The functioning of the ear

This chapter introduces the basic ideas behind the functioning of the human ear and the related physics questions on sound perception, intensity, frequency response and frequency discrimination. Hearing defects are also briefly discussed.

Objectives

By the end of this chapter you should be able to:

- describe the basic components of the *human ear*;
- define *sound intensity* and the *sound intensity scale* based on the decibel;
- perform calculations with intensity and the *decibel scale*;
- understand how the ear functions;
- describe how the ear separates sound according to frequency in the *cochlea*;
- state the meaning of the terms *threshold of hearing* and *audiogram*;
- understand basic *hearing defects* and say how they might be corrected.

The ear

The human ear is an exceptionally efficient and sensitive instrument that detects and analyses sound, and converts the mechanical energy carried by a sound wave into electrical energy that is fed into the brain. The ear is sensitive to sounds of frequency ranging from 20 Hz to 20 kHz. At 1000 Hz, the ear responds to sound that displaces the eardrum by only one-tenth of the diameter of the hydrogen atom! The ear can be divided into three main parts: the outer, middle and inner ear (see Figure I1.1).

Sound waves reaching the ear are fed into the *auditory canal* and fall on the *eardrum*, a membrane that begins to vibrate as a result. The eardrum forms the entrance to the middle ear, an air cavity of 2 cm^3 in volume that contains the *ossicles* (three small bones), the *malleus* (hammer), *incus* (anvil) and *stapes* (stirrup). This cavity is connected to the throat by the *Eustachian tube*, which is normally closed but can be opened by swallowing or yawning to equalize the pressure on each side of the eardrum. The ossicles act as a lever system, which transmits the energy falling on the eardrum onto the *oval window*, an opening marking the beginning of the inner ear. The tension in the muscles attached to the ossicles increases in the presence of a very loud sound, thus limiting the motion of the stapes and hence the energy transferred to the oval window and protecting the delicate inner ear from damage. This is known as the *acoustic reflex*; this takes about 10 ms to become effective, so it offers no protection in the case of sudden very loud sounds, such as gunfire.

The purpose of the lever system is to amplify the amplitude of the incoming sound wave. As seen in the example that follows, the lever system amplifies by a factor of 1.5. However, as a result of the difference in the areas of the oval window and the eardrum, the

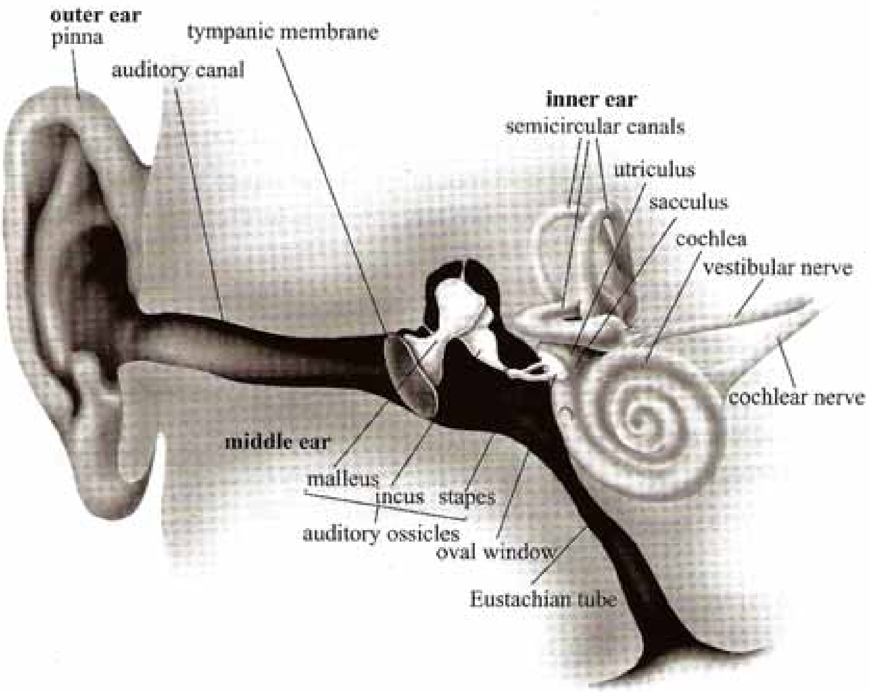

Figure I1.1 A diagram of the ear showing its division into outer, middle and inner ear. The cochlea is responsible for the resolution of the incoming sound into its various components. © 1995 Dorling Kindersley Multimedia.

amplification is increased further by a factor of about 13, resulting in a total amplification of about 20.

The inner ear has three main parts: the *vestibule* (the entrance cavity), the *semicircular canals* and the *cochlea*. The vestibule is surrounded by bone except for the opening to the middle ear (the oval window); this opening is sealed by the stapes and the round window, which is covered by a membrane. This cavity is filled with a liquid. The semicircular canals play no role in hearing – their function is to provide us with a sense of balance. The cochlea is a tube, coiled like a snail (making 2.5 turns) and has a length of about 3.5 cm.

The sound wave entering the oval window travels in the liquid-filled canals of the cochlea. First in the *vestibular*, then through the *helicotrema* and into the *tympanic canal*. It ends at the round window membrane, which acts as a pressure release point (a total length of about 2 cm). Between these two canals is a third, the *scala media* or cochlean duct. The *basilar membrane* separates the cochlean duct from the tympanic canal and this membrane contains a large number of nerve endings, whose purpose is to transmit the electrical signals to the brain. These signals are generated when vibrations in the basilar membrane are fed into the *organ of Corti*, which is attached to the membrane. Different parts of the basilar membrane are sensitive to different frequency ranges of the incoming sound wave.

The three parts of the ear are illustrated schematically in Figure I1.2, and Figure I1.3 is a schematic cross-section through the cochlea.

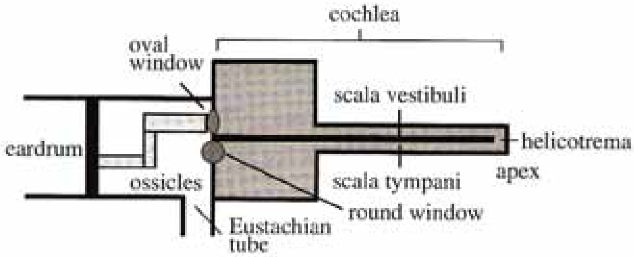

Figure I1.2 A schematic diagram showing the three parts of the ear. Here the cochlea is shown as a straight uncoiled tube. The cochlea is separated into three canals or scalae. The top, scala vestibuli, and the bottom, scala tympani, which communicate through an opening at the apex of the cochlea called the helicotrema. The scala vestibuli starts at the oval window while the scala tympani starts at the round window. The middle canal is known as scala media or cochlean duct.

The auditory pathways from the cochlea into the brain are highly complex and not yet fully understood. There is, however, a considerable amount of processing of the information along the pathway in processing centres, with

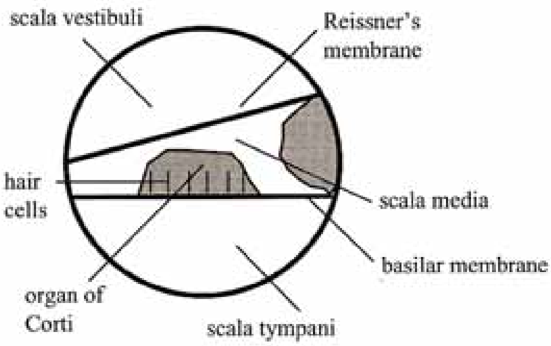

Figure I1.3 A schematic cross-section through the cochlea. The scala media is separated from the scala vestibuli by Reissner's membrane and from the scala tympani by the basilar membrane.

connections between the left ear and the right ear's main pathway and vice versa.

Mismatch of impedances

When a wave enters a new medium, only part of it will be transmitted into the new medium; part will be reflected back into the old medium. The fraction of the transmitted intensity depends on the *impedances* of the two media. The least amount of reflection takes place when the impedances of the two media are as close to each other as possible. For exactly equal impedances no reflection takes place at all.

▶ For sound transmission, the acoustic impedance of a medium is defined as the product of the speed of sound in the medium times the medium's density

$$Z = \rho c$$

The region up to the oval window is filled with air and its impedance is thus about 450 kg m^{-2} s^{-1}. The region behind the oval window is filled with the cochlear liquid and its impedance is about 1.5×10^6 kg m^{-2} s^{-1}. This large mismatch of impedances means that it is necessary to amplify the sound wave arriving at the oval window so that a substantial fraction of it can be transmitted.

Example question

Q1

The area of the eardrum is $A = 43\ \text{mm}^2$ and that of the oval window $a = 3.2\ \text{mm}^2$. Using Figure I1.4 as a model for the action of the ossicles, find the pressure amplification at the oval window.

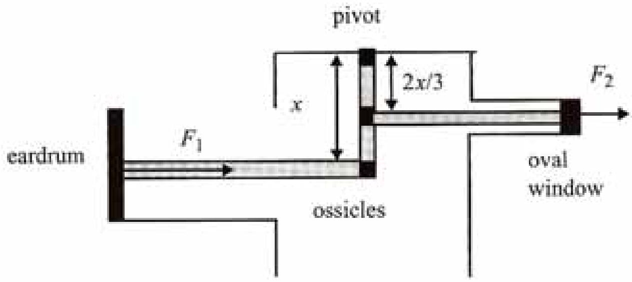

Figure I1.4.

Answer

Taking torques about an axis through the pivot we find

$$F_1 x = F_2 \frac{2x}{3}$$

$$\Rightarrow F_2 = \frac{3}{2} F_1$$

therefore, the pressure on the oval window is

$$\begin{aligned} P_2 &= \frac{F_2}{a} \\ &= \frac{3}{2}\frac{F_1}{a} \\ &= \frac{3}{2}\frac{P_1 A}{a} \\ &= \frac{3}{2}\frac{A}{a} P_1 \\ &= 20.2 P_1 \end{aligned}$$

that is, an amplification of about 20.

Complex sounds

Few sounds are as simple and pure as the single-frequency harmonic waves we discussed in Topic 4, Oscillations and waves. When a human voice is fed into an oscilloscope through a microphone, the trace that appears on the oscilloscope screen will not be a sine wave.

▶ Harmonic waves (i.e. waves of one specific frequency) are important, however, because of a powerful mathematical technique, called Fourier analysis, that is used to analyse complex sounds. It can be shown that any complex sound is a superposition of many (perhaps infinite) simple sine or cosine waves.

To illustrate this point, consider an extreme example, the graph of the function $y = t^2$ from $-\pi$ to $+\pi$ (see Figure I1.5).

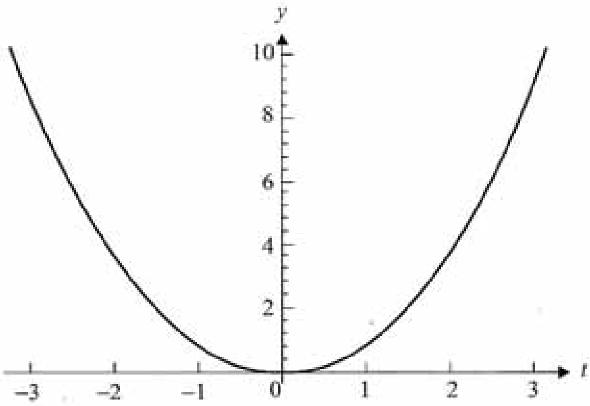

Figure I1.5 The graph of the function $y = t^2$.

This function can be well approximated by a sum of harmonic functions called a *Fourier series*:

$$\frac{\pi^2}{3} + 4\sum_{n=1}(-1)^n\frac{\cos(nt)}{n^2}$$

Figure I1.6 is the graph of this function for only the first five terms in the sum (*Fourier components*).

The approximation is quite good. It can be made even better by keeping more terms in the sum. The point here is that any periodic function can be written as a sum of harmonic functions. Thus, complex sounds entering the

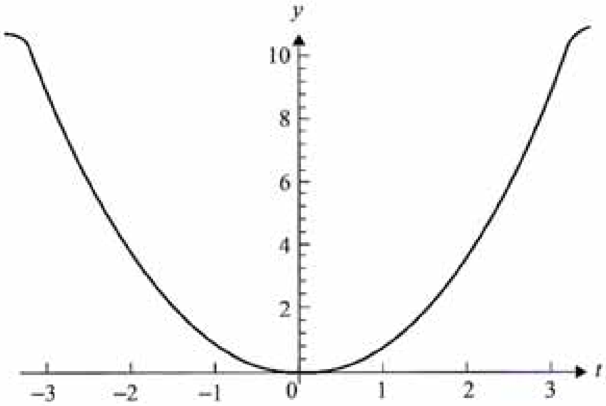

Figure I1.6 The approximation to $y = t^2$ obtained using the first five Fourier components (i.e harmonic waves).

ear can be decomposed into the component frequencies of the harmonic functions making up the complex sound. This happens in the cochlea and is explained in a later section.

Intensity of sound

Consider a source of sound S. The power of the source is the energy per second emitted by the source.

The energy emitted by the source can be thought of as being spread uniformly on the surface of an imaginary sphere of radius r centred on the source (see Figure I1.7). Thus, at a distance r from the source the energy received per second by an instrument of area A is

$$P\frac{A}{4\pi r^2}$$

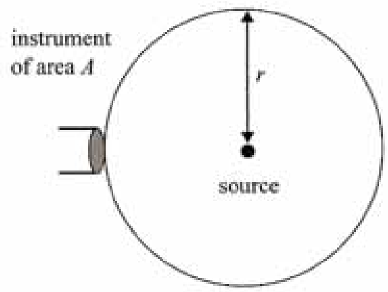

Figure I1.7 The energy emitted by the source goes through an imaginary sphere around the source.

▶ The energy per second per unit instrument area, that is

$$I = \frac{P}{4\pi r^2}$$

is known as the intensity of the sound wave at a distance r from the point source. The unit of intensity is thus $W\,m^{-2}$. The lowest intensity the human ear can perceive is called the threshold of hearing and is $I_0 = 1.0 \times 10^{-12}\,W\,m^{-2}$. The intensity is proportional to the square of the amplitude of the wave.

It is known that the 'sensation of hearing' (let us call this quantity β) does not increase linearly with intensity I. Rather, the ear has a logarithmic response to intensity: the increase in the hearing sensation is proportional to the *fractional* increase in intensity (the Weber–Fechner law). Mathematically this means that

$$\Delta\beta \propto \frac{\Delta I}{I}$$

We can exploit this law to define a scale of sound intensity level as follows. On this scale (the decibel scale) an increase of 10 units implies an increase of intensity by a *factor* of 10. The zero on this scale corresponds to the threshold of hearing. This implies that

$$\beta(\text{in decibels}) = 10\log\left(\frac{I}{I_0}\right)$$

Table I1.1 shows some typical intensity levels of common sounds along with their pressure amplitudes.

Source of sound	Intensity level/dB	Pressure amplitude/Pa
Threshold of hearing	0	0.00002
Falling leaves	10	0.00006
Radio studio	20	0.0002
Whisper	30	0.0006
Library	40	0.002
Office space	50	0.006
Conversation	60	0.02
Room with TV on	70	0.06
Noisy street	80	0.2
Rock concert	110	6
Pain setting in	120	20
Pneumatic hammer	130	60
Jet plane 3 m away	140	200

Table I1.1 Typical decibel values and the corresponding amplitude of the sound wave in pascal.

Example questions

Q2

The intensity of a sound increases from $10^{-10}\,W\,m^{-2}$ to $10^{-8}\,W\,m^{-2}$. By how much does the sound intensity level change?

Answer

The original level was

$$\beta_1 = 10\log\left(\frac{10^{-10}}{10^{-12}}\right)$$
$$= 10\log 10^2$$
$$= 20\,dB$$

The new sound intensity level is

$$\beta_2 = 10\log\left(\frac{10^{-8}}{10^{-12}}\right)$$
$$= 10\log 10^4$$
$$= 40\,dB$$

The increase is thus 20 dB.

Note: Since the increase in the intensity level is proportional to the fractional increase in intensity, we could have written directly

$$\Delta\beta = 10\log\left(\frac{10^{-8}}{10^{-10}}\right) = 10\log 10^2 = 20\,dB$$

Q3

One sound has intensity $2 \times 10^{-6}\,W\,m^{-2}$ and another has intensity $4 \times 10^{-8}\,W\,m^{-2}$. What is the difference of the sound intensity levels?

Answer

The first sound has intensity level

$$\beta_1 = 10\log\left(\frac{2 \times 10^{-6}}{10^{-12}}\right)$$
$$= 10\log(2 \times 10^6)$$
$$= 63\,dB$$

The second sound has intensity level

$$\beta_2 = 10 \log\left(\frac{4 \times 10^{-8}}{10^{-12}}\right)$$
$$= 10 \log (4 \times 10^4)$$
$$= 46 \text{ dB}$$

Their difference is thus 17 dB.

Q4

The sound intensity level in a room is 70.0 dB. A radio produces sound of intensity level 72.0 dB. What is the sound intensity level in the room now?

Answer

The intensity in the room is found from

$$70 \text{ dB} = 10 \log\left(\frac{I}{10^{-12}}\right)$$
$$\Rightarrow \frac{I}{10^{-12}} = 10^7$$
$$\Rightarrow I = 10^{-5} \text{ W m}^{-2}$$

Similarly, the intensity due to the radio is

$$72 \text{ dB} = 10 \log\left(\frac{I}{10^{-12}}\right)$$
$$\Rightarrow \frac{I}{10^{-12}} = 10^{7.2}$$
$$\Rightarrow I = 10^{-4.8} \text{ W m}^{-2}$$

The combined intensity is

$$I_{\text{total}} = (10^{-5} + 10^{-4.8}) \text{ W m}^{-2}$$

Thus, the new sound intensity level is

$$10 \log\left(\frac{10^{-5} + 10^{-4.8}}{10^{-12}}\right) = 74.1 \text{ dB}$$

Note: It is important to realize that it is intensities that must be added not the sound intensity levels in decibels.

Q5

A shouting voice has a power output of about 10^{-3} W. At what distance from the source is the sound intensity level 80 dB?

Answer

Using

$$\beta = 10 \log\left(\frac{I}{10^{-12}}\right)$$
$$= 80$$
$$\Rightarrow \log\left(\frac{I}{10^{-12}}\right) = 8$$
$$\Rightarrow I = 10^{-4} \text{ W m}^{-2}$$

Now using the definition of intensity

$$I = \frac{P}{4\pi r^2}$$

it follows that

$$10^{-4} = \frac{10^{-3}}{4\pi r^2}$$
$$\Rightarrow r = \sqrt{\frac{10}{4\pi}}$$
$$= 0.89 \text{ m}$$

Q6

Assume that in a football stadium 40 000 fans cheer their teams, each producing a power output of 10^{-3} W. If the average distance of the fans from the centre of the stadium is 150 m, find the sound intensity level there.

Answer

The combined power output is $40\,000 \times 10^{-3}$ W = 40 W. The intensity at the centre of the stadium is thus

$$I = \frac{P}{4\pi r^2}$$
$$= \frac{40}{4\pi (150)^2}$$
$$= 1.41 \times 10^{-4} \text{ W m}^{-2}$$

The sound intensity level is

$$\beta = 10 \log\left(\frac{1.41 \times 10^{-4}}{10^{-12}}\right)$$
$$= 81.5 \text{ dB}$$

Frequency response and loudness

The human ear has a threshold when it comes to the frequency of the sound wave. The lowest frequency that can be heard is about 20 Hz and the largest about 20 kHz. With age, the upper frequency threshold is reduced.

▶ The ear is not equally sensitive at all frequencies. This means that the ear responds differently to a sound of given intensity depending on the frequency of the sound.

Thus, the statement made earlier that the normal human ear has a threshold of hearing of 1.0×10^{-12} W m^{-2} is correct provided the sound has a frequency of 1000 Hz. Sounds of larger or smaller intensity than this may be heard or are barely audible depending on the frequency of the sound. The graph in Figure I1.8 shows the intensity level of barely audible sounds (the threshold of hearing) as a function of frequency. Thus, a sound at 100 Hz must have an intensity level of 35 dB to be barely audible, and a sound at 20 Hz must have an intensity level of 72 dB. All points on this curve are perceived by the ear to have the same loudness, even though they have different intensities. These points define the 'zero phon loudness curve'. The 'N phon loudness curve' consists of those sounds that the ear perceives to be equally loud as a sound of N dB at 1000 Hz.

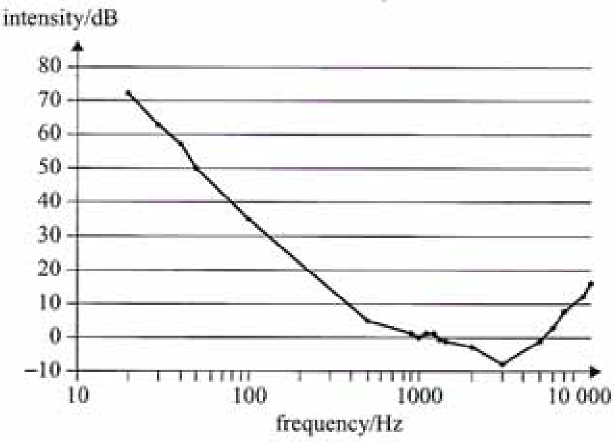

Figure I1.8 The threshold of hearing curve as a function of frequency. A sound of frequency 3 kHz is just audible even at −8 dB.

▶ The ear is most sensitive for frequencies around 3 kHz and least sensitive for frequencies less than (about) 50 Hz and higher than (about) 10 kHz.

The sensitivity of the human ear for frequencies around 3 kHz can be understood in terms of resonance in the ear canal. The canal can be thought of as a tube with one open and one closed end, so standing waves in this tube have a wavelength in the fundamental mode equal to $4L$, where L is the length of the canal – about 2.8 cm (see Figure I1.9).

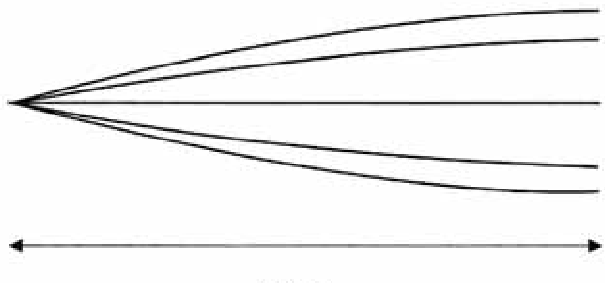

Figure I1.9 Schematic diagram of standing waves in the ear canal.

This means that the frequency corresponding to this wavelength is

$$f = \frac{340}{0.112} = 3036 \text{ Hz}$$

which is in good agreement with the most sensitive frequency.

Pitch

'Pitch' is a subjective characteristic of sound. It measures how high or low a sound is. It is determined *primarily by frequency* (which is why pitch is often taken to mean the same thing as frequency) *but also by intensity*. Imagine a pure tone of frequency 100 Hz that is sounded first at low intensity and then at high intensity. The louder sound will 'feel' lower in pitch than the soft sound.

Frequency separation in the cochlea

As mentioned earlier, a complex sound can, in general, be decomposed into its component harmonic waves.

▶ The ear analyses the different frequencies of a sound wave reaching it in the cochlea.

A wave entering the oval window travels along the basilar membrane, which separates the tympanic canal from the cochlean duct. The basilar membrane decreases in stiffness along its length (of about 35 mm). The velocity of the sound wave is thus high at the beginning of the membrane and drops along its length. In general, the response of a given point on the basilar membrane is small unless that part of the membrane is in resonance with the sound wave.

▶ From the pioneering work of Georg von Bekesy in the 1950s, we have learned that the beginning of the basilar membrane is in resonance with high frequencies and the end with low frequencies. The brain thus perceives frequency by locating the part of the basilar membrane that is vibrating.

This is done through the organ of Corti, which lies on the basilar membrane, and its hair cell receptors, which feed the information into the nerves that end at the base of the receptor cells. A wave of frequency 8 kHz will be in resonance at a point at about 2.5 mm from the beginning of the basilar membrane, a frequency of 1 kHz will be in resonance at a distance of about 22 mm and a frequency of 100 Hz will be in resonance a full 32 mm along the basilar membrane (see Figure I1.10).

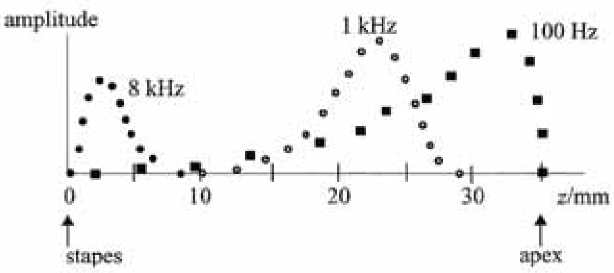

Figure I1.10 The amplitude of vibration of the basilar membrane for three different frequencies as a function of distance from the stapes. The point of largest amplitude shifts toward the apex of the basilar membrane as the frequency decreases.

Hearing defects

Two hearing losses may be distinguished.

- *Sensory nerve deafness,* which involves damage to the hair cells and neural pathways (e.g. due to tumours of the acoustic nerve or meningitis).
- *Conduction deafness,* in which damage to the middle ear prevents the transmission of sound into the cochlea. (This may be due to the plugging of the auditory canal by foreign bodies, such as wax, thickening of the ear drum because of repeated infections, destruction of the ossicles, or too rigid an attachment of the stapes to the oval window.)

Hearing can be monitored with an audiogram: the patient is supplied with very faint sounds of a specific frequency through earphones and their intensity is increased until they are just audible to the patient. A typical example is the audiogram of Figure I1.11 (consider first the data points represented by circles). At a frequency of 1000 Hz a sound must increase in intensity by 45 dB and at 4000 Hz by 70 dB if these sounds are to be audible to the patient. This defines what is called a *hearing loss* in dB. At 1000 Hz, therefore, the intensity of sound must be increased (for example, by a hearing aid) by a factor of

$$\begin{aligned}\Delta\beta &= 45\ \text{dB}\\ &= 10\log\frac{I}{I_0}\\ \Rightarrow I &= I_0 \times 10^{4.5}\end{aligned}$$

where I_0 is the intensity prior to amplification. This audiogram shows a substantial loss of hearing, especially in the higher frequencies. The damage was possibly caused by excessive exposure to loud sounds over very long periods of time. Ageing, which also results in hearing loss, would show in a more gently varying curve in the audiogram and with smaller loss in decibels.

Sound can reach the cochlea through the bones of the head and thus the audiogram is performed not only with earphones but also

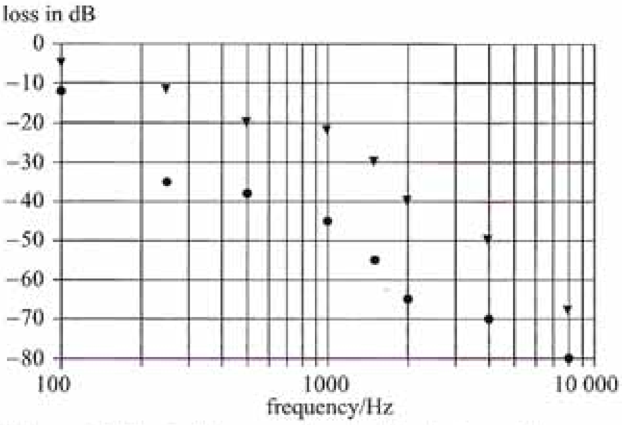

Figure I1.11 Audiogram for air (circles) and bone (triangles) conduction.

with an oscillator (placed at the bottom of the skull) that transmits the sound into the bones. The data points represented by triangles in the audiogram in Figure I1.11 are the results of bone transmission.

There is a *gap* between the two sets of data points of about 25 dB (at 1000 Hz). This is most likely an indication of a *conduction* hearing problem, the origin of which is in the middle or outer ear. Audiograms in which the data points for air and bone conduction *almost coincide* indicate *cochlea/nerve* problems in the inner ear.

In conductive hearing loss (where the sound *does* reach the inner ear) use of a hearing aid that amplifies the sound may help a patient. A hearing aid responds to and amplifies sound within a limited range of frequencies (the frequency range of human speech) but does not work well outside this range. For sensory nerve hearing loss, where the damage is to the hair cells, a *cochlear implant* may be useful. This is a device that consists of a microphone to pick up the sound, a signal processor to convert the sound into an electrical signal and a transmission system to transmit the electrical signal to electrodes. The electrodes are then surgically implanted in the cochlea. Different electrodes are stimulated according to the frequency of the sound so, in effect, the cochlear implant mimics the function of the cochlea. Assuming that enough healthy nerves are left near the electrodes, stimulation of the electrodes induces stimulation of the neighbouring nerves and the signal can be carried to the brain.

Speech recognition

All the information in speech is contained within the frequency interval from 200 Hz to 6 kHz. Noise and loss of frequency discrimination affect speech intelligibility. Generally, the meaning of a sentence can still be extracted from context but the intelligibility of isolated words is more strongly affected. The shorter the word, the bigger the loss of intelligibility. It has been found that, in cases of hearing loss at high frequencies (>3000 Hz), amplifying the sound does not help the patient to identify spoken syllables. Similarly, the inability of the cochlea to correctly identify frequencies also leads to errors in the identification of syllables.

Questions

1 A point sound source emitting uniformly in all directions is observed to have a sound intensity level of 70 dB at a distance of 5 m. What is the power of the source?

2 The sound level intensity of a screaming child in a room is 75 dB. What is the sound level intensity when three screaming children are put together in the same room?

3 The sound intensity level a certain distance from a source is 68 dB. If the distance to the source is halved, what is the new sound intensity level?

4 The sound level intensity of a given sound wave is 15 dB higher than that of another sound wave. What is the ratio of the intensities of the two waves?

5 If a radio creates sound of intensity level 70 dB, how many radios are required to create sound of intensity level 80 dB?

6 The audiogram of Figure I1.11 shows that the loss in dB for a person at 4000 Hz is 70 dB. What is the least intensity of sound this person

can hear at 4000 Hz? (Use the diagram of Figure I1.8 to find the threshold of hearing at 4000 Hz.)

7 Figure I1.12 shows the threshold of hearing curve for a patient. Explain what this means. What is the frequency at which this patient is most sensitive? What is the intensity of sound this patient receives at 200 Hz?

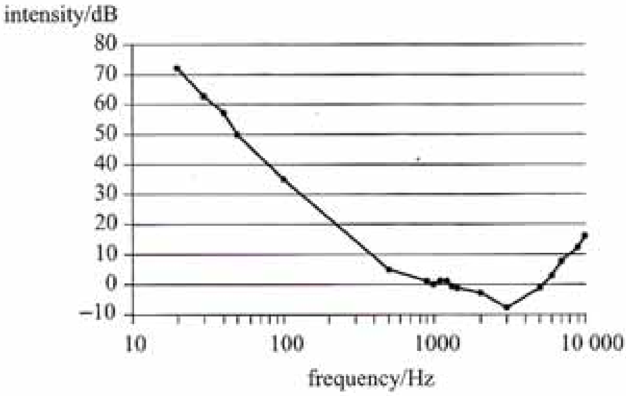

Figure I1.12 For question 7.

8 What is the reason for employing a logarithmic scale for sound intensity levels?

9 What can you conclude from the audiogram shown in Figure I1.13? Bone conduction is represented by triangles and air conduction by circles.

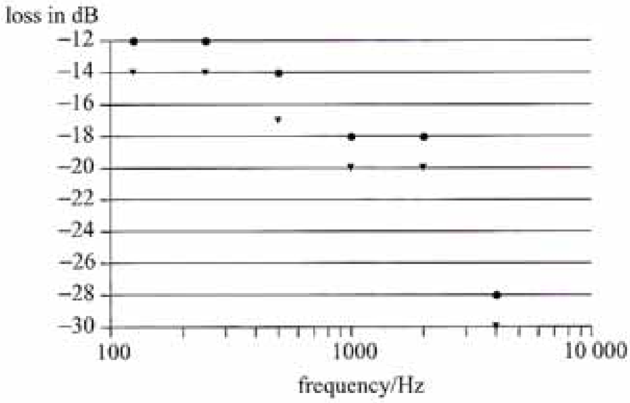

Figure I1.13 For question 9.