4.7 Hearing

Learning Objectives

- Distinguish among the qualities of sound perception including pitch, timbre, and loudness
- Define decibels and be able to calculate loudness in decibels given either pressure or intensity
- Describe the function of the outer ear, middle ear, and inner ear
- Describe the anatomic location of the eustachian tube and describe its function
- Briefly explain why the tympanic membrane is much larger than the oval window
- Indicate the following on an anatomic drawing of the cochlea: modiolus, perilymph, endolymph, Reissner's membrane, scala vestibuli, scala media, scala tympani, tectorial membrane, outer hair cells, inner hair cells
- Explain why "stereocilia" are not true cilia
- Describe the afferent and efferent innervation of outer and inner hair cells. Explain why this means that the inner hair cells form the sensory response to sound
- Identify the cochlea, helicotrema, oval window, round window. Explain what is meant by "tonotopic mapping"
- Define "evoked otoacoustic emissions" and identify their source
- Trace auditory sensory information from spiral ganglia to dorsal and ventral cochlear nuclei to inferior colliculus to medial geniculate nucleus to auditory cortex
- Identify the auditory cortex and indicate its tonotopic representation
- Identify Broca's area and Wernicke's area and associate with each deficits in oral communication
- Describe the basic operation of cochlear implants
- Define ABR and BAER and describe how they can be used to test hearing in newborns

THE HUMAN AUDITORY SYSTEM DISCRIMINATES AMONG TONE, TIMBRE, AND INTENSITY

We know from everyday experience that we can discriminate among different aspects of sound. We can discriminate the pitch, or the perceived highness or lowness (as opposed to loudness or softness) of a sound. Sound waves of a single frequency are called **tones**. **Pitch** is the perceived counterpart of tones. The greater the frequency, the higher the pitch. The human ear is most sensitive to tones between 500 and 5000 Hz, but can detect tones from 10 to 20,000 Hz with reduced sensitivity.

Sounds in general are a composite of many different frequencies. Pleasing sounds consist of mixtures of tones with frequencies in particular ratios. For example, two tones with frequencies in the ratio of 2:1 are said to be separated by an **octave**. Similar ratios form the basis of **intervals** in music, as described in Table 4.7.1. The presence of a mixture of frequencies is the **timbre** of a sound. This is due to the presence of overtones, which are frequencies imposed on a tone. A tuning fork produces a pure tone. If you were to sing a "C" or play it on a clarinet, the sound wave will have overtones that allow us to identify the source as voice or clarinet. Further, the overtones in voices are produced by the characteristics of vocal cords, pharynx, and nasal passages so that we are able to identify voices based on their characteristic overtones. The human ear is exquisitely sensitive to differences in frequency, being able to detect differences between separate tones of as little as 2 Hz. Musicians use this ability to tune their instruments. Simultaneous sounds of frequencies f_1 and f_2 produce an interference pattern with a beat frequency of $|f_1 - f_2|$. Musicians detect these beats and tune their instruments until the beats disappear.

Humans also discriminate sounds on the basis of their intensity or loudness. In physical terms, the intensity of sound is the amount of energy transmitted per second through a unit area perpendicular to the direction of propagation of the sound. The usual units are W cm^{-2} or W m^{-2}. We use a logarithmic transform of the intensity to represent the entire range of audible sounds. This is the **decibel**, abbreviated as d*B*. It is defined as

$$[4.7.1] \qquad \text{Loudness (decibels)} = 10 \log \frac{I_{\text{sound}}}{I_{\text{ref}}}$$

Here the intensity is in units of W cm^{-2}. The intensity of a sound requires a reference sound. Usually this is taken as the threshold of hearing, which varies with the frequency. For this reason, there are two decibel scales: one for **pressure sound level** and another for **hearing sound level**. They differ in using a physical intensity for a reference or threshold of hearing. The pressure sound level uses a reference of 10^{-16} W cm^{-2} = 10^{-12} W m^{-2}. Because the intensity is proportional to the square of

DOI: http://dx.doi.org/10.1016/B978-0-12-800883-6.00040-9

TABLE 4.7.1 Frequency Ratios in Musical Intervals

Interval	Frequency Ratio	Examples
Octave	2:1	Five octaves of C: 1024, 512, 256, 128, 64 Hz
Third	5:4	320 and 256 Hz (middle C)
Fourth	4:3	342 and 256 Hz
Fifth	3:2	384 and 256 Hz

the pressure amplitude (see Appendix 4.7.A1), the loudness can be given as

$$[4.7.2] \qquad \text{Loudness (dB)} = 20 \log \frac{\Delta P_{\text{sound}}}{\Delta P_{\text{ref}}}$$

Here ΔP is the increment in pressure over ambient pressure: the extra pressure that is propagated in the sound wave. The factor of 2 appears in Eqn [4.7.2] compared to Eqn [4.7.1] because of the proportionality of the intensity to the square of the pressure amplitude. The auditory threshold at 2000 Hz is about 2×10^{-5} Pa (1 Pa = 1 N m^{-2}). This corresponds to an intensity of about 0.5×10^{-12} W m^{-2}. Older texts sometimes use units of dyne cm^{-2} for units of pressure. Since 1 dyne = 10^{-5} N and 1 cm^{-2} = 10^{4} m^{-2}, the conversion is 1 dyne cm^{-2} = 0.1 N m^{-2}. The threshold for hearing as a function of sound intensity (expressed in d*B*, ΔP, and *I*) is shown in Figure 4.7.1. This threshold is not absolute: background noise raises the hearing threshold through a process called **masking**. Masking is a signal-to-noise ratio problem that explains why it is difficult to understand conversation in a crowded, noisy room.

THE AUDITORY SYSTEM CAN LOCATE SOURCES OF NOISE USING TIME DELAYS AND INTENSITY DIFFERENCES

In addition to the sensations of pitch, timbre, and loudness, we can also locate the sources of sound using a combination of the time delay in the sounds striking the right and left ear, and the sound shadow produced by the head.

THE EAR CONSISTS OF THREE PARTS: THE *OUTER EAR*, *MIDDLE EAR*, AND *INNER EAR*: EACH HAS A DEFINITE FUNCTION

THE OUTER EAR COLLECTS AND CHANNELS SOUND TO THE MIDDLE EAR

The outer or external ear consists of the **pinna**, also called the **auricle**, and the **external auditory meatus**, the ear canal. The ear canal ends, and the middle ear begins, at the **tympanic membrane**, the eardrum. The shape of the auricle helps focus sound into the ear. In some animals, the auricle is funnel shaped and can be moved independently to search for sounds. Humans' auricle lacks this ability. The external auditory meatus contains **sebaceous glands** and **ceruminous glands** that secrete oils and waxes, respectively. The wax is the ester of a fatty acid and a long chain alcohol. These sticky secretions trap dust, dirt, insects, and bacteria that enter the auditory canal. The combination of these secretions and sloughed-off epithelial cells from the canal forms ear wax, or cerumen (from the Latin "cera" meaning "wax"). Impacted ear wax can cause pain or deafness. It can be removed by dissolution with a warm bicarbonate-rich solution or by mechanical removal (Figure 4.7.2).

THE MIDDLE EAR TRANSFORMS AIR PRESSURE WAVES TO FLUID PRESSURE WAVES

The middle ear is an air-filled cavity between the tympanic membrane on one side and the promontory of the temporal bone on the other. The middle ear contains three **ossicles**, tiny bones that connect the tympanic membrane to the membrane covering the **oval window** on the inner ear. The oval window leads to a fluid-filled chamber which coils around in a structure called the **cochlea**. The three bones are the **malleus** (**hammer**), **incus** (**anvil**), and **stapes** (**stirrup**). They transfer

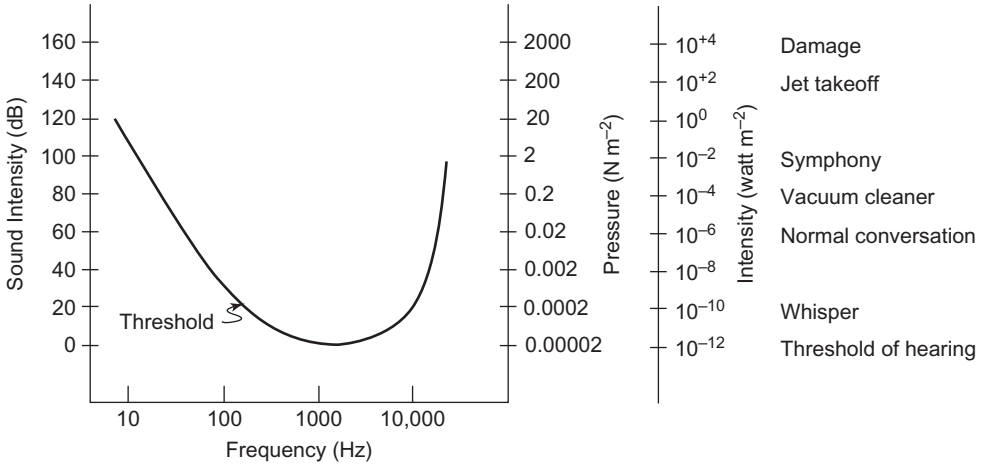

FIGURE 4.7.1 Threshold of hearing as a function of frequency. The ear is most sensitive to frequencies between 500 and 5000 Hz. The sound intensity can be described in decibels, or dB, as indicated on the scale to the left; as the maximum of the pressure amplitude, as given on the scale on the near right; and as energy per unit area per unit time, given on the scale to the far right. Typical sources of each of these sound intensities are indicated at the far right.

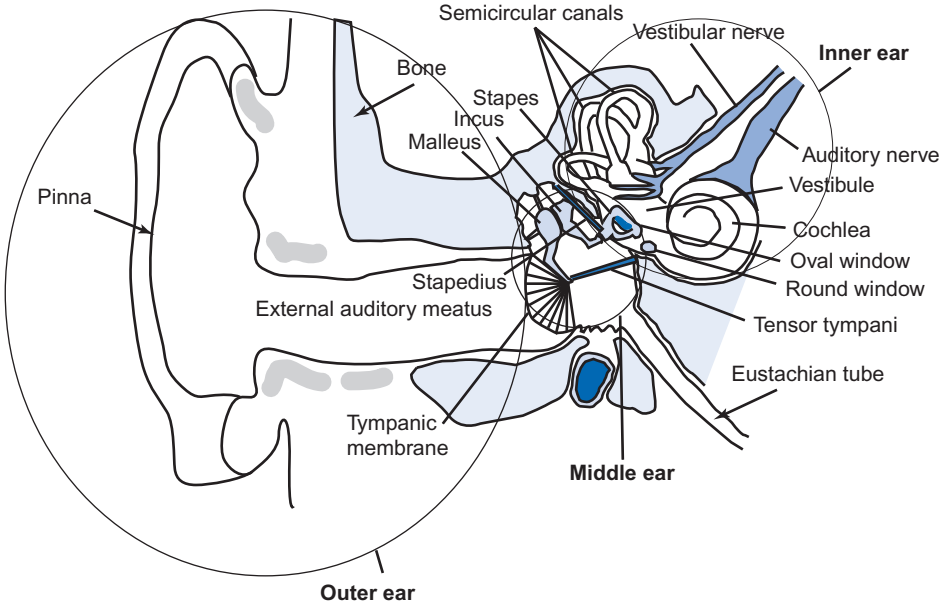

FIGURE 4.7.2 Structure of the ear. The outer ear includes the pinna and the external auditory meatus, the ear canal. These focus sound onto the tympanic membrane, which begins the middle ear. The air-filled middle ear contains three tiny bones, the ossicles (malleus, incus, and stapes), that transmit vibrations of the tympanic membrane to the inner ear. The stapes abuts the inner ear at the oval window. The inner ear consists of the cochlea for transduction of sound, and the semicircular canals for balance and detection of motion. The eustachian tube connects the middle ear to the atmosphere through the oral cavity. The auditory nerve transmits sound information to the CNS; the vestibular nerve transmits information about balance and movement. The two nerves combined make up cranial nerve VIII.

vibration of the tympanic membrane to vibration of the fluid in the cochlea.

The fluid in the external ear canal is air. Compressional waves in air cause the tympanic membrane to vibrate. This vibration must be transferred to the fluid in the cochlea of the inner ear. The cochlear fluid is a watery solution. Compressional waves in water require a completely different pressure because of the different inertia (the density) and different elastic properties (the bulk modulus: see Appendix 4.7.A1). Because of these differences, sound usually reflects off the air/water interface rather than being conducted into the water. The tympanic membrane and bones of the middle ear provide an impedance matching device to transfer sound energy to the cochlear fluid. The effective area of the tympanic membrane is about 0.4–0.6 cm^2; the area of the oval window is 0.03 cm^2. If there is no dissipation of energy in the bones themselves, then the sound energy can be concentrated some 15-fold where the stapes contacts the oval window.

The position of the tympanic membrane sets the position of the ossicles and therefore sets the transfer of vibration from air to cochlear fluid. If the pressure within the middle ear is higher than that in the external ear canal, the tympanic membrane will bulge outward and sensitivity to sound will decrease. If the pressure within the middle ear is lower than that in the external ear canal, the tympanic membrane will bulge inward and sound sensitivity will also decrease. This explains our loss of hearing sensitivity when we change altitude quickly, either ascending or descending. The **eustachian tube** connects the inner ear to the oral cavity and equilibrates the pressure on the two sides of the eardrum. Opening this tube by yawning wide or chewing equilibrates the pressure on the two sides of the membrane and pops the tympanic membrane back into place. This is sometimes accompanied by a popping or crackling auditory sensation and return of normal hearing sensitivity.

Thread-like ligaments and two muscles, the **tensor tympani** and the **stapedius**, keep the tiny ossicles in place. The two muscles contract reflexly with loud noises, body movements, stimulation of the ear canal, chewing, and contraction of facial muscles during vocalization. Contraction of these muscles reduces hearing sensitivity by making the ossicles more rigid. The latency of the reflex contraction to loud noises is about 40–60 ms, too long to prevent damage from brief intense sounds like explosions. However, these muscles may protect against prolonged, low-frequency sounds.

THE INNER EAR TRANSDUCES FLUID PRESSURE WAVES INTO NERVE IMPULSES THAT ARE TRANSMITTED TO THE BRAIN

The inner ear lies in a bony cavity in the temporal bone medial to the middle ear. The bony part of the inner ear is called the **bony labyrinth**, and it contains a variety of membranous ducts and sacs which are collectively called the **membranous labyrinth**. The membranous labyrinth floats in a fluid, the **perilymph**, that occupies the space between the membranous labyrinth and the bony labyrinth. The perilymph has low [K^+] and high [Na^+]. The part of the membranous labyrinth that makes contact with the stapes is the vestibule. The cochlea forms a spiral around a bony core called the **modiolus**. Two membranes, **Reissner's membrane** and the **basilar membrane**, divide the cochlea into three compartments. The outer two compartments contain perilymph and are continuous with the vestibule. The inner compartment contains **endolymph**, which has high [K^+] and low [Na^+]. The structures of the inner ear and details of the cochlea are shown in Figure 4.7.3.

The cochlea lies in a negative space—a hollow region of bone—making about 2 and 3/4 turns around a central core, the modiolus, that is shaped like a cone, with edges hollowed out to accommodate the cochlea. The nerves that receive input from the auditory receptor cells

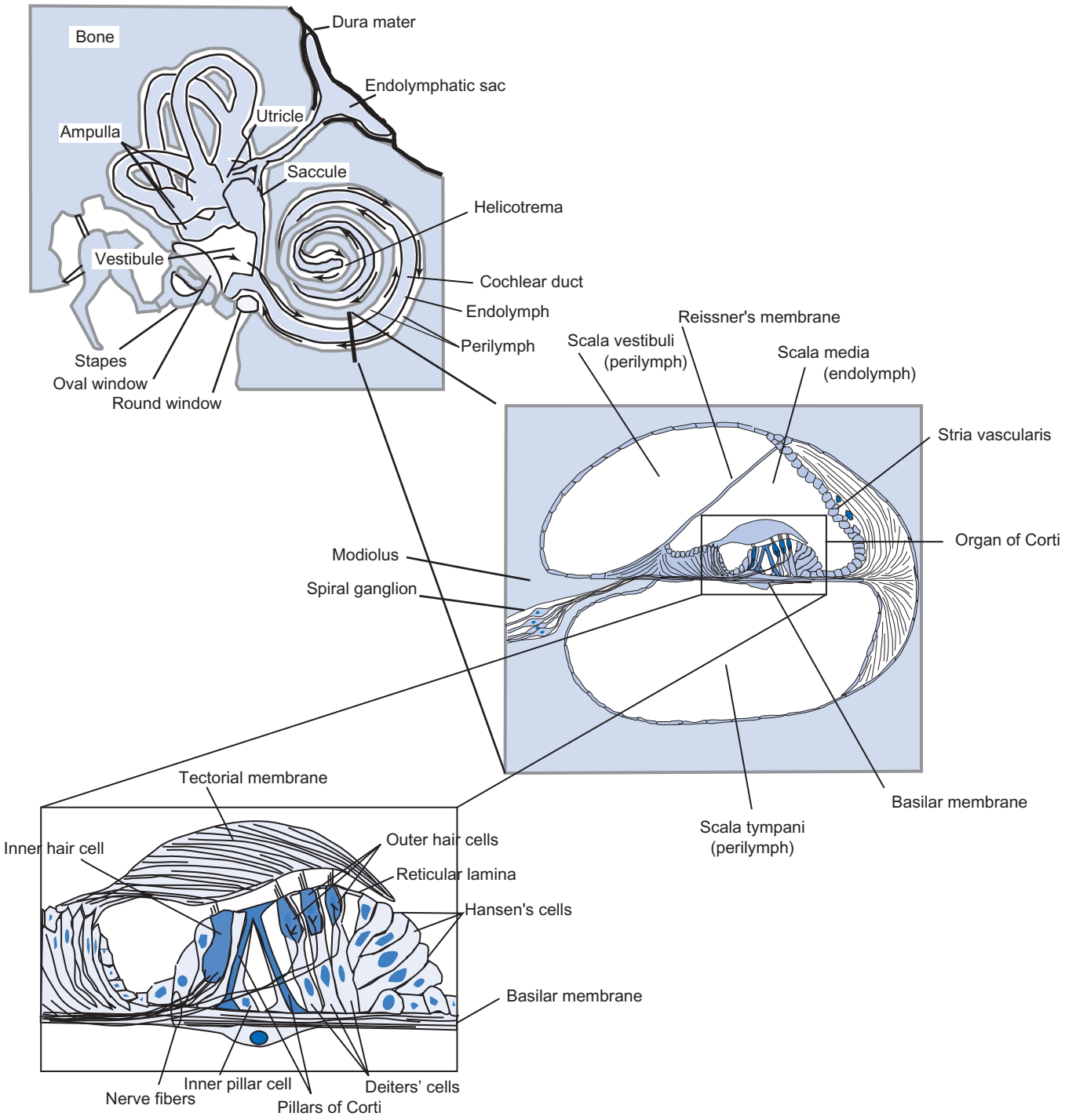

FIGURE 4.7.3 Structure of the inner ear. For explanation of cell types, see the text.

have cell bodies that reside in **spiral ganglia** within the modiolus. Their fibers collect to form the auditory nerve, which is part of cranial nerve VIII. Figure 4.7.3 shows cross sections of the cochlea.

The fluid in the scala vestibuli and scala tympani is perilymph. These two compartments are connected at the **helicotrema**, which is at the apex of the cochlea. At the base, the scala vestibuli is continuous with the vestibule. The foot of the stapes connects to the vestibule at the oval window and imparts pressure waves to the vestibule perilymph. These pressure waves are conducted all the way up the cochlea in the scala vestibuli, through the helicotrema and back down the scala tympani. Pressure waves at the far end of the scala tympani cause the round window to vibrate.

The **organ of Corti** is the specialized part of the cochlea that senses the vibrations of the basilar membrane and transduces these vibrations into electrical impulses. The vibrations are sensed by hair cells. There are two types of hair cells, as shown in Figure 4.7.3. There are three

Clinical Applications: Ear Infections and Tubes

Large numbers of children suffer from recurrent or chronic middle ear infections, called **otitis media**. This is an inflammation of the middle ear caused by viruses or bacteria, which invade the middle ear, usually by way of the eustachian tube, which connects the middle ear to the oral cavity. The resulting infection can be accompanied by fluid buildup in the middle ear that pushes on the eardrum. This causes a temporary loss of hearing sensitivity, and it hurts. Ordinarily the eustachian tube drains fluids into the oral cavity, but it may not open sufficiently due to congestion or a variety of other reasons. Most children are treated with antibiotics for bacterial infections. Children (or adults) with chronic or recurrent otitis media with effusions may require tubes to be inserted in their ears. The procedure of inserting tubes in the eardrum is called **tympanostomy**.

There are over 50 different designs for the tubes inserted into the tympanic membrane. The tubes can be made of rigid plastic or metal or silastic. Some are simple tubes whereas others have flanges to help secure them in the eardrum. The procedure requires making an incision in the membrane and inserting the tube in a place on the eardrum that will not interfere with the motion of the malleus. Because the surgery is delicate, the young patient is anesthetized with a general anesthetic. The procedure is very common. The National Institute for Child Health and Development estimates that in the United States in 1996, 280,000 children under the age of 3 years underwent tympanostomy.

The tubes provide an effective drain for the middle ear to the external ear canal. This extra drain allows the inner ear to become ventilated and reduce its fluid buildup. This reduces the temporary hearing loss and pain of acute otitis media, and reduces its recurrence. However, the tubes do not guarantee that ear infections will not recur. Much of the justification for the procedure relies on the argument that loss of hearing in young children harms their verbal and social development, but recent studies suggest that tympanostomy has no effect on children's scores on tests of expressive language or general cognition.

outer hair cells for every **inner hair cell**. The inner hair cells are innervated by about 20 unbranched afferent nerve fibers. Afferents from the outer hair cells form branches, with each neuron connecting to about 10 outer hair cells and each outer hair cell connecting to about 4 afferents. The hair cells also receive efferent nerve fibers, and most of these (80%) innervate the outer hair cells. The hair cells have tiny **stereocilia** that project from a structure called the **cuticular plate** on the apical aspect of the hair cells. The hair cells are overlaid by a gelatinous, acellular structure called the **tectorial membrane**. Shear between the tectorial membrane and stereocilia of the hair cells produces deformation of the stereocilia that eventually activates afferent nerve fibers. The inner hair cells generate afferent input to the auditory system, whereas the outer hair cells change their length to "tune" the inner hair cells' response.

The chief structural elements of the organ of Corti are provided by inner and outer **pillar cells**. These pillar cells contain trunks of dense, interlaced microfibrils, and microtubules that lock together at the apical side and contact the basilar membrane at widely separated points. They form a cell-free space, **the tunnel of Corti**, beneath the **reticular lamina**, a sheet of cells through which the apical surfaces of the hair cells penetrate. The thin reticular lamina is composed of **phalangeal processes** arising in part from **Deiters' cells**. These Deiters' cells support the outer hair cells only on their basal and apical surfaces, leaving the sides of the outer hair cells free of cellular contact.

HAIR CELLS OF THE COCHLEA RESPOND TO DEFORMATION OF STEREOCILIA TOUCHING THE TECTORIAL MEMBRANE

The inner and outer hair cells share the common feature of having their apical surfaces immersed in endolymph, the fluid of the scala media. In addition, their apical surfaces have a **cuticular plate** in which are embedded rows of **stereocilia.** The two types of hair cells differ substantially. The inner hair cell is flask shaped and is innervated mostly by afferent nerves. The outer hair cells are rod shaped and are innervated predominantly by efferent nerves. The inner hair cells form a spiral row running from oval window to apex of the cochlea; the outer hair cells form three parallel spiral rows over the same surface. Figure 4.7.4 shows these cells. The stereocilia of the inner hair cells are longer than those of the outer hair cells. Both sets of stereocilia form several rows, arranged either as a "V" shape of "W" shape, with the points of the "V" pointing away from the modiolus around which the cochlea spirals. The stereocilia on the side away from the modiolus are longer. Only the longest stereocilia contact the tectorial membrane that overlays the hair cells. These structures are called stereocilia because they resemble cilia without having the same basic structure of 9 double microtubules surrounding a central pair. Electron micrographs reveal thin filamentous connections between the tips of the short stereocilia and the taller ones. These **tip links** connect to a K^+ channel. Bending of the stereocilia away from the center of the modiolus excites the hair cell by stretching the tip links, opening a K^+ channel and depolarizing the cell.

As shown in Figure 4.7.4, the stereocilia protrude through a cellular layer called the **reticular lamina.** This structure divides the fluid in the scala media, which is filled with endolymph, from the scala tympani, which is filled with perilymph. The endolymph is the most unusual extracellular fluid in the body. This extracellular fluid has $[K^+]$ of about 150 mM and $[Na^+]$ of about 1 mM. In addition, it is maintained at a high positive potential, +50 to +100 mV. This highly unusual fluid is maintained by the **stria vascularis** (see Figure 4.7.3). The positive endolymph potential is maintained by active processes within the stria vascularis, and it collapses within minutes after interrupting the oxygen supply. The basolateral surfaces

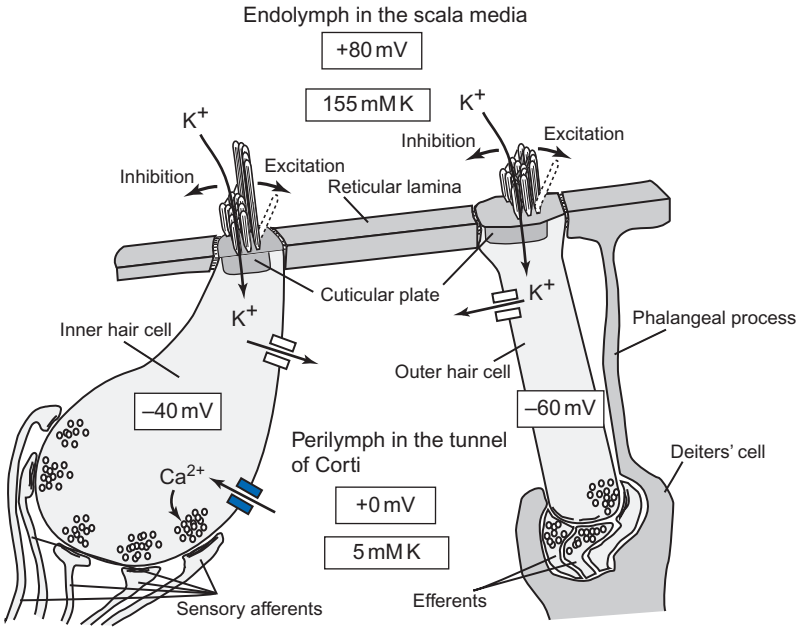

FIGURE 4.7.4 Diagram of the inner and outer hair cell arrangement. The inner hair cells have stereocilia whose first row is much longer than the others. The endolymph has a high $[K^+]$ and a positive potential that is actively maintained by the stria vascularis. Stretching of the tip links between the stereocilia open K^+ channels on the stereocilia membranes that results in K^+ inward flow and depolarization. These channels at rest have some conductance so that relaxation of the tip links produces a reduction in the K^+ current with resulting hyperpolarization and inhibition of the nervous activity on the sensory afferent nerves. *(Source: Modified from Geisler,* From Sound to Synapse, *Oxford University Press, 1998.)*

of both the inner and outer hair cells lie beneath the reticular lamina. The basilar membrane is freely permeable to small ions, whereas the reticular lamina seals the fluid in the scala media (endolymph) from that in the tunnel of Corti (perilymph). Thus, the fluid bathing the basolateral surfaces of the hair cells is perilymph, with high $[Na^+]$ and low $[K^+]$. The low $[K^+]$ and high $[Na^+]$ is necessary for the propagation of the action potential along the afferent sensory axon.

When the long stereocilia tug on the shorter ones through the tip links, K^+ channels are opened, causing K^+ ions to enter the hair cell, because the potential in the scala media is +80 mV whereas that in the hair cell is some −40 to −60 mV, and $[K^+]$ is comparable between the scala media and the hair cell. This flow depolarizes the cell, leading to opening of voltage-gated Ca^{2+} channels on the basolateral aspect of the hair cells. The depolarization of the hair cells is its **transduction potential**. Because there is a potential difference between the scala media, containing endolymph, and the scala tympani, containing perilymph, there is always some current passing through the reticular lamina, and movement of the stereocilia towards the modiolus reduces the conductance to K^+ and reduces the current. Thus, movement of the stereocilia causes either excitation or inhibition.

OUTER HAIR CELLS MOVE IN RESPONSE TO EFFERENT STIMULATION AND THEREBY TUNE THE INNER HAIR CELLS

Outer hair cells change their axial dimensions in response to electrical stimulation. Hyperpolarization lengthens the cells, and depolarizations shorten them. This **"electromotility"** depends on membrane potential, and it occurs extremely rapidly. Outer hair cells can follow sinusoidal electrical commands up to 24 kHz. The outer hair cells can shorten or lengthen only a few percent of their rest length of 30 μm, but this is about the same magnitude as sound evoked vibrations of the basilar membrane. The mechanism of electromotility is not established but requires the protein **prestin**. Prestin knockouts in mice result in 20–40 dB hearing loss.

Instead of sound vibration being dampened by the absorption of sound energy by the lamina, the electromotility response of the outer hair cells adds energy, something like negative damping. This phenomenon has been termed the **cochlear amplifier**. This allows for the exquisite response of the inner hair cells. Loss of outer hair cells produces a loss of sensitivity and frequency discrimination.

In 1978, Kemp reported that brief clicks directed to the ear resulted in faint sounds coming out of the ear. These are called **evoked otoacoustic emissions**. Later studies found that otoacoustic emissions could emanate spontaneously from unstimulated cochleas. Otoacoustic emissions from one ear can be modulated by sound presented to the contralateral ear, suggesting that otoacoustic emissions probably originate from the outer hair cells.

THE COCHLEA PRODUCES A TONOTOPIC MAPPING OF SOUND FREQUENCY

As described above, hair cells sense sound through deformation of the stereocilia on their apical membranes. How is this deformation produced? Georgé von Békésy first described vibration of the basilar membrane in a series of experiments performed on cochleas from cadavers during the 1940s (G. von Békésy, *Experiments on Hearing*, McGraw Hill, 1960). This remarkable achievement earned a Nobel Prize for von Békésy in 1961. He found that pressure waves applied to the oval window resulted in a "traveling wave" of displacement of the basilar membrane. The wave is conveyed from the base of the cochlea, near the oval window, toward the apex of the cochlea near the helicotrema. The amplitude of the displacement varies with the distance along the cochlea

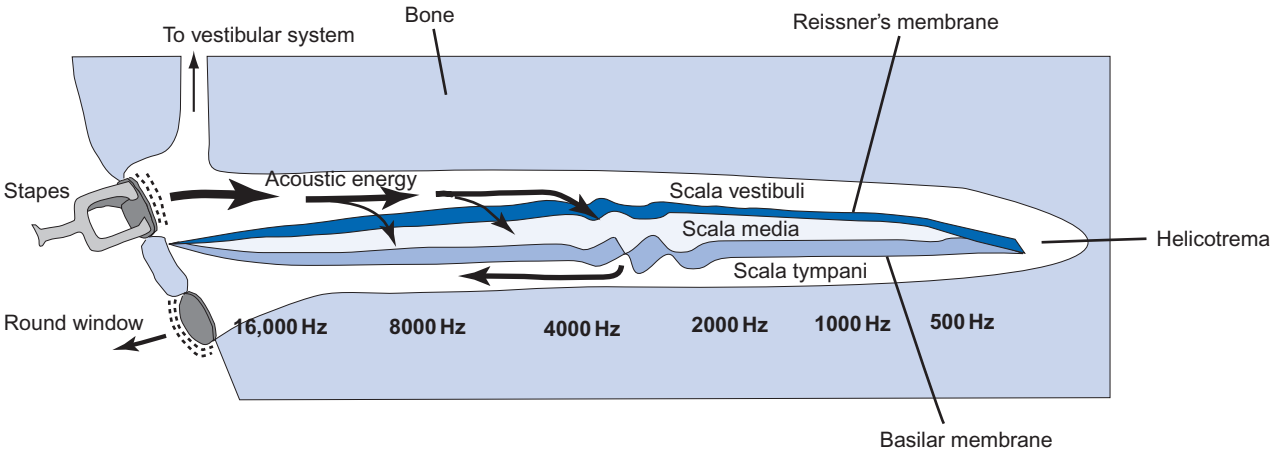

FIGURE 4.7.5 Highly schematic diagram of the cochlea, unwound from its spiral as if it were straight. The cochlear partition between scala vestibuli and scala tympani consists of Reissner's membrane and the basilar membrane, including the organ of Corti. The fluid bathing the scala vestibuli and scala tympani is perilymph, whereas the scala media is filled with endolymph. Acoustic energy produced from a loud tone is conveyed to the perilymph in the scala vestibuli by the vibration of the oval window by the stapes. Most of the energy is absorbed by an area of the cochlear partition that depends on the frequency of the sound, and this produces tuned vibrations of the cochlear partition. Some of the sound energy remains to displace the round window, resulting in sound emission from the cochlea back into the middle ear. *(Source: Modified from Geisler,* From Sound to Synapse, *Oxford University Press, 1998.)*

and with the frequency of the sound. For a particular tone, the maximum amplitude is located within a fairly narrow region of the cochlea. In Von Békésy's experiments, the displacement was broadly distributed along the cochlea because he measured traveling waves at high sound levels and he used dead cochleas. Later, more sensitive techniques with live cochleas showed a much more sharply tuned displacement of the basilar membrane.

At the base of the cochlea, near the oval window, the basilar membrane is narrow. As it winds around the modiolus, the basilar membrane becomes progressively wider, reaching its widest dimensions at the apex of the cochlea. Because of this, high-frequency sounds produce a maximum displacement of the basilar membrane near the base of the cochlea, whereas low-frequency sounds maximally displace the basilar membrane near the apex. The result is that inner hair cells located near the oval window, near the base of the cochlea, respond to high-frequency sounds and inner hair cells located near the apex respond to low-frequency sounds. Thus the cochlea is organized **tonotopically**: there is a relationship between frequency response and location on the cochlea. This idea is illustrated in Figure 4.7.5.

AUDITORY INFORMATION PASSES THROUGH THE BRAIN STEM TO THE AUDITORY CORTEX

The inner hair cells make synapses on the processes of **bipolar cells** whose cell bodies are located in the **spiral ganglion**, buried in the bone of the modiolus. There are about 30,000 of these cells in the human spiral ganglion, and the vast majority of these make contact with a single inner hair cell, and each inner hair cell contacts between 10 and 20 primary afferent auditory nerve fibers. The fibers leave the spiral ganglion and are collected in the **auditory nerve** that joins the vestibular nerve to form **cranial nerve VIII** (Figure 4.7.6).

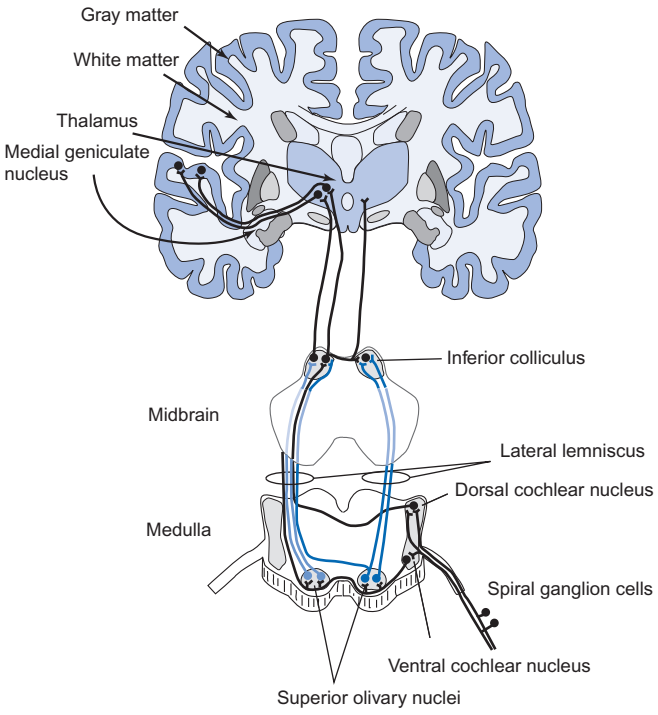

FIGURE 4.7.6 Schematic diagram of central auditory pathways. Spiral ganglion cells receive afferent signals directly from inner hair cells. The primary afferent processes enter in the medulla and make synapses on cells in the dorsal and ventral cochlear nuclei in the medulla. Second-order fibers ascend in the contralateral lateral lemniscus to make contact with cells in the inferior colliculus. Neurons in the ventral cochlear nucleus send collaterals to the superior olivary nucleus. Third-order cells in the olivary nuclei send ascending fibers to the inferior colliculus. Cells in the inferior colliculus in turn send fibers to the medial geniculate nucleus of the thalamus, which relays the information to the auditory region of the temporal lobe in the cerebral cortex. *(Source: Modified from Nicholls, Martin, Wallace, and Fuchs,* From Neuron to Brain, *Sinauer Associates, 2001.)*

Primary auditory nerve fibers in cranial nerve VIII make synapses on secondary afferent neurons in the **dorsal and ventral cochlear nuclei** which are located in the brain stem. The secondary afferent fibers originating in the

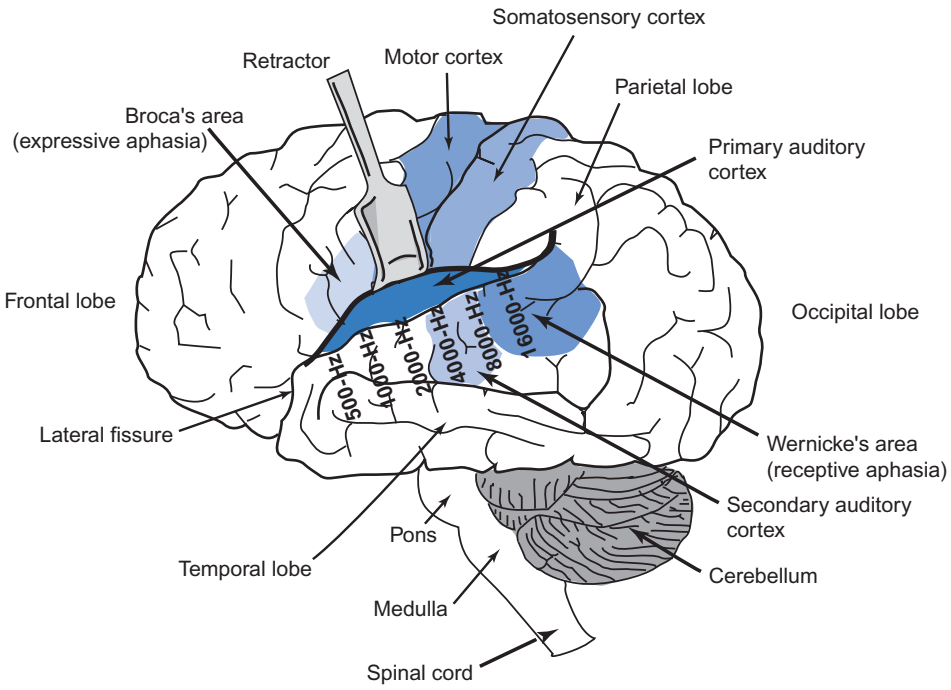

FIGURE 4.7.7 The primary auditory cortex (A1) and areas for producing and understanding speech. Afferents from the apex of the cochlea feed into anterior areas of A1, and these respond to low sound frequencies. Afferents from the base of the cochlea feed into posterior regions of A1, and these process high-frequency sounds. Broca's area lies in the frontal lobe adjacent to the areas of the motor cortex responsible for motor control of the face, lips, jaw, tongue, pharynx, and larynx. Damage to this area causes expressive aphasia. Damage to Wernicke's area, in the posterior part of the temporal lobe adjacent to the secondary auditory cortex, causes receptive aphasia. *(Source: Modified from Kolb and Whishaw,* An Introduction to Brain and Behavior, *Worth Publishers, 2000.)*

cochlear nuclei may take a variety of paths. Some synapse on neurons located in the **superior olivary complex**, both ipsilateral and contralateral. The superior olivary complex is a collection of nuclei located in the medulla. Fibers leaving the superior olive then synapse on neurons in the **inferior colliculus**. The inferior colliculus also receives input from the eyes, so these connections make possible complex behavior such as seeking the source of a sound.

Fibers from the inferior colliculus project to the **medial geniculate nucleus** (MGN) of the **thalamus** and also cross over to the contralateral side. These make contact with another set of neurons that send fibers directly to the **primary auditory cortex**, located laterally near the Sylvian sulcus. The primary auditory cortex is also called A1 and is located within Heschl's gyrus, the transverse temporal gyrus. The perception of pitch is mapped onto A1 **tonotopically**. The neurons in the cochlea code for frequency based on their location in the cochlea: cells at the base respond to high frequencies and cells at the apex of the cochlea respond to low frequencies. This spatial arrangement of sensitivity to tone is preserved throughout the neural connections so that there is a mapping of location on the A1 cortex and pitch: cells in the anterior part of A1 respond to low frequencies and cells in the posterior part respond to high frequencies (see Figure 4.7.7).

LANGUAGE IS PROCESSED IN AREAS NEAR THE PRIMARY AUDITORY CORTEX IN THE LEFT HEMISPHERE, BUT MUSIC IS PROCESSED IN THE RIGHT HEMISPHERE

Based on "natural experiments," Paul Broca (1824–1880) identified an area immediately in front of the central fissure and superior to the lateral fissure, now called **Broca's area**. Damage to this area results in **expressive aphasia**, the inability to speak despite being able to understand spoken language and despite having normal vocal apparatus. Karl Wernicke later identified a separate area, **Wernicke's area**, located in the posterior of the left temporal lobe. Damage to this area results in **receptive aphasia**, which is the inability to understand the spoken word. People with receptive aphasia can speak, but their speech is garbled and unintelligible.

Human experience is replete with anecdotes of persons whose brains have been damaged and have lost the ability to speak, and yet they can sing. This is possible because the processing of music is located in separate areas in the temporal lobe of the right hemisphere. Despite this, cooperation among many brain areas is needed for full function, as damage to the left hemisphere can interfere with the ability to read or write music.

PERCEPTION OF PITCH IS ACCOMPLISHED BY A COMBINATION OF TUNING AND PHASE LOCKING

As described earlier, the primary afferents for sound detection arise from the inner hair cells that respond to deformation of the stereocilia in response to movement of the tectorial and basilar membranes. Pressure waves transmitted to the oval window through the action of the ossicles produce a wave of basilar membrane deformation that has a maximum displacement that depends on the frequency, but it also depends on the magnitude of the pressure waves. The physical response of these membranes, along with the electrical properties of the inner hair cells, produces a "**tuning curve**" for the primary afferent fibers. These tuning curves are the threshold intensities for impulse generation as a function of

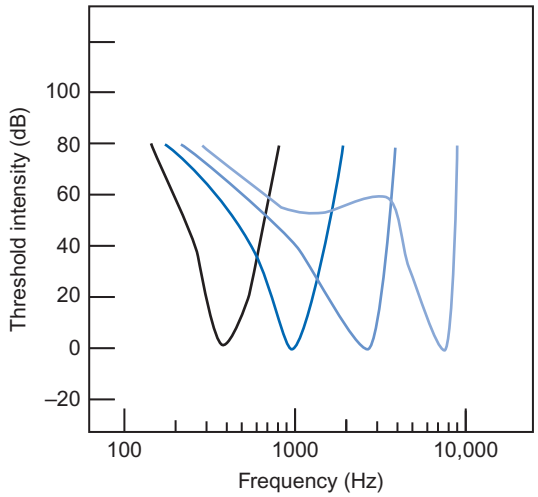

FIGURE 4.7.8 Tuning curves for primary afferents at different locations on the cochlea. Afferents from inner hair cells near the apex show a maximum sensitivity (lowest threshold) to sounds with low frequencies (black curve). Primary afferents from inner hair cells near the base show maximum sensitivity to higher frequencies (curve on far right). Cells in between show intermediate behavior (middle two curves). The sensitivity falls off sharply at higher frequencies, but lower frequencies can still excite the afferents off their maximum sensitivity. *(Source: Modified from Lieberman and Mulroy, in Hamernik, Henderson, and Davis, eds.,* New Perspectives on Noise-Induced Hearing Loss, *Raven Press, 1982.)*

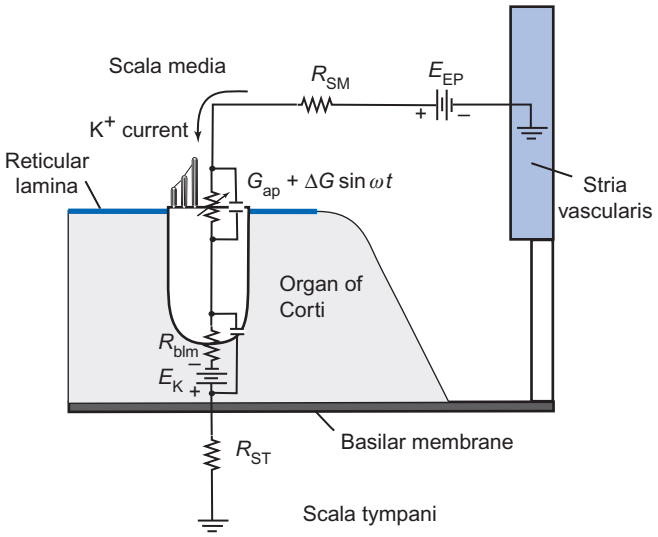

FIGURE 4.7.9 Schematic diagram of the resistances and potentials across the basilar membrane. Ground is taken as the vascular system. The stria vascularis produces an endolymphatic potential, E_{EP}, which is approximately +80 mV with respect to the scala tympani (ST). The scala media (SM) and scala tympani present resistances to current flow. Current flows across the cell membrane in accordance with a steady-state conductance G, which varies when K^+ channels are opened on the apical membrane by movement of the stereocilia. A sinusoidal variation in the apical conductance results in a sinusoidal variation in the K^+ current and a sinusoidal variation in the membrane potential. *(Source: Modified from Geisler,* From Sound to Synapse, *Oxford University Press, 1998.)*

frequency. Typical tuning curves for primary afferents of the auditory nerve are shown in Figure 4.7.8.

The tuning curves show a fairly sharp maximum sensitivity (lowest threshold) at a **characteristic frequency**. The sensitivity falls off sharply at higher frequencies but generally shows a "tail" at lower frequencies. This is because the traveling wave for high frequencies travels only so far up the cochlea, whereas all low-frequency sounds must pass through the inner hair cells' location to reach their place of maximum basilar membrane displacement. Louder sounds produce larger deformations of the basilar membrane, so that a louder sound might be confused with one of higher pitch. There is psychophysical evidence for a shift in the perceived pitch with sound intensity. The auditory system deals with this problem by using a phase-lock coding of sound frequency in addition to the tuning curves of the primary afferents.

Because a pure tone transmitted to the perilymph induces a sinusoidal displacement of the basilar membrane, the primary afferent neurons tend to fire in phase with the sound wave. Thus, the neurons fire at the same part of the sound wave, usually its maximum pressure. Thus, their firing rate encodes the original frequency of the sound. However, each primary afferent has a limited firing rate of something less than 200 Hz; they cannot keep up with tones having a higher frequency than this. Because each inner hair cell is innervated by 10 or more primary afferents, the ensemble does indeed track the stimulus waveform. The convergence of these primary afferent neurons on targets within the cochlear nucleus suggests that summation within the cochlear nucleus allows a frequency coding of sound pitch as well as a place coding (the tonotopic map). This requirement probably explains the curious multiple innervation of the inner hair cells. These combined codes allow for simultaneous discrimination of sound intensity and pitch.

THE *COCHLEAR MICROPHONIC* SHOWS THAT THE INNER HAIR CELLS HAVE AN AC RESPONSE THAT CAN KEEP UP WITH MODERATE FREQUENCY VIBRATIONS

In a rather bizarre experiment, Weaver and Bray in 1930 reported that the electrical activity recorded from a cat's auditory nerve could be amplified and hooked up to speakers. If you then spoke into the cat's ear, you could simultaneously hear yourself over the speaker! In 1931, Adrian found that the source of the AC signal was not the nerve, but the cochlea, and he coined the term **cochlear microphonic** to describe it. Figure 4.7.9 shows an equivalent electrical circuit for a hair cell in which a sinusoidal perturbation of the hair cells' stereocilia, at the frequency of the sound, would produce a similar sinusoidal variation in the current across the organ of Corti and a sinusoidal variation in the membrane potential.

Note that all of the current that passes through the hair cell in Figure 4.7.9 also passes through the resistances R_{SM} and R_{ST}, and gives rise to a transduction voltages in the scala media and scala tympani. As currents from all hair cells share these pathways, the total current is the sum of all the individual hair cell currents. The voltage

Clinical Applications: Cochlear Implants

The first cochlear implant was developed by Graeme Clark at the University of Melbourne, Australia, and implanted in 1978. Cochlear implants are now commercially available. The devices are not hearing aids, which merely amplify sound. Instead, the devices themselves receive sound and distribute the power into a series of bandwiths and stimulate the hearing nerve tonotopically by bypassing the ossicles and the oval window. It can restore partial hearing to the deaf.

Cochlear implants consist of both internal and external parts. The internal parts are surgically implanted, whereas the external parts can be detached at any time. The external parts consist of a microphone, a speech processor, and a transmitter. The microphone picks up sounds and converts them into electronic signals. These are sent over a thin wire to the speech processor, which divides the sound into frequency bands and sends the coded signal to a transmitter worn outside the skull. The transmitter sends the coded signal across the skin by radio waves.

The internal parts consist of a receiver and an electrode array. The receiver converts the radio waves into electrical signals that it sends down an electrode array implanted in the cochlea. Because the cochlea detects sound frequency according to the location of the inner hair cells along the cochlea, with high frequency at the base of the cochlea and low frequency at the apex, the electrode array can stimulate local regions of the cochlea and elicit specific perceptions of sound frequency. The Nucleus® 24 divides sound into a maximum of 22 channels. By dividing sound into 22 frequency bands and stimulating 22 different regions of the cochlea, the implant approximates normal hearing.

Fourier analysis divides any function into an infinite series of sines:

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \sin \omega_n x$$

In the case of sound, $f(x)$ is the pressure as a function of time, which can be approximated by a series of 20 terms of different angular frequencies, ω. Ideally the power of received sound could be divided up according to a Fourier analysis, but this takes too much time to be done on the fly, in real time, by a small device. The speech processors divide the interesting part of the sound spectrum into 20 different frequency bands by using a system of bandpass filters. It then stimulates parts of the cochlea according to the power of the sound signal in each of these bands. A variety of speech coding strategies can be used.

Sound perception with cochlear implants is not normal. Wearers detect a "metallic" sound to their perception and it takes time and motivation to learn to hear with the device. The benefits vary with the user and their motivation to learn. Over half of the recipients can distinguish speech with no visual cues.

Clinical Applications: Hearing Tests for Newborns

Until recently, parents became aware of their child's deafness only when the child failed to respond to noises or failed to make sounds or develop speech. Modern instruments allow testing of hearing in infants as soon as 6 h after birth.

Two major methods are available: otoacoustic emissions and auditory brain stem response (ABR), also called brain stem auditory evoked response (BAER).

As mentioned in the text, the ear not only hears sounds but emits sounds. The otoacoustic emissions tests the ear for the sounds it emits as a consequence of hearing sounds. Clicks presented to the ear generate transient evoked otoacoustic emissions from a variety of regions of the cochlea corresponding to different frequencies of emitted sounds. Analysis of the spectrum of otoacoustic emissions can diagnose problems in hearing.

The ABR is a kind of evoked electroencephalogram. Each time a brief click is presented to the ear, neurons are activated sequentially in the pathway from ear to brain. The activity of these neurons causes minute potential changes that can be detected on the surface of the skin overlying the brain stem. The signals from these neurons are usually lost in the sea of noise produced by all of their neighboring neurons. The signal-to-noise ratio is improved by averaging many sweeps. Each sweep is triggered on a signal when a click is presented into the ear of the newborn. Since each response of the auditory system follows sequentially, signals from the auditory system occur at definite intervals following the click, but ongoing brain activity is not synchronized in this way. By taking many sweeps and averaging them, the random background noise of the brain cancels out, whereas the tiny signals from the auditory system do not cancel out. Movement of the head produces large signals (electromyograms) which typically do not happen often enough to cancel out during signal averaging, so the patient must be still. The test is usually performed during sleep.

Figure 4.7.10 shows a typical trace during a BAER. Each peak is attributed to a part of the neuronal chain linking ear to brain.

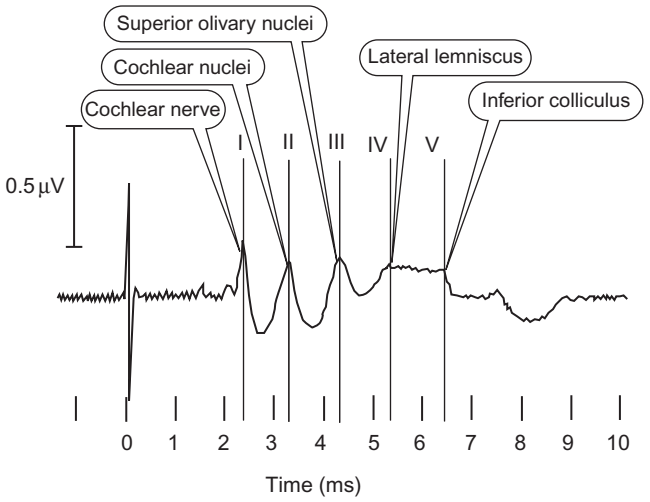

FIGURE 4.7.10 Tracing of ABR. Wave I is attributed to activity along the auditory nerve, cranial nerve VIII. This activates neurons in the cochlear nuclei (wave II) and then those in the superior olivary nuclei (wave III). The activity then passes through the lateral lemniscus (wave IV) to the inferior colliculus (wave V).

registered in the scala tympani is the sum of contributions from many cells. The time-varying portions, the AC response, is the cochlear microphonic. There is also a DC component that occurs at higher amplitude with greater frequency of sound stimulus.

SUMMARY

The sense of hearing informs us of the loudness, tone, and timbre of sounds. Loudness is measured in decibels, the logarithmic ratio of intensity or pressure relative to a reference. Each 20 decibel unit is a 10-fold increase in pressure level and a 100-fold increase in sound intensity. The ear is most sensitive to sounds between 500 and 5000 Hz.

The outer ear or pinna collects sound and funnels it into the external auditory meatus. The tympanic membrane at the end of the ear canal divides the outer ear from the inner ear. Sound makes the tympanic membrane vibrate, and this vibration transfers to three tiny bones—the malleus, incus, and stapes (hammer, anvil and stirrup)—that focuses the vibration onto the oval window, which marks the beginning of the inner ear. The pressure on the two sides of the tympanic membrane equilibrates through the eustachian tube that connects the inner ear with the oral cavity. Yawning or chewing can open the tube and pressure can be equilibrated.

The relevant part of the inner ear for hearing is the cochlea. The entire inner ear is surrounded by a bony labyrinth. The cochlea forms a spiral around a bony core called the modiolus. The cochlea itself consist of three compartments: the scala vestibuli, the scala media, and the scala tympani. The scala vestibuli and scala tympani are connected at the apex of the cochlea at the helicotrema, and both are filled with perilymph. The scala media is filled with endolymph, which has a high $[K^+]$ and a high voltage (+80 mV) relative to the perilymph. The stria vascularis in the walls of the scala media maintains this unusual fluid. Reissner's membrane separates the scala vestibuli from the scala media, and the basilar membrane separates the scala tympani from the scala media. Vibration of the oval window is transferred to the scala vestibuli and eventually returns via the scala tympani to the round window. The receptor cells that respond to sound vibrations reside in the organ of Corti, which lies on the basilar membrane. This complicated device consists of a single row of inner hair cells and three rows of outer hair cells, supported and arranged by supporting structures, all overlaid by a gelatinous tectorial membrane. The inner hair cells are richly innervated with afferent nerves whose cell bodies reside in the spiral ganglia in the modiolus. The outer hair cells receive a rich supply of efferent nerves. The outer hair cells move in response to nervous stimulation. It is believed that the inner hair cells respond to sound, while the outer hair cells tune the response of the inner hair cells.

The hair cells owe their name to a field of stereocilia on their apical surface. This field has a polarity, with stereocilia becoming progressively larger as one moves away from the center of the spiral. The long sterocilia connect to shorter ones by tip links. Movement of the stereocilia away from the modiolus tugs on the tip links that open K^+ channels in the apical membrane, which depolarizes the cells and activates the spiral ganglion cells.

Sensory information from the cochlea joins the vestibular nerve in cranial nerve VIII and enters the brain stem to make synapses with secondary afferent neurons in the dorsal and ventral cochlear nuclei. These neurons project to neurons in the superior olivary nucleus both ipsilaterally and contralaterally, and cross over to ascend to the inferior colliculus. Neurons in the inferior colliculus project to the medial geniculate nucleus in the thalamus, which then projects to the primary auditory cortex. The auditory cortex is found beneath the lateral fissure. Sound tone is represented tonotopically: the base of the cochlea responds to high-frequency sound and the apex responds to low frequencies. This mapping is preserved through the auditory pathways so that the anterior region of the auditory cortex senses low frequencies and the posterior region senses high frequencies.

REVIEW QUESTIONS

1. How are decibels related to sound intensity? What is the relationship between sound intensity and pressure amplitude? What is timbre?
2. Where is the eustachian tube and what does it do? What bone connects to the eardrum? What bone connects to the oval window?
3. What is the function of the tensor tympani and stapedius muscles? Why do you suppose you can sometimes hear your heartbeat?
4. What cells form the major afferent component of hearing? What do the outer hair cells do?
5. Where are the scala tympani and scala vestibuli? Where is the scala media? What fluid does each contain? What is so unusual about the endolymph?

6. Where are the cell bodies of auditory nerve primary afferents? Where do their axons project? Where is the primary auditory cortex?
7. What is meant by tonotopic mapping? What regions of the cochlea respond to high frequencies? Low frequencies? What areas of A1 respond to high frequencies? Low frequencies?
8. Where is Broca's area? What deficits result from its damage. Where is Wernicke's area? What deficits result from its damage?

APPENDIX 4.7.A1 THE PHYSICS OF SOUND

THE SPEED OF SOUND IS RELATED TO AIR DENSITY AND BULK MODULUS

Consider a right cylinder with cross-sectional area, A, that is filled with a gas of density ρ. This gas has two characteristics that are required for the transmission of sound: it possesses inertia and it is elastically deformable. Gases are characterized by their **bulk modulus**, defined as

$$B = \frac{\Delta P}{-\Delta V/V} \qquad [4.7.A1.1]$$

where ΔP is the change in pressure, V is the volume, and ΔV is the change in volume. The bulk modulus is always taken as being positive. What this equation says, then, is that a positive pressure increment results in a negative volume increment. That is, applying a pressure to a gas causes it to compress; reducing the pressure causes it to expand.

Suppose now that the cylinder of gas is fitted with a piston by which we can deliver a rapid force to the right. When this occurs, the piston will compress the gas immediately adjacent to it. This increased pressure will cause the gas molecules to move to the right, transmitting the pressure to the next volume element. This repeated process generates a compressional wave pulse that moves to the right (see Figure 4.7.A1.1).

For simplicity we assume that the piston moves with constant velocity v, and this causes all particles in the pressure pulse to move to the right at velocity v. If the piston moves for the time Δt, the trailing edge of the pressure pulse will have moved a distance $v\Delta t$. During this interval, the leading edge of the wave pulse will have advanced a distance $c\Delta t$, where c is the **wave speed**.

The longitudinal displacement of the particles is taken as y, and their original location is designated as x. Thus, after the interval Δt, the longitudinal displacement y will vary with x, as shown in Figure 4.7.A1.1.

The mass of gas originally at rest in the volume $V = c\Delta tA$ is just the undeformed density, ρ, times this volume, or $m = \rho c\Delta tA$. The change of momentum of the gas molecules provided by the piston, and present in the wave pulse, is

$$\Delta(mv) = \rho c\Delta tAv \qquad [4.7.A1.2]$$

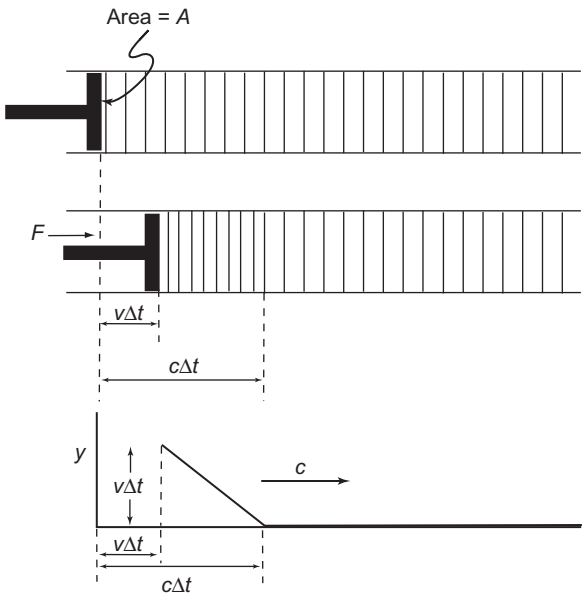

FIGURE 4.7.A1.1 Cartoon of a right circular cylinder fitted with a piston that moves with velocity v for time Δt. This produces a change in volume $\Delta V = -v\Delta tA$; this compression pulse travels at velocity c, so that the effective volume element which is being compressed is $V = c\Delta tA$.

The average force, F, times the time applied gives the momentum change. This is the impulse–momentum theorem, which is a direct result of Newton's third law. Insertion of this result into Eqn [4.7.A1.2] gives

$$\begin{aligned} F\Delta t &= \Delta(mv) = \rho c\Delta tAv \\ F &= \rho cAv \\ \frac{F}{A} &= \rho cv \\ \Delta P_0 &= \rho cv \end{aligned} \qquad [4.7.A1.3]$$

The force divided by the area, A, gives the longitudinal stress, which is equal to the pressure increment, ΔP_0, which is the maximal pressure that is produced at the end of the piston's excursion at $t = \Delta t$. This is the increase in pressure over the pressure that was present before the pulse. We will insert this into the definition of the bulk modulus, but first we want to solve for v in terms of volume and c. We recognize that the undisturbed volume was $V = Ac\Delta t$, and the change in volume was $\Delta V = -Av\Delta t$. The minus sign indicates that the volume decreased upon application of the positive longitudinal stress. This allows us to solve for v:

$$\begin{aligned} V &= cA\Delta t \\ \Delta V &= -vA\Delta t \\ v &= -\frac{\Delta V}{V}c \end{aligned} \qquad [4.7.A1.4]$$

Inserting this last result into Eqn [4.7.A1.3] gives

$$\Delta P_0 = \rho c^2\left(-\frac{\Delta V}{V}\right) \qquad [4.7.A1.5]$$

We can solve for c using the definition of the bulk modulus, to obtain

[4.7.A1.6] $$c = \sqrt{\frac{B}{\rho}}$$

THE SPEED OF SOUND DEPENDS ON TEMPERATURE AND THE MOLECULAR WEIGHT OF THE AIR MOLECULES

The bulk modulus, B, in Eqn [4.7.A1.6] is that defined in Eqn [4.7.A1.1]. It contains terms of dP, V, and dV. We have a relation that links these variables, and that is the ideal gas law: $PV = nRT$. However, we do not know the relation between P and V unless we know the temperature. The propagation of sound waves involves compression and expansion of gas. Compression of gas heats it and expansion of gas cools the gas. This is the basis for refrigeration and air-conditioning. Newton solved the problem of determining c by assuming that the temperature was constant. He reasoned that the propagation was so rapid that the temperature could not rise or fall. His results give the isothermal speed of sound, and it is wrong. Laplace made the argument that the propagation of sound occurred **adiabatically**, or without heat flow from or to the surroundings. What we need is the adiabatic relation between P and V.

The Molar Specific Heat at Constant *V* Is Not the Same as the Molar Specific Heat at Constant *P*

The specific heat of a substance is defined as the amount of heat energy required to raise the temperature of 1 g of the substance by 1 K. For gases, this is generally expressed in terms of moles rather than grams, and so it is called the **molar specific heat**. There are two common ways of determining the molar specific heat of gases: either by keeping the volume constant or by keeping the pressure constant. The molar specific heat at constant volume is denoted as C_V and the molar specific heat at constant pressure is denoted as C_P. These molar specific heats have the units of joules mol^{-1} K^{-1}.

The first law of thermodynamics is equivalent to the conservation of energy. We can write it as

[4.7.A1.7] $$dQ = dU + dW$$

where dQ is the increment of heat energy that enters a system from the surroundings, dU is the change in the internal energy of the system, and dW is work done *by* the system. In the case of a gas, the work it will do will be $dW = \int F dx = \int F/A \; A dx = \int P dV$. Expansion involves positive work; compression involves negative work (work is done *on* the system by the surroundings). Consider the case where volume is constant. Under these conditions, $dV = 0$ and so $dW_V = 0$. By the definition of C_V, we have

[4.7.A1.8] $$C_V = \frac{dQ_V}{n\,dT}$$

where dQ_V is the increment of heat energy at constant V, n is the number of moles of gas being heated isovolumetrically, and dT is the temperature increment. Solving for dQ_V and inserting into the conservation of energy, Eqn [4.7.A1.7], we get

[4.7.A1.9]
$$dQ_V = dU_V + dW_V$$
$$dQ_V = nC_V\,dT$$
$$dW_V = 0 \text{ (condition of isovolumetric heating)}$$
$$dU_V = nC_V\,dT$$

Now let us suppose that we heat the same amount of gas by the same increment, dT, but this time we do it at constant pressure rather than constant volume. Here the work increment, dW, will not be zero. As we heat the gas it will expand. Here the heat absorbed is given from the definition of the molar specific heat:

[4.7.A1.10] $$C_P = \frac{dQ_P}{n\,dT}$$

We insert this and the work term into the conservation of energy:

[4.7.A1.11]
$$dQ_P = dU_P + dW_P$$
$$dQ_P = nC_P\,dT$$
$$dW_P = P\,dV_P$$
$$dU_P = nC_P\,dT - P\,dV_P$$

Now the internal energy of a gas depends on the temperature alone and not on P or V singly. Thus for the same temperature increment, the internal energy increment is the same. Therefore, we can equate the last line in Eqns [4.7.A1.9] and [4.7.A1.11] to obtain:

[4.7.A1.12] $$nC_V\,dT = nC_P\,dT - P\,dV_P$$

Here we make use of the Ideal Gas Law, $PV = nRT$, which at constant pressure gives $P\,dV_P = nR\,dT$.

Substitution into Eqn [4.7.A1.12] gives

[4.7.A1.13] $$nC_V\,dT = nC_P\,dT - nR\,dT$$

which finally gives us

[4.7.A1.14] $$C_P - C_V = R$$

The value of the gas constant here should be in heat units: 8.314 J mol^{-1} K^{-1} or 1.99 kcal mol^{-1} K^{-1}. The reason Eqn [3.8.A1.14] is true is that all of the heat energy entering a gas at constant volume goes into raising the gas's temperature. When the gas is allowed to expand, some of this energy also goes into the work of expansion.

For Adiabatic Processes, PV^{γ} = constant, Where $\gamma = C_P/C_V$

We next consider an adiabatic exansion of a gas. Here **adiabatic** means that there is no heat transfer during

the process. The gas performs positive work during expansion, which is $dW = \int P\,dV$. The energy for this work comes from the internal energy. Thus, the internal energy decreases by the amount $dU = nC_V\,dT$. We use C_V here because we are writing only the internal energy change; we use C_P to include both the work term and the internal energy term. This should be clear from inspecting Eqns [4.7.A1.11] and [4.7.A1.12], from which we see that $dU_P = nC_V\,dT$. We insert this result into the First Law of Thermodynamics, Eqn [4.7.A1.7], to find

$$\begin{aligned} dQ &= dU + dW \\ 0 &= nC_V\,dT + P\,dV \end{aligned} \qquad [4.7.A1.15]$$

We can insert the Ideal Gas Law, $PV = nRT$, into this equation, separate the variables, and integrate:

$$\begin{aligned} 0 &= nC_V\,dT + \frac{nRT}{V}dV \\ nC_V dT &= -\frac{nRT}{V}dV \\ \int \frac{dT}{T} &= -\frac{R}{C_V}\int \frac{dV}{V} \\ \ln T &= -\frac{R}{C_V}\ln V + K \\ TV^{\frac{R}{C_V}} &= K_1 \end{aligned} \qquad [4.7.A1.16]$$

Since PV is proportional to T for an ideal gas, and $R = C_P - C_V$, we can rewrite the last equation as

$$\begin{aligned} PVV^{\frac{R}{C_V}} &= K \\ PVV^{\frac{C_P - C_V}{C_V}} &= K \\ PVV^{\frac{C_P}{C_V}}V^{-1} &= K \end{aligned} \qquad [4.7.A1.17]$$

The final relation is written as

$$PV^{\frac{C_P}{C_V}} = PV^{\gamma} = K \qquad [4.7.A1.18]$$

Inserting the Adiabatic Relation $PV^{\gamma} = K$ Gives the Speed of Sound as a Function of Temperature

Our three key equations are outlined above: they are the definition of the bulk modulus, B; the dependence of the speed of sound on the square root of B/ρ; and the relation between P and V during adiabatic expansion (and compression). These can be united by taking the total differential of Eqn [4.7.A1.18] and incorporating the result into Eqns [4.7.A1.1] and [4.7.A1.6]. The total differential of Eqn [4.7.A1.18] can be rearranged to obtain

$$\begin{aligned} \Delta PV^{\gamma} + P_{\gamma}V^{(\gamma-1)}\Delta V &= 0 \\ \Delta P = -PV^{-1}\Delta V \\ \frac{\Delta P}{-\Delta V/V} &= \gamma P \end{aligned} \qquad [4.7.A1.19]$$

This last equation gives an expression for B, the bulk modulus, in terms of the heat capacities and the pressure. Inserting it into our equation for the velocity of sound, we find

$$C = \sqrt{\frac{\gamma P}{\rho}} \qquad [4.7.A1.20]$$

Here we can substitute in again for P from the ideal gas law, $PV = nRT$. Recall that ρ is the density of the gas. This is its mass, m, divided by its volume, V. The mass in any volume is the number of moles, n, times the molecular weight, M, in g mol^{-1}. Substituting in $P = nRT/V$ and $\rho = nM/V$ into Eqn [4.7.A1.20], we get

$$C = \sqrt{\frac{\gamma nRT/V}{nM/V}} \qquad [4.7.A1.21]$$

$$C = \sqrt{\frac{\gamma RT}{M}} \qquad [4.7.A1.22]$$

This final equation shows that the speed of sound varies with the square root of the temperature and inversely with the molecular weight of the gas. For air at 0°C, $T = 273$ K, $M = 28.8$ g mol^{-1}, $R = 8.314$ J mol^{-1} K^{-1}, $C_P = 29.72$ J mol^{-1} K^{-1}, and $C_V = 21.41$ J mol^{-1} K^{-1}. Plugging in these values, we calculate a speed of sound to be 331 m s^{-1}. This is equal to the measured speed of sound in dry air at 0°C.

THE SPEED OF SOUND IS RELATED TO THE AVERAGE SPEED OF AIR MOLECULES

The equipartition theorem of thermodynamics states that each degree of freedom of movement has an equilibrium energy of $kT/2$, where k is Boltzmann's constant (the gas constant per molecule) and T is the temperature in K. In a diatomic molecule such as N_2, which is the major constituent of air, there are three translational degrees of freedom (the x-, y-, and z-directions) so that the average kinetic energy is $3kT/2$. In addition, the molecule can rotate about its center of mass with three axes of rotation. One of these, however, passes directly through the nuclei of the N atoms and so the energy associated with this rotation is negligibly small. Thus the remaining two degrees of rotation have a total energy of kT. The total energy is thus $5kT/2$. This is the equilibrium energy per molecule in these modes of movement. For a mole of material, the internal energy content is $U = 5RT/2$. The molar heat capacity for these diatomic molecules at constant volume (where all the heat energy is converted to internal energy) is $C_V = 5R/2$. Equation [4.7.A1.14] tells us that $C_P = C_V + R = 7R/2$. This simple calculation suggests that the molar heat capacity for air should be approximately $3.5 \times R = 3.5 \times 8.314$ J mol^{-1} K^{-1} = 29.10 J mol^{-1} K^{-1}.

The observed value, 29.72 J mol^{-1} K^{-1}, is close to this theoretical value.

In addition, if $C_V = 5R/2$ and $C_P = 7R/2$, then their ratio $\gamma = C_P/C_V = 1.40$. This is close to the 1.39 observed for air.

According to the equipartition theorem, then, the average kinetic energy of gas molecules should be given by

[4.7.A1.23] $$\frac{1}{2}m\mathbf{v}^2 = \frac{3}{2}kT$$

Multiplying through by Avogadro's number, $N_0 = 6.02 \times 10^{23}$ particles mol^{-1}, we convert m, the mass per molecule, to the molecular weight and we convert k, Boltzmann's constant, to R:

[4.7.A1.24] $$M\mathbf{v}^2 = 3RT$$

We can substitute in for RT in Equation [4.7.A1.22] to obtain

[4.7.A1.25] $$c = \sqrt{\frac{\gamma(M\mathbf{v}^2/3)}{M}}$$
$$c = \sqrt{\frac{\gamma}{3}}\mathbf{v}$$

The value of γ is about 1.4. This equation predicts that the speed of sound is on the same order of magnitude as the average velocity of air molecules. This makes intuitive sense, because it is the motion of the air molecules that propagates the sound.

THE INTENSITY OF SOUND IS PROPORTIONAL TO THE SQUARE OF THE PRESSURE

The intensity of sound can be derived in a variety of ways, usually by writing a wave equation for the propagating sound wave and summing the energy of the oscillating particles making up that wave, treating them all as harmonic oscillators. We can derive it perhaps more easily from the energy delivered by the piston that began the wave, realizing that, in this model with no dissipation of energy, the power in must be equal to the power out.

The intensity of sound is defined as the energy per unit area per unit time, or the power per unit area. It is usually given in units of W m^{-2} = J s^{-1} m^{-2}. The piston that produces the sound moved at a constant velocity, v, for the time interval Δt. The intensity of the resulting sound wave, transmitting the power of the piston without loss, is just the total energy delivered by the piston divided by the time and area. The total energy delivered is the force times the distance traveled in that interval, which is $v\Delta t$. The piston moves at a constant velocity but does not deliver a constant force because the gas is compressed more at the end than at the beginning. The energy transferred from piston to gas is

[4.7.A1.26] $$E = \int \Delta PA\, dx$$

where ΔP is the excess pressure, A is the area, and dx is the distance increment. This is just the force × distance work. The intensity, I, is given as

[4.7.A1.27] $$I = \frac{E}{A\Delta t} = \int \frac{\Delta PA\, dx}{A\Delta t}$$

The piston advances at a constant velocity, v, so that we can transform dx into dt:

[4.7.A1.28] $$x = vt$$
$$dx = v\, dt$$

Similarly, the longitudinal displacement, y, is linearly related to t. We can substitute in v dt for dx in Eqn [4.7.A1.28]. In addition, we can substitute in for ΔP from Eqn [4.7.A1.5] to get

[4.7.A1.29] $$I = \int_0^{\Delta t} \frac{\rho c^2(-\Delta V/V)V\, dt}{\Delta t}$$

Lastly, we substitute in for $V = cA\Delta t$ and $\Delta V = -vAt$. We use t here instead of Δt because the longitudinal displacement is varying with time as the piston moves. This gives

[4.7.A1.30] $$I = \int_0^{\Delta t} \frac{\rho c^2(Avt/Ac\Delta t)v\, dt}{\Delta t}$$
$$I = \int_0^{\Delta t} \frac{(\rho c/\Delta t)v^2 t\, dt}{\Delta t}$$

Evaluation of the integral gives

[4.7.A1.31] $$I = \frac{\rho c v^2}{2}$$

From our initial analysis of the pressure and velocity, from Eqn [4.7.A1.3], we have $v = \Delta P_0/\rho c$. Inserting this value for v, we get our final equation for the intensity:

[4.7.A1.32] $$I = \frac{{\Delta P_0}^2}{2\rho c}$$

Thus the intensity of sound depends on the square of the pressure amplitude. The intensity has units of J s^{-1} m^{-2} = W m^{-2}.

PROPAGATION OF SOUND CAN BE WRITTEN AS A WAVE EQUATION FOR WHICH THE INTENSITY IS PROPORTIONAL TO THE SQUARE OF THE PRESSURE AMPLITUDE

The transmission of a force to the air produces a longitudinal displacement of the gas molecules, which propagates through the air. We can write a wave equation for a traveling sinusoidal wave through a compressible medium like air as

[4.7.A1.33] $$y = A_0 \sin(\omega t - kx)$$

Here A_0 is the amplitude, ω is the angular frequency, in rad s^{-1}, and k is the wave number, in units of rad cm^{-1}. Thus $\omega = 2\pi f = 2\pi/T$, where f is the frequency in

cycles s^{-1} and T is the period of the cycle, in s $cycle^{-1}$. We can see from these relationships that each addition of T to the time t in the argument adds 2π to the argument of the sin, making a complete cycle. Similarly, the addition of a wavelength, λ, to the value of x in the argument also adds 2π to the argument of the sin; therefore, $k = 2\pi/\lambda$. Since the velocity of the wave is the frequency times the wavelength, we have

$$\begin{aligned} c &= \lambda f \\ &= \frac{2\pi}{k}\frac{\omega}{2\pi} \\ c &= \frac{\omega}{k} \end{aligned} \qquad [4.7.A1.34]$$

From the definition of the bulk modulus (Eqns [4.7.A1.1] and [4.7.A1.4]) we have

$$\Delta P = B\frac{v}{c} \qquad [4.7.A1.35]$$

The ratio of v/c can be discerned from Figure [4.7.A1.1]. The slope of y, the longitudinal displacement, as a function of x, the location, is given as

$$\frac{\partial y}{\partial x} = \frac{v\Delta t}{c\Delta t - v\Delta t} \approx \frac{v}{c} \qquad [4.7.A1.36]$$

since $v \ll c$. When v approaches or exceeds c, a new phenomenon results, the shock wave. The equations for sound are valid only when $v \ll c$. Thus Equation [4.7.A1.35] becomes

$$\Delta P = B\frac{\partial y}{\partial x} \qquad [4.7.A1.37]$$

Taking the partial derivative of Eqn [4.7.A1.33], we obtain

$$\frac{\partial y}{\partial x} = -kA_0 \cos(\omega t - kx) \qquad [4.7.A1.38]$$

where A_0 is the amplitude of the oscillation. From Eqn [4.7.A1.6], we write

$$B = c^2\rho \qquad [4.7.A1.39]$$

Inserting both of these last two equations into Eqn [4.7.A1.37], we arrive at

$$\Delta P = -kA_0c^2\rho \cos(\omega t - kx) \qquad [4.7.A1.40]$$

This equation can be rewritten as

$$\Delta P = -\Delta P_0 \cos(\omega t - kx) \qquad [4.7.A1.41]$$

where

$$\Delta P_0 = kA_0c^2\rho \qquad [4.7.A1.42]$$

Equations [4.7.A1.41] and [4.7.A1.33] show that the pressure wave and the longitudinal displacement are 90° out of phase: the pressure is greatest (and least) when the displacement is zero. We can use the relation $ck = \omega$ from Eqn [4.7.A1.34] to write

$$\Delta P_0 = \omega A_0 c\rho \qquad [4.7.A1.43]$$

The intensity of a sound wave is the energy per unit area per unit time. The unit area is oriented at a right angle to the direction of propagation of the wave. We can determine this by determining the energy in a single wavelength and dividing by the time it takes for that wavelength to be completed. This time is $T = 1/f$. Each particle in the sound wave can be viewed as undergoing simple harmonic oscillation at the angular frequency of the sound. It undergoes this oscillation in response to a mechanical stimulus which is related to the pressure. The energy of a simple harmonic oscillator is given as

$$E = \frac{1}{2}k'A_0^2 \qquad [4.7.A1.44]$$

where k' is an equivalent force constant. For mechanical systems, $k' = \omega^2 m$, where ω is the angular velocity in rad s^{-1} and m is the mass. In this case, all points on a single wavelength have the same total energy, and so the energy in one wavelength is the energy per unit mass times the total mass. The total mass of air that is moving is the density of the air times its volume. The volume is λA, where λ is the wavelength and A is the area. Thus, the total energy in a wavelength is given as

$$E = -\frac{1}{2}\omega^2\rho\lambda A A_0^2 \qquad [4.7.A1.45]$$

Now the intensity is the energy per unit area per unit time. Thus we divide E by A and by T to obtain the intensity:

$$I = \frac{E}{AT} = \frac{1}{2}\omega^2\rho\frac{\lambda}{T}A_0^2 \qquad [4.7.A1.46]$$

Recognizing that $\lambda/T = c$, we write

$$I = -\frac{1}{2}\frac{\omega^2\rho^2c^2A_0^2}{\rho c} \qquad [4.7.A1.47]$$

We write it in this way to better recognize the substitution of Eqn [4.7.A1.43] into Eqn [4.7.A1.47]. The result is the same as what we obtained before:

$$I = \frac{{\Delta P_0}^2}{2\rho c} \qquad [4.7.A1.48]$$

Note that the intensity of sound is independent of its frequency. This result is less apparent when the derivation was performed from the integration of energy delivered to the air, as opposed to the analysis of the resulting sound wave.