Basic Hearing - I

- Hearing is, like vision, a sensory system for “remote” sensing.
- In hearing, the proximal stimulus is sound, changes in air pressure over time.
- Pressure changes vary in their intensity (loudness) and their rate of occurrence - frequency (pitch).
- The ear contains specialized mechanisms to amplify sound and convert it into a neural code.
- The auditory pathway and auditory cortex contain cells to extract basic features from this neural code.

Hearing and Speech

- The auditory pathway (ear to cortex) is a relatively old system. It is similar in the mammals.
- The speech/motor system for articulation and related physiology is a "young" system.
- The evolution of speech probably represents a modification of the articulatory/control system to produce a signal robust to noise and matched to the auditory system.

The Role of Hearing

We can hear events that we can not see. Thus, we might hear a predator before we saw it -or- hear our prey before we saw it. We can also tell where a sound is coming from.

Hearing (and speaking) are the primary means for communication via language. In humans, language is an essential tool for communication.

Hearing is also essential for the perception of music. The role of music in human cognition and culture is complex and beyond our scope in this course.

Outline

In considering the basics of hearing, we will cover:

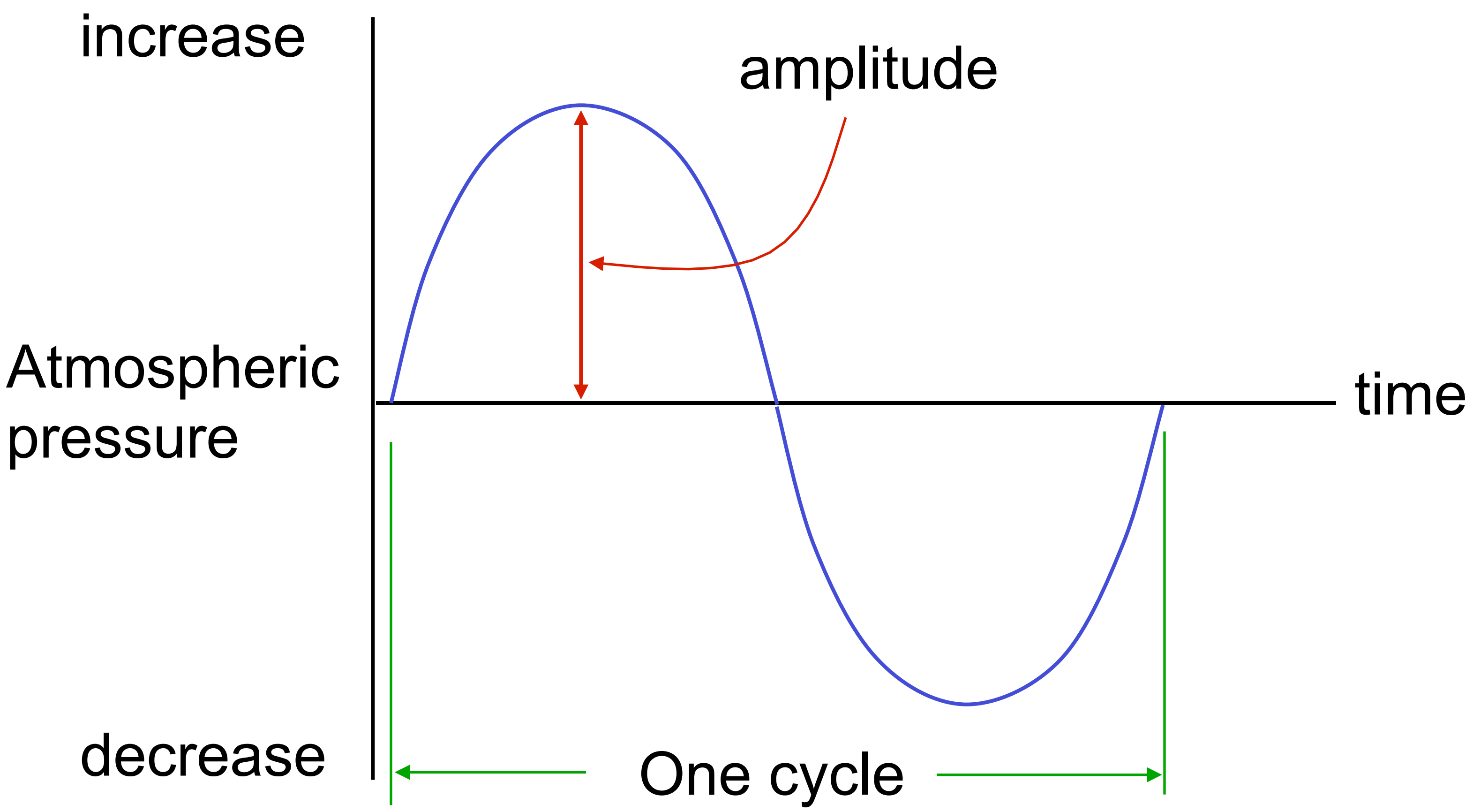
1) The nature of sound
2) The psychological dimensions of hearing
3) The physiology of the auditory system
    a) The ear (outer, middle, inner)
    b) The auditory pathway and auditory cortex
4) The neural coding of pitch

Sound

Sound is a change in air pressure over time. If the change in pressure occurs at a rate between 20 and 20,000 times per second, we can hear it. Pressure changes that take place more rapidly than 20,000 per second are ultra-sound.

When an object moves, it displaces the air around it. If the air motion is repetitive, 20 to 20k times per second, then we can hear it. The particular qualities of this repetitive air motion depend upon the object that produced the air motion. Thus, the qualities of the sound tell us about the object that produced it.

Sound - Diagram of a Pure Tone



Sound - Physical Dimensions

The physical dimensions of sound are amplitude (intensity), frequency, and phase.

Amplitude - The magnitude of the pressure change.

Frequency - The number of cycles in pressure change per unit time. For time in seconds, the number of cycles per second is called Hertz (Hz).

Phase - The starting point of the air pressure. In the example, the starting point is atmospheric pressure. It could be any point between the maximum and minimum pressure. The starting point is specified in degrees (as if it were an angle). The example has a phase of 0º.

Physical Dimensions - Amplitude

Amplitude is measured as the logarithm of pressure relative to a reference pressure.

Humans can hear sounds over a very wide range of amplitudes. If we make the zero point on our scale the threshold for hearing, then a normal conversation is 1,000 times higher in amplitude, loud sounds are 10,000 to 100,000 times higher and the threshold for pain is 1,000,000 times higher. This number scale is difficult to use, so we convert to a scale called decibels.

The Decibel Scale

Sound amplitude, in decibels (dB) is determined by the equation:

$$dB = 20 \log p/p_0$$

In this equation, p_0 is the reference pressure. For the SPL scale, this is 20 micropascals. This is approximately the threshold for a 1,000 Hz tone in a young adult with normal hearing. A sound with a pressure of 20 micropascals would have an amplitude of 0 dB.

A normal conversation, at 1,000 times this pressure, would be at 60 dB. Sound above 80 dB is considered loud. A jet engine, at full thrust, would be about 140 dB.

The Decibel Scale - 2

The SPL (Sound Pressure Level) scale is one intensity scale. There are others, including dBA and dBC. They are all logarithmic scales, computed like the dB SPL scale. They differ with respect to the pressure used as the reference pressure (p_0) and whether the same reference pressure is used for all frequencies.

Converting between scales is not trivial and we will use the SPL scale for all measurements.

Physical Dimensions - Frequency

Most sounds are not sine-waves (pure tones). Instead, they have many different frequencies. Sounds whose pressure changes over time are "random" are heard as noise. Sounds whose pressure changes over time are regular (recurring) are called periodic.

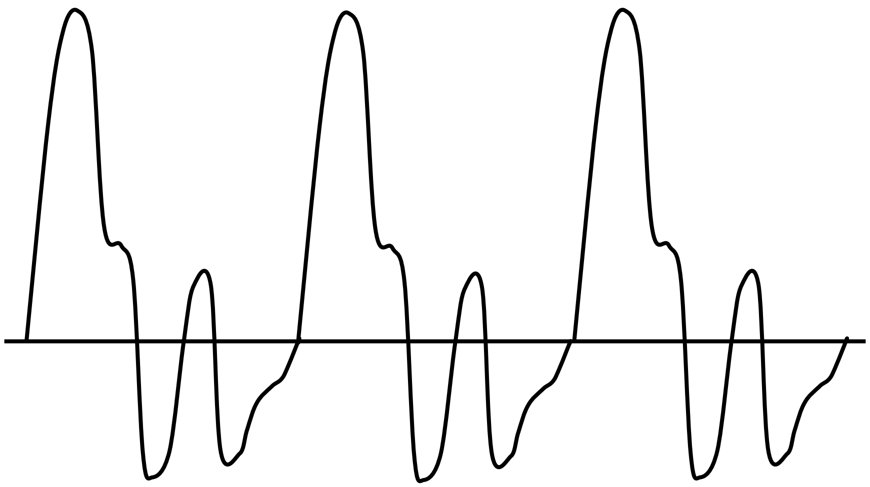
In both cases, these sounds can be thought of as composed of many pure-tones with different amplitudes and phases that have been added together. Fourier analysis can be applied to sounds to break them down into their component parts.

Frequency - 2

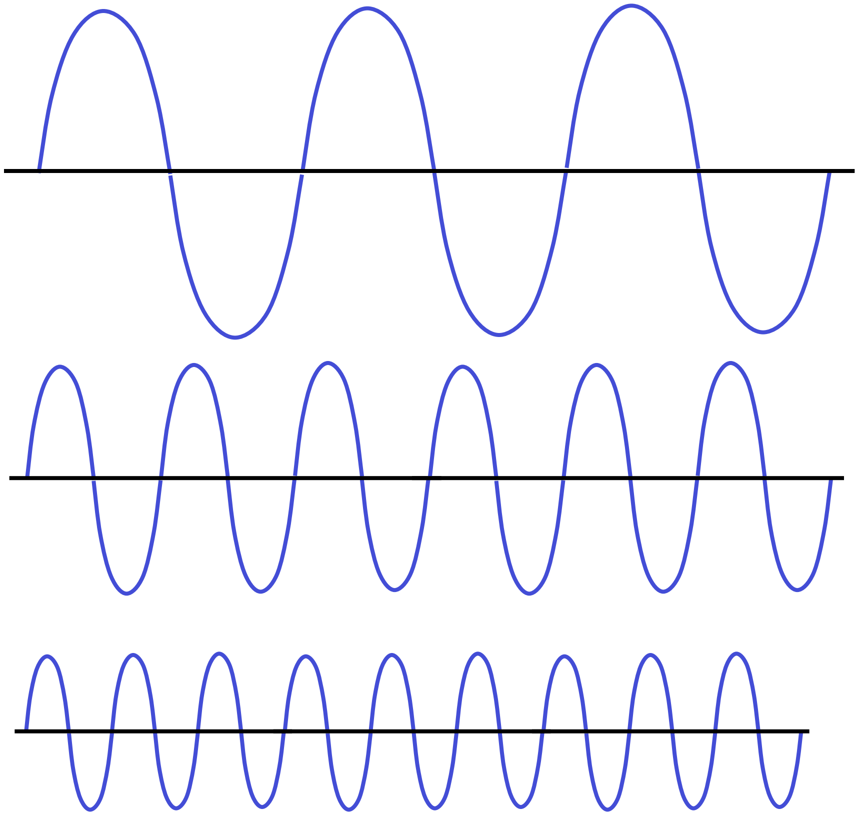
If we record a flute playing the note A4 and then use Fourier analysis to de-compose it, we discover that most of the energy is at three frequencies: 440 Hz, 880 Hz, and 1320 Hz.

The lowest frequency of this periodic sound is called the fundamental frequency. In this musical note, the other two frequencies are integer multiples of the fundamental and are called harmonics.

Frequency - 3



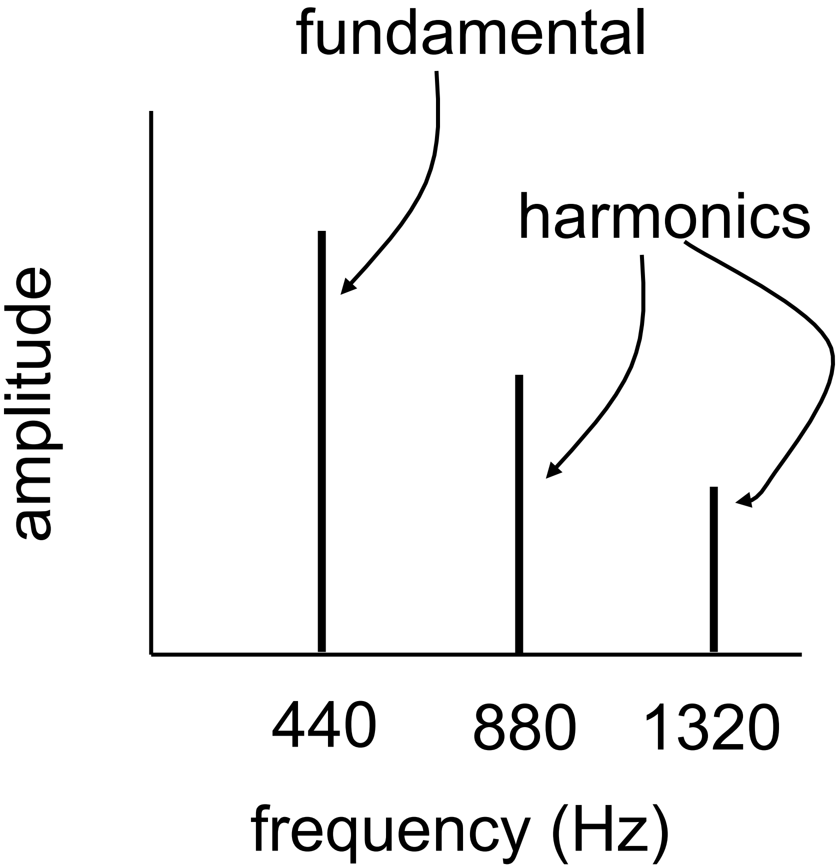
Complex waveform



Component sine-waves

Frequency - 4

The graph at right shows the amplitude of each of the component frequencies in our A4 note on the flute.



Frequency - Summary

Any complex sound can be de-composed into its component frequencies. This includes environmental sounds, musical notes, and speech.

When we examine the workings of the ear, we will discover that one of the basic operations performed by the ear is to de-compose sound into its component frequencies.

Sound - Psychological Dimensions

The physical dimensions of amplitude and frequency correspond to the psychological dimensions of loudness and pitch. In addition, the dimension of timbre becomes important for complex sounds.

Loudness

Threshold	0
Leaves rustling	20
Quiet residential street	40
Conversation	60
Heavy traffic	80
Express subway train	100
Jet engine at takeoff	140
Space Shuttle	180

Here are the intensities (in dB SPL) for some sounds in our environment. Note that the pain threshold is a bit below the intensity of the jet engine.

Pitch

Pitch is the psychological dimension that corresponds to frequency. As we move up the piano keyboard, the notes have increasing pitch.

This dimension is not “linear” like frequency measured in Hz. In general, human abilities to distinguish between different frequencies is proportional to the frequency. Put another way, perception is more fine grained at low frequencies and coarser at higher frequencies.

For frequencies from 20 – 500 Hz (about), pitch perception is linear with frequency. For 500 – 20 kHz, resolution is proportional to the frequency.

Pitch - 2

For music, we break pitch down into two qualities: tone height and tone chroma.

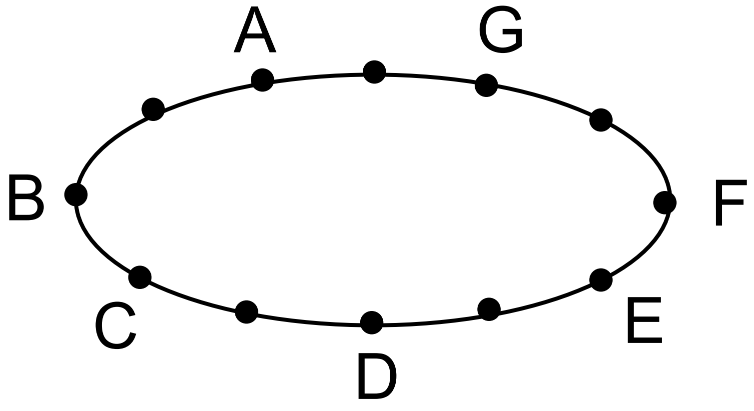
On the keyboard, the notes are labeled: A, B, C, D, E, F, and G. Then the letters repeat. The increase in pitch with the letters (moving up the keyboard) is tone height. The letters repeat because the notes that they represent sound similar. The place of the tone within the musical scale is tone chroma.

Octaves

When two notes have the same tone chroma as one another, the fundamental frequency of one is an integer multiple of the other. That is, these similar sounding notes are separated by one or more octaves. An octave is a doubling of the frequency.

On the western musical scale, note A3 has a fundamental of 220 Hz, A4 is 440 Hz, and A5 is 880 Hz. These notes all have the same tone chroma even though they differ in tone height.

Musical Scale



One way to think of this is that the musical scale is a spiral. Notes at the same point around, but on different layers, have the same chroma. This is a psychological dimension. Notes with the same chroma sound similar.

Timbre

Two sounds with the same fundamental frequency have the same pitch. However, they can still sound quite different. This is a difference in timbre.

For example, if a piano, violin and trumpet all play A4, the notes have the same pitch, but sound quite different. This difference is the psychological dimension of timbre.

We will return to this later when we consider the perception of complex sounds.

The Ear

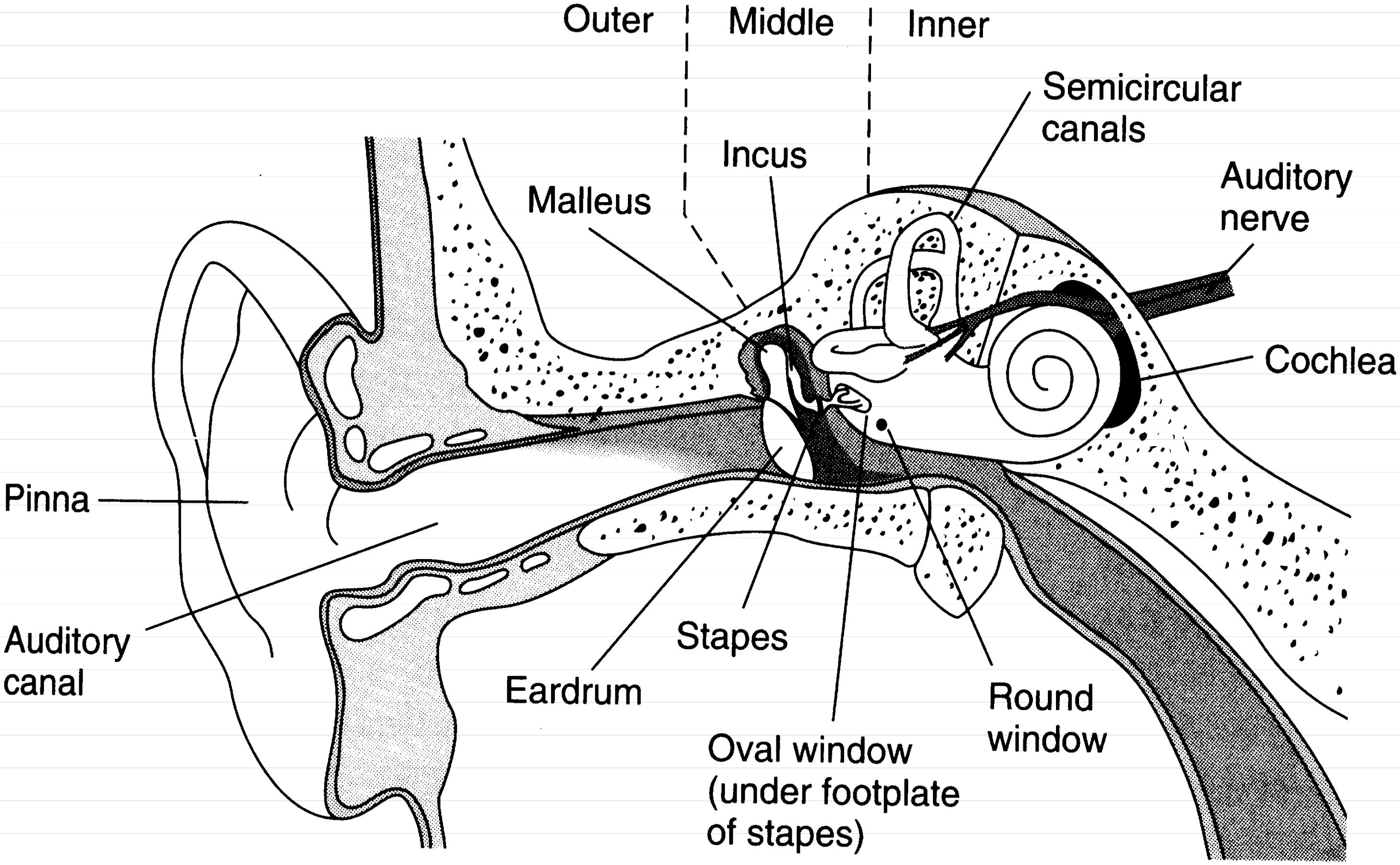
We divide the ear into three parts: outer, middle and inner.

The outer ear consists of the pinna and the auditory canal which ends at the eardrum (tympanic membrane).

The purpose of the pinna is, in part, to help us locate sounds in space.

The ear canal, about 3 cm long, amplifies some sound and allows the delicate tympanic membrane to be protected from the outside world.

Ear Diagram


Outer
Middle
Inner
Semicircular canals
Incus
Malleus
Auditory nerve
Cochlea
Pinna
Auditory canal
Stapes
Eardrum
Round window
Oval window (under footplate of stapes)

Middle Ear

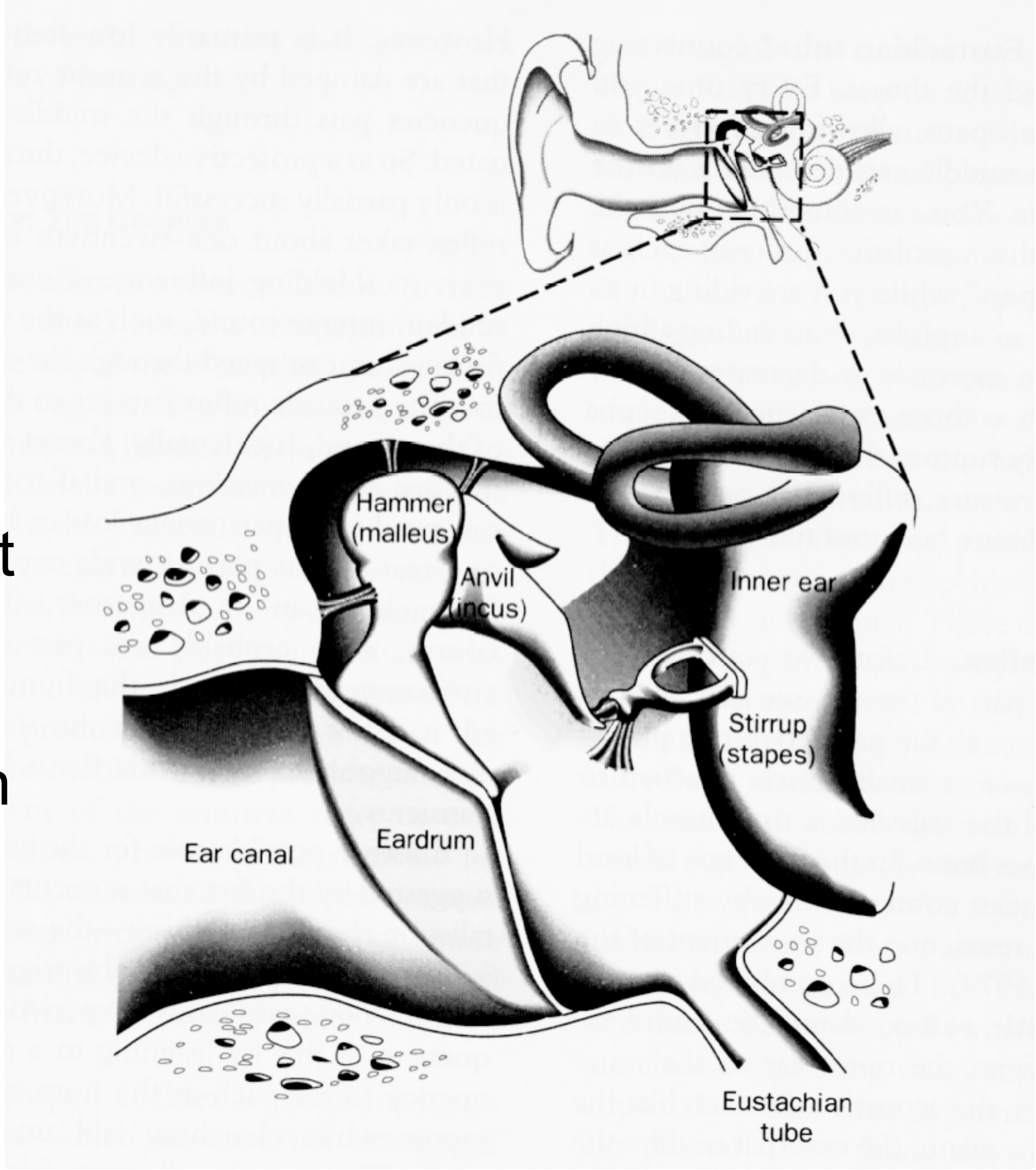
The middle ear contains three small bones. These bones connect the tympanic membrane (eardrum) to the inner ear.

The bones act as a lever system. This serves to amplify sound. They also take vibration on a large surface (the tympanic membrane) and concentrate it on a small surface.

Together, these two aspects of the middle ear amplify the sound by a factor of 22 to 1 or more. This amplification is necessary to set the fluid of the inner ear in motion (the auditory receptors and inner ear are filled with fluid).

Middle Ear Diagram

The middle ear also contains muscles that can act to restrict the motion of the three tiny bones to dampen loud sounds and protect the ear from damage.



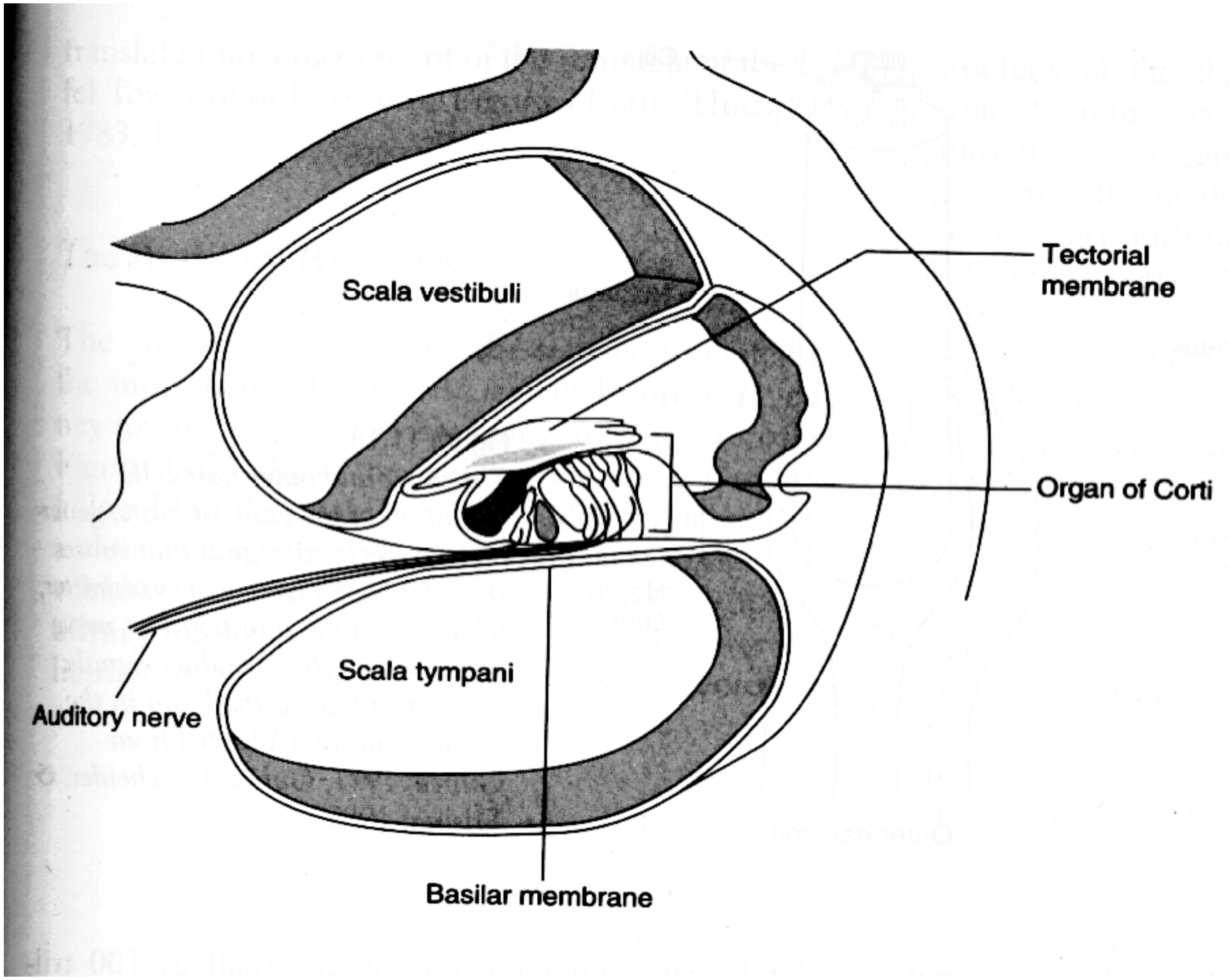
Inner Ear

The inner ear is filled with fluid. It has a coiled shape, like a snail. In cross-section, it has three chambers: the scala vestibuli and the scala tympani are separated by the cochlear partition.

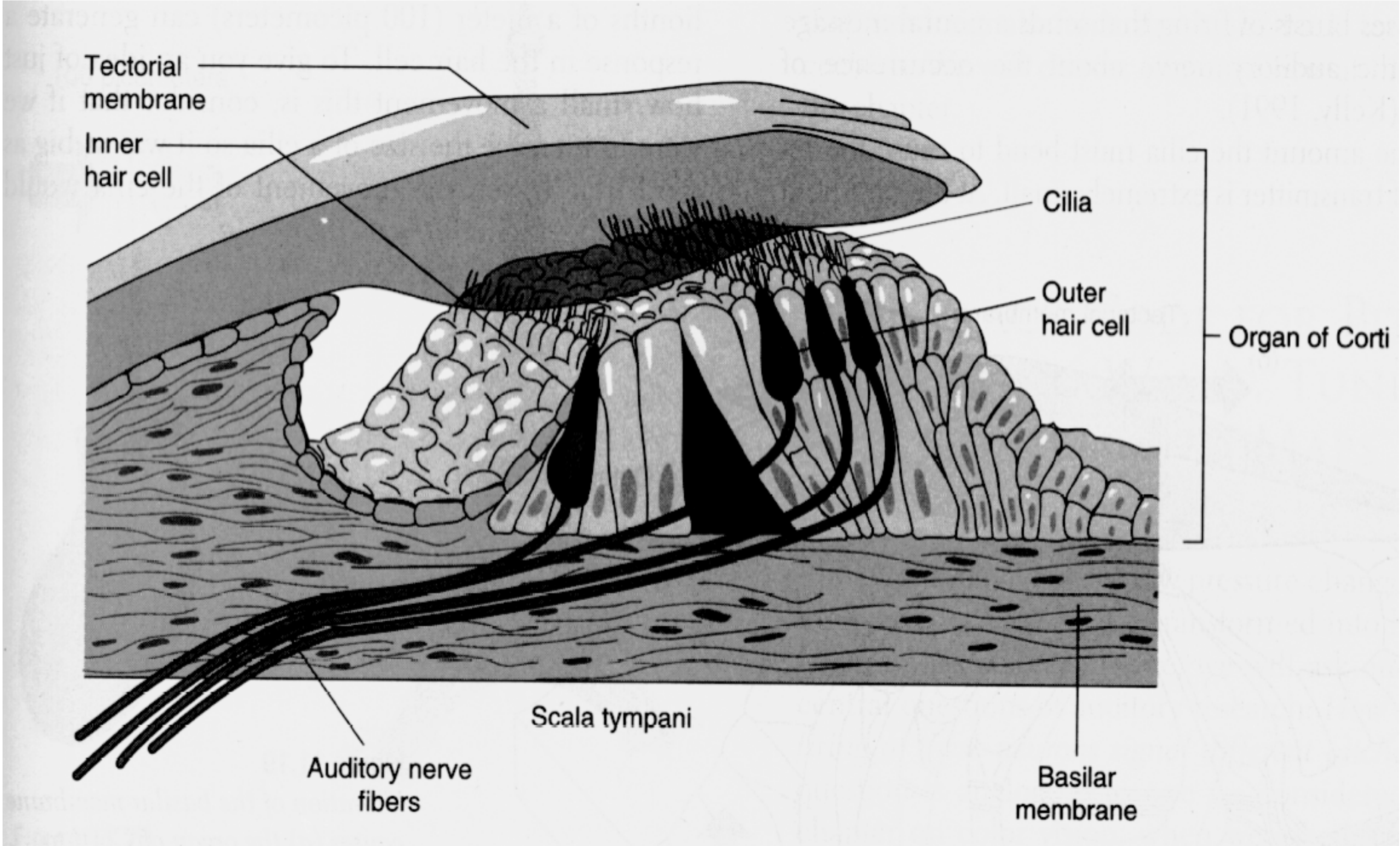
Within the cochlear partition is the organ of Corti. The organ of Corti sits on top of the basilar membrane. The tectorial membrane sits on top of the organ of Corti.

Within the organ of Corti are the hair cells which are the receptors for hearing. There are two sets of hair cells: inner hair cells and outer hair cells.

Inner Ear Cross Section



Inner Ear Cross Section - 2



Receptors

The hair cells have cilia whose ends make contact with the tectorial membrane. When sound enters the inner ear, it sets the fluid in motion. This causes the basilar membrane to move up and down. As it does, the tectorial membrane moves back and forth.

The motion of the tectorial membrane causes the cilia of the inner and outer hair cells to bend. This changes the hair cell electrical activity and causes the neurons that synapse on the hair cells to fire.

Auditory Pathway

The axons of the neurons leaving the cochlea form the auditory (8th) nerve. They project to the cochlear nucleus. Neurons from the cochlear nucleus project to the superior olivary nucleus in the brain stem. In turn, these neurons project to the inferior colliculus in the midbrain. These neurons project to the medial geniculate nucleus (part of the thalamus).

MGN neurons project to the primary auditory cortex (A1). Next to this is the secondary auditory cortex (A2) and around the auditory cortex are other regions involved in auditory perception. Auditory cortex is in the temporal cortex.

Neural Coding

How does the operation of the inner ear result in a neural code for pitch and loudness? What is the neural code for pitch and loudness.

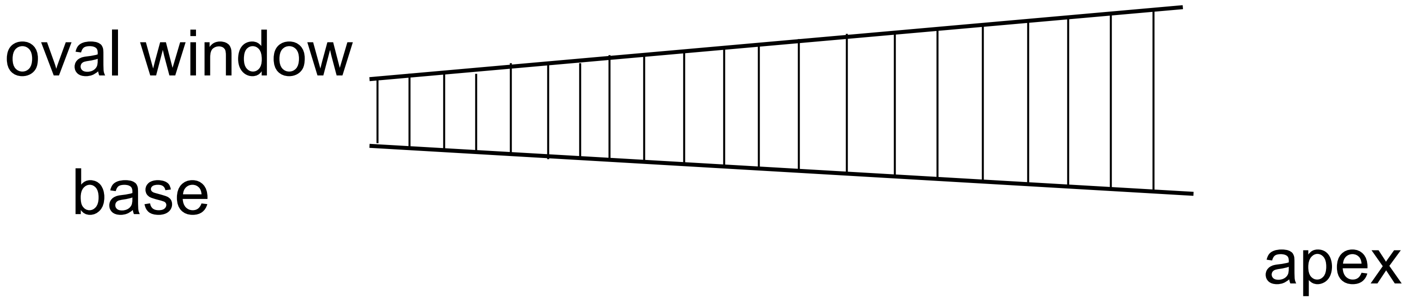
First, consider pitch. There are two basic theories: place coding and frequency coding.

According to place coding, different parts of the basilar membrane (and different hair cells) code different frequencies.

According to frequency coding, neurons fire once for each cycle (of amplitude change) in the sound.

Place Coding

Helmholtz proposed that the basilar membrane was made up of transverse fibers that vibrated to different frequencies. The short, thin fibers near the base would respond to high frequencies while the long, wide fibers near the apex would respond to low frequencies. Hair cells at different places on the membrane would thus respond to different frequencies.

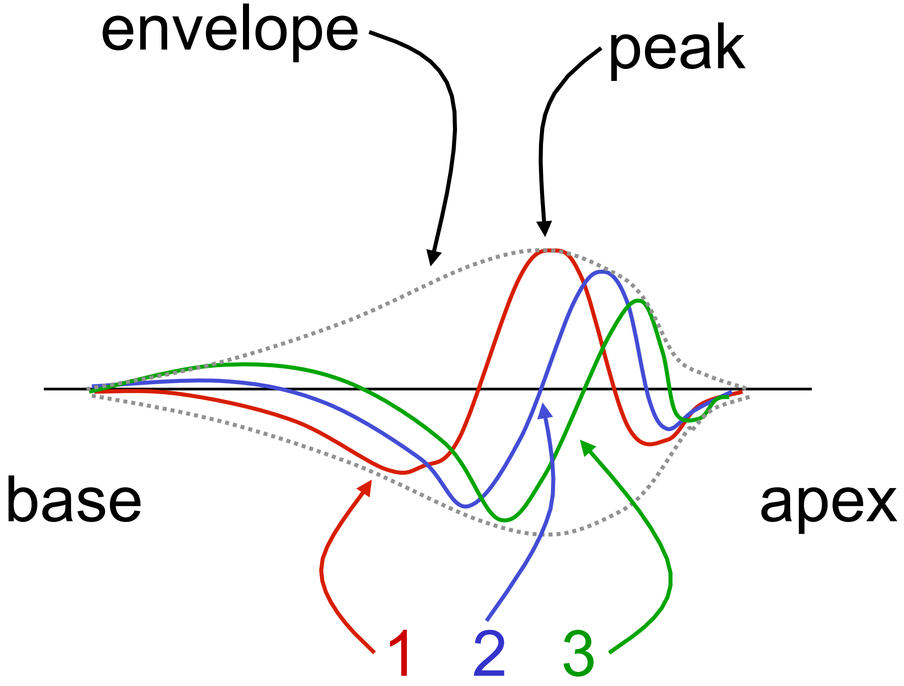


Place Coding - 2

Helmholtz was partly correct. However, the motion of the basilar membrane is more complex than described by place theory.

The vibration of the basilar membrane was described by Békésy as a traveling wave. According to this view, sound may cause a substantial part of the basilar membrane to vibrate, with the vibration moving down the membrane from the base to the apex. However, for every frequency, there will be a point on the membrane with a maximal vibration. The hair cells at this point will have the “largest” response. This is the basis for a place code for frequency.

Traveling Wave



As the vibration starts at the base and travels down the membrane, as shown at times 1, 2, and 3. The envelope represents the maximum movement at each point on the membrane. The peak in the envelope is the point of maximum movement on the whole membrane.

Traveling Wave - 2

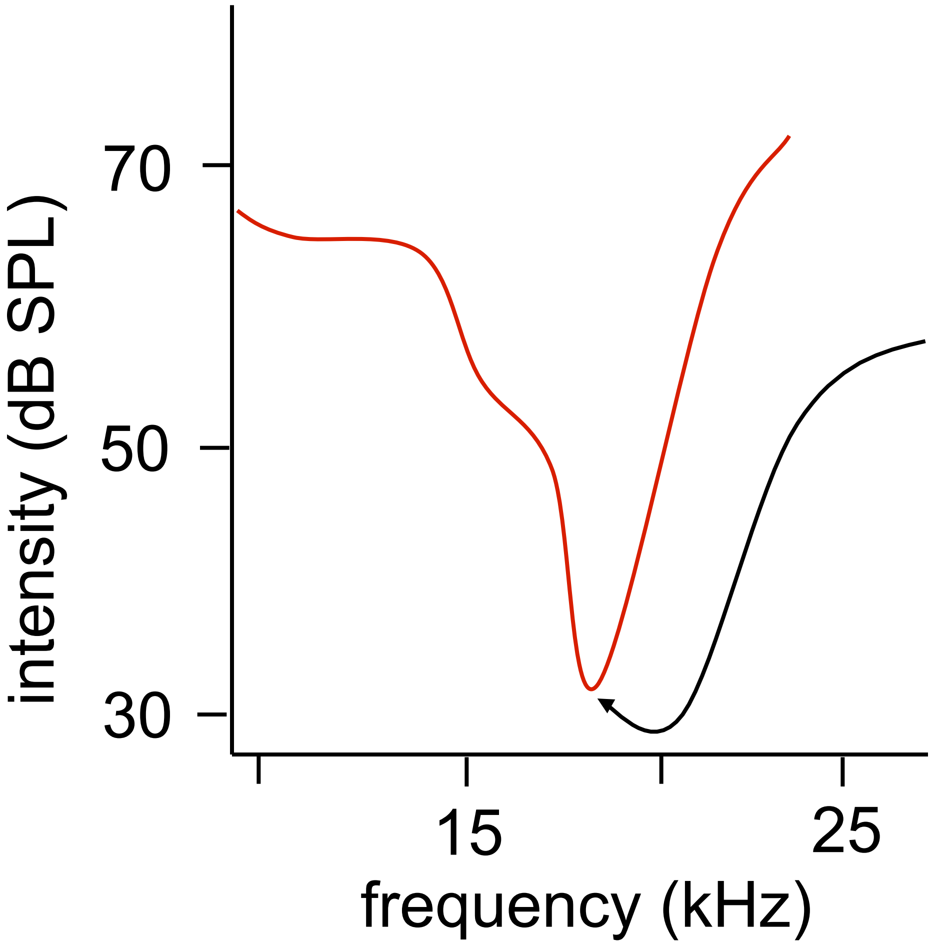
For different frequencies, the peak in the traveling wave occurs at different points on the basilar membrane. For high frequencies, is is near the base. As frequencies get lower, is moves toward the apex. For low frequencies, it is at the apex.

Physiological Evidence

1) If we record from different points along the cochlea, we find maximum response for different frequencies occurs at different places.

2) If we record from a single inner hair cell and measure the intensity required to elicit the same, small response at each frequency, we get the tuning curve for this cell. The point at which the lowest intensity elicits this response is the characteristic frequency for this receptor.

Tuning Curve - 1



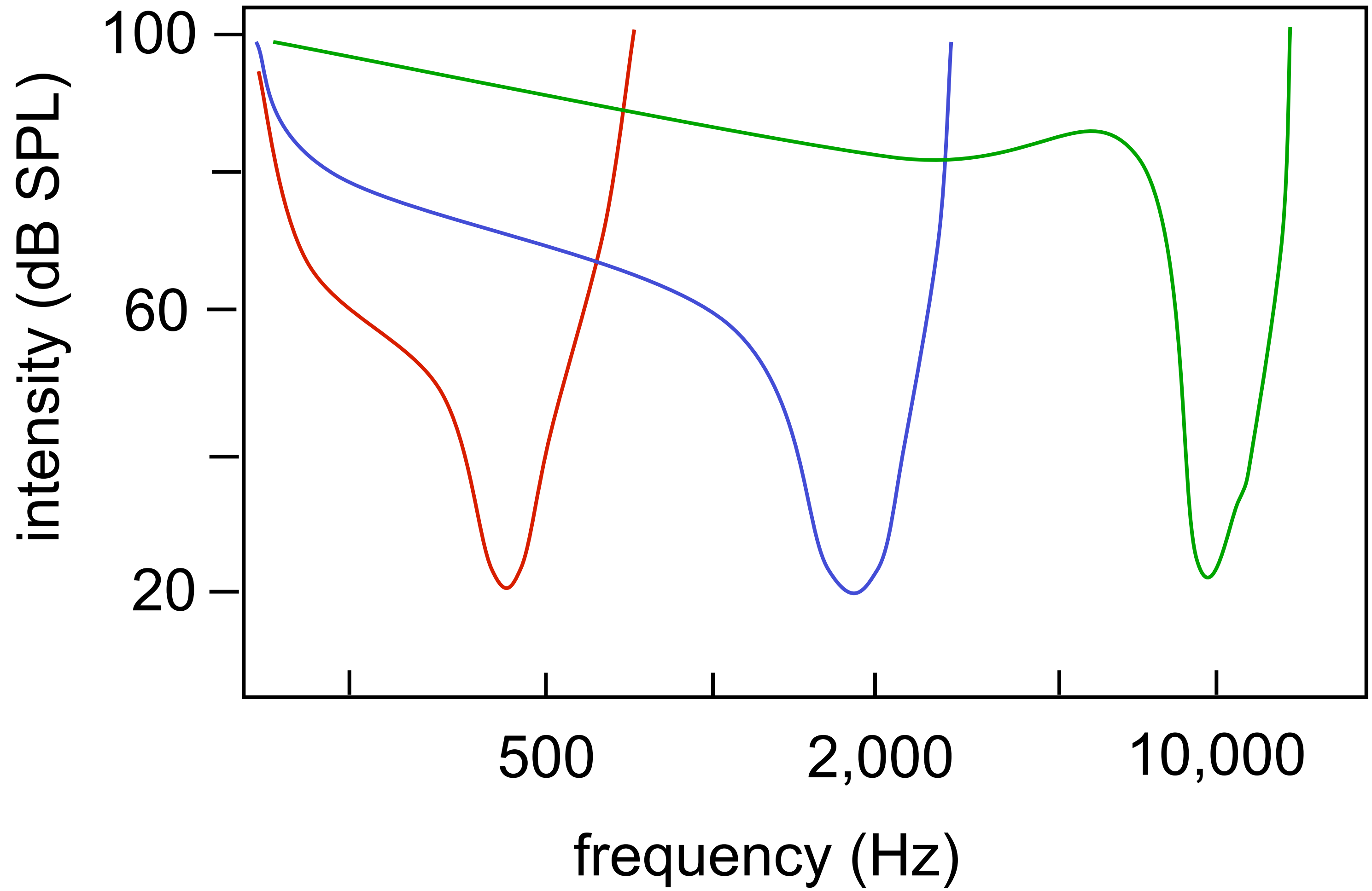
For this hair cell, the characteristic frequency is about 18 kHz.

Tuning Curve - 2

3) For hair cells at different points along the basilar membrane, the tuning curves show different characteristic frequencies. Low frequencies toward the apex and high frequencies toward the base.

4) In humans we can do a psychophysical experiment. We play a test tone of a fixed frequency just above the threshold. Simultaneously, we play a second tone (the masker). We vary the masker intensity to make the test tone barely audible (threshold). We determine this masker intensity for different frequency masker tones.

Psychophysical Tuning Curves



Outer Hair Cells

The outer hair cells act to sharpen the tuning curves of the inner hair cells. These cells also show place coding. When stimulated, the cilia tilt and get longer. This, in turn, seems to modify the motion of the basilar and tectorial membranes. This change seems to sharpen the traveling wave and thus narrow the region of maximal response.

Place Coding Summary

The basilar membrane and the hair cells act as a frequency analyzer. For complex sounds, each frequency present in the sound is coded by a peak in the traveling wave and the response of the inner hair cells at that place along the basilar membrane.

The A4 note of the flute with a fundamental of 440 Hz and harmonics at 880 and 1320 Hz produces neural activity in fibers with characteristic frequencies of 440, 880, and 1320 Hz.

Frequency (Timing) Coding

In the 1880’s, Rutherford proposed frequency coding. The basic idea was that neural units responded once for each cycle. The frequency of neural firing signaled the frequency of the sound.

This means that units would fire 3,000 times per second for a 3,000 Hz tone. Since neural units can not fire more rapidly than about 500 times per second, this is clearly not workable.

Volleying

A variation of frequency coding was proposed by Wever and Bray. The idea is that while individual neurons can not fire faster than 500 times per second, a group of neurons could have their aggregate firing rate be higher. If different neurons fired to different cycles, then as a group, at least one would fire to each cycle. This is **volleying**.

Phase Locking

A second key idea is **phase locking**. The central idea is that when a neuron fires to a sound, it fires at a particular phase of the signal (e.g., when the pressure goes from maximum to minimum).

Combined, the ideas of volleying and phase locking would allow a group of neurons to signal frequency by their firing rate. Recordings from the auditory nerve show evidence of volleying up to about 5,000 Hz. Also, neurons do show phase locking up to about 5,000 Hz.

The perception of pitch is partly based on frequency coding.

Periodicity Pitch

If we play a combination of tones with frequencies of 300, 600, 900, and 1200 Hz, our listener will match the pitch to that of a 300 Hz tone. Listeners match pitches based on the fundamental frequency.

What frequency would they match to a combination of 600, 900, and 1200 Hz?

They also match this to 300 Hz. They report it as having the same pitch as the 300, 600, 900, and 1200 Hz tone combination but with a slightly different timbre.

Periodicity Pitch - 2

Presenting a 1,600 Hz tone to the left ear and a 1,700 Hz tone to the right ear results in a pitch percept of 100 Hz.

In both of these cases, listeners report a pitch percept when there is no energy present at the fundamental frequency. This is the missing fundamental or periodicity pitch.

Periodicity Pitch - 3

These two results pose a problem for a place theory of pitch perception. The lack of energy at the fundamental means that the place on the basilar membrane corresponding to the fundamental is not being stimulate.

The creation of this pitch by tones presented to different ears means that the effect is taking place in the auditory pathway after information from the two ears comes together.

This means that while the code for pitch may be based on place and frequency coding, the perception of pitch occurs at a higher level that analyzes the pattern of harmonics.

Periodicity Pitch Applications

The fundamental for a male speaker is generally in the 100 - 160 Hz range. The fundamental for a female speaker is generally in the 160 - 280 Hz range. However, the "telephone" only transmits frequencies of 300 to 3,000 Hz. How then do we accurately perceived the pitch of an adult speaker's voice?

To make a compact pipe organ, a 55 Hz note is "produced" by combining 110 and 165 Hz notes. The brain interprets the combination as the harmonics of 55 Hz.

Neural Coding Above the Cochlea

In general, cells in each area are arranged tonotopically. Cells that respond to similar frequencies are near one another.

At the level of the cochlear nucleus and the superior olivary nucleus, we find cells that phase lock to the sound. At the auditory cortex, there is relatively little phase locking above 500 Hz.

At the cochlear nucleus, cells respond to pure tones. At the level of the cortex and beyond, more complicated response properties are found.

Two Ears

Starting at the superior olivary nucleus, the auditory regions all receive input from both ears. The pathway ascending on the same side is the ipsilateral pathway. The pathway that crosses from the right ear to the left side (and vice versa) is the contralateral pathway. The interconnections between the two pathways are important to our ability to locate sound in space.

Auditory Cortex

In the auditory cortex, a range of different response properties are found.

1) Frequency sweeps. Movement of a frequency from low to high. Other cells respond for high to low sweeps.

2) Combinations of tones.

3) Tones but only when part of a sequence.

4) In squirrel monkeys, cells that respond to con-specific calls. The analogous brain region of the temporal cortex in humans is involved with speech.

Auditory Cortex - 2

In humans, neuro-imaging shows regions of the brain specialized for spoken language and involved in melody perception. For non-musicians, the melody perception seems to take place more on the right side of the brain. For speech, it is the left side of the brain.

The Efferent Pathway

In addition to the afferent or ascending pathway, there is a descending or efferent pathway. It projects all the way from the cortex back to the hair cells.

1) This may serve as a form of “gain control”, allowing us to hear over a wide range of intensities.

2) It may be involved in helping us separate a particular sound from the background. For example, listening to one conversation in a room with other conversations.