electrical interference and amplifying the signal with a high-gain, low-noise amplifier.
- For a normal heart, the shape of the ECG has three important features:
  - the P wave, which occurs during depolarisation and contraction of the atria
  - the QRS wave, which is due to the depolarisation and contraction of the ventricles
  - the T wave, which is due to the repolarisation and relaxation of the ventricles.
- The shape of the ECG can be used to diagnose abnormalities in the function of the heart.

PRACTICE QUESTIONS

1. a. An ECG trace is to be obtained for a patient. State and explain the procedure and some design features of the equipment needed to ensure a good trace is obtained. The quality of your written communication will be assessed in this question.

b. Figure Q1 shows an ECG trace for a healthy person.

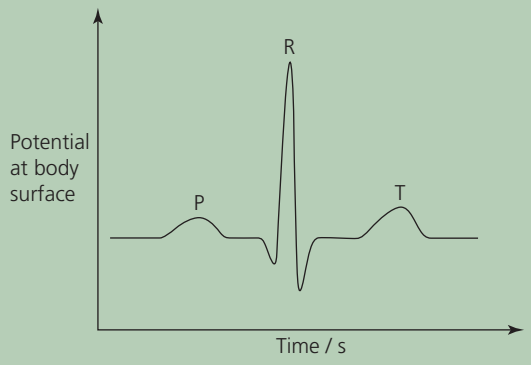


Figure Q1

i. Sketch the trace and add a suitable scale and unit to the potential axis.

ii. Add a suitable scale to the time axis on your sketch.

c. State the electrical events which give rise to the points P, R and T.

AQA June 2011 Unit 5B Q4

3 NON-IONISING IMAGING

PRIOR KNOWLEDGE

You will need to know that sound travels as waves that can be reflected at boundaries between different media, and are attenuated as they pass through a medium. You will need to recall from *Chapter 7 in Year 1 Student Book* how optical fibres carry light by repeated total internal reflection. You will need some knowledge of the structure of the hydrogen atom. You should know from Chapter 7 of this book how magnetic fields can be produced from current-carrying coils and will understand, from Chapter 2, the phenomenon of resonance.

LEARNING OBJECTIVES

In this chapter you will learn how non-ionising radiation – that is, radiation that does not have sufficient energy to ionise matter – can produce images of the internal structures of the human body without risk of harm. The methods of imaging using non-ionising radiation that you will study here are ultrasound scanning, magnetic resonance imaging (MRI) and optical imaging, including endoscopy, which uses optical fibres to produce images both for diagnostic purposes and during surgical treatment.

(Specification 3.10.4)

3.1 ULTRASOUND IMAGING

Ultrasound is used to produce detailed images in real time at relatively low cost and without risk to the patient. It is suitable for diagnosing a wide range of conditions, from heart valve disorders to tumour detection, as well as for monitoring fetal development. It works by transmitting sound pulses through the body and detecting the returning echoes.

Ultrasound consists of sound waves of such high frequency that they cannot be detected by the human ear. Strictly speaking, the ultrasound range starts at about 20 kHz, but medical applications use much higher frequencies, typically between 1 MHz and 20 MHz.

Generating ultrasound

Sound waves are longitudinal pressure waves that are produced by a vibrating object (*see section 5.1 in Chapter 5 in Year 1 Student Book*).

The ultrasonic generators used for medicine need to vibrate several million times per second. They use **piezoelectric transducers** to convert electrical signals to pressure waves. Figure 1 shows the principle.

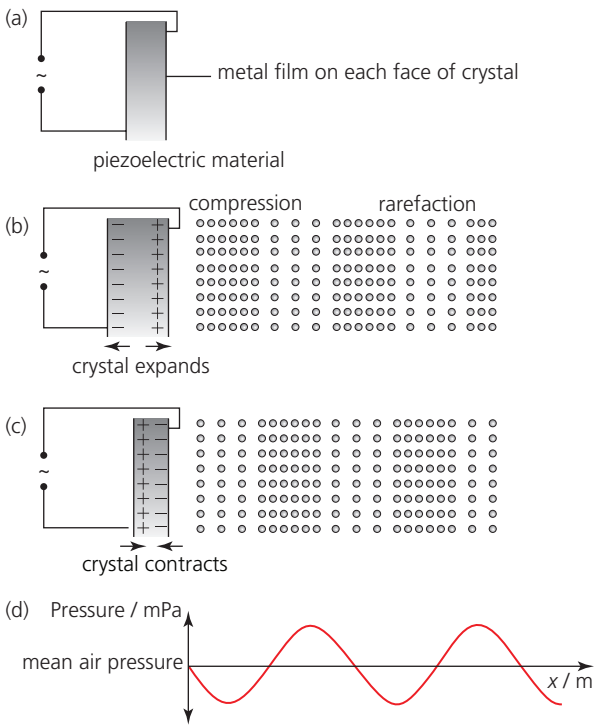


Figure 1 *Generation of a pressure wave by the piezoelectric effect. (a) An ac signal is applied across a crystal of piezoelectric material. (b) One polarity causes the crystal to expand, pushing the air in front of it. (c) When the polarity reverses, the crystal contracts. (d) This sequence is repeated millions of times each second, causing a pressure wave (ultrasound) to be transmitted.*

Piezoelectric materials are crystalline solids that deform when a potential difference (pd) is applied across them. This is called the **piezoelectric effect**. An *alternating* pd applied across the material causes it to vibrate at the same frequency as the electrical signal. A piezoelectric crystal has a natural frequency that is determined by the material and by its physical dimensions. When the frequency of the ac signal matches the natural frequency of the crystal, there is **resonance**: the ultrasonic wave has its largest amplitude for a given input of electrical energy (see section 2.4 of Chapter 2).

Figure 2 shows the structure of a piezoelectric transducer.

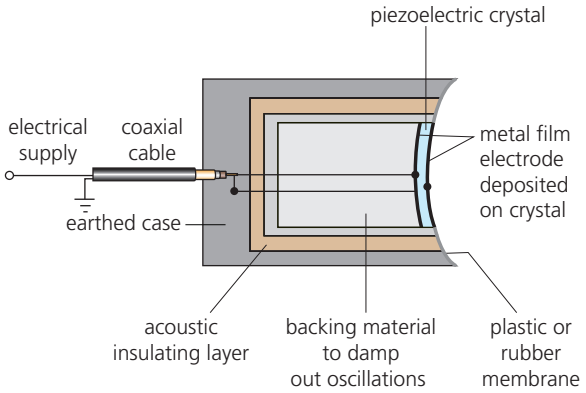


Figure 2 *A piezoelectric transducer*

The piezoelectric effect also works in reverse. If a piezoelectric crystal is deformed, by an incident ultrasonic wave, for example, a potential difference is generated across the crystal. A piezoelectric transducer can therefore be used to detect, as well as to generate, ultrasound pulses.

There is a special backing material in the transducer to dampen excessive vibrations. This allows short pulse lengths, and therefore better resolution, to be achieved – so that the emitted pulse ceases before the echo pulse arrives. Otherwise, because of the very short return time, there would be overlapping at the transducer between the transmitted and reflected pulses.

Ultrasound waves in the body

Ultrasound travels as longitudinal waves through the body, moving through different tissues at different speeds. The speed of sound, c, in a material depends on the elasticity of the material and its density approximately according to

$$c = \sqrt{\frac{K}{\rho}}$$

where K is the **bulk elastic modulus** and ρ is the average density of the material. The frequency of the sound also has a slight effect on its speed. The bulk elastic modulus is a quantity that indicates the resistance of a substance on uniform compression and is measured in pascal (Pa). It is similar to the Young modulus except that the bulk elastic modulus is defined for compression where the pressure is applied from all directions uniformly. The Young modulus is defined only for one axis of the material, such as the length of a wire. So the bulk elastic modulus measures the change in volume when a pressure is applied, whereas the Young modulus measures the change in length when a tensile stress is applied (*see section 12.3 in Chapter 12 in Year 1 Student Book*).

As ultrasound travels through the body, some of its energy is reflected as the wave passes from one material to another. These 'echoes' from the boundaries between different materials are used to build an image of the internal structure of the body. The amount of energy that is reflected at each interface depends on each material's **acoustic impedance**, Z, defined as the product of the density, ρ, and the speed of sound, c, in that material:

$$Z = \rho c$$

Z has the unit $\mathrm{kg\,m^{-2}\,s^{-1}}$.

The acoustic impedance is a measure of the opposition that a material presents to an acoustic wave. In ultrasound imaging, a lower acoustic impedance means that an ultrasonic pulse can travel through the tissue with greater ease. Some typical acoustic impedances for body materials are listed in Table 1.

Material	Speed of sound / $\mathrm{m\,s^{-1}}$	Average density / $\mathrm{kg\,m^{-3}}$	Acoustic impedance / $\mathrm{kg\,m^{-2}\,s^{-1}}$
Bone (dense)	3500	1850	6.48×10^6
Fat	1450	952	1.38×10^6
Muscle	1580	1080	1.71×10^6
Skin	1730	1150	$1.99 \times 10^6\ \mathrm{kg\,m^{-2}\,s^{-1}}$

Table 1 *Some typical values of acoustic impedance*

QUESTIONS

1. The speed of sound in skin is $1730\,m\,s^{-1}$ and the density of skin is $1150\,kg\,m^{-3}$. Show that the value of the acoustic impedance for skin in Table 1 is $1.99 \times 10^6\,kg\,m^{-2}s^{-1}$.
2. A gel used in performing ultrasound scans has a density of $1060\,kg\,m^{-3}$ and a bulk elastic modulus of 2.03 GPa. Show that the acoustic impedance of the gel is $1.65 \times 10^6\,kg\,m^{-2}s^{-1}$ if the speed of sound in the gel is 1192 m.

Stretch and challenge

3. The speed of sound in a material is given approximately by $c = \sqrt{K/\rho}$, where K is the material's bulk elastic modulus (unit Pa) and ρ is its density. Show that $\sqrt{K/\rho}$ has the unit $m\,s^{-1}$.

When a plane wave front strikes the boundary between two materials, at normal incidence (Figure 3), the proportion of the incident intensity, I_i (in $W\,m^{-2}$, or $W\,cm^{-2}$) that is reflected is given by

$$\frac{I_r}{I_i} = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2 = \alpha$$

where I_r is the intensity of the reflected wave, and Z_1 and Z_2 are the acoustic impedances of the two media. This formula only applies for normal incidence. The ratio $\frac{I_r}{I_i}$ is known as the **reflection coefficient**, symbol α.

The intensity of the *transmitted* wave, I_t, is the difference between the intensities of the incident and reflected waves:

$$I_t = I_i - I_r$$

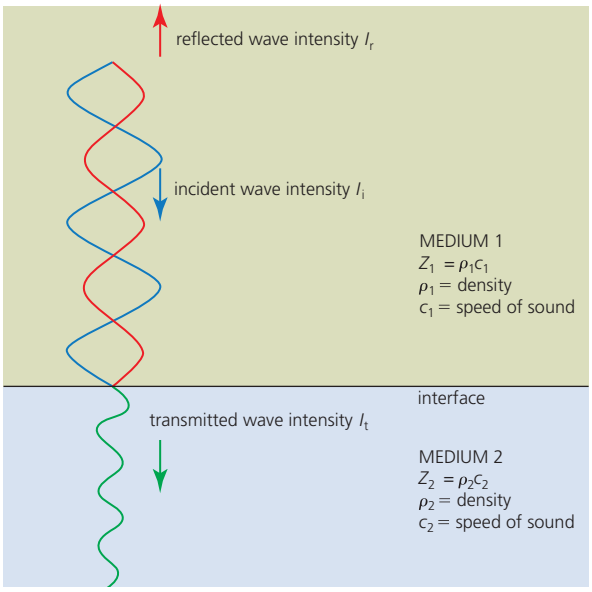


Figure 3 *Reflection of ultrasound at a boundary*

Worked example

The intensity of a diagnostic ultrasound wave is $30\,mW\,cm^{-2}$.

a. Find the reflection coefficient α for the ultrasound wave as it strikes the boundary between fat and muscle, at normal incidence.

b. Find the intensity transmitted into the muscle.

a. Using the formula for $\frac{I_r}{I_i}$ and substituting values for acoustic impedance from Table 1 gives

$$\alpha = \frac{I_r}{I_i} = \left(\frac{Z_{muscle} - Z_{fat}}{Z_{muscle} + Z_{fat}}\right)^2$$

$$= \left(\frac{1.71\times10^6 - 1.38 \times 10^6}{1.71\times10^6 + 1.38 \times 10^6}\right)^2 = \left(\frac{0.33}{3.09}\right)^2 = 0.011$$

b. The value found in part **a** gives the reflected intensity as

$$I_r = \alpha I_i = 0.011 \times 30 = 0.33\,W\,cm^{-2}$$

So the transmitted intensity is

$$I_t = I_i - I_r = 30 - 0.33 = 29.67\,W\,cm^{-2}$$

As ultrasound travels through a material, its intensity is reduced because the material absorbs some of the wave's energy, transferring it to internal energy. In medical applications, this causes a heating effect in the body tissue. In addition, some of the wave's energy is scattered from its original path, which also reduces the intensity.

The reduction of intensity as the wave travels through a material is known as **attenuation**. Equal thicknesses of the same material will attenuate the intensity of the wave by the same proportion. For example, travelling through the first centimetre may reduce the wave's intensity by one-half, the second centimetre will reduce the intensity by a one-half again, so that overall the intensity will drop to one-quarter of its original value. This is the characteristic 'constant ratio' pattern of an exponential drop. The graph of intensity versus distance (Figure 4) is an exponential decay curve, like that for radioactive decay (Chapter 9) and that for capacitor charging/discharging current (Chapter 6).

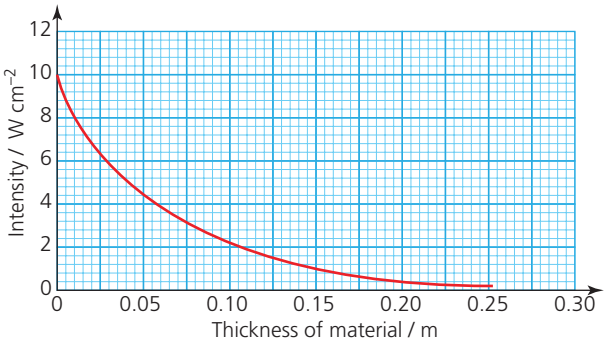


Figure 4 *Exponential attenuation of ultrasound intensity in a material*

The intensity at a distance x in the material, I_x, is given by

$$I_x = I_0 e^{-\mu x}$$

where I_0 is the original intensity on entering the material (at $x = 0$) and μ is the **intensity attenuation coefficient** for that material, measured in m^{-1}. In practice, the intensity ratio I_x/I_0 is often expressed in decibels (see Medical Physics section 1.5).

For most soft tissue in the body, μ is almost proportional to frequency, so high-frequency ultrasound is attenuated more than low-frequency ultrasound.

QUESTIONS

4. a. Calculate the fraction of ultrasound reflected at an air–skin boundary, using the value for the acoustic impedance of skin used in question 1 (take the acoustic impedance of air to be $4.29 \times 10^2 kg m^{-2} s^{-1}$) and show that its value is 0.999.

b. A gel is used between the ultrasonic transducer and the skin. Using your values for acoustic impedance for skin and gel calculated in questions **1** and **2**, respectively, calculate the fraction of ultrasound reflected at a gel–skin boundary and show that its value is 8.73×10^{-3}.

KEY IDEAS

- Ultrasound is generated and detected by a piezoelectric transducer.
- Images are generated by ultrasound reflections from boundaries between tissues with different acoustic impedances.
- The acoustic impedance of a tissue, Z, is given by
$$Z = \rho c$$
where ρ is the density of the material and c is the speed of sound through it.
- The reflection coefficient is the fraction of the incident ultrasound intensity reflected back from a boundary, and is given by
$$\alpha = \frac{I_r}{I_i} = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2$$
- Within a material, the intensity of the ultrasound undergoes exponential attenuation. The greater the frequency, the greater the attenuation.

A-scans and B-scans

Early applications of ultrasound used simple echoes to measure the position of body parts. This technique is known as an **A-scan** or **amplitude scan** (Figure 5). In fact, one of the first applications of ultrasound was to measure the position of the mid-line of the brain. A transducer placed on the skin emits a single, short pulse. Echoes return from boundaries between different body tissues. The output is displayed as a graph on an oscilloscope, showing how the echo amplitude changes with time. If you know the speed of sound in body tissue, it is possible to calculate the distance to any boundaries. A-scans are used in ophthalmology to measure distances in the eye.

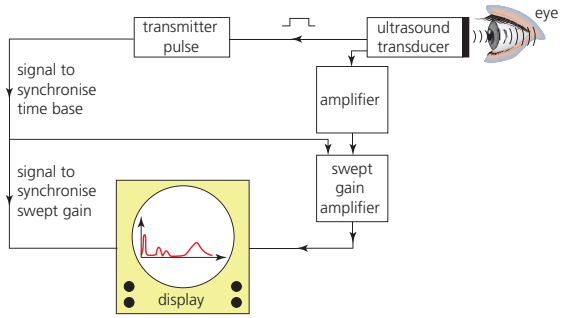


Figure 5 *A simplified A-scan system and a typical output. The 'swept gain amplifier' increases the amplification of echoes that return from deeper inside the body to compensate for attenuation.*

A **B-scan** or **brightness scan** creates an image by scanning the ultrasound beam over the patient. A rapid sequence of A-scans is carried out as the beam sweeps across (Figure 6). The strength of the echo signal is used to control the brightness of a picture element, rather than the height of a trace on an oscilloscope. This is combined with information about the position of the transducer to create an image (Figure 7).

The number of complete images, or frames, that can be produced each second depends on how long it takes for the echoes from each pulse to return, and on how many pulses are used to create the full picture.

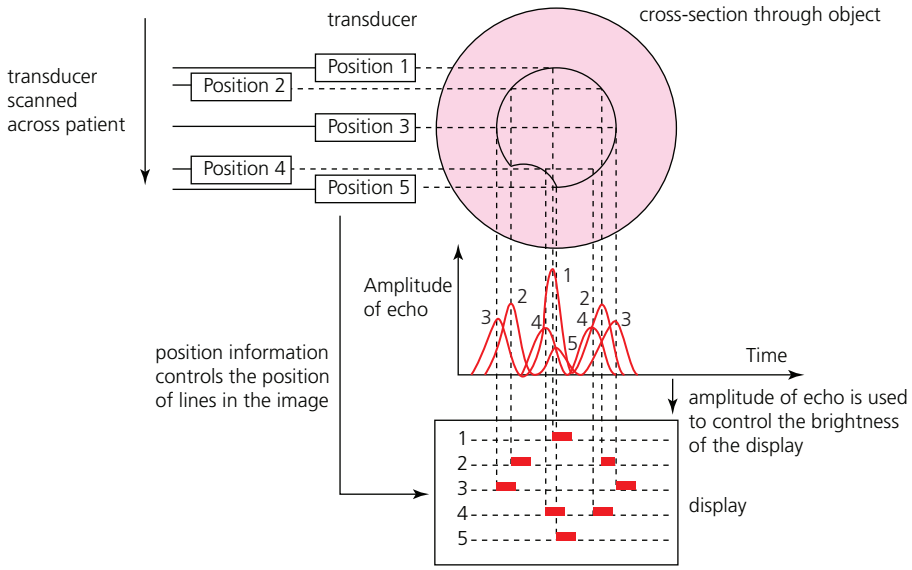


Figure 6 *The B-scan technique. A B-scan image is built up of many lines. Each line represents one pulse–echo sequence. The brightness of the image represents the strength of the echo.*

There is a compromise between the number of A-scans (lines) in each image and the number of frames per second. More A-scans will improve the resolution of the image, but will limit the frame rate.

Higher ultrasound frequencies are diffracted less and so give higher-resolution images, allowing finer structure to be imaged. But higher frequencies are attenuated more. A solution is to use high frequency but make the transducers small (Figure 8), so that they can be placed close to the tissue being studied, thus reducing attenuation. Ultrasound of 2 MHz is used for adults and 7 MHz for infants.

Ultrasound scans of the developing fetus (Figure 7) are used to determine gestational age and to detect multiple pregnancies or fetal abnormalities. Echocardiography uses ultrasound to provide a real-time image of the heart, allowing valve disorder to be seen.

Figure 7 *Image of a fetus during a B-scan of a woman's uterus. A gel is applied to the woman's skin to 'couple' the ultrasound probe to the body.*

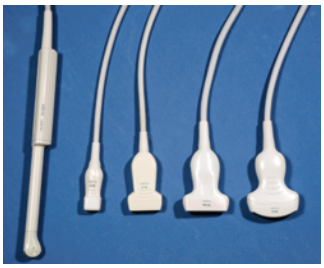


Figure 8 *Ultrasound transducers of different size and shape for optimum positioning*

A skilled operator is needed to produce good images, which require expert interpretation.

Advantages and disadvantages of ultrasound for diagnostic imaging

- Ultrasound imaging is non-invasive and causes very little discomfort for the patient. At the low intensities used in diagnostic ultrasound, the big advantage over X-ray techniques is that there are no known hazards to the patient, nor to the fetus in fetal scanning, and no risks for the operator.
- Ultrasound is also generally better than X-rays at imaging soft tissue, though structures behind the lungs or behind bone cannot be seen, due to the large reflections.
- Using high-frequency ultrasound transducers, the spatial resolution is comparable with other imaging techniques – ultrasound can resolve structures as small as 0.5 mm at 3.5 MHz. However, it does not give the same quality of image in terms of contrast as a CT scan (see Medical Physics section 4.6) or an MR scan (see Medical Physics section 3.3).
- Moving anatomical parts can be imaged in real time.
- Ultrasound equipment is relatively inexpensive compared with MR scanners or CT scanners, and much more portable.
- A high level of skill and experience by the operator is needed to acquire good-quality images and make accurate diagnoses.

QUESTIONS

5. Higher-frequency ultrasound gives better resolution images. Suggest why 7 MHz can be used for infants but only 2 MHz for adults.
6. By considering your answers to question **4 a** and **b** explain why a gel is used between the ultrasonic transducer and the skin when scanning.
7. During an A-scan, an ultrasound pulse reflects from the front and back edges of an organ in the body. The time interval between the two reflected pulses is 170 µs. The speed of the ultrasound in the organ is 1200 m s^{-1}. Using distance = speed × time calculate the total distance travelled by the ultrasound in the organ and hence calculate its thickness.

KEY IDEAS

- There are no known hazards to the patient, nor to the fetus, in ultrasound scans.
- An A-scan displays ultrasound echoes as pulses that allow the relative position of the boundaries of different tissues in the body to be determined.
- A B-scan is made up of a sequence of A-scans as the ultrasound beam is swept over the body. The amplitude of the echo controls the brightness of a display and, combined with positional information, creates an image of body structures.
- The higher the frequency of ultrasound used in a B-scan, the better the resolution of the image.

3.2 FIBRE OPTICS AND ENDOSCOPY

In 'keyhole surgery' only a small cut is made in the patient's body, and video cameras enable the surgeon to see the operation on a screen. Light is carried into and out of the body by bundles of **optical fibres**. Optical fibres are very fine, flexible strands of glass that carry light using total internal reflection. Rays of light that strike the inside surface of a glass fibre at an incident angle greater than the critical angle are internally reflected along the fibre, as if the fibre had mirrored internal walls (*see sections 7.4 and 7.5 in Chapter 7 in Year 1 Student Book*).

In a bundle of thousands of fine optical fibres packed close together, light can leak from one fibre to another. To prevent this, the fibres are coated with a second layer of glass of slightly lower refractive index, called the **cladding**. The total internal reflection occurs at the core–cladding interface. The use of cladding also serves to protect the core from surface scratches, which would also lead to a loss of light.

Endoscopes

An **endoscope** (Figure 9) is a sophisticated medical instrument commonly used to view the gastrointestinal tract. It is a flexible tube that, in addition to bundles of optical fibres, carries air, water and suction channels, and can also carry tools with which to take samples of tissue for analysis (*see Assignment 5 in Chapter 5 in Year 1 Student Book*).

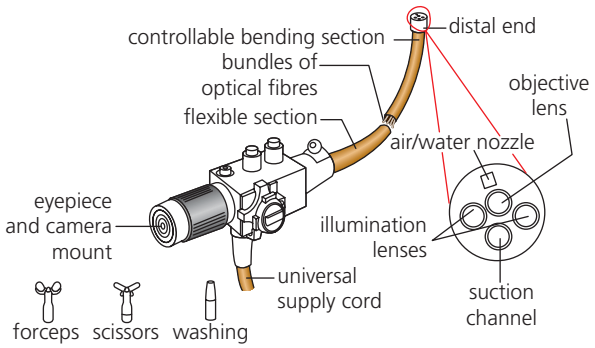


Figure 9 *The endoscope and its tools*

Light is carried to the site of the examination through an **incoherent bundle** of glass fibres. As many as 30 000 individual fibres make up the bundle. An incoherent bundle cannot be used to form an image because the ends of the individual fibres are arranged randomly. In a **coherent bundle**, the fibres have the same spatial position at both ends of the bundle (Figure 10). The light emitted from the end of the bundle is an exact copy of the incident light and an image can be reproduced. Coherent bundles are expensive to manufacture, so incoherent bundles are used for illumination.

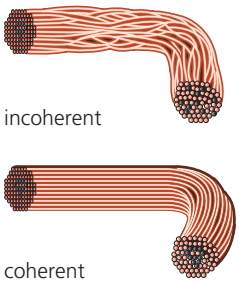


Figure 10 *Light is carried down an endoscope in an incoherent bundle, and back up in a coherent bundle.*

Endoscopes are used to examine the upper digestive tract (gastroscopy) or the rectum and colon (colonoscopy). One common application is to look for tumours, without the need for surgical intervention.

Arthroscopy is a similar technique used to examine joints, like the shoulder or knee, to diagnose problems such as arthritis. The arthroscope can also be used to treat damaged cartilage or to take small tissue samples (biopsies). Arthroscopy is performed through small incisions. It is much less painful than open surgery, carries less risk of infection and has a faster recovery time.

The laparoscope, a rigid form of endoscope, is used for examining the body through the small incisions made in keyhole surgery. There is often an extra optical fibre used for transmitting laser light, which may be used instead of a scalpel.

QUESTIONS

8. Explain the difference between incoherent and coherent bundles of optical fibres.
9. What advantages does keyhole surgery using optical fibres have over other forms of surgery for the patient?

KEY IDEAS

- An endoscope is a medical device that makes use of optical fibres to see inside the human body.
- The fibres in an endoscope may be incoherent bundles, which provide illumination, or coherent bundles, which transmit images.
- Surgical versions of endoscopes can be used in keyhole surgery to perform surgical procedures that minimise physical trauma to the patient.

3.3 MAGNETIC RESONANCE SCANNING

Magnetic resonance (**MR**) scanning, also known as magnetic resonance imaging (MRI), is used to produce an image of a cross-section through a patient. These images can be used to diagnose a wide range of different conditions, including cancer. MR scanning requires the patient to be subjected to a strong, uniform magnetic field. In practice, this means that the patient has to lie along the axis of a large solenoid (cylindrical coil), as in Figure 11.

Figure 11 *Preparing a patient to be scanned by an MR scanner*

MR scans are considered to be very safe. The patient is not subjected to any ionising radiation, as in the case of CT scans (see Medical Physics section 4.6), which use X-rays. However, MR scanning takes longer and requires the patient to remain still in an enclosed space for a longer period, and it can be very noisy. Some patients find this claustrophobic and unpleasant. MR scans are also more expensive than CT scans.

MR scans produce more detailed images of soft tissue than CT scans do. The brain, spinal cord and nerves, as well as muscles, ligaments and tendons, are seen much more clearly. For this reason, MR scanning is often used to image knee and shoulder injuries (Figure 12). See also the introductory page of Chapter 2.

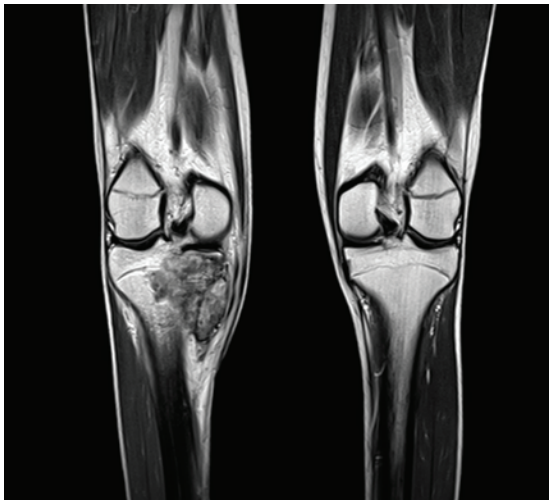


Figure 12 *MR scan (from front) of knees, showing bone cancer (centre of left image) of the upper end of the patient's right shinbone*

The principle behind an MR scan

MR scanning relies on the magnetic properties of the proton. A proton forms the nucleus of every hydrogen atom in our bodies. Protons spin and, because they are charged, they generate their own magnetic field and act like tiny magnets (Figure 13a).

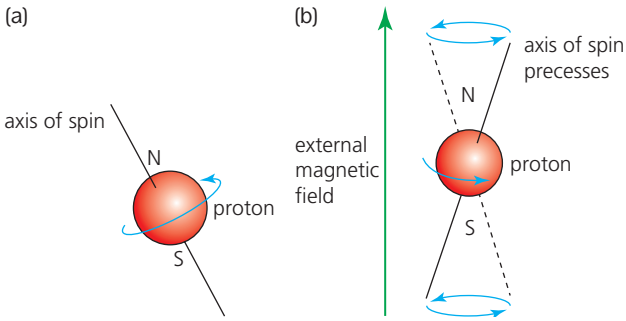


Figure 13 *(a) The protons forming the hydrogen nuclei spin and so behave like tiny magnets. (b) When an external magnetic field is applied, the spin axis of a proton precesses around the external field lines.*

When an external magnetic field is applied to the hydrogen atoms, the protons align themselves with the field, so that their spin axes **precess** about the direction of the applied field (Figure 13b). The frequency of the precession is called the **Larmor frequency**, and this motion produces a very small electromagnetic signal at the Larmor frequency.

A pulse of radio-frequency (RF) energy at the Larmor frequency is then directed at the patient. The protons absorb this energy and resonance occurs, causing them to become excited to higher energy states. This is called **nuclear magnetic resonance (NMR)**. When the radio pulse is stopped, the protons de-excite: they return to their original states and emit RF signals. The strength of the RF signal depends on the proton density in different regions of the body, which is itself linked to the amount of water in the tissue. It is these signals that are used by a computer to reconstruct an image of a section through the patient's body (like that in Figure 12).

High magnetic fields of the order of a few tesla (T) are needed to produce observable effects in the atomic nuclei for magnetic resonance imaging. The field is produced by large superconducting coils, which are able produce a flux density of between 1 and

3 T – about 10 000 times greater than the Earth's magnetic field (*see section 14.4 in Chapter 14 in Year 1 Student Book*).

Figure 14 shows the structure of an MR scanner. The patient lies inside the bore of the superconducting magnet, which causes the protons in the body to align with its magnetic field. The RF transmitter coil transmits the radio pulse, which excites the protons in the body. As the protons de-excite, RF signals from them are picked up by receiver coils and then processed by a computer.

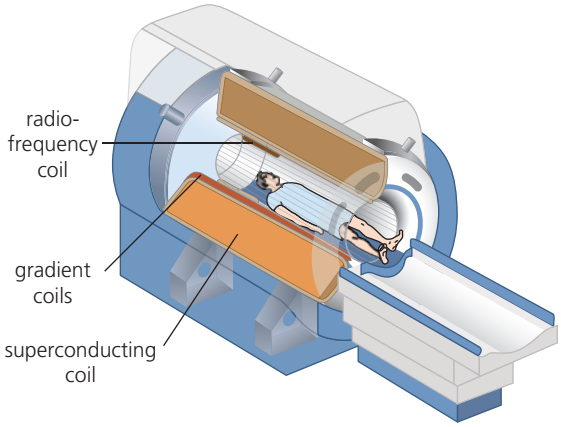


Figure 14 *Inside an MR scanner*

The cross-sectional imaging of the MR scanner relies on three **gradient coils**, which enable the location of the position of the source of emitted RF. The gradient coils are loops of wire or thin conductive sheets on a cylindrical shell lying just inside the bore of the scanner. When current is passed through these coils, a secondary magnetic field is created by each, which has a gradient of flux density in one of three dimensions. These 'gradient fields' slightly vary the flux density of the superconducting magnet in a predictable pattern, causing the resonant (Larmor) frequency of the protons, and hence the RF signal emitted, to vary as a function of position in the *X*, *Y* and *Z* directions (Figure 15).

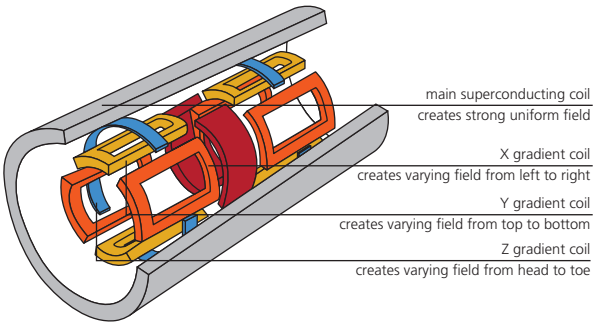


Figure 15 *Gradient coils in an MR scanner*

QUESTIONS

10. Sketch the shape of the main magnetic field produced by an MR scanner's superconducting coil.

11.a. Describe what happens to the nuclei of hydrogen atoms in the body when exposed to a high magnetic field in an MR scanner.

b. Explain how the nuclei in the body can be stimulated to emit radio frequencies in an MR scanner.

12. What is the purpose of the gradient coils in an MR scanner?

13. What types of tissue are best imaged by an MR scanner?

KEY IDEAS

- An MR (magnetic resonance) scanner uses a strong magnetic field to produce images of the internal soft-tissue structures of the human body.
- Hydrogen nuclei (protons) in the body align their spin axes with the applied uniform magnetic field but precess around its field lines.
- The protons are then excited to higher energy states by resonance, by applying a radio-frequency (RF) pulse at the precession (Larmor) frequency.
- As the protons de-excite, they emit RF energy, which can be detected and processed by a computer to determine the density of hydrogen nuclei (and hence water) in the body.
- Gradient coils in the MR scanner vary the flux density in three dimensions, so varying the RF signal frequency from different parts of the body. This allows images of cross-sections of the body to be created.

ASSIGNMENT 1: WATCHING THE BRAIN THINK

(PS 1.2)

Functional magnetic resonance imaging (fMRI) is a technique that uses magnetic resonance (Medical Physics section 3.3) to measure the activity of the brain. As blood flows through the brain, its oxygen content changes in response to different neural activities. fMRI works by detecting those changes.

When a region of the brain is more active, it uses more energy and needs more glucose and oxygen. Blood contains a molecule called haemoglobin, which carries oxygen and glucose to tissues in the body, including brain tissue. Haemoglobin's structure changes slightly depending on whether or not it is carrying oxygen. In addition, haemoglobin has magnetic properties that are different depending on whether or not it is carrying oxygen. When the haemoglobin is not carrying oxygen, it becomes a *paramagnetic* substance. When it is carrying oxygen, it becomes a *diamagnetic* substance. Paramagnetic materials become magnetised in an external magnetic field but their magnetism disappears when the field is removed. Diamagnetic materials are unaffected by external magnetic fields and are not attracted by them.

In the 1990s, a Japanese physicist named Seiji Ogawa realised that these differences in magnetic field response in haemoglobin when oxygenated and deoxygenated could be used to map images of brain activity using a normal MR scan.

In an MR scan the radio-frequency (RF) signal from hydrogen nuclei varies in strength depending on the type of tissue. This is what enables a picture to be formed of different types of biological material in the brain, such as grey matter, white matter and the cerebral spinal fluid (see Figure 1 of Chapter 2). When a particular area of the brain is active, increased numbers of neurons in that region are activated and have an increased demand for oxygen. The neurons are fed oxygen by a network of blood capillaries containing haemoglobin. The difference in magnetism between oxygenated and deoxygenated blood affects the strength of the RF signal from hydrogen nuclei in haemoglobin. In this way, fMRI can be used to produce activation maps showing which parts of the brain are involved in a particular mental process (Figure A1).

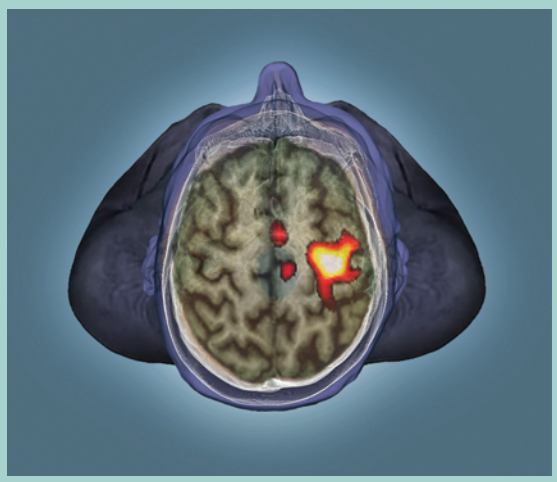


Figure A1 *An fMRI scan of the brain of someone doing a fairly simple task (like talking, finger tapping or listening). When performing different tasks, different parts of the brain 'light up' due to increased blood flow. So the fMRI scan can be used to determine which parts of the brain are experiencing increased neuron activity and which parts have been impaired because of, for example, stroke.*

QUESTIONS

A1 Describe briefly what is meant by magnetic resonance imaging.

A2 What is the fundamental difference between MR imaging and fMRI?

A3 Neurons are a type of nerve cell that uses electrical energy (see Medical Physics section 2.2). Suggest a biological reason why the blood supply should be increased to parts of the brain where neural activity has increased.

A4 A stroke is a medical condition where part of the brain has been permanently damaged due to lack of oxygen. In some cases, it is possible for stroke patients to be rehabilitated, because different parts of the brain can take over functions originally undertaken by the damaged part. In light of this, explain whether you think that fMRI can tell doctors definitively which parts of the brain are responsible for different bodily functions.

PRACTICE QUESTIONS

1. **a.** Figure Q1 shows an ultrasound transducer used in an A-scan.

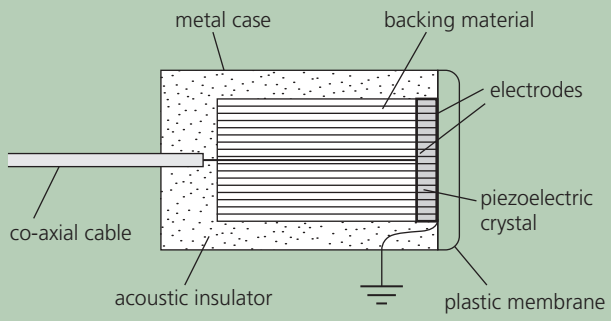


Figure Q1

Outline, with reference to the diagram, the process by which the transducer produces a short pulse of ultrasound.

b. Ultrasound is incident on the boundary between two materials. Some of the ultrasound is reflected at the boundary and the remainder is transmitted across the boundary. The ratio of the intensity of the reflected ultrasound, I_r, to the intensity of the incident ultrasound, I_i, is given by the equation

$$\frac{I_r}{I_i} = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2$$

where Z_1 and Z_2 are the acoustic impedances of the two materials.

i. Calculate the percentage of the incident ultrasound which would be transmitted into the skin when incident on an air–skin boundary.

[Acoustic impedance of air $= 4.29 \times 10^2\,\text{kg m}^{-2}\,\text{s}^{-1}$

Acoustic impedance of skin $= 1.65 \times 10^6\,\text{kg m}^{-2}\,\text{s}^{-1}$]

ii. When obtaining the ultrasound image of an unborn fetus, a coupling gel is used. Explain why a coupling gel is needed and state the property of the gel that ensures a good quality image.

AQA June 2013 Unit 5B Q3

2. **a.** Describe an endoscope and how it is used.

b. Describe the purpose of the coherent and incoherent fibre bundles in an endoscope.

c. Each fibre in a bundle has a core surrounded by cladding. Calculate the critical angle at the core–cladding interface of a fibre in the bundle.

[refractive index of core = 1.60; refractive index of cladding = 1.55]

3. Describe how a magnetic resonance scanner produces an image of the body, including the function of the gradient coils. The quality of your written communication will be assessed in your answer.

4. **a.** Explain the difference in ultrasound imaging between and an A-scan and a B scan.

b. Figure Q2 shows an A-scan of an organ in the body detected by an ultrasound transducer.

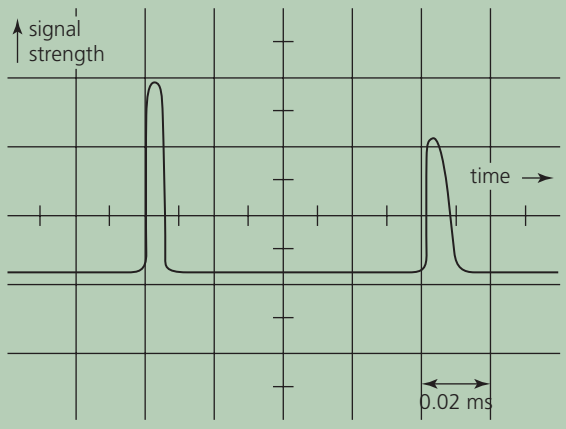


Figure Q2

The ultrasound operator identifies the A-scan pulses as coming from the front and rear surfaces of a potential tumour. Estimate the thickness of the tumour if the speed of ultrasound in the tissue is $1300\,\text{m s}^{-1}$ and the horizontal scale is $0.02\,\text{ms/cm}^{-1}$.

c. Why is the height of the second pulse smaller than that of the first pulse?

4 X-RAY IMAGING

PRIOR KNOWLEDGE

You will know that X-rays are short-wavelength high-frequency electromagnetic radiation. You should understand that electromagnetic radiation can described as photons with energy $E = hf$. From Chapter 5 you should know the definition of the electronvolt and how electrons gain energy when moving through a potential difference. You will need to recall from *Chapter 8 in Year 1 Student Book* how electrons can be ejected from materials by photons in the photoelectric effect. You should be familiar with the electron shell structure of atoms and how they can be excited to higher states. You will also need to be familiar with using equations involving exponentials and how to manipulate them.

LEARNING OBJECTIVES

In this chapter you will learn how X-rays are produced and used to image the human body for diagnostic purposes. You will see how they can be detected, how the images can be enhanced, and how X-rays are attenuated as they pass through matter. You will find out about the operation of a CT scanner, which uses X-rays to form highly detailed medical images.

(Specification 3.10.5)

4.1 THE PRODUCTION OF DIAGNOSTIC X-RAYS

X-rays were discovered by Wilhelm Röntgen in 1895. Within a month of the discovery, X-rays were being used by doctors to examine patients with fractures or gunshot wounds. At that time, scientists were unaware of the dangers of ionising radiation, and many of the early patients suffered from high X-ray doses. Modern techniques have made the medical use of X-rays much safer.

Röntgen was experimenting with high-voltage discharge tubes when he discovered X-rays. He used an induction coil to apply a high potential difference, about 35 000 V, between two electrodes in an evacuated glass tube (Figure 1). Electrons from the cathode were accelerated across the tube towards the anode. The evacuated tube meant that the electrons could reach high speeds, because they were not impeded by collisions with air molecules. Röntgen noticed that a fluorescent screen placed a few metres from the tube produced light when the tube was operating. The fluorescent glow persisted even when Röntgen shielded his tube with thick black paper. He concluded that invisible, penetrating rays were being produced by the tube. He named these X-rays.

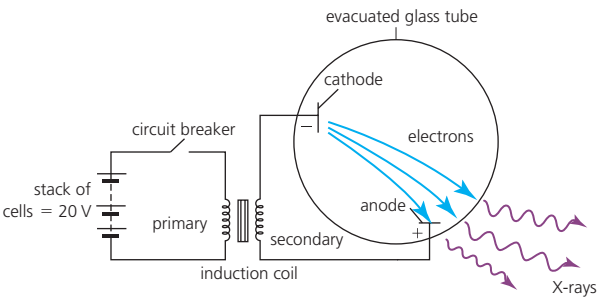


Figure 1 *A diagram of the apparatus with which Röntgen discovered X-rays*

We now know that X-rays are very high-frequency electromagnetic waves, produced when high-energy electrons lose energy as they collide with atoms.

In a modern X-ray tube (Figure 2), the cathode is heated by an electric current through a filament. The cathode then emits electrons by **thermionic emission**. Some of the free electrons gain an amount of thermal energy sufficient to overcome the attractive electrical force bonding them to the positive ions of

the metal crystal lattice and they can escape from the metal surface.

The electrons are accelerated towards the anode by a potential difference of around 150 kV. When the electrons reach the target, which is mounted on the anode, they collide with the target atoms. As the electrons slow down, they emit a continuous spectrum of X-rays, known as **bremsstrahlung** or 'braking' radiation.

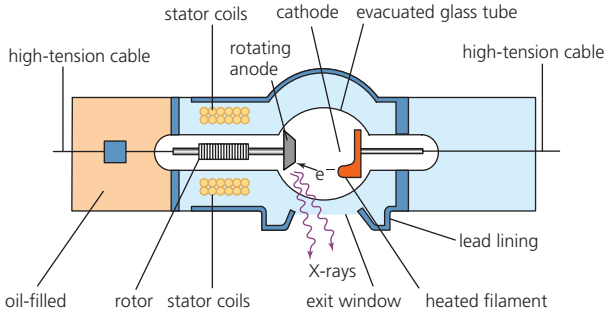


Figure 2 *A modern X-ray tube has a rotating anode and is mounted in a lead-shielded, oil-cooled case. The anode rotates to reduce local heating problems.*

Some of the high-speed electrons will collide with electrons in the target atoms. If they have sufficient energy, they will knock atomic electrons out of their orbits and ionise the atoms. As a result, **characteristic X-rays** are emitted when other electrons in the target atoms drop from higher energy levels to fill the vacancy (Figure 3). The wavelengths emitted depend on the spacing of the energy levels in the target atoms. This means that a line spectrum (*see section 8.1 in Chapter 8 in Year 1 Student Book*), characteristic of the target material, is superimposed on the continuous bremsstrahlung spectrum (Figure 4).

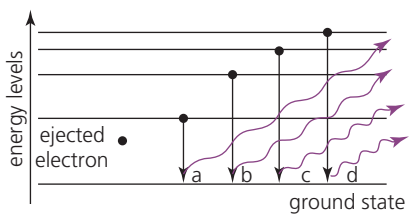


Figure 3 *X-rays of frequency characteristic of the target atoms are emitted as electrons fall from higher energy levels to fill vacancies. Four characteristic emissions are shown, which correspond to the lines in Figure 4.*

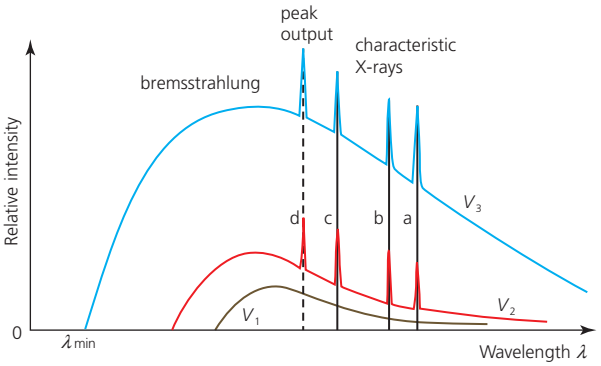


Figure 4 *The shape of an X-ray spectrum depends on the potential difference, V, across the tube. Here the three curves are for tube voltages V_1, V_2 and V_3, with $V_3 > V_2 > V_1$. A higher value of V will increase the total intensity emitted (this is approximately proportional to V^2), the peak output shifts to a shorter wavelength, the minimum wavelength decreases, and more characteristic lines may appear.*

The maximum possible energy of an X-ray photon depends on the potential difference, V, across the tube. As an electron is accelerated, it gains energy, $E = eV$, where e is the charge of the electron. If the electron were to lose all of this energy in one collision with a target atom, the X-ray would be emitted with energy $E = eV$. The energy of a photon is given by $E = hf$, where h is the Planck constant and f is the frequency of the photon. The maximum photon energy from an X-ray tube is therefore

$$E = hf = eV$$

And so, since $c = f\lambda$, the minimum wavelength of the emitted photons is

$$\lambda_{min} = \frac{c}{f} = \frac{hc}{E} = \frac{hc}{eV}$$

For example, if the potential across an X-ray tube is 150 kV, the shortest wavelength of the X-rays, λ_{min}, is

$$\lambda_{min} = \frac{hc}{eV} = \frac{6.63 \times 10^{-34} \times 3.00 \times 10^{8}}{1.60 \times 10^{-19} \times 150 \times 10^{3}} = 8.29 \times 10^{-12}\ \text{m}$$

In practice, less than 1% of the energy in the electron beam is transferred to X-ray radiation. Most of the energy is transferred to internal energy in the target, raising its temperature. The target has to be made of a metal with high thermal conductivity, high specific heat capacity and high melting point. Tungsten meets these criteria well and is often used as the target in diagnostic X-ray tubes. Tungsten has a high proton

number, so a tungsten nucleus has a large mass and a high positive charge. This increases the probability that collisions with the high-energy electrons will lead to X-ray emission.

The tungsten target is mounted on the anode, which is able to rotate (see Figure 2). This means that the electron beam strikes different areas of the target and reduces local heating problems. The anode has a bevelled edge so that the target can be wide without increasing the width of the X-ray beam (Figure 5).

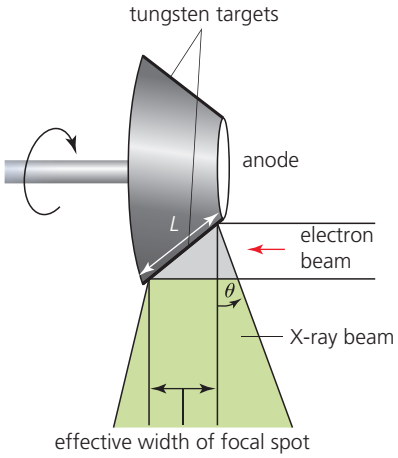


Figure 5 *The tungsten target is mounted on the anode, which has a bevelled edge. Here the width of the target is L, but the width of the X-ray beam at the target is L sin θ. For most applications, θ is about 17°.*

The intensity of the X-ray beam is the total energy emitted at all wavelengths per second through unit area. It is proportional to the area under the curve in Figure 4. For a particular anode voltage, the X-ray intensity is controlled by the number of electrons striking the target every second, since more electrons leads to more collisions, which means that more X-rays are produced. Therefore, the intensity of the X-ray beam depends on the tube current, which can be controlled by varying the current to the cathode filament.

QUESTIONS

1. Explain the shape of a typical X-ray spectrum from an X-ray tube.
2. X-rays are produced by an X-ray tube with anode voltage 120 kV.
   a. What is the maximum energy of an X-ray photon produced?
   b. What is the minimum wavelength of an X-ray photon produced?
3. a. What controls the intensity of an X-ray beam?
   b. An X-ray tube has a current of 50 mA and potential difference of 140 kV between its anode and cathode. The X-ray photons are produced from the tungsten target with an efficiency of 0.7%. What is the power of the X-ray beam?
4. Why does the target material need to have high values of
   a. thermal conductivity
   b. specific heat capacity
   c. melting point?

KEY IDEAS

- X-rays are produced by a beam of high-energy electrons striking a metal (tungsten) target.
- X-ray tubes produce an X-ray energy spectrum made up of a continuous (bremsstrahlung) component and characteristic X-ray lines.
- The maximum photon energy from an X-ray tube is

$$E = hf = eV$$

  where V is the anode voltage.
- The X-ray beam intensity, for a given anode voltage, is controlled by the current to the cathode.

4.2 IMAGE SHARPNESS AND CONTRAST

There are two main ways to create images with X-rays. The first is to pass an X-ray beam through a section of the body and project a shadow image onto a detector. The second method, used in computed tomography (CT) (see Medical Physics section 4.6), uses a computer to reconstruct an image from narrow X-ray beams that are scanned over the patient's body by measuring the levels of attenuation with a detector.

All Images produced by X-rays are subject to blurring or **unsharpness**, which has three causes.

1. **Geometric unsharpness** is caused by the X-rays originating from a **source focal spot** that has an extended area (Figure 6) rather than being a point. To keep the image sharp, the width of the X-ray beam has to be kept as small as possible, which is why in the X-ray tube the target is inclined to the beam, as shown in Figure 5. It is also important to keep the detector close to the patient.

   A sharper image can be achieved by moving the X-ray source further from the patient, but the inverse square law for radiation (see Chapter 9) means that the intensity of the X-ray radiation at the patient is also decreased, and longer exposure times are necessary.

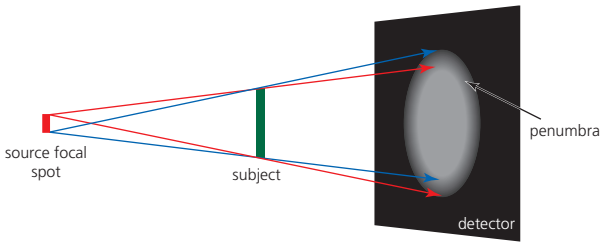


Figure 6 *The effect of blurring due to the finite size of an X-ray source*

2. **Motion unsharpness** is caused by movement of the patient, of the detector or of the X- ray source during the exposure. Movement of the patient is the most common, and patients are usually asked to keep still and to hold their breath in the short time they are exposed to X-rays. Sometimes they need to be sedated. Figure 7 shows a blurred, unsharp X-ray image of the teeth caused by motion of the patient's jaw when the image was taken.

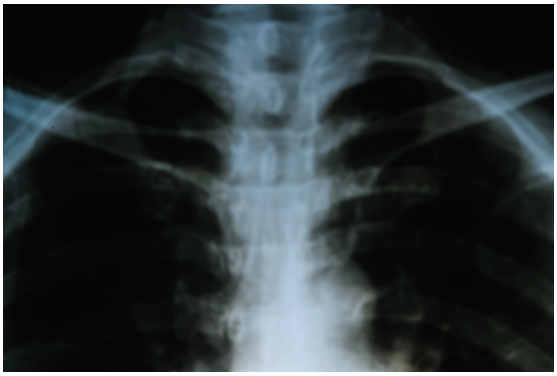


Figure 7 *Dental X-ray showing image unsharpness caused by motion of the patient.*

3. **Detector unsharpness** is dependent on the type of X-ray detector used. Each type has a limiting factor that determines its ability to resolve fine detail. For X-ray film, it is the size of the photographic grains in the film material. In digital radiography, it is the size of the individual detector elements. There is no way to improve this other than to use a detector with better resolution.

Image **contrast** is the difference in visibility between two areas in an image. In X-ray imaging, the contrast is the difference between the image densities of different parts of the body. The better the contrast, the easier it is to distinguish different features, and abnormalities are more likely to be identified. Contrast depends on the relative number of X-ray photons transmitted through dissimilar regions of the body, and also on the response of the detector. Different tissues in the body absorb X-rays to different extents – the amount of absorption increases with the proton number of the material. (We will look at this in Medical Physics section 4.5.) In practice, only bone, soft tissue and air are sufficiently different from each other to produce noticeable contrast (Figure 8).

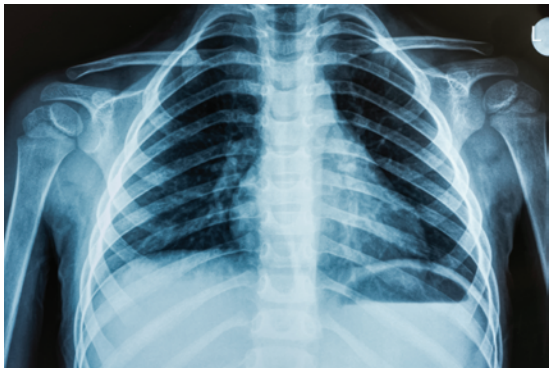


Figure 8 *Contrasting structures in a chest X-ray. Bones, which have high X-ray absorption, are white, meaning that fewer X-ray photons have passed through to the detector. The air in the lungs is black (very low absorption) and soft tissues like the heart, muscle and skin are similar shades of grey. There is little contrast between similar types of soft tissue.*

KEY IDEAS

- Factors that influence the sharpness of an X-ray image are:
  - the size of the X-ray focal spot
  - any movement of the patient or detector while the exposure is being made
  - the image resolution of the detector.

4.3 X-RAY DOSE

As with all forms of ionising radiation, it is necessary to limit the **dose** to the patient as much as possible. Diagnostic X-rays are responsible for most of our exposure to man-made radiation. It is important to keep the X-ray dose for any particular procedure as small as possible. The actual dose delivered needs to be weighed with the clinical benefit to the patient.

The X-ray dose depends on the exposure time and the intensity of the beam. The exposure time can be reduced by using a more sensitive detector (that is, an electronic detector rather than film) and/or by intensifying the image (see Medical Physics section 4.4). The intensity depends on the tube voltage and the cathode current. Since using a high tube voltage to produce shorter-wavelength and hence more penetrating X-rays can produce higher-contrast images, the beam intensity is reduced by lowering the cathode current.

Another way of reducing the dose is to filter out the low-energy component of the X-ray spectrum. Low-energy photons are unlikely to reach the X-ray detector because they are absorbed by the skin, so they cause a dose to the patient without providing any diagnostic information. It is desirable therefore to remove these low-energy photons. To do so, a filter is made from a few millimetres of aluminium. This absorbs a greater proportion of low-energy than high-energy photons. This increases the minimum energy of the photons to which the patient is exposed. The process of increasing the average energy level of an X-ray beam by filtering out the low-energy photons is called **beam hardening**. Figure 9 shows the X-ray spectrum with and without a filter.

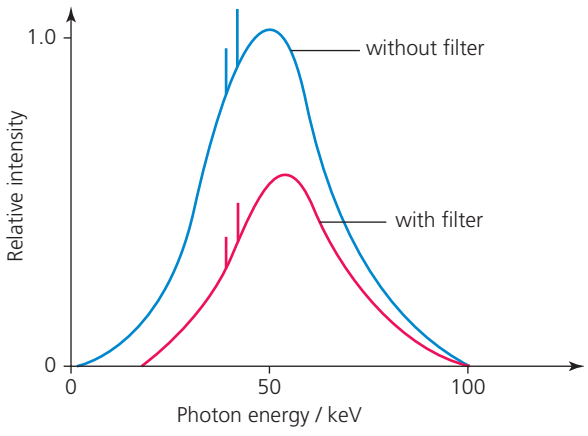


Figure 9 *The effect of beam hardening on an X-ray beam, showing the increase in average energy when a filter is used. (Note that, unlike Figure 4, the x-axis here is energy not wavelength.)*

Some organs are more sensitive to radiation than others, so suitable shielding, a lead apron for example (Figure 10), is used to cover these organs. It is also important to reduce the dose to the radiographer. Radiographers should stand behind suitable shielding during the exposure. The inverse square law means that increasing their distance from the X-ray tube will also significantly reduce the dose.

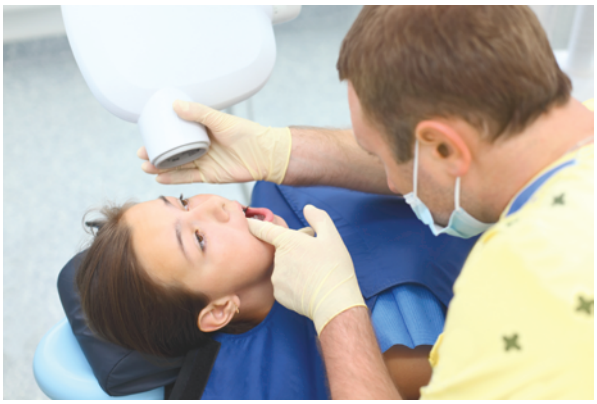


Figure 10 *Lead is used in shielding because its high proton number, and hence density, means that it is a good absorber of X-rays.*

KEY IDEAS

- The X-ray dose to a patient may be reduced by:
  - using sensitive detectors to reduce exposure time
  - beam hardening to filter out unwanted low-energy X-rays
  - reducing the intensity of the beam by lowering the cathode current.

4.4 IMAGE DETECTION AND ENHANCEMENT

Traditionally, X-ray images were produced using photographic film. X-ray photons cause ionisation in silver halide grains in the film. When the film is developed, those grains that were exposed to X-rays turn black. The greater the exposure, the more grains are developed and the darker the film. The variation in darkness (or optical density) between areas of the film representing soft tissue, as compared to those representing bone, is the contrast of the image.

Photographic film is not very sensitive to X-ray radiation. A typical film will absorb less than 0.1%

of the X-ray energy that reaches it. Consequently, a relatively long exposure time is needed to get a well-exposed image. But increasing the exposure time increases the dose to the patient, as well as increasing the likelihood of blurring due to movement.

Digital detection

Flat-panel (FPD) detectors are modern digital X-ray detectors that use scintillators, photodiodes and electronic scanning to produce X-ray images.

Advances in semiconductor arrays for display screens means that flat-panel displays like the ones in televisions and laptops can be used to produce an X-ray image sensor. This type of FPD sensor uses thin-film transistors (TFTs) integrated with arrays of **photodiodes**. The photodiodes are bonded to a **scintillator** material such as caesium iodide (CsI). A scintillator is a material that emits prompt brief flashes of light when a charged particle or a high-energy photon passes through it.

When X-rays are incident on the panel, visible light is emitted from the scintillator that is proportional to the X-ray energy and detected by the photodiodes, which act as individual pixels (picture elements) and determine the image resolution. Charge carriers are created inside the photodiode, and the signal from each diode is read out through a TFT switch, which is converted to a digital output and interpreted by a computer to produce a digital image (Figure 11).

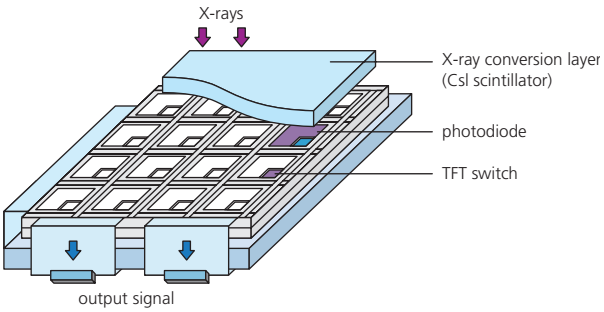


Figure 11 *Flat-panel (FPD) X-ray digital imaging sensor*

The advantages of using FPD detectors are as follows:

- they allow 'instant' X-ray pictures to be produced without the time needed for processing and developing film
- owing to their sensitivity, an adequate image can be obtained while subjecting the patient to a much lower dose of X-rays than would be the case with film
- the digital images processed by an FPD can be image-processed, stored easily and communicated quickly to medical staff for diagnosis and evaluation.

Electronic scanning

Computed radiography (**CR**) is an X-ray technique that uses an imaging plate consisting of phosphors that are sensitive to X-rays. The phosphors store the X-ray energy by exciting electrons in them to higher energy states. The imaging plate is housed in a special cassette and placed under the part of the body to be imaged. After exposure, a 'latent image' is formed on the plate in proportion to the X-ray intensity it receives. A CR reader then scans the imaging plate using a laser and causes the stored energy in the phosphors to be released as visible light (see Figure 12). This is detected and read using a photomultiplier (see Medical Physics section 5.3) and a digital image is created that can be displayed on a monitor or printed out. When the image has been captured and stored, then the imaging plate is 'erased' by subjecting it to intense light ready for it to be used again. As with FPD-generated images, the digital image may be image-processed, stored and sent electronically.

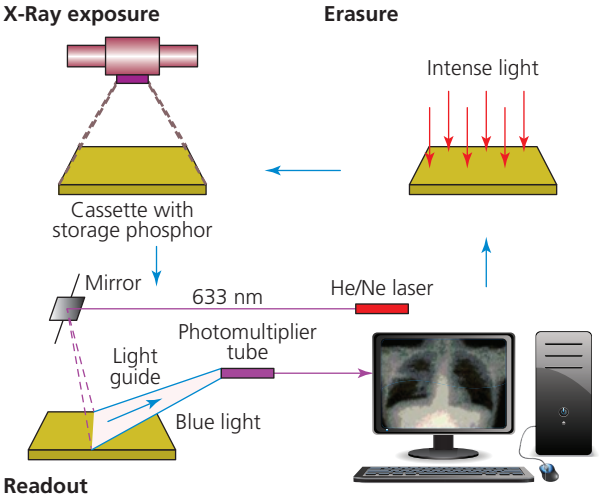


Figure 12 *Computed radiography (CR scanning) using an image plate*

Contrast enhancement

The contrast between images of different soft tissue, or between normal tissue and a tumour, is often very small owing to their similar proton numbers. One way to visualise internal organs in which the proton numbers are similar is by using a **contrast medium**. This is a substance that has a high proton number and absorbs X-rays to a greater degree than the surrounding tissue and is thus **X-ray opaque**. Patients who need an X-ray of their digestive tract often eat a 'barium meal', a thick suspension of

barium sulfate that is taken orally. Barium sulfate is X-ray opaque, and areas coated with it will appear white on the X-ray image (Figure 13) and be easier to see. Barium sulfate is toxic, but it is safe to use because it is insoluble and so is prevented from entering the blood.

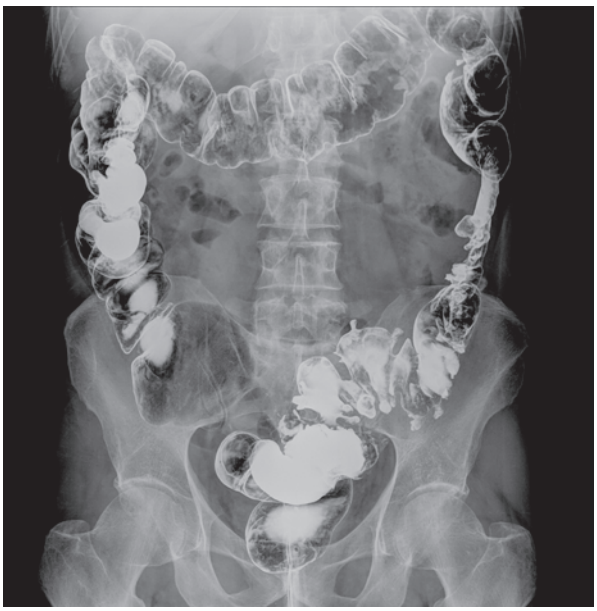


Figure 13 *Outline of the lower intestine using a barium contrast X-ray*

The passage of the barium sulfate through the gastrointestinal tract can be observed by a radiologist on a screen. The radiologist takes a series of individual X-ray images at timed intervals depending on the areas to be studied.

An **intensifying screen** is used in X-ray imaging to reduce the time needed for X-rays to expose photographic film. An intensifying screen is made up of crystals, such as zinc sulfide, which absorb X-ray photons. The electrons in the atoms of the crystal become excited to higher states and return to their equilibrium states, emitting light in all directions. Double-sided photographic film is sandwiched between two intensifying screens as shown in Figure 14. The film is in close contact with the screens, so most light photons are captured, forming a sharp image.

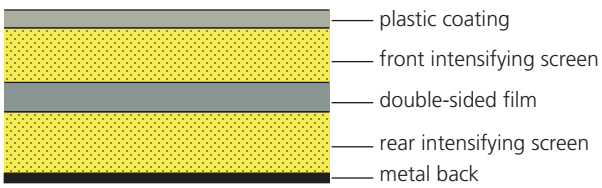


Figure 14 *The structure of an intensifying screen*

Since the film is more sensitive to light than to X-ray photons, the imaging is speeded up and the X-ray exposure time is less, reducing the dose to the patient.

Doctors often need to see moving images to aid diagnosis, perhaps to see the heart in motion, for example, or to view the passage of a barium meal through the gut. **Fluoroscopy** uses a fluorescent screen instead of a film, and this allows the X-ray image of the patient to be produced in real time. An **image intensifier** is used in conjunction with the fluorescent screen (Figure 15). A continuous X-ray beam strikes a phosphor, often caesium iodide, which then emits light. The light photons strike a photocathode, releasing electrons, which are focused by electrodes inside a vacuum tube. The electrons accelerate across the tube until they strike the fluorescent screen, which provides a visible image. The output can be recorded via a video camera.

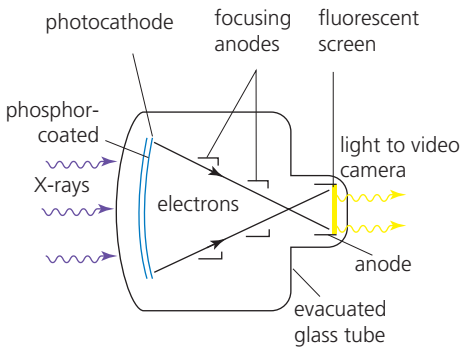


Figure 15 *A fluoroscopic image intensifier*

The output screen is smaller than the input screen. This reduction, and the acceleration of the electrons across the tube, makes an image up to 1000 times brighter than the original. This means that much lower-intensity X-rays can be used, reducing the dose to the patient. But even with the use of an image intensifier, fluoroscopy gives a relatively high dose because a continuous beam of X-rays is used. The dose is often 15 times that of a conventional X-ray procedure.

QUESTIONS

5. What is a scintillator?
6. What key advantages do flat-panel (FPD) detectors have over X-ray photographic film?

7. How can contrast be enhanced in an X-ray image where there are tissues of similar proton number?
8. What is fluoroscopy used for?

KEY IDEAS

- Flat-panel detectors (FPD) are electronic detectors that use scintillators to give instant X-ray images.
- FPD detectors have the advantages of greater sensitivity and so lower X-ray dose to the patient. They also allow images to be stored and communicated more efficiently. Their image resolution is limited by the size of individual photodiode elements on which the image is recorded.
- Computed radiography is a method of detecting X-ray images on an image plate that can be scanned and read out using a laser. Image plates can be re-used.
- Contrast enhancement of tissues with similar proton number can be achieved using a contrast medium such as barium sulfate.
- Intensifying screens are used to reduce the time needed for X-rays to expose photographic film, and so reduce the X-ray dose.
- Fluoroscopy is a method of X-ray imaging that can show movement in real time, using a fluorescent screen together with an image intensifier. The dose to the patient is much higher than in X-ray imaging that involves a single exposure.

4.5 DIFFERENTIAL ABSORPTION OF X-RAYS

X-rays interact with matter in a number of ways: they can be scattered by atomic electrons or they can be totally absorbed. These processes act to reduce the intensity of an X-ray beam as it passes through the material. This reduction in intensity is called **attenuation**. The amount of attenuation depends on the proton number and density of the material, as well as on the energy of the photons. Bone is denser than soft tissue. It also has a higher effective proton number (because bone contains significant amounts of calcium and phosphorus). These factors mean that bone absorbs more X-ray energy than an equivalent thickness of soft tissue. This is what allows us to differentiate between bone and soft tissue in X-ray images.

Exponential attenuation

For a narrow, monoenergetic X-ray beam, the attenuation is exponential. The intensity I of the beam after passing through a thickness x of a material is given by

$$I = I_0 \, e^{-\mu x}$$

where I_0 is the original intensity of the beam and μ is the **linear attenuation coefficient**, measured in m^{-1} or, more commonly, cm^{-1}. The equation above can be rearranged to give the linear attenuation coefficient:

$$\mu = \frac{1}{x} \ln\left(\frac{I_0}{I}\right)$$

This is a constant that describes the rate of energy loss by a photon beam per unit thickness within a medium.

A given thickness of the material will reduce the X-ray intensity by half. This thickness is referred to as the **half-value thickness** (**HVT**).

Worked example

Use the equation $I = I_0 \, e^{-\mu x}$ to calculate the half-value thickness (HVT) for bone.

Rearranging the equation, we have

$$\frac{I}{I_0} = e^{-\mu x}$$

At the half-value thickness, $X_{\frac{1}{2}}$, we have

$$\frac{I}{I_0} = \frac{1}{2}$$

so at $X_{\frac{1}{2}}$ we can write

$$\ln\left(\frac{I_0}{I}\right) = -\mu x_{\frac{1}{2}}$$

$$\ln\left(\frac{1}{2}\right) = -\mu x_{\frac{1}{2}}$$

$$x_{\frac{1}{2}} = \frac{-\ln(\frac{1}{2})}{\mu} = \frac{\ln 2}{\mu}$$

For bone, the linear attenuation coefficient for 150 keV X-rays is about 0.6 cm^{-1}. So, for bone, the HVT is

$$x_{\frac{1}{2}} = \frac{\ln 2}{0.6} = 1.1\,\text{cm}$$

The linear attenuation coefficient μ depends on the density of the material, ρ. A higher-density material absorbs more energy from X-rays, simply because the X-ray photons encounter more atoms (or more massive atoms) in the same volume. The linear attenuation coefficient for water vapour is less than that for ice because the water molecules are further apart. If μ is divided by the density of the material, it will produce a value that is constant for that particular element or compound. The **mass attenuation coefficient**, μ_m, is

$$\mu_m = \frac{\mu}{\rho}$$

This is now independent of density, so that the value of μ_m is the same for water whether it is in liquid, solid or gaseous form. Since linear attenuation coefficient is dependent on density, the mass attenuation coefficient is often reported for convenience, and is usually stated in units of $cm^2\,g^{-1}$.

QUESTIONS

9. a. What is the difference between the linear attenuation coefficient and the mass attenuation coefficient of an X-ray absorbing material?

b. The mass attenuation coefficient for bone with a density of 1.8 $g\,cm^{-3}$ is 0.2 $cm^2\,g^{-1}$ for 80 keV photons. What is the linear attenuation coefficient for bone at this density?

10.a. The thickness of aluminium required to reduce the intensity of an X-ray beam by one-half is 1.5 mm. Calculate the linear attenuation coefficient of aluminium.

b. In order to reduce the dose to a patient, the lower energy X-rays are removed from an X-ray beam using an aluminium filter of thickness 2.5 mm. By what factor is the intensity reduced when passing through the filter?

11.a. The HVT of lead for 90 keV X-ray photons is 12 mm. What is the linear attenuation coefficient of 90 keV electrons in lead?

b. It is required to reduce the X-ray intensity of a beam 90 keV X-rays using lead shielding. What thickness of lead is required to reduce the beam intensity to 2% of its original value?

KEY IDEAS

- X-rays are attenuated exponentially when they pass through a material.
- The half-value thickness (HVT) is the thickness of the material that will reduce the initial X-ray intensity by half.
- The linear attenuation coefficient (μ) is a constant that describes the rate of energy loss by a photon beam per unit thickness within a medium and depends on the density.
- The mass attenuation coefficient is found by dividing μ by the density of the material and is a constant value for a given element/compound and given X-ray energy.

4.6 CT SCANNER

Conventional X-ray images are shadow pictures, that is, two-dimensional projections of a three-dimensional object, so they carry no information about depth. The limited contrast, especially between soft tissues, is also a problem. X-ray **computed tomography** (**CT**) offers a solution to both of these problems. Modern **CT scanners** can produce high-contrast images of a cross-section through the head or body (Figure 16).

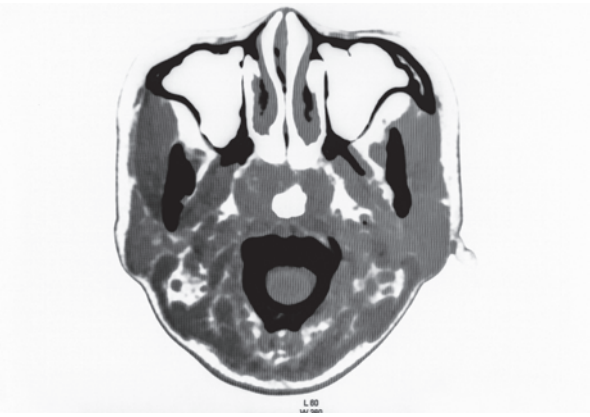


Figure 16 *An image of the head produced by a CT scanner. The white area on the right is bleeding on the brain caused by a cerebral haemorrhage.*

In CT scanning, the X-ray tube is rotated around the patient. A narrow, monoenergetic X-ray beam passes through the body at different orientations, and an array of detectors (which also rotates) records the intensity transmitted through the patient at each position (Figure 17). The results are digitised and the image is generated by computer.

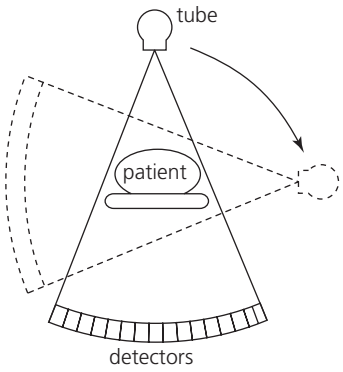


Figure 17 *The set-up of a CT scanner. The X-ray tube emits a finely collimated fan-shaped beam of almost monoenergetic X-rays towards the patient. The X-rays are detected by an array of detectors. The tube and array is rotated, often through 360°, around the patient.*

Image contrast depends on different intensities of X-radiation reaching the detector. This in turn depends on the average linear attenuation coefficient in the path of the X-ray beam. For a typical diagnostic X-ray beam, muscle has a linear attenuation coefficient of about $0.180\,m^{-1}$. For each centimetre of muscle:

$$\frac{I}{I_0} = e^{-\mu x} = e^{-0.180 \times 1} = 0.835$$

Similar calculations for bone and blood give values for I/I_0 of 0.619 and 0.837, respectively. So, the contrast between bone and muscle is significant, but the difference between blood and muscle is small, around 0.2%. Contrast media containing iodine may be used intravenously to highlight blood vessels.

In practice, the X-ray beam passes through a number of different structures and tissues as it travels through the body. The intensity recorded by the detector depends on the thickness and the value of μ of each different material. The image is reconstructed by analysing the intensity measured in many different directions (Figure 18).

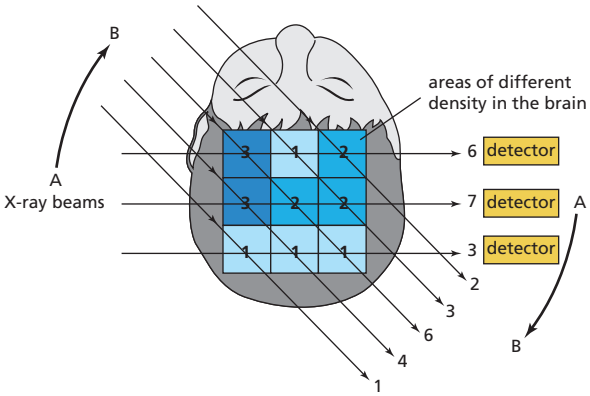


Figure 18 *A simplified representation of the analysis in CT scanning*

CT scanners are relatively expensive and they give a significantly higher radiation dose to the patient compared with a traditional X-ray. These disadvantages are often outweighed by the medical benefits of a detailed three-dimensional image. Modern CT scanners can reveal detail as small as 1 mm with density differences of less than 1%.

QUESTIONS

12. Why does the beam used in a CT scanner need to be monoenergetic?

KEY IDEAS

- In a CT scanner, the X-ray tube and the array of detectors are rotated around the patient.
- From the detected beam attenuation in different directions, an image is built up by a computer.

PRACTICE QUESTIONS

1. In an X-ray tube, electrons are accelerated from rest through a pd of 72.4 kV before they hit the target anode.
   - **a. i.** Calculate the kinetic energy (in J) of an electron as it reaches the anode. Give your answer to an appropriate number of significant figures.
   - **ii.** Assuming that the electron gives up all this energy to form an X-ray photon, calculate the wavelength of the photon.
   - **b.** X-rays are used in a CT scanner. Describe briefly how a CT scanner produces an image.

AQA June 2012 Unit 5B Q5

2. An X-ray tube gives rise to an X-ray spectrum as shown in Figure Q1. Explain the process that produces the spikes at certain photon wavelengths.

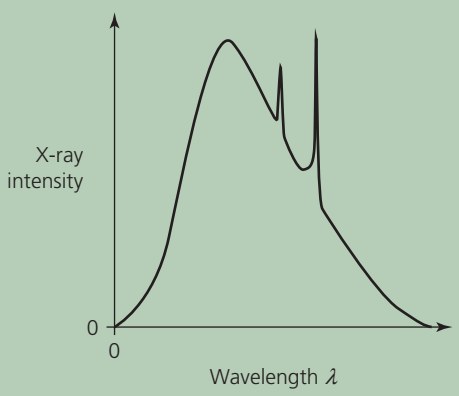


Figure Q1

3. State and explain **two** advantages of using an MR scanner to scan a patient's brain compared with a CT scanner.

AQA June 2010 Unit 5B Q4 (part)

4. A film cassette, placed under a patient being X-rayed, is shown in Figure Q2.

Figure Q2

Explain how the intensifying screens in the film cassette achieve their purpose and state their benefit to the patient.

AQA June 2011 Unit 5B Q5 (part)

5. State and explain a medical procedure when
   - **a.** an endoscope might be used instead of an X-ray for imaging
   - **b.** an X-ray might be used instead of an endoscope for imaging.

6. **a. i.** Explain why it is difficult to get good contrast in soft tissue using X-rays.
   - **ii.** How could the contrast of soft tissue in the body be increased?
   - **b. i.** Explain what is meant by *half-value thickness* and *linear attenuation coefficient*.
   - **ii.** In order to take an X-ray photograph, the X-ray beam is passed through an aluminium filter to remove low-energy X-ray photons before reaching the patient. Why it is desirable to remove these low-energy X-rays?
   - **iii.** The linear attenuation coefficient for X-rays that penetrate an aluminium filter is 250 m^{-1}. The intensity of an X-ray beam after travelling through 3.0 cm of aluminium is 250 $W m^{-2}$.

   What is the initial X-ray intensity that was incident on the aluminium?

5 RADIONUCLIDE IMAGING AND THERAPY

PRIOR KNOWLEDGE

You will need to have an understanding of the physics of the nuclear atom from Chapter 9, including isotopes, nuclear equations, radioactive decay, activity and half-life, and the properties of alpha, beta and gamma radiations and their ionisation and penetrating powers. You may want to refer back to Chapter 10 regarding the emission of gamma photons from nuclei in excited states. You should also recall the properties of X-rays and how they are produced, described in Medical Physics Chapter 4.

LEARNING OBJECTIVES

In this chapter you will learn how radioactive materials can be used in medicine to diagnose and determine the severity of a wide range of medical conditions, including cancers, heart disease, and gastrointestinal, endocrine and neurological disorders. You will see how radionuclides can be used to determine molecular activity and metabolic processes occurring in the body and offer the potential to detect disease at an early stage. You will also look at how radionuclides and high-energy X-rays can be used therapeutically to treat patients who have cancer or other medical conditions.

(Specification 3.10.6)

5.1 RADIONUCLIDES IN MEDICINE

A **radionuclide** or **radioisotope** is a nuclide that is radioactive, that is, an isotope of an element with an unstable nucleus. The radionuclide will decay, resulting in the emission of gamma rays and/or subatomic particles such as alpha or beta particles. Many radionuclides occur naturally, and others are produced artificially in nuclear reactors and cyclotrons (see section 7.3 of Chapter 7).

The choice of radionuclides for medical imaging and for therapy depends on their chemical suitability, their biological behaviour and their physical properties.

- Radionuclides must be chemically suitable to be produced in a form that can be bonded to molecules as a pharmaceutical product. In imaging, the resulting radiopharmaceutical product is said to be 'labelled' with a **radioactive tracer**, which means it can be tracked and detected, by its emissions, with an external radiation detector.
- The radiopharmaceutical must be biologically suitable so that it does not interfere with the normal function of the body but simply accumulates in the organ to be examined.
- The radionuclide selected must have the right physical properties, such as the type of radiation emitted, its energy and the length of the half-life. It is important that the half-life is not too long, so as to minimise radiation dose, and not too short, so that there is not sufficient time to carry out diagnostic imaging or therapy. In imaging, gamma-emitting tracers are used, as the radiation can easily escape from the body and can be readily detected.

Gamma-emitting radioactive tracers

Three important gamma-emitting radionuclides used in nuclear medicine are technetium-99_m, iodine-131 and indium-111.

1. **Technetium-99_m** (Tc-99_m or $^{99_m}_{43}Tc$) is a **metastable** isotope of technetium-99. The subscript 'm' stands for metastable, which means that the nucleus is **excited** (see section 10.2 of Chapter 10). In order for the nucleus to de-excite to a more stable state, it emits a gamma ray, of photon energy 140.5 keV, and has a half-life of 6 hours.

Figure 1 shows the decay scheme. Molybdenum-99 (Mo-99) decays to technetium-99 (Tc-99) via the metastable stateTc-99$_m$ and then to ruthenium-99 (Ru-99), which is stable.

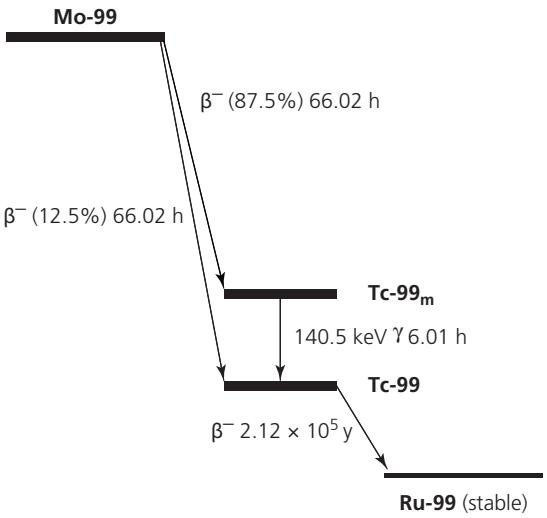


Figure 1 *Decay scheme involving Tc-99$_m$*

Technetium-99$_m$ is commonly used as a medical tracer because the gamma radiation it emits has sufficiently low energy and its short physical half-life makes it ideal for imaging internal body organs with minimal radiation dosage. The energy of 140.5 keV is suitable for detection with a gamma camera (see Medical Physics section 5.4). In addition, it can be easily separated from its parent Mo-99 radionuclide, and its chemical properties are such that it can be used for labelling a wide variety of compounds that target different organs within the body. Common uses for Tc-99$_m$ include studies of the brain, heart, thyroid, lungs, liver, gall bladder, kidneys, skeleton, blood and tumours.

2. **Iodine-131** (I-131 or $^{131}_{53}$I) is a radionuclide primarily used as a tracer to study the functioning of the thyroid gland (and also for treating thyroid disorders). It is also used to locate tumours of the brain and of the liver. I-131 decays with a half-life of 8 days with beta-minus and gamma emissions. The energy of the main gamma emission is 364 keV and that of the main beta emission is 606 keV.

   Iodine-131 is injected into the bloodstream and integrates with molecules that comprise thyroid hormones; consequently, it accumulates in the thyroid. Gamma ray scans of the thyroid can thus monitor thyroid activity and assess the function and size of the thyroid gland (Figure 2). In recent years, iodine-131 has been largely superseded by another isotope, iodine-132, a gamma emitter with a half-life of only 13 hours, to reduce patient dose.

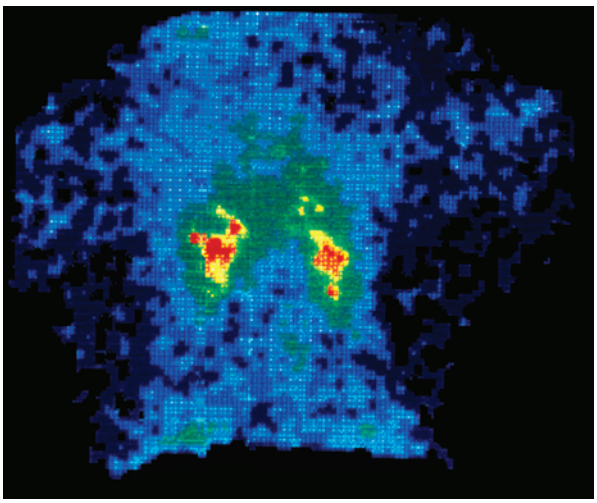


Figure 2 *Radionuclide imaging of the thyroid using I-131*

3. **Indium-111** (In-111 or $^{111}_{49}$In) is an isotope of indium with a half-life of 2.8 days. Its main emissions are gamma rays at 245 keV and 171.2 keV and X-ray photons at 23 keV. Indium-111 is used as a tracer for studies of the bone marrow, as a tumour-localising tracer, as a tracer in the cerebrospinal fluid and as a labelling agent for white blood cells.

The molybdenum–technetium generator

The short half-life (6 hours) of technetium-99$_m$ means that it has a short shelf life. It therefore has to be produced on demand. This is done using a **molybdenum–technetium generator**. These machines produce Tc-99$_m$ from the decay of its parent isotope molybdenum-99. Molybdenum-99 has a longer half-life (66 hours, see Figure 1) and can therefore be transported to hospitals and then remain useful for up to a week. Molybdenum-99 is produced in nuclear reactors that are specifically devoted to the production of nuclear isotopes for medicine. A highly enriched uranium target is bombarded with neutrons, causing fission and forming Mo-99 (and other isotopes).

The molybdenum produced is fixed or *adsorbed* onto the surface of alumina in a glass or plastic column as ammonium molybdenate and placed into the molybdenum–technetium generator (Figure 3). The molybdenum decays over the next few days into

technetium-99_m as water-soluble pertechnate ions in a column, which is then a ready source of technetium-99_m. The column is washed through with a saline (sodium chloride) solution or **eluted**, which means that the technetium-99_m ions exchange places with the chloride atoms. The saline is thus made rich with the technetium-99_m, which is drawn off through a stainless-steel needle into a sterilised vial, and is in a form that can be injected intravenously into a patient.

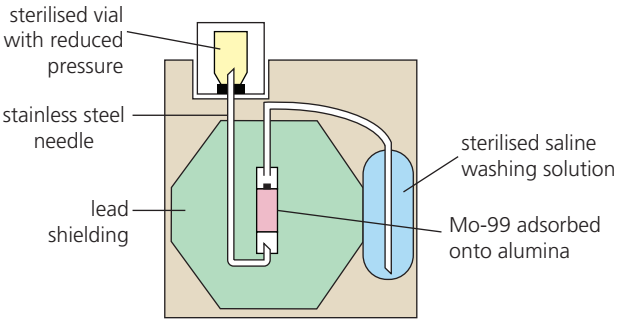


Figure 3 *A molybdenum–technetium generator. The technetium-99_m is in a saline solution and drawn into a sterilised vial ready for use.*

The ion column can be re-used because the concentration of technetium-99_m builds up again as the molybdenum continues to decay (see Figure 4). After several days, though, the generator will need to be replenished with a fresh source of molybdenum-66 from the reactor.

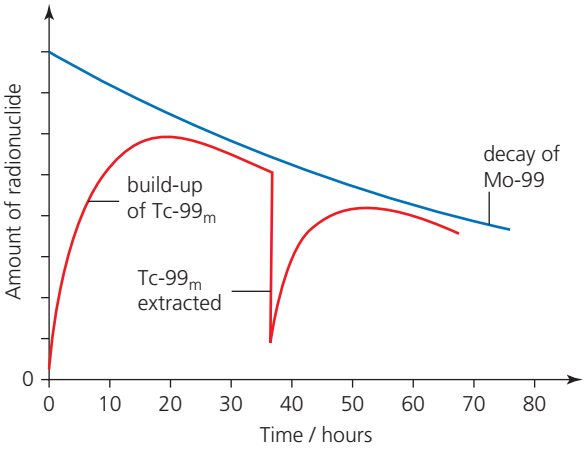


Figure 4 *Graph showing the production of technetium-99_m in a molybdenum–technetium generator. The amount of technetium-99_m that can be extracted decreases as the molybdenum decays*

QUESTIONS

1. Explain what is meant by a radioactive tracer in medicine.
2. Give **four** reasons why technetium-99_m is widely used as a radionuclide in gamma imaging.
3. **a.** Write the nuclear equation for the decay of technetium-99_m to technetium-99.
   b. Technetium-99 is not stable. Why is its decay by beta emission (see Figure 1) considered *not* to be a hazard to the patient?

Stretch and challenge

4. A molybdenum–technetium generator is replenished with molybdenum-99 that has a half-life $T_{1/2}$ = 66 hours and an initial activity A_0 = 10.0 GBq. Technetium-99_m is extracted from the generator at 36 hour intervals. At each extraction, the activity from the generator falls by 80% . How many extractions can be made from the generator if the activity of the extracted technetium must not fall below 5% of the initial activity of the molybdenum?

KEY IDEAS

- Radioactive isotopes may be used as tracers to image parts of the human body.
- A chemical with an affinity for a particular part of the body, but no adverse effect on the body, is labelled with the radioactive tracer and injected into the body.
- The radionuclide used must emit a suitable type of radiation (usually gamma), have a suitable photon energy to enable detection and have a suitable half-life.
- Three important gamma-emitting tracers are technetium-99_m, iodine-131 and indium-111.
- Technetium-99_m is produced on demand using a molybdenum–technetium generator.

5.2 EFFECTIVE HALF-LIFE OF RADIONUCLIDES IN MEDICAL USES

Radionuclides decay over a period of time according to an exponential relationship (see section 9.5 of Chapter 9). The rate at which a radionuclide decays is independent of its physical state (solid, liquid or gas), the chemical compound in which the nucleus

is bonded and any external environmental factors such as temperature or pressure. The rate of decay depends only on its nuclear properties. In medicine, the half-life of the radionuclide, $T_{1/2}$ (the time taken for the number of nuclei of a radionuclide, or its activity level, to reduce by one-half of its initial value), is called its **physical half-life** T_P.

Values of physical half-life for radionuclides commonly used in nuclear medicine are shown in Table 1.

Radionuclide	Physical half-life T_P
Technetium-99$_m$, $^{99}_{43}$Tc	6.01 hours
Indium-111, $^{111}_{49}$In	2.81 days
Iodine-123, $^{123}_{53}$I	13.27 hours
Iodine-125, $^{125}_{53}$I	59.41 days
Iodine-131, $^{131}_{53}$I	8.02 days
Potassium-40, $^{40}_{19}$K	1.49×10^9 years
Tritium, $^{1}_{1}$H	12.3 years
Radium-226, $^{226}_{88}$Ra	1620 years
Fluorine-18, $^{18}_{9}$F	110 minutes
Yttrium-90, $^{90}_{39}$Y	2.69 days
Gallium-67, $^{67}_{31}$Ga	3.26 days

Table 1 *Values of physical half-life for common medical radionuclides*

In medical procedures, the radionuclide is combined into a chemical compound to form a radiopharmaceutical. Any chemical compound introduced into the body will eventually be excreted by the normal processes of respiration, urination and defecation. Thus, over time, its concentration in the body will naturally decrease, also exponentially with time, with a **biological half-life** T_B.

When a radiopharmaceutical is taken into the body, there will be an **effective half-life** T_E, which takes into account the biological excretion and the radioactive decay, given by

$$\frac{1}{T_E} = \frac{1}{T_B} + \frac{1}{T_P}$$

Rearranging this gives the effective half-life as

$$T_E = \frac{T_P \times T_B}{T_P + T_B}$$

Whereas the physical half-life of a radionuclide is constant, the effective half-life can vary because the biological half-life depends on the patient's individual metabolism and state of health. However, the effective half-life of a radionuclide is always less than the physical half-life, because the biological removal of the nuclide from the body means that the nuclide is removed more quickly overall.

Worked example

The biological half-life of indium-111 is 5.5 hours and the physical half-life is 2.81 days. What is the effective half-life of indium-111 in the body?

Converting the physical half-life to hours ($2.81 \times 24 = 67.44$) and using the expression given in the text for the effective half-life gives

$$T_E = \frac{T_P \times T_B}{T_P + T_B}$$
$$= \frac{67.44 \times 5.5}{67.44 + 5.5}$$
$$= \frac{370.92}{72.94} = 5.08 \text{ h}$$

QUESTIONS

5. **a.** Explain the difference between biological half-life and physical half-life.
   b. If iodine-131 has an average biological half-life of 21 days and a physical half-life of 8.0 days, what is its effective half-life?
   c. Why is the effective half-life for a radiopharmaceutical taken into the body not constant?
6. A radioactive isotope has a physical half-life of 14.3 days. It is injected into a patient for therapeutic purposes and concentrated in a specific organ. The maximum detectable activity in the organ occurs 10 hours after injection. One week after the injection, the activity is 35% of what it was when it reached its maximum detectable activity. Calculate, for this patient,
   a. the effective half-life of the isotope
   b. the biological half-life.

KEY IDEAS

- Biological half-life T_B is the time taken for a chemical compound to be excreted from the body to the point where only one-half of the initial amount is left. It is not a constant value and depends on the health and metabolism of individual patients.
- Effective half-life T_E takes into account the processes of radioactive decay and biological removal. It is the actual time that a radiopharmaceutical's activity reduces by one-half of its initial value:

$$\frac{1}{T_E} = \frac{1}{T_B} + \frac{1}{T_P}$$

where T_P is the physical half-life, that is, $T_{1/2}$ for the decay of the radionuclide, which is constant.

5.3 GAMMA CAMERAS

A **gamma camera** (see Figure 13 in Chapter 10) is used to detect the radiation from a gamma-emitting tracer introduced into the patient's body (Figure 5). Gamma cameras can provide diagnostic images of the brain, thyroid (see Figure 2), liver, kidneys and other major organs and structures in the body. Gamma scans differ from X-ray imaging techniques in that they map the functions and metabolic processes of the body rather than providing simply structural images. The tracer's progress through the body can be tracked by the camera. Typical scans take about 30 minutes to complete.

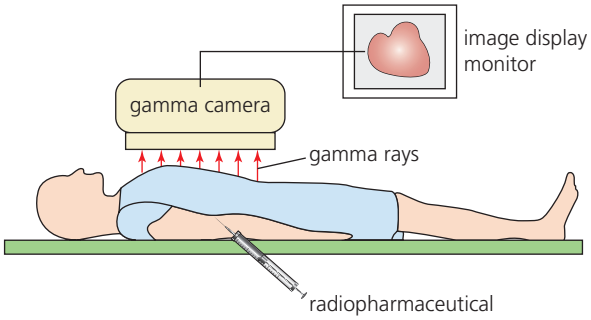


Figure 5 *A patient undergoing a gamma scan. The gamma camera is placed over the patient and detects gamma rays coming from the body.*

A gamma camera (Figure 6) consists of an array of **photomultipliers** coupled to a large scintillator crystal of sodium iodide (NaI). Gamma rays from the patient's body are incident on the scintillator crystal and the light flashes are detected by the photomultipliers, which convert the light flashes to voltage pulses.

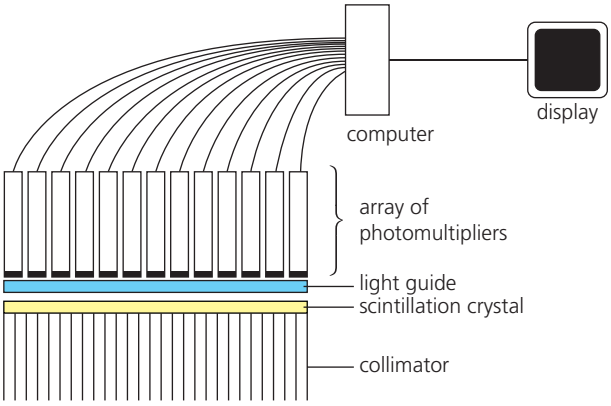


Figure 6 *The main components of a gamma camera*

Between the patient and the scintillator crystal is a collimator made up of a honeycomb of lead tubes. This ensures that any gamma ray photons not parallel to the collimator are absorbed by the lead walls and are not detected by the scintillator crystal. This also allows a sharper image to be formed. The flashes produced in the crystal from the gamma rays are detected and amplified by the photomultipliers, and the intensity of the pulses is processed by a computer to determine their location in the body and so form an image.

Figure 7 shows how a photomultiplier works. It consists of an evacuated glass tube containing a number of electrodes, called **dynodes**, which are at increasing electric potential down the tube. Photons from the scintillator crystal enter the photomultiplier tube and strike the **photocathode**. Electrons are produced as a result of the photoelectric effect and are accelerated towards the first dynode. When one electron hits the dynode, this causes several electrons to be ejected, a process known as **secondary emission**. The secondary electrons are then accelerated to the next dynode, which is at a higher potential than the previous one, producing even more electrons. This action is repeated through the series of dynodes, resulting in **electron multiplication** taking place. A large number of electrons is thus collected at the anode and the original signal is amplified by a factor of about 10^6. The output signal is proportional to the number of electrons ejected by the photocathode, which is proportional to the number of visible photons that strike it, which in turn is proportional to the energy of the incident gamma ray photon.

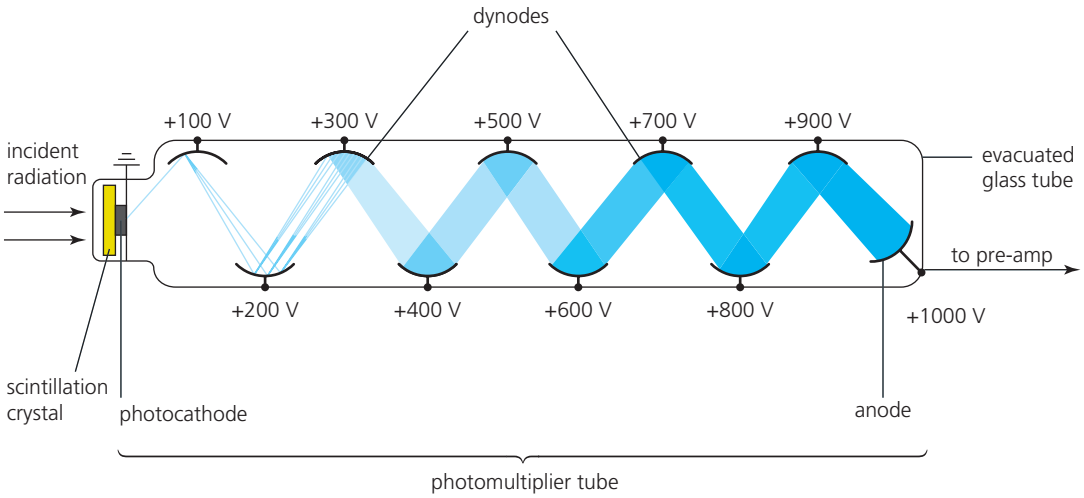


Figure 7 *A photomultiplier attached to a scintillation crystal*

The most common radiopharmaceutical tracer used in gamma camera imaging is technetium-99_m.

QUESTIONS

7. a. Why is it necessary to collimate the gamma ray photons before they are detected by the scintillator in a gamma camera?

b. In what important respect does gamma camera imaging differ from conventional X-ray imaging?

c. Why is it necessary to use a tracer with a half-life of hours rather than seconds or minutes in a gamma camera scan?

KEY IDEAS

- A gamma camera detects the gamma rays from a radioactive tracer introduced into the body using a large scintillator crystal connected to an array of photomultipliers.
- A photomultiplier detects flashes of light from the scintillator and converts them to voltage pulses, which are amplified.
- The variation in intensity of the signal from the photomultipliers can be processed by a computer to form an image.
- Gamma cameras are able to determine the functional state of organs and metabolic processes in the body.

ASSIGNMENT 1: IMAGING THE HEART

(MS 0.1, MS 0.2, MS 0.5, PS 1.1, PS 1.2)

Statistics from the British Heart Foundation show the following facts about coronary heart disease (CHD):

- CHD is the UK's single biggest killer. It is also the leading cause of death worldwide.
- In the UK, more than one in seven men and nearly one in ten women die from CHD.
- CHD is responsible for nearly 70 000 deaths in the UK each year, an average of 190 people each day, or one death every eight minutes.
- Most deaths from CHD are caused by a heart attack.
- Over 22 000 people under the age of 75 in the UK die from CHD each year.

For these reasons, it is important to ensure that your heart is well and working properly. Doctors may use an imaging technique called a myocardial perfusion scan (MPS) to examine the functioning of the heart. This looks at the pumping action of the heart muscle (called the myocardium) and the flow of blood to the heart. It can help diagnose CHD and

may also help in deciding if corrective surgery to improve the heart function is possible.

During an MPS a radiopharmaceutical carries a gamma-emitting radioactive tracer (commonly technetium-99_m) to the heart through the bloodstream where it is detected by a gamma camera. As the tracer moves through the heart muscle, areas that have good blood flow (perfusion) absorb the tracer (Figure A1). Areas that do not absorb tracer may not be getting enough blood or may have been damaged by a heart attack and become scar tissue.

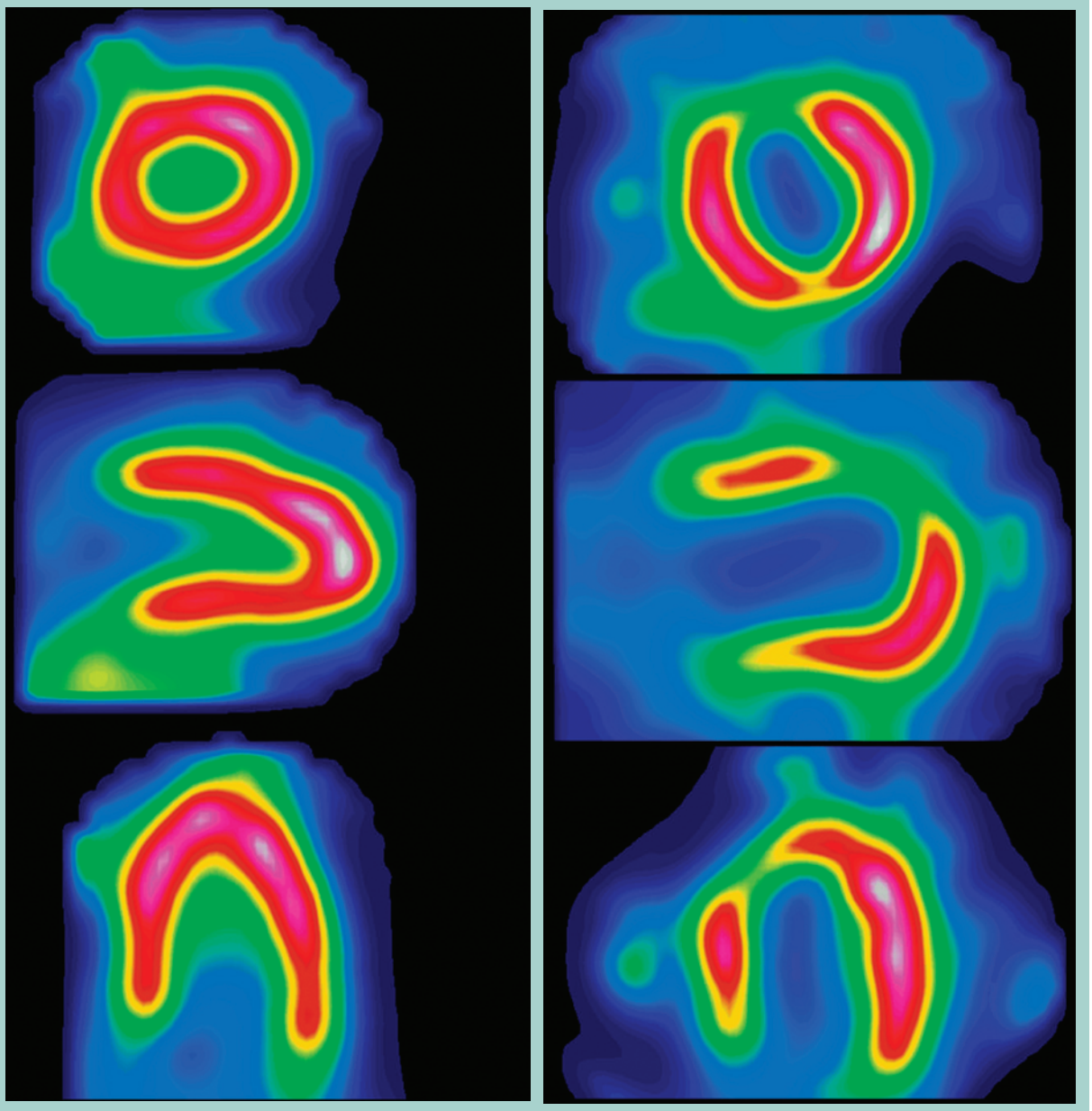


Figure A1 *Myocardial perfusion images of the heart. The red-orange colour represents a higher (and blue a lower) concentration of tracer detected by the gamma camera.*

Two sets of pictures may be made during the scan. One set is taken while the patient is resting, and another set is taken after the heart has been stressed, possibly by exercise. The resting pictures are then compared with the stressed images. The heart rate and blood pressure are closely monitored throughout the procedure.

An MPS uses an imaging technique called 'single-photon emission computed tomography' (SPECT). This is similar to CT scanning (see Medical Physics section 4.6) except that gamma rays are used instead of X-rays. The patient lies still, and a gamma camera records the scintillations coming from technetium-99_m in the heart. The gamma camera is rotated around the patient, taking measurements at different angles. These are sent to a computer that reconstructs an image of the heart.

SPECT imaging reveals the distribution of the radiopharmaceutical (Figure A1), and therefore the relative blood flow to the different regions of the myocardium.

QUESTIONS

A1 What is a radiopharmaceutical?

A2 A quantity of Tc-99_m with activity 10 GBq and half-life 6 h is injected into a patient in preparation for an MPS. The patient then exercises for 20 min as part of the stress test for the scan and then lies on a couch with a gamma camera. The scan takes 45 min. Assuming that the patient has not excreted in this time, what activity would be recorded immediately after the scan?

A3 If Tc-99_m has a biological half-life of 1 day, what is its effective half-life?

A4 How much activity is left in the patient one week after the scan? Why is this only likely to be an estimate?

A5 The scintillations in the gamma camera's scintillation crystal are detected by photomultipliers. Each photomultiplier has a chain of 10 dynodes. Electrons are produced at the photocathode by the photoelectric effect. For every electron that hits the dynode surface, five others are ejected by secondary emission.

a. How many electrons are ejected from:

i. the first dynode

ii. the second dynode

iii. the third dynode?

b. If a single photon from the scintillator in the gamma camera produces a single electron at the photocathode of the photomultiplier, how many electrons are produced at the output of the photomultiplier?

A6 The images shown in Figure A1 are from a patient with coronary heart disease. In order to improve blood flow, the patient underwent coronary bypass surgery, which improves blood flow to the heart muscle by diverting the flow of blood around a section of a blocked artery in your heart. The image on the left shows the heart before surgery and the one on the right after the coronary bypass surgery. Comment as to whether you think the surgery was successful in restoring blood flow to the patient's heart.

A7 What other diagnostic techniques could doctors perform to give information on the heart?

5.4 POSITRON EMISSION TOMOGRAPHY (PET)

Positron emission tomography (PET) is a different method of producing images of the body using a radioactive tracer. It is based on the process of annihilation of positrons and electrons when they come into contact with each other (*see section 3.3 in Chapter 3 in Year 1 Student Book*).

Some radionuclides used as medical tracers decay by beta-plus emission (see section 9.3 of Chapter 9), in which a positron (the antiparticle of an electron) is emitted. When injected into the body, the positron only travels a short distance before it encounters an electron and they annihilate one another, producing back-to-back gamma rays (Figure 8). The gamma photons produced must travel in opposite directions in order to conserve momentum.

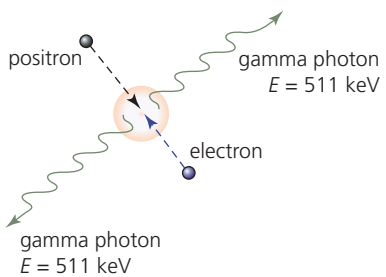


Figure 8 *When a position and electron meet, they annihilate, producing two gamma rays moving in opposite directions.*

In a PET scan, a positron-emitting radioactive tracer is injected into the body. Positron-emitting radionuclides of elements such as carbon and oxygen, which are taken up in biological processes, can be used. Alternatively, biological compounds such as the sugar glucose can be labelled with a tracer such as radioactive fluorine-18, a positron emitter with a half-life of about 110 minutes, becoming fluorodeoxyglucose (FDG). Glucose is preferentially taken up by the brain (the brain uses glucose as its main source of energy) and is metabolised by cancer cells, so FDG is ideal for detecting brain tumours. It can also be used to map the activity of the brain and diagnose conditions such as Alzheimer's disease.

A PET scanner consists of a circular ring of detectors that completely surround the patient (Figure 9).

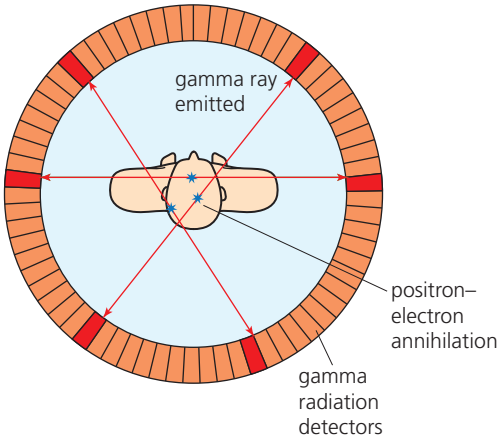
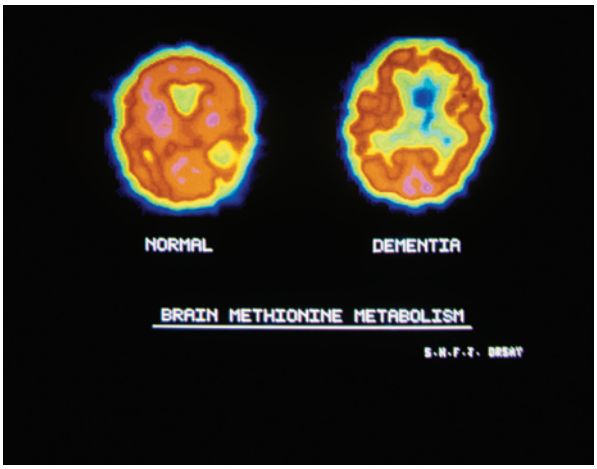
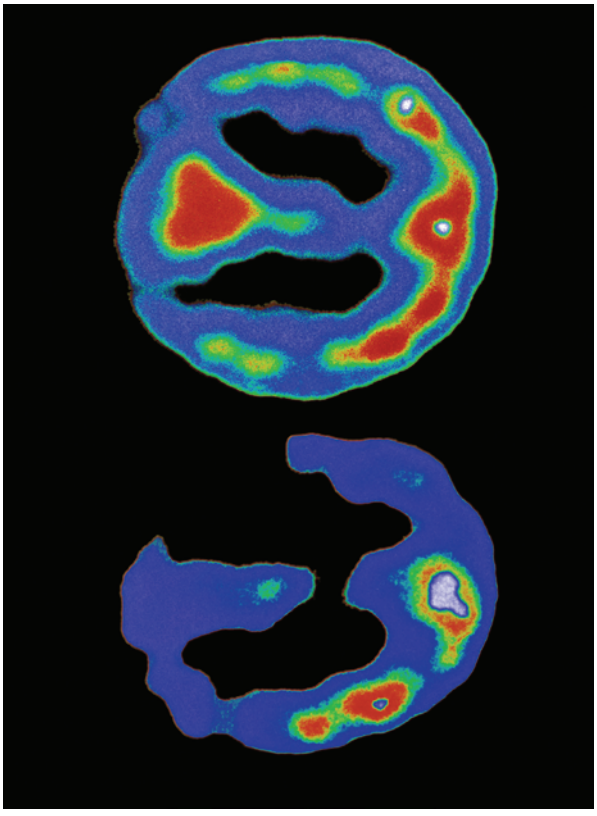


Figure 9 *During a PET scan, the patient is surrounded by a ring of detectors that detect gamma rays from positron–electron annihilation*

When a positron is emitted by a nucleus from the tracer, it annihilates almost instantly with an electron and two gamma rays emerge, travelling at 180° to one another. A simultaneous detection ('coincidence detection') of gamma ray photons in two detectors places the source on a line between those detectors. By taking a number of coincidence measurements of simultaneous detection at different angles, the position and distribution of the radionuclide can be determined. The data are processed by a computer to form an image of the region of the body where the radioactive tracer has accumulated (Figure 10).



(a)



(b)

Figure 10 *PET scans. (a) A PET scan of the brain of a patient with Alzheimer's disease (right) and a normal brain (left). (b) The location of a brain tumour in another patient. The scans show the relative concentrations (red being the highest) of the tracer FDG. In an Alzheimer's patient, there is a reduction in brain activity and therefore less uptake of glucose (larger blue area at bottom right). The tumour consumes glucose and therefore concentrates it (yellow area).*

PET scans are primarily used in the diagnosis and treatment of cancer (see Figure 10b). An area of abnormally high glucose concentration might indicate a fast-growing malignancy. A PET scan can also tell whether subsequent treatment with radiotherapy or chemotherapy has been effective and whether the cancer is still metabolically active or if its growth has been arrested.

See also *Assignment 1 at the end of section 3.3 in Chapter 3 in Year 1 Student Book*.

QUESTIONS

8. Explain how the emission of pairs of gamma rays can determine the location of the source of radiation in the body.
9. Why is the tracer fluorodeoxyglucose (FDG) particularly useful in assessing metabolic activity?

KEY IDEAS

- In a PET scan, the process of annihilation of positrons (from a radioactive tracer) and electrons produces simultaneous gamma rays at 180° that can be detected by a ring of detectors.
- PET scanners can determine the position, distribution and metabolic activity of cancer cells and can also measure brain activity.

5.5 IMAGING COMPARISONS

Table 2 shows a comparison of the different imaging techniques that have been considered in Medical Physics Chapters 3 to 5: ultrasound, magnetic resonance (MR), conventional X-ray, computed tomography (CT) and radionuclide imaging.

	Ultrasound	Magnetic resonance (MR)	X-ray	Computed tomography (CT)	Radionuclide
Imaging method	Reflected sound waves	Magnetic fields and radio waves	Absorption of X-rays	Absorption of X-rays	Emission of gamma rays
Radiation exposure	None	None	Moderate to high exposure to ionising radiation	Moderate to high exposure to ionising radiation	Moderate to high exposure to ionising radiation
Safety risk	None	None	Exposure to ionising radiation increases risk of cancer, especially in the fetus	Exposure to ionising radiation increases risk of cancer, especially in the fetus	Exposure to ionising radiation increases risk of cancer, especially in the fetus
Time for complete scan	Not used for whole head or body scanning. Real-time imaging of moving parts, e.g. heart	Scanning can take around 30 minutes	Usually completed in a few seconds	Usually completed within 5 minutes	Can take 30 minutes or more
Image resolution of different tissues	Less anatomical detail than MR or CT scans. Useful for imaging soft tissue	Much higher detail in soft tissues. Able to adjust contrast by changing the magnetic field and the radio pulses	Not as good at resolving differences between soft tissues. Need to use contrast agents to which patients may react badly. Excellent for imaging bony structures	Not as good at resolving differences between soft tissues. Need to use contrast agents to which patients may react badly. Excellent for imaging bony structures	Able to examine functional effectiveness and metabolic processes of body organs

Cost	Much cheaper than either CT or MR scanners	MR scanners are very expensive. Cost per scan is usually more than for CT scans	Cheaper than CT scanning	CT scanners are expensive, but cheaper than MR scanners	Expensive
Other comments	Portable	MRI machines can produce images in any plane without moving the patient	X-ray images can be processed quickly using electronic detectors	Lower scan times reduce motion blur. Less problematic for claustrophobic patients	Requires a lot of pre-planning and preparation, including the manufacture of radioisotopes

Table 2 *Comparison of different imaging techniques*

5.6 THERAPY USING HIGH-ENERGY X-RAYS

X-rays, as well as being used for imaging (see Medical Physics Chapter 4), can be used to treat cancers using **external beam radiotherapy** (**EBRT**). Cells in general are most sensitive to damage by X-rays when they are dividing. For cancer cells, cell division occurs at a faster rate than for healthy cells, and exposing them to X-rays will kill them at a faster rate. EBRT exploits this feature of cancer cells to reduce the size of tumours and other cancerous growths in the body.

The X-ray photon energies used for EBRT are much higher than those used in imaging. At low photon energies of 30–50 keV, bone absorbs up to 11 times more energy than the surrounding tissue, so more damage is done to bone than to cancerous soft tissue. At higher photon energies, between 0.5 and 5 MeV, the dominant absorption mechanism does not depend on the proton number of the material but occurs by the scattering of X-rays from electrons in the material, with no preferential absorption for either bone or soft tissue.

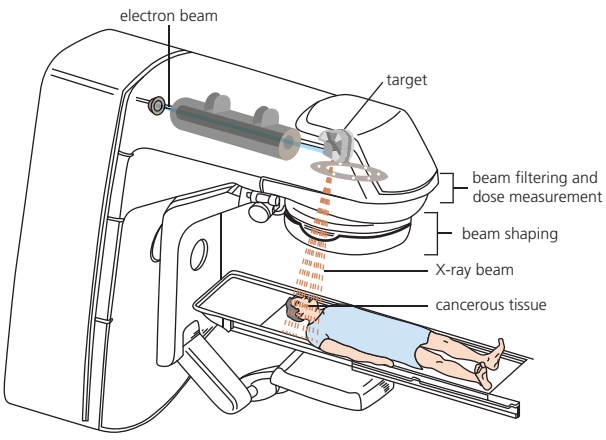


Figure 11 *A medical linac for external X-ray beam therapy*

High-energy X-rays for EBRT are produced using linear accelerators or linacs (see Chapter 5). These machines accelerate electrons to high energies using radio-frequency electromagnetic waves. The electrons are made to collide with a heavy-metal target and high-energy X-ray photons are emitted. The X-ray beam is shaped and directed to the cancerous tissue (Figure 11).

In order to reduce the damage to healthy cells and maximise the damage to the cancerous cells, the tissue to be irradiated is accurately located and then the X-ray beam is aimed precisely at it from different directions. Most medical linacs have a rotating gantry by which the beam can be rotated around the patient with the centre of the cancerous tissue as the centre of rotation (Figure 12). The beams from different angles will overlap at the centre, ensuring that the cumulative dose to the cancer is much greater than that to the surrounding tissue.

Figure 12 *The rotating gantry limits the dose to healthy tissue.*

Damage to healthy cells is also limited by delivering X-rays in small doses called **fractions** over a period of days or weeks. The healthy cells recover faster than the cancer cells, so spreading the treatment over a period of time gives the surrounding healthy issue time to recover.

Typical radiotherapy sessions with high-energy X-rays last only a few minutes. It is important that the patient is perfectly still during the procedure, so a mould is made to immobilise the part of the body that is targeted.

QUESTIONS

10. Why are high-energy X-rays (0.5–5 MeV) used for X-ray therapy rather than low-energy ones (30–50 keV)?

11. Explain how the following limit the exposure of healthy cells to radiation.

a. The X-ray source is aimed at a cancerous region from different angles.

b. X-ray therapy is given in fractions and not all at once.

KEY IDEAS

- Cancers can be treated using high-energy X-rays produced by linear accelerators (linacs).
- High-energy (0.5–5 MeV) X-rays are used, as they are not preferentially absorbed either by bone or by tissue.
- In order to minimise the dose to healthy tissue, the X-ray beams are aimed at a cancerous region from different angles by rotating the X-ray source around the patient.
- X-ray therapy is delivered in fractions to allow healthy tissue an opportunity to recover.

ASSIGNMENT 2: USING PROTONS TO TREAT CANCER

(MS 0.1, MS 0.2, MS 3.1, PS 1.2)

The use of high-energy X-rays in radiotherapy to treat cancer by destroying or arresting the growth of cancer cells has two significant disadvantages.

- The dose of X-rays must be delivered into the tumour to ionise the cancer cells and so damage them. Even with all the precautions discussed in Medical Physics section 5.6, healthy intervening and surrounding tissue is likely to be damaged. A lower-than-desired dose is therefore frequently used in order to minimise the damage to healthy tissues and avoid unwanted side effects.
- Secondly, because X-rays lack charge and mass, they interact less with matter than charged particles, and hence their ionising power is not as great.

A different approach to destroying cancer cells is to use *proton therapy* (Figure A1). This uses beams of protons at specific energies to target tumours in the body. The protons are produced in particle accelerators such as cyclotrons or synchrotrons, which produce protons at defined energies in the range 70–250 MeV.

Figure A1 *A proton therapy machine*

Proton therapy has some relative advantages over X-ray therapy to treat cancer.

- Protons, having mass and charge, interact more strongly with electrons in matter.
- Protons that have been accelerated to a specific energy will deposit their maximum energy in the body at a *specific depth* (where they have slowed to optimum speed) and not beyond the tumour site.

Figure A2 shows a graph of the relative dose to tissue from beams of electrons, beams of protons, and photons (X-rays) as a function of depth in the body.

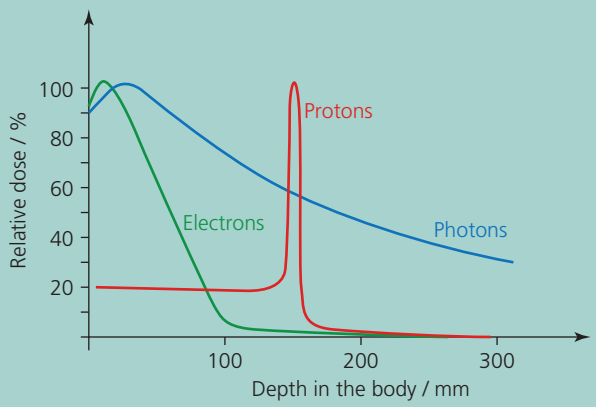


Figure A2 *Relative dose versus body depth for electrons, protons and photons*

The maximum dose delivered by protons occurs as a peak (called the 'Bragg peak') at one particular depth that depends on the proton energy. As a proton enters the body, it is travelling very fast and deposits only a small dose on its way to the tumour site. At the Bragg peak, the absorbed dose suddenly increases to a maximum and the proton is stopped. In X-ray therapy, it can be seen from the curve for photons that the relative dose is continuous throughout the body and has no 'cut-off' depth.

In this way, proton therapy allows doctors to target tumours with much greater precision inside the body, using maximum dosage while at the same time reducing unnecessary damage to healthy cells. A further advantage is that fewer treatment sessions for the patient are needed – a high dose can be delivered to a tumour in a single treatment, whereas the X-ray therapy would need to be delivered in fractions (see Medical Physics section 5.6).

The technology for cyclotron-based or synchrotron-based proton therapy facilities owes much to that developed for particle physics research. The first centres for proton therapy were located in or near particle physics research laboratories, where particle accelerators were already in use and could be adapted for medical use. There are now increasing numbers of dedicated proton therapy facilities solely for the treatment of cancer.

QUESTIONS

A1 **a.** Define an electronvolt (eV).

b. Calculate the energy in joules of a 250 MeV proton.

A2 Explain why, in high-energy X-ray radiotherapy, it is necessary to give doses in fractions, in several treatment sessions over a period of days or weeks.

A3 What is meant by the Bragg peak in proton therapy, and what is its significance in the planning of treatment?

A4 Tumours often have extended sizes, sometimes being several centimetres across. With reference to Figure A2, suggest how a clinician could modify the proton therapy beam to ensure that all the tumour is irradiated.

A5 With reference to Figure A2, suggest why electrons are not used in the same way as protons for cancer therapy.

A6 Suggest why there are only relatively few proton therapy treatment centres around the world and they are not in general hospitals?

5.7 THERAPY USING RADIOACTIVE IMPLANTS

Radiation sources may be implanted inside the body in order to treat cancer; this is known as **brachytherapy**. Unlike external beam therapy using high-energy X-rays, brachytherapy allows doctors to deliver higher doses of radiation to smaller, more specific areas of the body. Radioactive implants can be used to treat cancers of the uterus, cervix, breast, prostate and lung.

Radioactive implants are sealed sources made of small metal wires or 'seeds' (Figure 13).

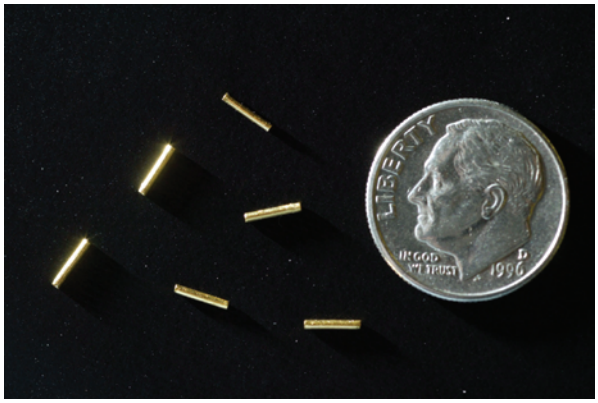


Figure 13 *Radioactive implant 'seeds'*

The implant can be placed in cavities in the body using a special applicator, or surgically implanted directly into a tumour. They can be left in permanently until the radiation has decayed fully, or may be removed after a few minutes or days (in the case of a body cavity). The radiation source maintains its position in relation to the tumour when the patient moves or if there is any movement of the tumour within the body. This means that the risk is reduced to healthy areas of the body around the tumour, and a uniform dose can be maintained.

An internal radiation source can present an external radiation hazard to those around the patient, particularly if it is a gamma source. The patient may need to be isolated, and medical staff who treat the patient need to take precautions so they are not exposed to too much radiation.

The suitability of a radionuclide for implantation is determined by its half-life, the radiation type, its penetrating power, ionising ability and energy. Alpha emitters are not often used, as they have very little penetrating power in tissue, although they can be very energetic. Radionuclides emitting beta and gamma rays are usually employed in brachytherapy, as they have greater penetrating power. Beta radiation does not penetrate more than 3–4 mm, but they are useful for the treatment of surface areas like those of the skin and eye. They are also more ionising than gamma rays, meaning that more damage is done to cells in a small cancerous region, leaving surrounding healthy tissue intact.

Radioactive implants may cause fewer side effects than external beam radiation, and the overall treatment time is usually shorter.

QUESTIONS

12.a. What is the main advantage that radioactive implants have over external radiation therapy?

b. What precautions need to be employed when using this method of treatment?

c. Why are beta-emitting implants often the most suitable for treating small areas of cancerous tissue?

KEY IDEAS

- Radioactive implants are small sealed sources of radiation that are implanted inside the body close to cancerous tissues.
- Implants enable higher doses of radiation to be delivered to smaller areas of the body, with less risk to surrounding healthy tissue.
- Safety precautions must be observed when using implants to avoid unnecessary exposure of other people to radiation.

PRACTICE QUESTIONS

1. Table Q1 gives the properties of two radionuclides used in medicine.

	Technetium-99_m	Caesium-137
Emitted radiation	Gamma	Beta minus
Half-life	6.0 hours	30 years
Energy of gamma ray / keV	140	662

Table Q1

Suggest which one of the nuclides in Table Q1 is suitable as

a. a tracer for medical diagnosis

b. a radioactive implant.

2. a. Explain the difference between *effective half-life*, *physical half-life* and *biological half-life* when using radioactive materials for diagnostic imaging in medicine.

b. Why is the effective half-life always less than the physical half-life?

3. A radiopharmaceutical contains a tracer with a half-life of 14 days and is injected into a patient. The count (corrected for background radiation) from a specific part of the body was initially 3500 counts s^{-1}. Seven days later, the count from the same part of the body is measured (corrected) at 1800 counts s^{-1}. What is the biological half-life of the tracer?

4. a. Explain briefly how positron emission tomography (PET) uses gamma rays to locate the position of a tumour in the body.

b. Describe **two** ways of limiting exposure to healthy cells when treating a patient with a cancerous tumour using high-energy X-ray radiotherapy.

5. Compare the use of CT and ultrasound scans in terms of patient safety, convenience and diagnostic information. The quality of your written communication will be assessed in your answer.

ANSWERS TO IN-TEXT QUESTIONS

1 PHYSICS OF THE EYE AND THE EAR

1. **a.** As the light goes from the cornea ($n = 1.38$) into the aqueous humour ($n = 1.33$).

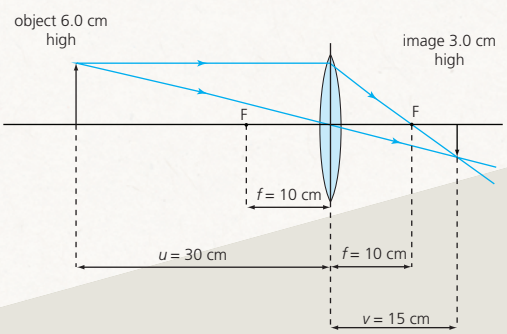
b. It subtracts from the initial deviation, because light is going from a medium of high refractive index to one of lower refractive index.

2. **a.** At low levels of illumination, the cones will not be stimulated. Scotopic vision using rod cells will dominate and the wavelength range of detected light will shift towards the blue end of the visible spectrum, with maximum sensitivity at about 500 nm and no response above about 670 nm.

b. The response of rods does not allow the brain to distinguish colours, so in very dim light colours are 'washed out'.

3. The density of cones in this region is very high and so the visual angle that can be resolved is very small. Also there are no large blood vessels present, giving an unobstructed view.

4.



The image height is 3.0 cm.

5. $\frac{1}{v} = \frac{1}{f} - \frac{1}{u} = \frac{1}{10} - \frac{1}{75} = \frac{13}{150}$

so $v = \frac{150}{13} = 11.5$ cm, which is positive, so the image is real.

Magnification $= \frac{v}{u} = \frac{11.5}{75} = 0.15$

Image height $= 0.15 \times 2 = 0.3$ cm

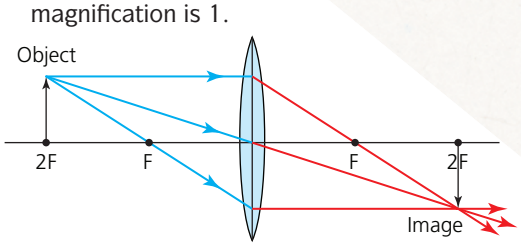
6. The lens is diverging, so the focal length is negative:

$$\frac{1}{v} = \frac{1}{f} - \frac{1}{u} = \frac{1}{-12} - \frac{1}{35} = \frac{-47}{420}$$

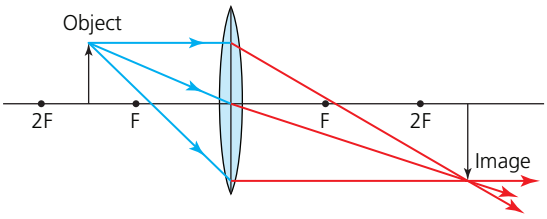
So $v = -8.9$ cm. This has a negative value, therefore the image is virtual and upright.

Magnification $= \frac{v}{u} = \frac{-8.9}{35} = -0.26$

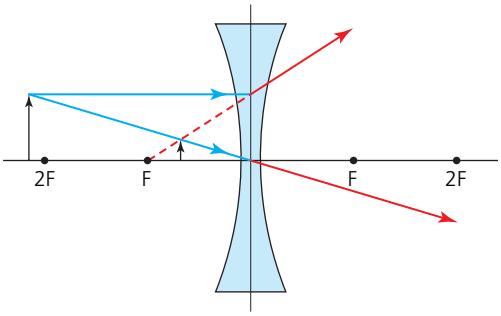
7. **a.** The image is real and inverted and the magnification is 1.



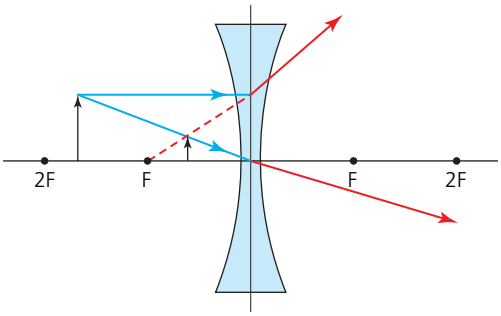
b. The image is real and inverted and the magnification is > 1.



8. **a.** The image is virtual and upright and the magnification is < 1.



b. The image is virtual and upright and the magnification is < 1.



(The ray diagrams show that, for a diverging lens, as the object distance is decreased, the image distance is decreased and the image size is increased. So as an object approaches the lens, its virtual image on the same side of the lens approaches the lens as well; and at the same time, the image becomes larger.)

9. To work out lens power, the focal length must be in metres.

a. In question 5, convex lens, focal length 10 cm = 0.1 m, power $= \frac{1}{0.1} = +10\,\text{D}$

b. In question 6, concave lens, focal length –12 cm = –0.12 m, power $= \frac{1}{-0.12} = -8.3\,\text{D}$

10. The power of the *unaided* eye when viewing an object at the near point of 70 cm is

$$P = \frac{1}{f} = \frac{1}{u} + \frac{1}{v} = \frac{1}{0.70} + \frac{1}{0.017} = 60.3\,\text{D}$$

(Remember that u and v must be expressed in metres when calculating lens powers.)

The power of the *corrected* eye when viewing an object at the near point of 20 cm is

$$P = \frac{1}{f} = \frac{1}{u} + \frac{1}{v} = \frac{1}{0.20} + \frac{1}{0.017} = 63.8\,\text{D}$$

So, in order to bring the power of the unaided eye to that of the corrected power, the overall lens power must be increased by 63.8 D – 60.3 D = +3.5 D. This means that corrective lens of power +3.5 D must be added to correct the defect of vison. The positive sign indicates that the lens is a convex lens.

11. **a.** $\frac{1}{f} = \frac{1}{u} + \frac{1}{v} = \frac{1}{0.50} + \frac{1}{0.017} = 60.8\,\text{m}^{-1}$, so f = 16.4 mm

b. $P = \frac{1}{f} = 60.8\,\text{D}$

c. $P = \frac{1}{f} = \frac{1}{\infty} + \frac{1}{v} = 0 + \frac{1}{0.017} = 58.8\,\text{D}$

d. Power of eye required to view object normally = power of unaided eye when viewing object + power of correctives lens. So power of corrective lens = 58.8 – 60.8 = –2.0 D to correct the defect of vison. The negative sign indicates that the lens is a concave lens.

12. **a.** Harry has 2.00 D of short-sightedness in his right eye, with 1.50 D of astigmatism and an axis of 180°.

b. Harry has 3.50 D of long-sightedness in his left eye, with 3.00 D of astigmatism and an axis of 45°.

13. The cochlea contains hair cells along the basilar membrane. Different regions of the membrane respond to sound waves of different frequencies.

14. The ossicles of the middle ear (anvil, malleus and stirrup) transfer vibrations from the eardrum to the oval window. The bones act as a lever, producing an increase in force. The area of the oval window is much smaller than that of the eardrum, and since pressure = force/area, the pressure at the eardrum is amplified at the oval window.

15. a. Sound intensity $= \frac{10^8}{4\pi \times (8000)^2} = 0.1\,\text{W m}^{-2}$

b. Using $I_r = \frac{I_{source}}{4\pi r^2}$, $r = \sqrt{\frac{I_{source}}{4\pi I_r}} = \sqrt{\frac{10^8}{4\pi \times 1}} = 2.8\,\text{km}$

16. The decibel rating increases by 10 for every 10-fold increase in the sound intensity.

a. 100 = 10 × 10, so 20 dB

b. 1000 = 10 × 10 × 10, so 30 dB

17. We have $+6\,\text{dB} = 10\log_{10}\left(\frac{I}{I_0}\right)$. Therefore

$\frac{I}{I_0} = 10^{0.6} = 4.0$, so $I = 4 \times I_0$, which is a quadrupling of the intensity.

18. a. Sound intensity $= 10\log_{10}\left(\frac{I}{I_0}\right)$

$= 10\log_{10}\left(\frac{50 \times 10^{-5}}{1 \times 10^{-12}}\right) = 87\,\text{dB}$

b. Intensity level (dB) $= 10\log_{10}\left(\frac{I}{I_0}\right)$

$= 10\log_{10}\left(\frac{1}{1 \times 10^{-12}}\right) = 10 \times 12 = 120\,\text{dB}$

19. a. Intensity level $= 10\log_{10}\left(\frac{I}{I_0}\right)$ so

$$60\,\text{dB} = 10\log_{10}\left(\frac{I}{1 \times 10^{-12}}\right)$$
$$6 = \log_{10}\left(\frac{I}{1 \times 10^{-12}}\right)$$
$$I = 10^6 \times 10^{-12}$$
$$= 1 \times 10^{-6}\,\text{W m}^{-2}$$

b. Intensity at 20 dB $= 10^2 \times 10^{-12}$
$= 1.0 \times 10^{-10}\,\text{W m}^{-2}$
Intensity at 70 dB $= 10^7 \times 10^{-12}$
$= 1.0 \times 10^{-5}\,\text{W m}^{-2}$

Fractional change

$= \frac{1.0 \times 10^{-5} - 1.0 \times 10^{-10}}{1.0 \times 10^{-10}} = 99\,000$

20. Intensity level (dB) $= 10\log_{10}\left(\frac{I}{I_0}\right)$ so

intensity level 1 $= 10\log_{10}\left(\frac{I_1}{I_0}\right)$

intensity level 2 $= 10\log_{10}\left(\frac{I_2}{I_0}\right)$

For a 10 dB increase:

intensity level 1 − intensity level 2 = 10

$$= 10\left[\log_{10}\left(\frac{I_1}{I_0}\right) - \log_{10}\left(\frac{I_2}{I_0}\right)\right]$$

So

$$\log_{10}\left(\frac{I_1}{I_0}\right) - \log_{10}\left(\frac{I_2}{I_0}\right) = 1$$

Taking inverse logs:

$$\frac{I_1 / I_0}{I_2 / I_0} = \log_{10} 1$$
$$\frac{I_1}{I_2} = \frac{1}{10}$$

So $I_2 = 10\,I_1$

21. a. 60 phons, by definition

b. Just under 50 phons

22. a. The sensitivity of the human ear varies according to the frequency of the sound. The dBA scale simulates the frequency response of the ear. Frequencies at which sounds would be sensed by the human ear as quieter than a sound at 1kHz have the level increased to that loudness, and those which would be detected by the human ear as louder than a sound at 1kHz have the level of the loudness reduced.

b. Compared with graph A, graph B shows a decrease in hearing level with increasing frequency, which is characteristic of hearing loss with old age.

c. Compared with graph A, graph C shows a decreasing hearing loss with a 'notch' at 4 kHz, which is characteristic of deafness due to exposure to noise.

2 BIOLOGICAL MEASUREMENT

1. a. The resting potential is an electrical potential across the nerve cell membrane caused by an imbalance of ions. This leads to a potential difference across the nerve cell membrane and the cell is polarised.

The action potential occurs when the nerve is stimulated. Movement of ions causes the potential difference across the nerve cell membrane to reverse. The cell is depolarised and subsequently repolarised.

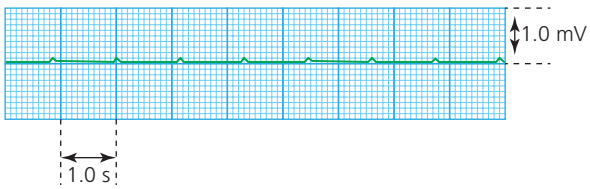
b. The action potential in heart muscle is similar to that in a nerve cell – depolarisation from the resting state and then repolarisation – but the depolarised state lasts longer and the time for the whole cycle is longer.

c. **i.** contraction
ii. relaxation

2. **a.** There are about 10 pulses in 7 s, so pulse rate is (10/7) per second = (600/7) = 86 per minute

c. **i.** The peaks would be closer together.
ii. The peaks would be further apart.

3. There would be only P waves.



3 NON-IONISING IMAGING

1. $Z = \rho c$

$Z_{skin} = 1150\,kg\,m^{-3} \times 1730\,m\,s^{-1}$
$= 1.99 \times 10^6\,kg\,m^{-2}\,s^{-1}$

$Z_{air} = 1.30\,kg\,m^{-3} \times 330\,m\,s^{-1}$
$= 4.29 \times 10^2\,kg\,m^{-2}\,s^{-1}$

Z_{skin} is about 4600 times greater

2. From the equations $Z = \rho c$ and $c = \sqrt{K/\rho}$, the acoustic impedance is

$$Z = \rho c = \rho\sqrt{K/\rho} = \sqrt{K\rho}$$
$$= \sqrt{2.03 \times 10^9 \times 1060} = 1.47 \times 10^6\,kg\,m^{-2}\,s^{-1}$$

3. The unit of $\sqrt{K/\rho}$ is $Pa^{1/2}\,kg^{-1/2}\,m^{3/2}$
$Pa = N\,m^{-2} = kg\,m\,s^{-2}\,m^{-2} = kg\,m^{-1}\,s^{-2}$

So this gives for the unit of $\sqrt{K/\rho}$

$kg^{1/2}\,m^{-1/2}\,s^{-1}\,kg^{-1/2}\,m^{3/2} = m\,s^{-1}$

4. **a.** $Z_1 = Z_{air} = 4.29 \times 10^2\,kg\,m^{-2}\,s^{-1}$
$Z_2 = Z_{skin} = 1.99 \times 10^6\,kg\,m^{-2}\,s^{-1}$
Fraction reflected

$$= \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2} = \frac{(1.99 \times 10^6 - 4.29 \times 10^2)^2}{(1.99 \times 10^6 + 4.29 \times 10^2)^2} = 0.999$$

b. $Z_1 = Z_{gel} = 1.47 \times 10^6\,kg\,m^{-2}\,s^{-1}$
$Z_2 = Z_{skin} = 1.99 \times 10^6\,kg\,m^{-2}\,s^{-1}$
Fraction reflected

$$= \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2} = \frac{(1.99 \times 10^6 - 1.47 \times 10^6)^2}{(1.99 \times 10^6 + 1.47 \times 10^6)^2}$$
$$= 2.26 \times 10^{-2}$$

i.e. just over 2%

5. Attenuation is greater for higher frequencies. There is less body tissue to travel through in an infant, so attenuation is not such a problem

6. When the gel is applied between the transducer and the skin, most of the incident ultrasound enters the skin. If there is an air gap between the transducer and the skin, then almost all of the ultrasound is reflected and very little is transmitted. For this reason, gel is applied to make sure there is no air gap, and enough acoustic energy is present to reflect off internal structures in the body.

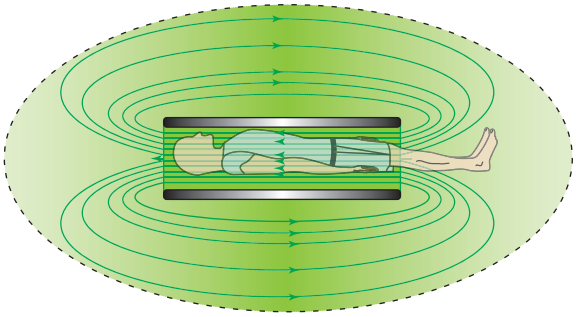
7. The difference in path length taken by the two pulses is twice the thickness of the organ, so

$$\text{thickness of organ} = \frac{\text{time between pulses} \times c}{2}$$
$$= \frac{170 \times 10^{-6} \times 1200}{2} = 0.102\,m$$
$$= 10.2\,cm$$

8. Incoherent bundles are used for illumination purposes and cannot transmit images. Coherent bundles are bundles where the optical fibres are precisely lined up at both ends of the fibre, so that an image can be transmitted.

9. Keyhole surgery using optical fibres results in smaller operation wounds, meaning less pain, less risk of infection and a faster recovery. The scars produced by insertion of the endoscope are smaller, resulting in a better cosmetic result. For some procedures, keyhole surgery means that an earlier discharge home from hospital is also possible.

10. The shape of the field is that of a solenoid. It is uniform in the area where the patient lies.



11. **a.** The hydrogen nuclei (protons) align themselves with the direction of the external magnetic field and precess around its direction at the Larmor frequency.

b. An external radio pulse equal to the Larmor frequency excites the nuclei to higher energy states. As they de-excite, they emit radio frequencies.

12. The gradient coils allow determination of the position of the source of RF emitted from the body, by varying the frequency. This allows cross-sectional scanning.

13. Soft (non-bony) tissues.

4 X-RAY IMAGING

1. An X-ray spectrum consists of a continuous spectrum of bremsstrahlung radiation, produced by electrons being decelerated in the target material and giving a full range of photon energies above a minimum. On this there is superimposed a line spectrum, corresponding to electrons in excited atoms of the target material dropping down to lower energy states.

2. **a.** Maximum energy $= 120\,\text{keV}$
$= 120 \times 10^3 \times 1.6 \times 10^{-19}$
$= 1.92 \times 10^{-14}\,\text{J}$

b. Minimum wavelength

$$= \frac{hc}{eV} = \frac{6.6 \times 10^{-34} \times 3.00 \times 10^8}{1.92 \times 10^{-14}} = 1.03 \times 10^{-11}\,\text{m}$$

3. **a.** The X-ray tube current

b. Power consumption of the tube $= I \times V$
$= 50 \times 10^{-3} \times 140 \times 10^3 = 7000\,\text{W}$
$= 7\,\text{kW}$. The efficiency of conversion of energy into X-ray photons is 0.7%, so the power of the X-ray beam is
$7000 \times 0.7/100 = 49\,\text{W}$.

4. **a.** As much as 99% of the energy of the incident electrons will be converted to internal energy in the target. The target needs to be able to conduct this heat away.

b. A large specific heat capacity limits the temperature rise.

c. So much heat is produced that a material of lower melting point could melt.

5. A scintillator is a material (often in crystal form) that promptly emits brief flashes of light when a charged particle or high-energy photon passes through it.

6. FPD detectors allow instant X-ray images to be obtained and require lower X-ray doses than photographic film. They can be stored and sent electronically and can be image-processed to reveal fine detail.

7. X-ray contrast can be enhanced by using a contrast medium of higher proton number than surrounding tissue and coating the tissue of interest with it.

8. Fluoroscopy is used for imaging moving tissues in the patient rather than making static images.

9. **a.** The linear attenuation coefficient $\mu = \frac{1}{x}\ln\left(\frac{I_0}{I}\right)$ is the fraction of a beam of X-rays that is absorbed or scattered per unit thickness of the absorber. It has the unit cm^{-1}. It depends on the density.

The mass attenuation coefficient is found by dividing μ by the density of the element or compound, and will produce a value that is constant for a particular element or compound. It has the unit $\text{cm}^2\,\text{g}^{-1}$.

b. $\mu = \mu_m \times \rho = 0.2\,\text{cm}^2\,\text{g}^{-1} \times 1.8\,\text{g}\,\text{cm}^{-3}$
$= 0.36\,\text{cm}^{-1}$

10. a. $0.5 = 1.0 \times e^{-\mu x}$ where $x = 1.5\,\text{mm}$

So $\mu = -\ln(0.5)/1.5 = 0.46\,\text{mm}^{-1}$

b. $I/I_0 = e^{-0.46 \times 2.5} = 0.32$

11. a. At the HVT the intensity drops to half, so
$0.5 = 1.0 \times e^{-\mu x}$

$\ln 0.5 = -\mu x$

$\mu = -\ln(0.5)/0.012 = 58\,\text{m}^{-1}$
or $0.58\,\text{cm}^{-1}$

b. $0.02 = 1.0 \times e^{-\mu x}$

$\ln 0.2 = -\mu x$

$x = -\ln(0.02)/58 = 0.067\,\text{m}$ or $6.7\,\text{cm}$

12. Because the computed image depends on the calculated attenuation. This would vary for different wavelength X-rays and so confuse the processing.

5 RADIONUCLIDE IMAGING AND THERAPY

1. A radioactive tracer is a radioactive element or compound added to a substance and injected into the body to monitor the distribution and accumulation of the substance in the body.

2. It emits gamma rays of an energy that makes them easily detected by a gamma camera. Its half-life of 6 hours is long enough for imaging procedures but short enough for the patient not to receive a significant dose for too long. It is easily obtained from its parent Mo-99 radionuclide. It can be used to label a wide range of compounds for different tracer procedures.

3. a. ${}^{99_m}_{43}\text{Tc} \rightarrow {}^{99}_{43}\text{Tc} + \gamma$

b. Because its half-life of 2.1×10^5 years means that its decay constant, and hence its activity level, are very low (see section 9.5 of Chapter 9).

4. Minimum activity of technetium = 5% of that of the molybdenum at start

$= 0.05 \times 10.0 \times 10^9 = 5.00 \times 10^8\,\text{Bq}$

Activity of the extracted technetium is 80% of the activity of the molybdenum at the extraction time, so the minimum allowable activity of molybdenum for extraction is

$(100/80) \times 5.00 \times 10^8 = 6.25 \times 10^8\,\text{Bq}$

Decay constant $\lambda = \dfrac{\ln 2}{T_{1/2}} = \dfrac{0.693}{66} = 0.0105\,\text{h}^{-1}$

Time t (in hours) to reach minimum activity is found as follows:

$$\text{activity } A = A_0 e^{-\lambda t}$$
$$6.25 \times 10^8 = 10.0 \times 10^9 \times e^{-0.0105t}$$
$$e^{0.0105t} = \frac{10.0 \times 10^9}{6.25 \times 10^8} = 16$$
$$0.0105t = \ln(16)$$
$$t = \frac{\ln(16)}{0.0105} = 264\,\text{h}$$

Extractions are taken every 36 hours; giving $264/36 = 7.3$

So seven extractions of technetium-99_m can be made.

5. a. Biological half-life is the time taken for the body to process and excrete half the amount of a substance introduced into it. Physical half-life is the time taken for a radionuclide to reduce to one-half of its initial activity.

b. Working in days,

$$T_E = \frac{T_P \times T_B}{T_P + T_B} = \frac{8.0 \times 21}{8.0 + 21} = \frac{168}{29.0} = 5.8\,\text{days}$$

c. While the physical half-life remains constant, the effective half-life also depends on the biological half-life, which can vary due to the state of health of the individual.

6. **a.** One week = $7 \times 24 = 168$ hours. After $(168 - 10) = 158$ hours the activity is 0.35 of its maximum after it was injected. For radioactive decay

$$A = A_0 e^{-\lambda t}$$
$$0.35 = 1.0 \times e^{-\lambda \times 158}$$

The *effective* decay constant is

$$\lambda = -\ln(0.35)/158 = 6.64 \times 10^{-3}\,h^{-1}$$

The *effective* half-life is then

$$\frac{\ln 2}{\lambda} = \frac{0.693}{6.64 \times 10^{-3}} = 105\,h \ (= 4.4\ days)$$

b. Physical half-life = 14.3 days = 343 h. So using the equation given in the text and rearranging

$$\frac{1}{T_B} = \frac{1}{T_E} - \frac{1}{T_P} = \frac{T_P - T_E}{T_P \times T_E}$$
$$T_B = \frac{T_P \times T_E}{T_P - T_E} = \frac{343 \times 105}{343 - 105}$$
$$= \frac{36015}{238} = 151\,h \ (= 6.3\ days)$$

7. **a.** Collimation is needed to ensure that the photons captured by the scintillator are travelling in more or less the same direction, resulting in a sharper image.

b. A gamma camera scan permits investigation of the functional and metabolic processes of the body.

c. The preparation and actual scan time for a gamma camera scan can take at least 30 minutes, or even longer if the tracer needs time to travel in the body and accumulate in the regions of interest. The half-life needs to be sufficiently long so that enough radiation can be detected during the scan and has not decayed away to levels that are hard to detect.

8. A simultaneous detection of gamma ray photons in two detectors places the source on a line between those detectors, because a pair of photons is simultaneously emitted from the source at 180° to one another. If other simultaneous detections can be made at different angles, then the position of the source can be determined (the intersection of those lines).

9. FDG is a radioactive version of the sugar glucose, which is metabolised by the brain and by cancer cells, meaning that it accumulates where there is pronounced brain activity or fast growth.

10. Low-energy X-rays are absorbed preferentially by bone, so bones would be damaged more than surrounding tissue. High-energy X-rays are not absorbed preferentially by bone or tissue.

11. **a.** This maximises the dose to cancerous tissue and minimises the dose to surrounding healthy tissue.

b. Healthy cells recover faster from damage than cancer cells, so delivering X-rays in fractions gives surrounding healthy tissue time to recover.

12. **a.** Higher doses of radiation can be delivered to smaller, more specific areas of the body, with less risk to surrounding healthy tissue.

b. A patient with an implant may have an external radiation hazard and may need to be kept isolated. Medical staff also need protect themselves from radiation exposure when treating the patient.

c. Beta emitters are strongly ionising and have a range of a few millimetres in tissue, meaning that they damage cells over a small area without affecting the surrounding healthy tissue.

GLOSSARY

Accommodation The ability of the eye to adjust its focal length.

Acoustic impedance (Z) The product of the density, ρ, and the speed of sound, c, in that material.

Action potential The change in electrical potential that occurs across a cell membrane during the passage of a nerve impulse.

Arcminute (arcmin) A unit of angular measurement equal to one-sixtieth (1/60) of one degree.

A-scan (amplitude scan) A range-measuring system, which records the time taken for an ultrasonic pulse to travel to an interface in the body and be reflected back.

Astigmatism A defect in the eye or in a lens caused by a deviation from spherical curvature, which results in distorted images.

Attenuation The reduction of intensity as a wave travels through a material.

Audiogram A graphic representation of the relation of vibration frequency and the minimum sound intensity for hearing.

Audiometer An apparatus used in audiometry for testing the hearing.

Auditory nerve Carries auditory sensory information from the cochlea of the inner ear directly to the brain.

Auditory tube The tube that runs from the middle ear to the pharynx, also known as the Eustachian tube.

Beam hardening The process of increasing the average energy level of an X-ray beam by filtering out the low-energy photons.

Biological half-life (T_B) The time required for the body to eliminate half of an administered dose of any substance by regular physiological processes.

Brachytherapy The treatment of cancer, especially prostate cancer, by the insertion of radioactive implants directly into the tissue.

Bremsstrahlung ('braking' radiation) Electromagnetic radiation produced by the acceleration or especially the deceleration of a charged particle after passing through the electric and magnetic fields of a nucleus.

B-scan (brightness scan) A diagnostic test used in optometry and ophthalmology to produce a two-dimensional, cross-sectional view of the eye and the orbit.

Bulk elastic modulus The tendency of an object to deform in all directions when uniformly loaded in all directions.

Characteristic X-rays Emitted when outer-shell electrons fill a vacancy in the inner shell of an atom, releasing X-rays in a pattern that is 'characteristic' to each element.

Ciliary muscle One of the muscles around the lens of the eye that can change the shape of the lens in order to produce a clear image.

Cladding The outer covering of an optical fibre, which confines light to the core.

Cochlea Part of the inner ear where sound vibrations are converted into electrical signals.

Coherent bundle An ordered collection of optical fibres that can carry images.

Computed radiography (CR) The production of a digital image by using a phosphor imaging plate (IP) in place of conventional film.

Computed tomography (CT) Radiography in which a three-dimensional image of a body structure is constructed by computer from a series of plane cross-sectional images made along an axis.

Concave lens A lens that spreads a parallel beam into a divergent emergent beam.

Cones Light-sensitive cells in the retina, responsible for colour vision.

Contrast In X-ray image: the variation in brightness of the image due to differential absorption of X-rays in the patient.

Contrast medium A material of high density and atomic number put into the body to enable soft organs to be visualised on an X-ray image.

Convex lens A lens that causes a parallel beam to converge to a point called the focus.

Cornea The outer covering of the front of the eye, where most of the refraction of light takes place.

CT scanner Medical scanner that uses beams of X-rays to produce high-contrast images of a cross-section through the body.

dBA dBA-weighted decibels are a measurement of the relative loudness of sounds in air as perceived by the human ear. In the A-weighted system, the decibel values of sounds at low frequencies are weighted in such a way that they are reduced compared with unweighted decibels (dB). This weighting is applied because the human ear is less sensitive at low and high audio frequencies and the A weighting matches more closely the human ears frequency response to sound levels.

dB scale The logarithmic scale used to measure intensity of sound.

Decibel (dB) The unit used to measure the intensity of a sound.

Depolarisation A sudden change within a cell, during which the cell undergoes a dramatic electrical change.

Detector unsharpness The loss of detail caused by the resolving power of the detector.

Dioptre (D) Unit used to measure the refractive power of a lens or surface (equivalent to m^{-1}).

Dose A measure of the extent to which matter has been exposed to ionising radiation.

Dynode An intermediate electrode which emits additional electrons in a photomultiplier or similar amplifying device.

Eardrum The membrane of the

middle ear, which vibrates in response to sound waves.

Effective half-life (T_E) The rate of accumulation or elimination of a biochemical or pharmacological substance in an organism.

Electrocardiograph (ECG) An instrument used in the detection and diagnosis of heart abnormalities that measures the electrical activity of the heart.

Electron multiplication The process that occurs in a radiation detector, such as a photomultiplier tube, where secondary electrons are produced by a single photoelectron creating a large number of electrons at the detector output producing a signal gain.

Eluted Removed (an adsorbed substance) by washing with a solvent.

Endoscope An instrument that can be introduced into the body to give a view of its internal parts.

Excited Being at an energy level higher than the ground state.

External beam radiotherapy (EBRT) Radiotherapy in which radiation is directed at a tumour from outside the body.

Far point The farthest point from the eye at which images are clear.

Flat-panel (FPD) detectors A class of solid-state X-ray digital radiography devices similar in principle to the image sensors used in digital photography and video.

Fluoroscopy An X-ray technique used to produce real-time images.

Focal length (f) The distance between the principal focus of a lens and its centre.

Focus The point where light rays originating from a point on the object converge.

Fovea A region on the retina with a high concentration of cone cells.

Fractions The radiation dose used to destroy tumours and other cancerous cells in the body is administered in portions or fractions of the total prescribed dose over a period of time, rather than all at once. This enables healthy cells that may have been damaged by the radiation to recover between radiotherapy sessions and is more beneficial to the patient.

Gamma camera A device used to image gamma radiation emitting radioisotopes.

Geometric unsharpness The loss of definition that is the result of geometric factors of the radiographic equipment and setup.

Gradient coils Loops of wire or thin conductive sheets inside an MR scanner, which create a secondary magnetic field when current is passed through them.

Half-value thickness (HVT) The thickness of the material at which the intensity of radiation entering it is reduced by one-half.

Hypermetropia Long-sightedness: the inability to focus on objects close to the eye.

Image intensifier A device used in the X-ray technique fluoroscopy to produce a visible image using a lower dose of X-rays.

Incoherent bundle A bundle of filaments of optical glass or other transparent materials that transmit only light, not optical images.

Indium-111 (In-111 or $^{111}_{49}$In) An isotope of indium with a radioactive half-life of 2.80 days, making it useful as a radioactive tracer.

Intensifying screen Absorbs the energy in the X-ray beam that has penetrated the patient and converts this energy into a light pattern.

Intensity The power of a wave transmitted through unit area perpendicular to the direction of travel of the wave; unit: watts per square metre, $\mathrm{W m^{-1}}$

Intensity attenuation coefficient (μ) The fraction of a beam of X-rays or gamma rays that is absorbed or scattered per unit thickness of the absorbing material.

Intensity level The relative sound intensity at any point in a sound field as compared with a specified standard intensity.

Iodine-131 (I-131 or $^{131}_{53}$I) A heavy radioactive isotope of iodine that has the mass number 131 and a half-life of eight days; used in the diagnosis of thyroid disease and the treatment of goitre.

Iris A thin, circular structure in the eye, responsible for controlling the diameter and size of the pupil and thus the amount of light reaching the retina.

Isotope One of two or more atoms of the same element that have the same number of protons in their nucleus but different numbers of neutrons.

Larmor frequency The rate of precession of the magnetic moment of the proton around the external magnetic field.

Lens A transparent, biconvex structure in the eye that, along with the cornea, helps to refract light to be focused on the retina.

Lens formula Equation used to calculate the image distance for either real or virtual images and for either positive or negative lenses.

Linear attenuation coefficient (μ) A measure of a material's ability to attenuate X-rays.

Magnetic resonance (MR) The excitation of particles (as atomic nuclei or electrons) in a magnetic field by exposure to electromagnetic radiation of a specific frequency.

Magnification The ratio of image size to object size; for a lens it is equal to the ratio of image distance v to object distance u.

Mass attenuation coefficient (μ_m) a measure of a material's ability to attenuate X-rays that is independent of the material's density; equal to the linear attenuation coefficient divided by the density

Metastable An excited state of a nucleus with a relatively long half-life.

Molybdenum-technetium generator A device used to extract the metastable isotope ^{99m}Tc of technetium from a source of decaying molybdenum-99.

Motion unsharpness Caused by movement of the patient, the detector or the source of X-rays, during the exposure.

Myopia Short-sightedness: the inability to focus on objects at large distances from the eye.

Near point The closest distance from the eye at which an object can be brought into focus.

Nuclear magnetic resonance (NMR) The absorption of electromagnetic radiation at a

precise frequency by a nucleus in an external magnetic field.

Optic nerve A paired nerve that transmits visual information from the retina to the brain.

Optical fibres Thin transparent fibres of glass or plastic that transmit light throughout their length by internal reflections; used in medical instruments to view otherwise inaccessible parts of the body.

Ossicles Three bones in the middle ear, malleus, incus and stapes, that are among the smallest bones in the human body.

Phon A unit of loudness level for pure tones.

Photocathode A negatively charged electrode in a light-detection device such as a photomultiplier or phototube that is coated with a photosensitive compound.

Photodetector A device that detects or responds to incident light by using the electrical effect of individual photons.

Photodiode A semiconductor diode that, when exposed to light, generates a potential difference or changes its electrical resistance.

Photomultiplier An instrument containing a photoelectric cell and a series of electrodes, used to detect and amplify the light from very faint sources.

Photopigment A pigment whose chemical state depends on its degree of illumination, such as those in the retina of the eye.

Physical half-life (T_P) The period of time required to reduce the radioactivity level of a source to exactly one-half its original value due solely to radioactive decay.

Piezoelectric effect The ability of certain materials to generate an electric charge in response to applied mechanical stress.

Piezoelectric transducer A device that generates a potential difference when it is compressed or expanded; used to generate and detect ultrasound.

Polarised A transverse wave is polarised when the vibration of the wave is confined to one direction.

Power The rate at which energy is transferred or at which work is done, measured in joules per second, or watts, W.

Precess To change the orientation of the rotational axis of a rotating body.

Principal axis An imaginary line drawn at right angles to a lens passing through the optical centre, used in constructing ray diagrams.

Principal focus A particular point on the optical axis of a lens where light is focused.

Pulse rate The number of times your heart beats per minute.

Pupil The aperture of the eye.

Radioactive tracer A radioactive isotope used for medical imaging and which is tracked and detected by an external radiation detector.

Radioisotope A form of a nucleus (see isotope) that is radioactive.

Radionuclide An atom that has excess nuclear energy, making it unstable.

Real image An image formed by the convergence of rays of light; it can be formed on a screen.

Reflection coefficient A parameter that describes how much of an electromagnetic wave is reflected by an impedance discontinuity in the transmission medium.

Refractive index A measure of the refraction as light passes from one medium to another; given by the ratio of the speed of light in medium 1 to the speed of light in medium 2; also equal to the ratio of the sine of the angle of incidence to the sine of the angle of refraction.

Repolarised Describes the change in the cell membrane potential that returns it to a negative value just after the depolarisation phase of an action potential has changed the membrane potential to a positive value.

Resonance Occurs when a system accepts energy from a driving source at its natural frequency – the amplitude increases greatly.

Resting potential The electrical potential of a neuron or other excitable cell relative to its surroundings when not stimulated or involved in passage of an impulse.

Retina A tissue of light-sensitive cells lining the back of the eye.

Rhodopsin A photopigment present in rod cells, which disassociates when it absorbs a photon of light.

Rods Light-sensitive cells in the retina, which respond well in low light conditions.

Scintillator A material that fluoresces when struck by a charged particle or high-energy photon.

Secondary emission A phenomenon where primary incident particles of sufficient energy, when hitting a surface or passing through some material, induce the emission of secondary particles.

Source focal spot The small area of the target of an X-ray tube on which the electron beam is focused; it is an important factor in the ability of an X-ray machine to produce a sharp image.

Spectral response The variable output of a light-sensitive device that is based on the colour of the light it perceives.

Technetium– 99_m (Tc-99_m or $^{99m}_{43}Tc$) A metastable nuclear isomer of technetium-99 (itself an isotope of technetium) that is used in medical diagnostic procedures.

Thermionic emission The process by which electrons are released from the surface of a heated metal.

Threshold of hearing (I_0) The lowest sound intensity that can be detected by a healthy human ear, defined as 1×10^{-12} Wm^{-2}.

Ultrasound Sound waves with a frequency above 20 kHz.

Unsharpness The lack of quality of an X-ray image (blurring), due to a region of partial shadow.

Virtual image An image caused by rays that do not converge; the image can be seen by the eye but not formed on a screen.

Visual acuity The ability of the eye to resolve separate images.

X-ray opaque A bodily material that does not permit the passage of X-rays through it and appears light in colour on an X-ray detector.

INDEX

ACKNOWLEDGEMENTS

The publishers wish to thank the following for permission to reproduce photographs. Every effort has been made to trace copyright holders and to obtain their permission for the use of copyright materials. The publishers will gladly receive any information enabling them to rectify any error or omission at the first opportunity.

Chapter 1

p1, left: Equinox Graphics/Science Photo Library; p1, right: Phillippe Psaila/Science Photo Library; p1: itsmejust/Shutterstock; p18, Fig 31: Maria Platt-Evans/Science Photo Library; p19, Fig 32: Astier/BSIP/Science Photo Library; p21, Fig A1: John Bavosi/Science Photo Library

Chapter 2

p24, Fig 1a: Science Photo Library; p24, Fig 1b: Dr P. Marazzi/Science Photo Library

Chapter 3

p33, Fig 7: Saturn Stills/Science Photo Library; p33, Fig 8: Public Health England/Science Photo Library; p36, Fig 11: Monty Rakusen/Getty Images; p36, Fig 12: Centre Jean Perrin - ISM/Science Photo Library; p38, Fig A1: Zephyr/Science Photo Library

Chapter 4

p43, Fig 7, canonzoom/Shutterstock; p43, Fig 8: OZMedia/Shutterstock; p44, Fig 10: Pavel L Photo and Video/Shutterstock; p46, Fig 13: thailoei92/Shutterstock; p49, Fig 16: Chaikom/Shutterstock

Chapter 5

p52, Fig 2: Eye of Science/Science Photo Library; p57, Fig A1, left: Centre Jean Perrin, ISM/Science Photo Library; p57, Fig A1, right: Centre Jean Perrin, ISM/Science Photo Library; p59, Fig 10a: NINCDS/NIH/Jessica Wilson/Science Photo Library; p59, Fig 10b: CEA - ORSAY/CNRI/Science Photo Library; p61, Fig 12: Mark Kostich/Getty Images; p62, Fig A1: Amelie-Benoist/BSIP/Science Photo Library; p63, Fig 13: Gary Friedman/Getty Images