The Human Ear — Hearing, Sound Intensity and Loudness Levels

We've been discussing the ***generation*** of sounds, so now we'll discuss the ***perception*** of sounds.

Human Senses:

The astounding ~ 4 billion year evolution of living organisms on this planet, from the earliest single-cell life form(s) to the present day, with our current abilities to hear / see / smell / taste / feel / *etc*. – all are the result of the evolutionary forces of nature associated with "survival of the fittest" – *i.e*. it is evolutionarily{very} _beneficial_ for us to be able to hear/perceive the natural sounds that do exist in the environment – it helps us to locate/find food/keep from becoming food, etc., just as vision/sight enables us to perceive objects in our 3-D environment, the ability to move /locomote through the environment enhances our ability to find food/keep from becoming food; Our sense of balance, via a stereo-pair (!) of semi-circular canals (= inertial guidance system!) helps us respond to 3-D inertial forces (*e.g*. gravity) and maintain our balance/avoid injury, *etc*. Our sense of taste & smell warn us of things that are bad to eat and/or breathe…

Human Perception of Sound:

* The human ear responds to disturbances/temporal variations in <u>pressure</u>. Amazingly sensitive! It has more than 6 orders of magnitude in dynamic range of pressure sensitivity (12 orders of magnitude in sound intensity, $I \propto p^2$) and 3 orders of magnitude in frequency (20 Hz – 20 KHz)!
* Existence of 2 ears (stereo!) greatly enhances 3-D localization of sounds, and also the determination of pitch (*i.e*. frequency resolution)!
* <u>Pinpoint</u> accuracy for 3-D localization of sounds in the $f \sim 100\ Hz \rightarrow \sim 1.5\ KHz$ range; good sound localization accuracy up to ~ few KHz, and ~ reasonable, below ~ 100 Hz!
* Mechanical & auditory sensory structure of ear preserves/***is*** sensitive to/utilizes ***phase*** information over the $f \sim 100\ Hz \rightarrow \sim 1.5\ KHz$ frequency range.
* Our brains process/use frequency/timing, amplitude/loudness ***and*** phase information in different frequency ranges for enhanced/improved localization of sound sources…

The Human Ear has Three Basic Parts:

* **Outer Ear** – pinna - concentrates sound waves into the ear canal (*aka* meatus)

* **Middle Ear** – eardrum (tympanium) transforms pressure variations into mechanical displacements ($p = F/A$); the ossicles (hammer, anvil, stirrup = malleus, incus, stapes) also mechanically <u>amplify</u> the sounds!

* **Inner Ear** – cochlea (& semi-circular canals – for balance/orientation) hair cells convert pressure signals into neural signals, send them to _various_ centers in brain for processing via auditory nerve(s)
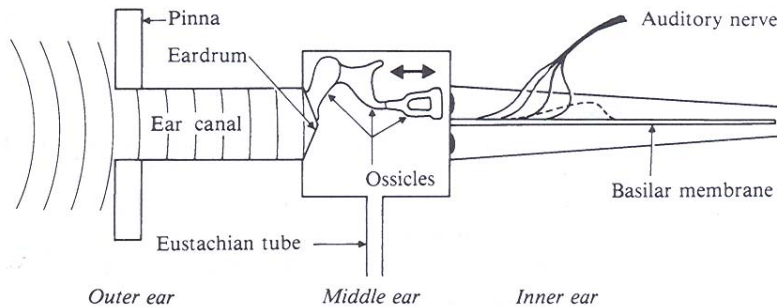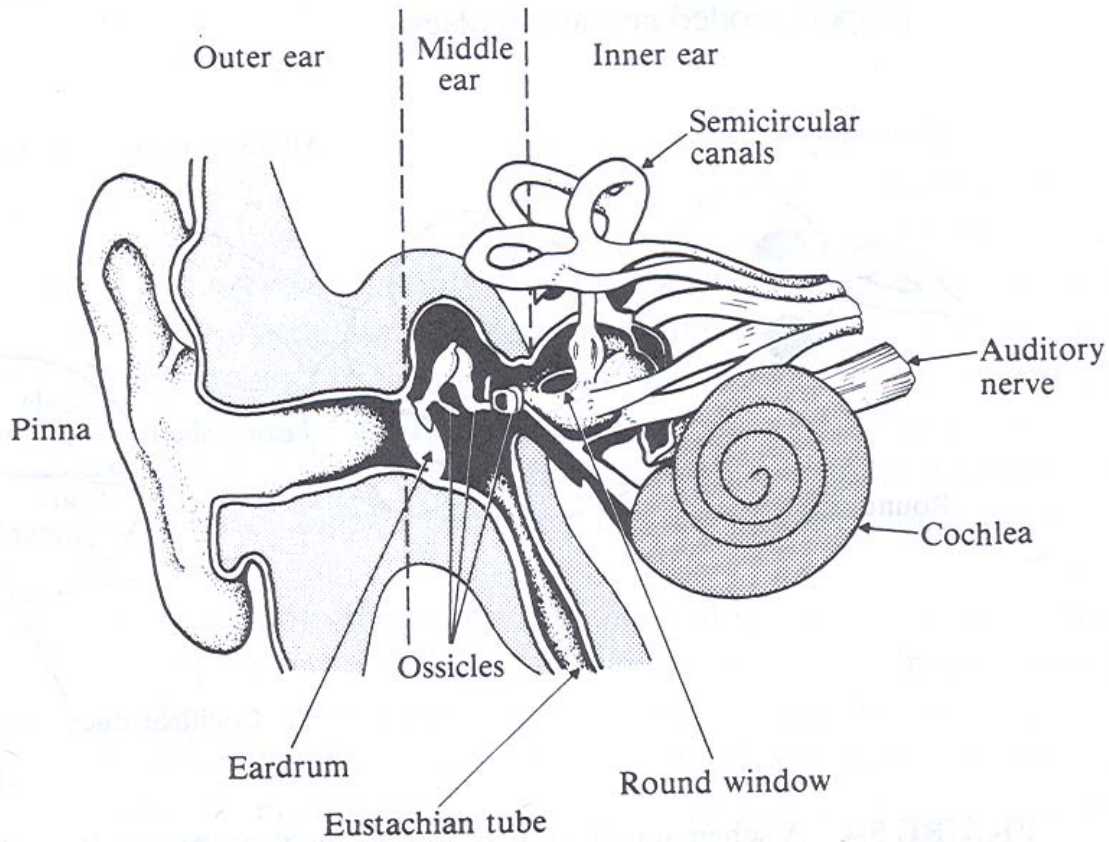
The Outer Ear:
* Pinna (ear flap) concentrates/focuses sound waves into ear canal.
* We used to have **_moveable_** ears (like donkeys), not much moveable, nowadays!!!
* Ridges/ruffles in pinna – unique to _you_ (!) improves sound localization at higher frequencies!

* Auditory/ear canal is $L \sim 3$ *cm* long, closed at the inner ear at the eardrum – a membrane.
  Auditoy canal = organ pipe ($L \sim 3$ *cm* long), with one end open and one end closed
  (Thus, there will be standing-wave <u>resonances</u> in the ear canal at: $f_n \sim nv/4L, \quad n = 1, 3, 5, 7...$
    $\Rightarrow$ Boosts our hearing sensitivity in the $f \sim$ 2-5 KHz frequency range!!!).

The Middle Ear:
* Ossicular chain – hammer, anvil, stirrup – transmit sound vibrations on ear drum to oval
  window of cochlea. Ossicles are a lever system, w/ $\sim 1.3\times$ mechanical advantage.
* Eardrum is $\sim 20\times$ bigger area than oval window – gives overall amplification factor of $\sim 26\times$!
* Performs important function of *impedance-<u>matching,</u>* efficiently transferring the mechanical
  vibrations of the ear drum/tympanium/tympanic membrane ($\sim$ low mechanical impedance) to
  the oval window ($\sim 20\times$ higher mechanical impedance) and into cochlear fluid.
  $Z_{mech}(f) \equiv F(f)/v(f)$, mechanical force $F(f)$ (Newtons) $= p(f) \times A =$ pressure $\times$ area.
* The <u>specific</u> longitudinal acoustic impedance of cochlear fluid $\sim$ that of sea water:
  $\mathbb{Z}_{ac}^{sw} \sim 1.5 \times 10^6 \ Pa\text{-}s/m \left(= N\text{-}s/m^3\right)$. <u>Note:</u> $Z_{mech} \equiv F/v\left(N\text{-}s/m\right) = \mathbb{Z}_{ac}\left(N\text{-}s/m^3\right) \times A_{\perp}\left(m^2\right)$
* Ossicles also <u>*protect*</u> the inner ear from <u>*very*</u> loud noises – via the so-called <u>*acoustic reflex*</u> –
  which triggers <u>*two*</u> sets of muscles – one tightens the eardrum, another pulls the stirrup away
  from the oval window!
* Also has a "safety valve" – the eustachian tube - for pressure equalization and fluid drainage.

The Inner Ear:
* Cochlea – coiled/rolled up ($\sim 2 \frac{3}{4}$ turns), filled with *perilymph* fluid.
* Cochlea is divided down its length by a soft partition known as the *<u>basilar membrane</u>*,
  forming 2 long chambers connected together by an opening at the far end called the *<u>helicotrema</u>*.
* $\sim$ 15000 hair cells (connected to $\sim$ 30,000 nerve fibers embedded in the basilar membrane)
  sense acoustic disturbances in perilymph fluid; transmit information to brain via the auditory nerve.
* Amplitude of sound waves in perilymph fluid reaches a maximum at a particular point
  along the cochlea, for a particular frequency! (see graph(s) below…)
* Sense of pitch (human perception of frequency) depends (in part) on location along the cochlea.
* Additionally, 3 semi-circular canal(s) attached to cochlea = 3-axis (x, y, z) accelerometers (!)
  used for balance/orientation determination/sensing – *i.e.* an *<u>inertial guidance</u>* system!
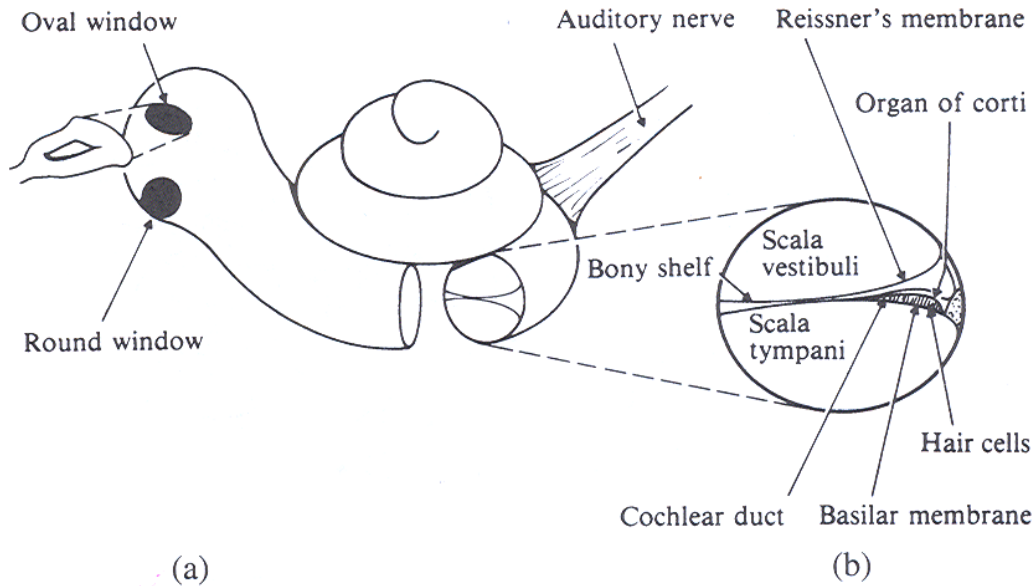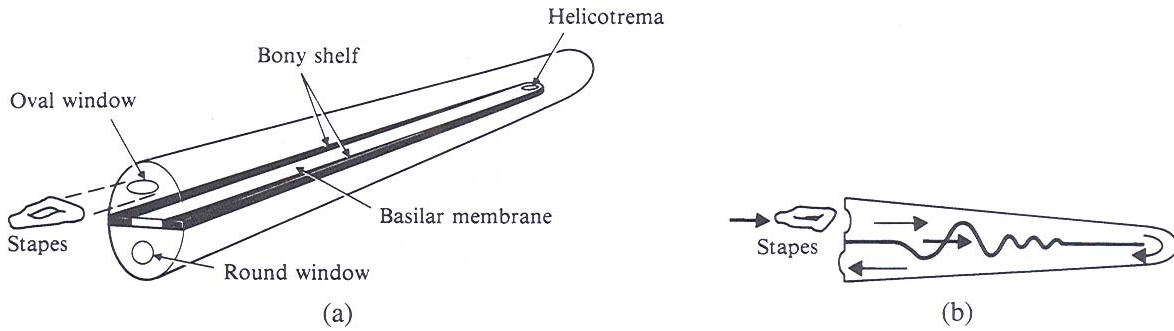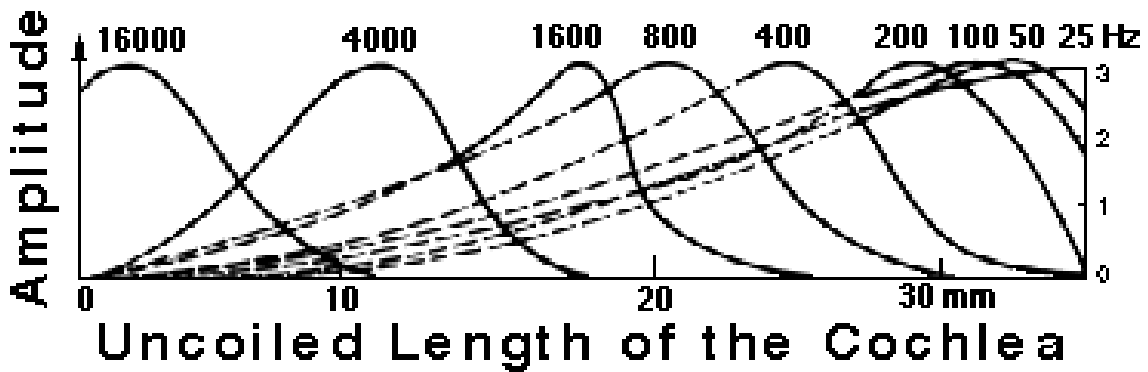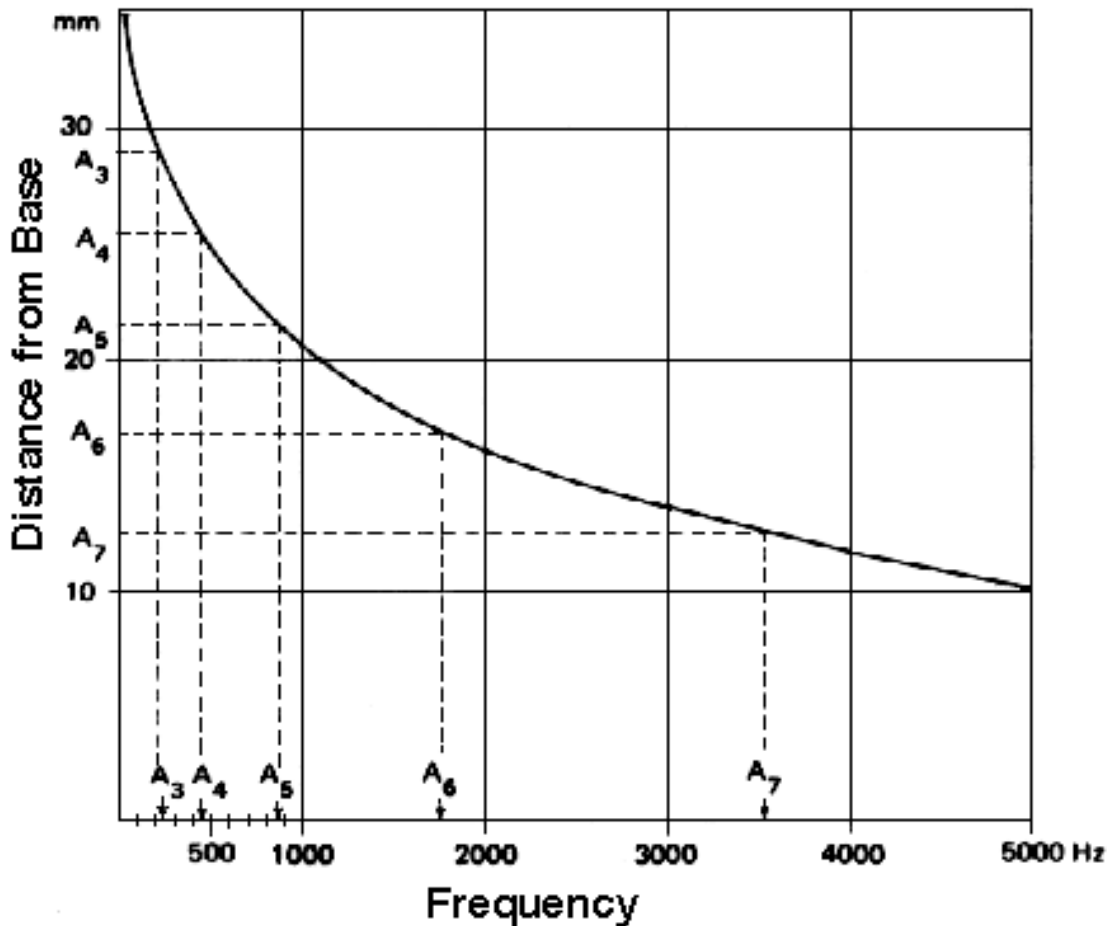
FIGURE 5.4   A schematic diagram of (a) the cochlea; (b) a section cut from the cochlea.

Unrolled human cochlea, and schematic of sound pressure pulse propagating in perilymph fluid:

Where frequencies of sounds peak along basilar membrane (*n.b.* <u>nonlinear</u> relationship!):

* The <u>spiral</u> <u>shape</u> of the cochlea ***enhances*** sensitivity to low frequency sounds by ~ 20 *dB*!

    Energy propagating in acoustic waves of perilymph fluid thus accumulates against the outside edge of the cochlear chamber. This effect is the strongest at the far end (*i.e.* center) of cochlea – where low frequency sounds are sensed. Analogous/related to the so-called "whispering gallery mode" effect – whispers travel along the curved walls of a large chamber, and remain strong enough to be heard clearly on the opposite side of the room!
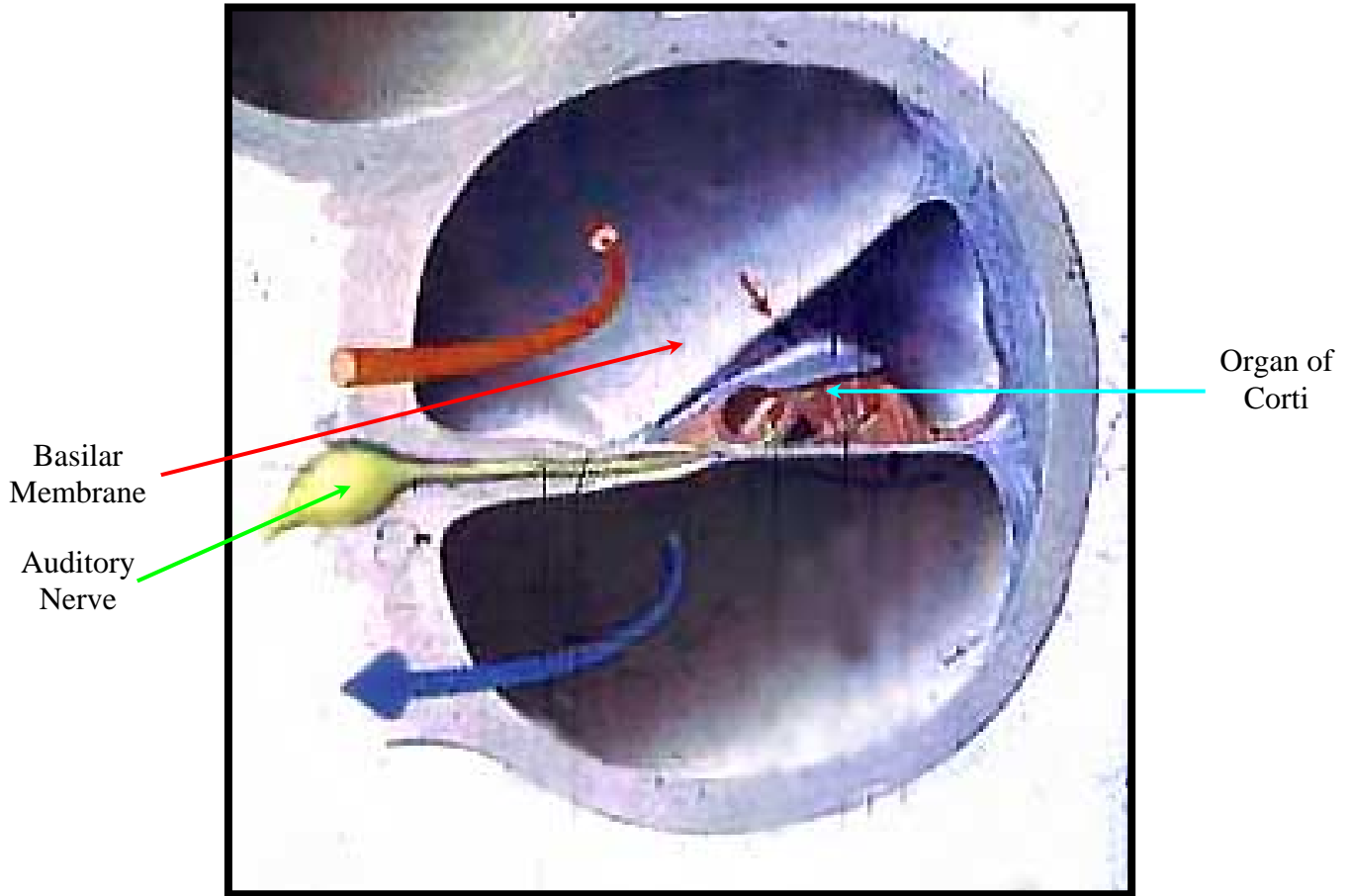
    D. Manoussaki, *et al*., Phys. Rev. Lett. 96, 088701 (2006).

* <u>Note that</u>:

$$20 \, dB = 10 \log_{10}\left(G_s\right) \;\Rightarrow\; \log_{10}\left(G_s\right) = 20/10 = 2 \;\Rightarrow\; G_s = 10^2 = 100 \,.$$

i.e. 20 *dB* corresponds to a signal gain $G_s = S_{out}/S_{in} = 100\times$ in the low frequency region!!!

Cross-Sectional/Cut-Away View of One Portion of the Human Cochlea:

The human ear has _two_ rows of hair cells in the Organ of Corti, which generate electrical signals in response to pressure signals in the perilymph fluid along the basilar membrane.

The primary function of the _inner_ row of ~ 4000 hair cells is to generate the electrical signals sent to the brain via the auditory nerve.

The _outer_ triple-row of ~ 12,000 chevron-shaped hair cells function as biological amplifiers, boosting the sensitivity level of the human ear by ~ 40 _dB_!
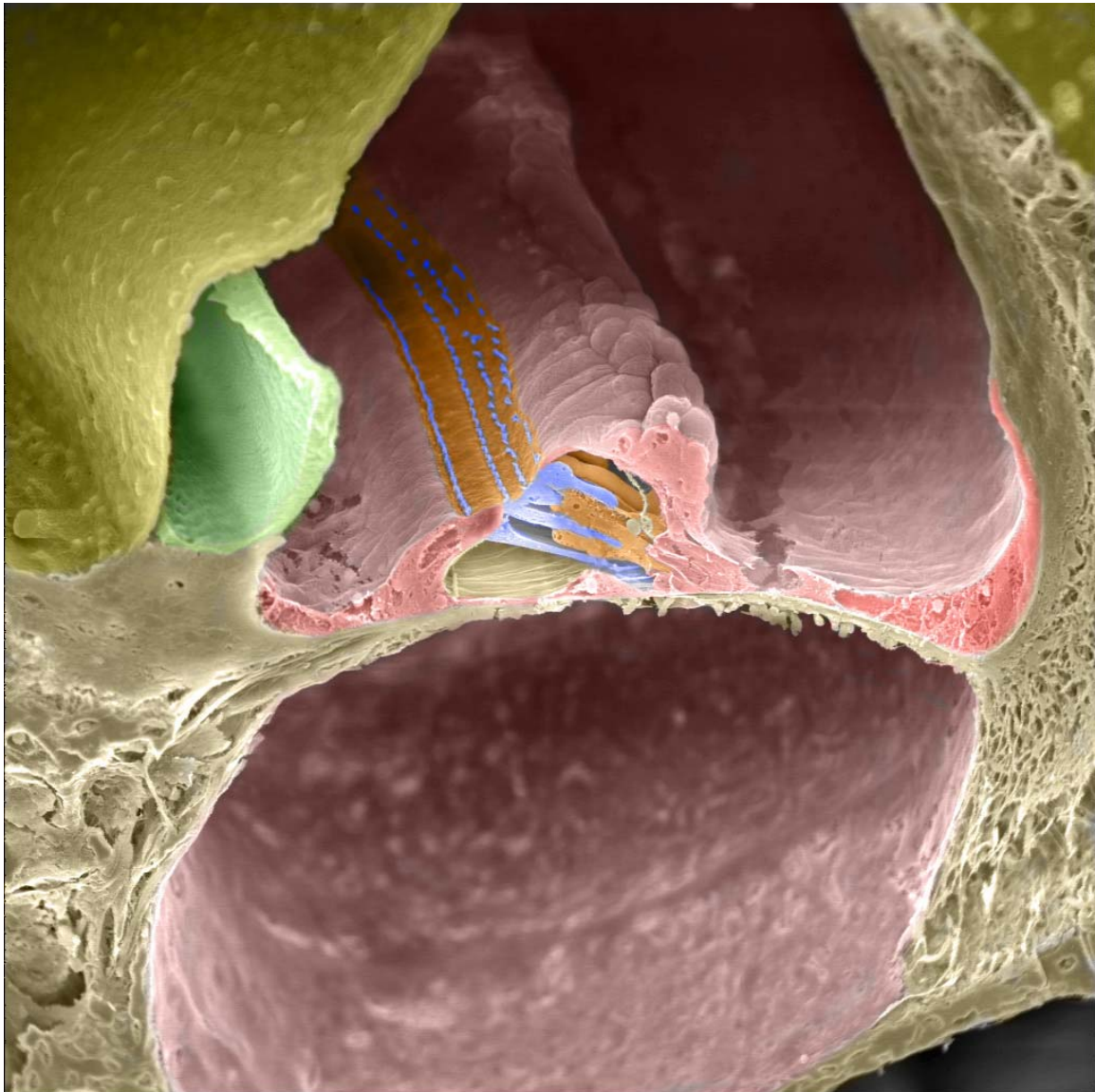
Note that:

$$40 \, dB = 10 \log_{10}(G_s) \ \Rightarrow \ \log_{10}(G_s) = 40/10 = 4 \ \Rightarrow \ G_s = 10^4 = 10,000 \, .$$

_i.e._ 40 _dB_ corresponds to a signal gain $G_s = S_{out}/S_{in} = 10,000 \times$!!!
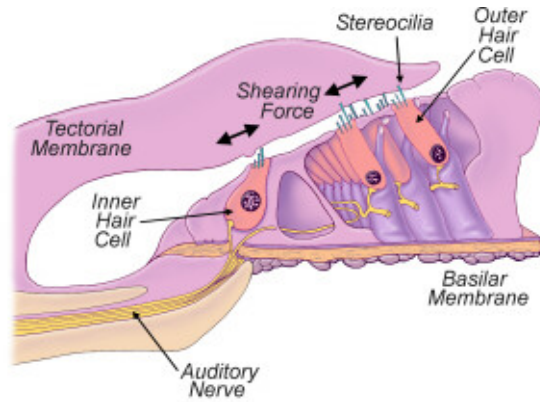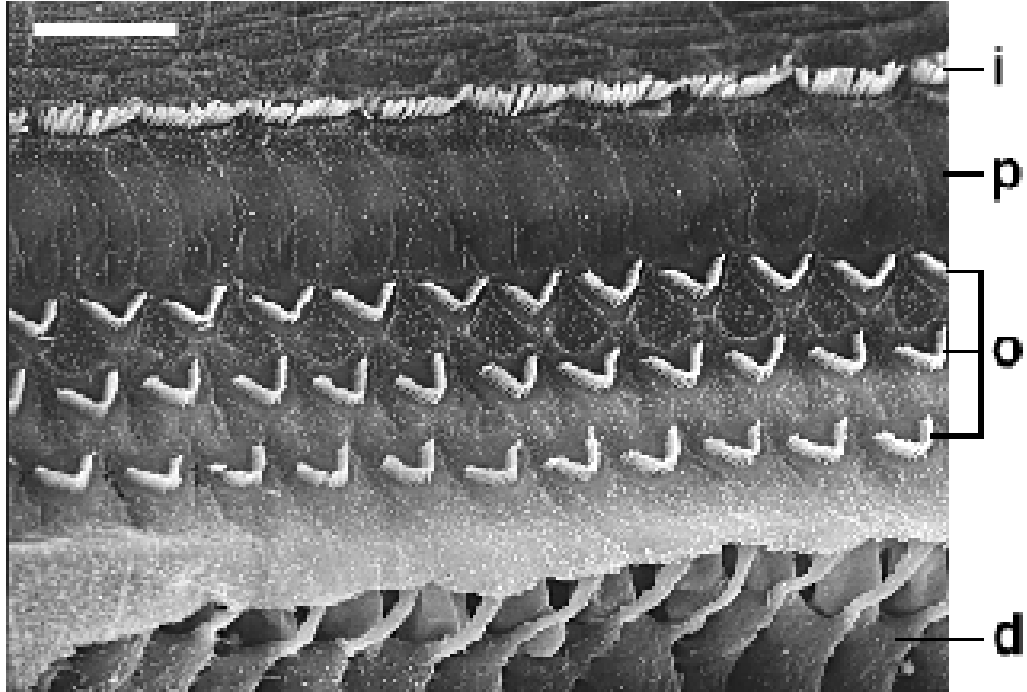
The Organ of Corti

Colored electron micrograph of the Organ of Corti:

Inner and Outer Hair Cells in the Organ of Corti of the Human Ear:

Each hair cell has many hairs (stereocilia) that are bent/vibrated when the basilar membrane responds to sound waves in the perilymph fluid. The bending of the stereocilia stimulates the hair cells, which in turn excite neurons in the auditory nerve.
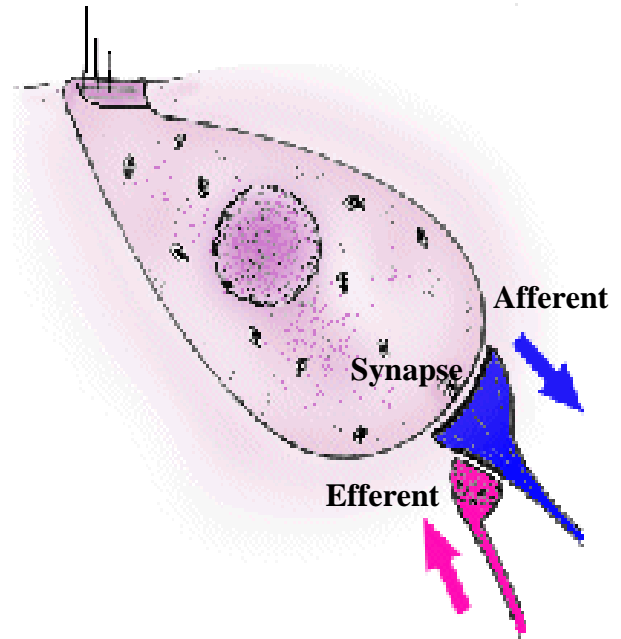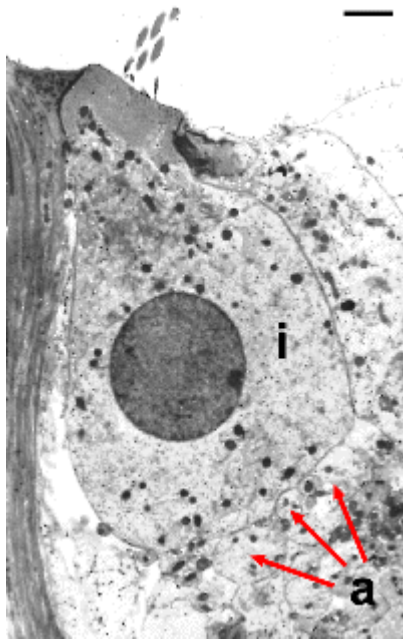
The neuron firing/impulse rate on the auditory nerve depends on both the sound intensity I _and_ the frequency f of the sound – _e.g._ neurons do _not_ fire on _every_ oscillation cycle of frequency f for very faint sounds. Neutrons _do_ tend to fire on the peaks of/in phase with a cycle, however.
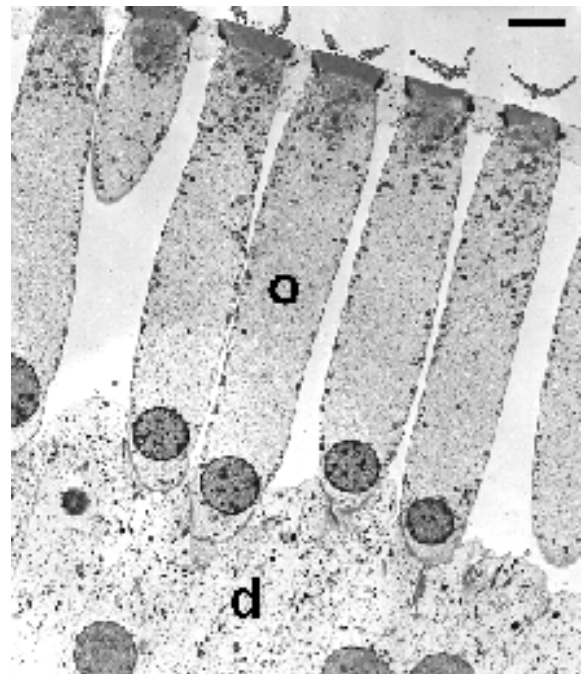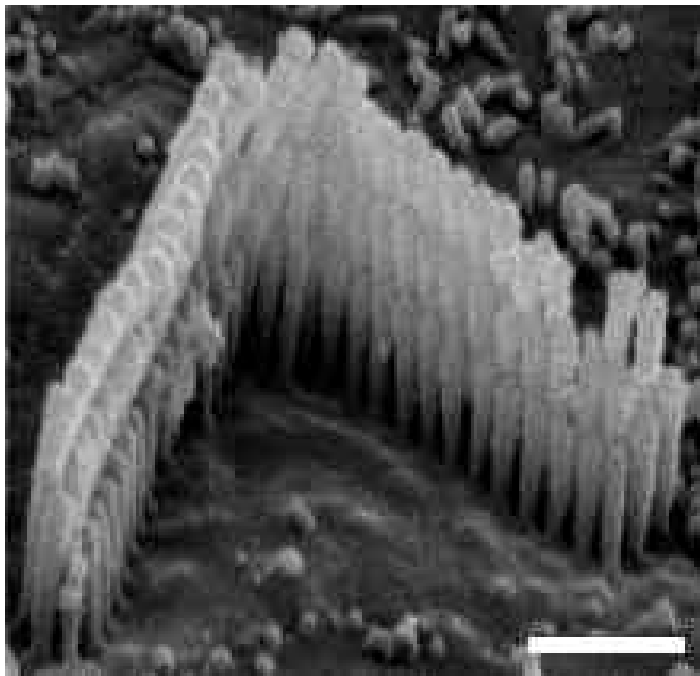
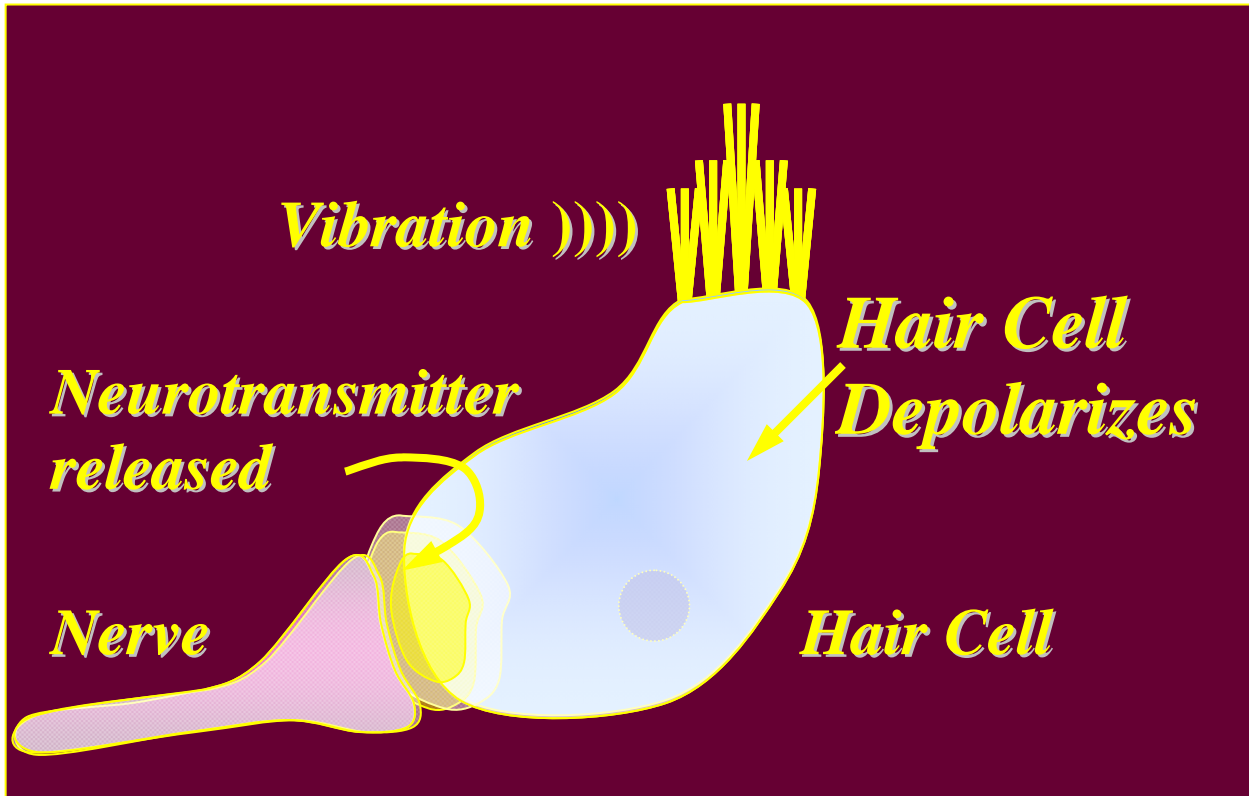**Individual Stereocilia**
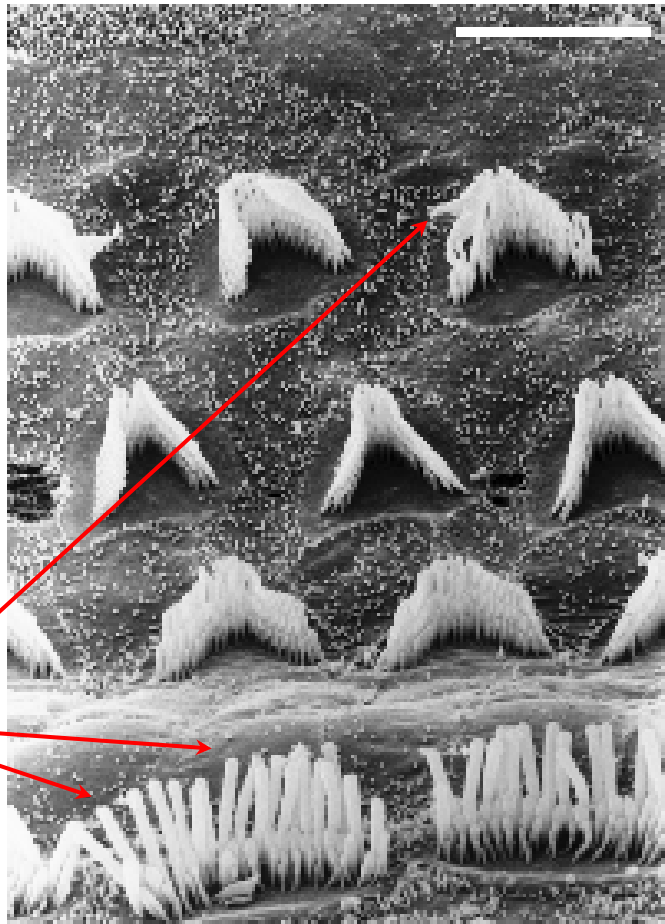
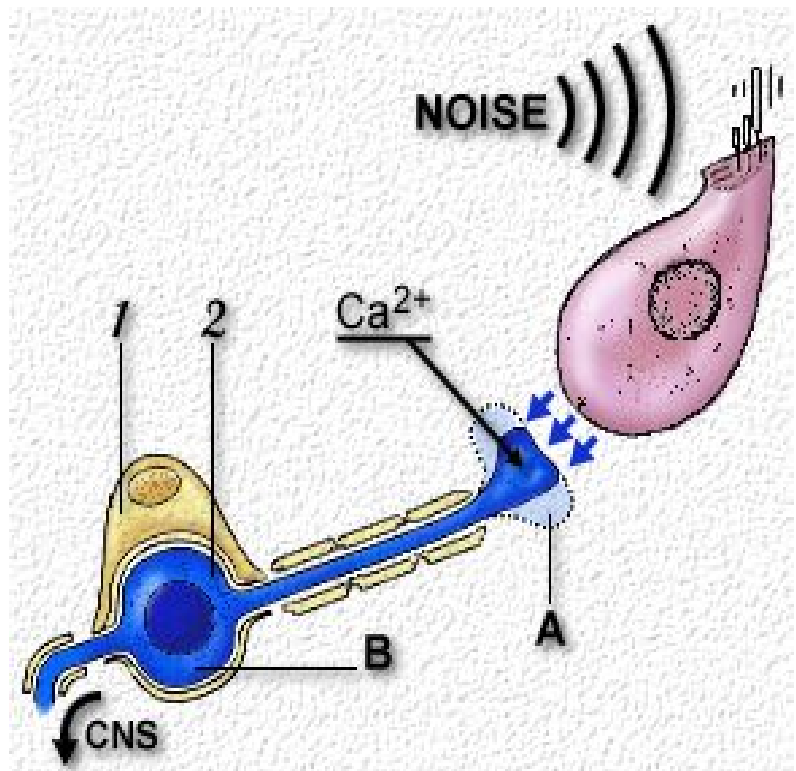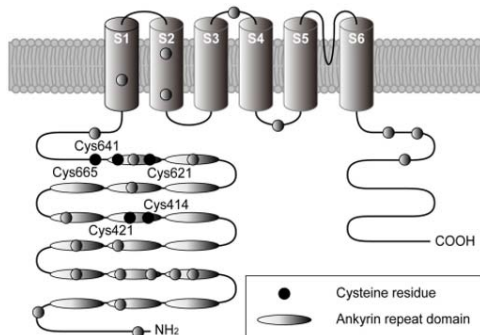**Inner Hair Cells:**

Outer Hair Cells:

Action of an Inner Hair Cell:

Function of a Stereocilia:

An auditory sound vibration / stimulation of stereocilia in a hair cell causes the release of a neuro-transmitter (TRPA1 protein – see below) which in turn stimulates a neuron in the auditory nerve.

TRPA1 Structural Model:

Firing of Auditory Nerve Fibers:

~ 95% of the auditory nerve fibers (type I) are connected to the ~ 4000 inner hair cells. Each type I axon innervates only a single hair cell, but each hair cell directs it output up to ~ 10 type I auditory nerve fibers. The type I auditory nerve fibers are bipolar and are _mylenated_ (*i.e.* have a nerve sheath – this protects the nerve fiber and also increases the transmission speed of action potentials along nerve fiber by up to ~ 300× over non-mylenated nerves – evolutionarily very important – for our auditory startle reflex!). ~ 5% of the auditory nerve fibers (type II) are connected to the ~ 12000 outer hair cells, are monopolar and are not mylenated.

   Each auditory nerve fiber responds over a certain range of frequency and sound pressure, and has a characteristic frequency f_C at which it has maximum sensitivity. Auditory nerve fibers having high characteristic frequency have a rapid roll-off in response above their characteristic frequency f_C, however, they have a long "tail" in response below it.

   A 90 *dB* sound stimulus with a single frequency, *e.g.* $f = 500$ Hz induces voltage signals on _several_ adjacent nerve fibers associated with the frequency band centered on 500 Hz. The time between successive voltage signals on an auditory nerve fiber almost always corresponds to one or two or more periods $\Delta t = \tau = 1/f$ of the frequency, firing on the peak of a vibration cycle. An axon of a nerve fiber does not fire at the peak of _every_ vibration cycle of the basilar membrane, but it rarely fires at any other time in the cycle. When two or more complex tones are present, things get a bit more complicated, however the pattern of electrical signals from the auditory nerves firing still carries accurate information about the frequency spectrum of the complex auditory/tonal stimulus.

   A complex auditory stimulus of two pure tones C4 (523 *Hz*) and C5 (1046 *Hz*) – *i.e.* an octave apart – do not have much overlap in terms of auditory neural tuning curves (*i.e.* frequency response curves) as shown in the figure (a) below, because very few hair cells respond to _both_ of these frequencies. However, as the interval between the two pure tones decreases, the situation changes – more and more overlap occurs, an increasing number of hair cells are stimulated by both tones, as shown in figures (b) and (c), leading to many interesting phenomena.
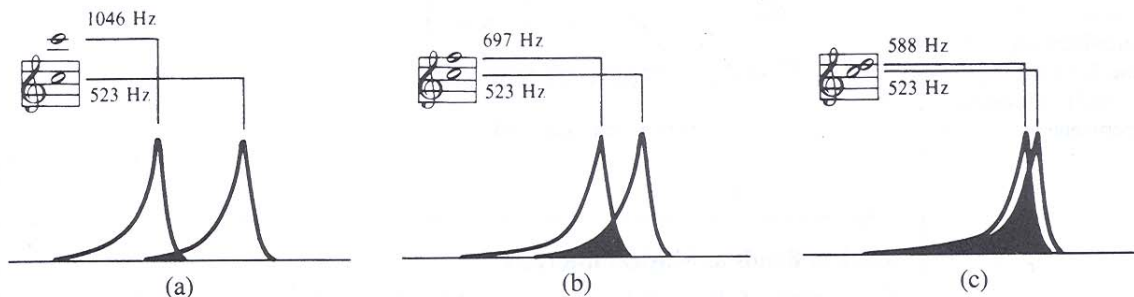
FIGURE 5.9   Frequency response curves for pairs of pure tones. As the interval between them decreases, their response curves show increasing overlap.
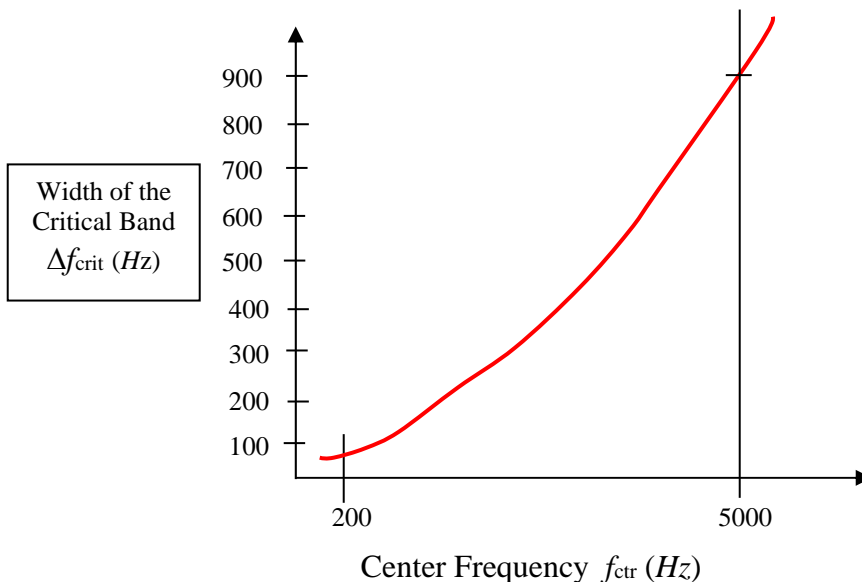
The Critical Band:

Two pure-tone sounds, which are slightly different in frequency f_1 and f_2 are _not_ heard as separate notes by a single human ear if they are too close together in frequency.

Reason: The mechanical vibrational behavior of basilar membrane, and the firing & wiring of hair cells $\Rightarrow$ auditory nerve has finite bandwidth associated with each…
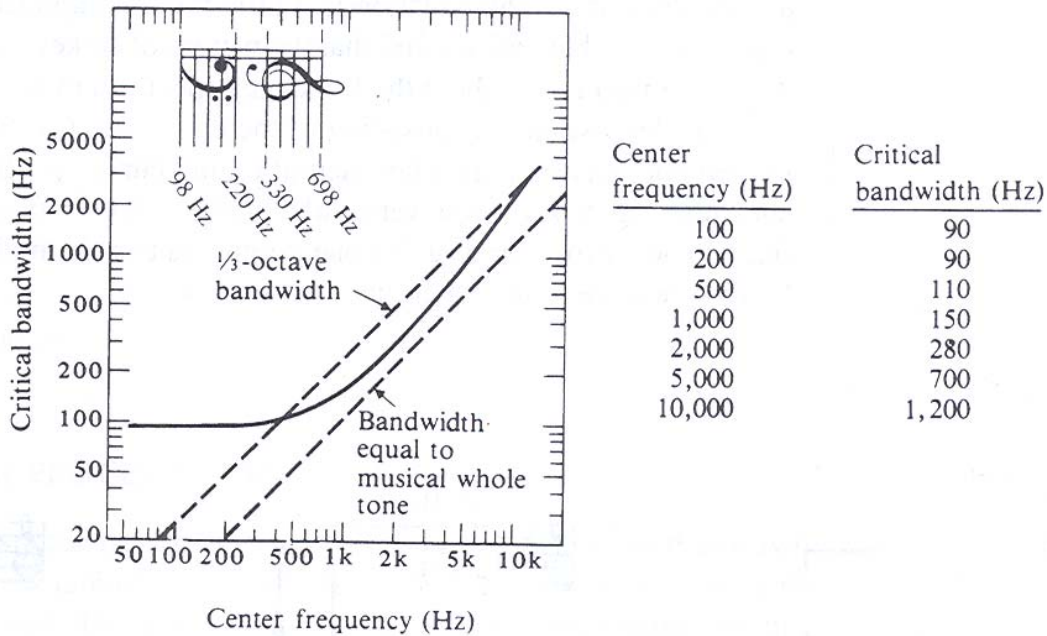
The sound of a particular frequency produces a traveling wave - which propagates along the basilar membrane. The pressure amplitude of this wave propagating in the perilymph fluid is _not_ a constant - it peaks somewhere along the basilar membrane; the position of where it peaks depends on frequency of sound wave in the perilymph fluid (see lecture notes above). The pressure amplitude of this wave is not infinitely sharply peaked at this location, the disturbance produced by the wave is _spread out_ over a certain length of basilar membrane – i.e. it has a finite spatial extent/width along the basilar membrane.

The hair cells/nerve endings on the basilar membrane are excited over a narrow region on either side of maximum amplitude of motion of basilar membrane. The range or band of frequencies affected is known as the **_critical band_**. At center frequencies of $f_{ctr} \leq 200\ Hz$, the width of the critical band is ~ constant at $\Delta f_{crit} \sim 90\ Hz$ (n.b. $\Delta f_{crit}/f_{ctr} \sim 50\%$!), above this frequency, the width of the critical band increases ~ linearly to $\Delta f_{crit} \sim 900\ Hz$ @ $f_{ctr} \sim 5000\ Hz$, ($\Delta f_{crit}/f_{ctr} \sim 20\%$) as shown in the figure below:

Center Frequency f_{ctr} (Hz)

n.b. The width of the critical band is _also_ dependent on sound intensity.

This effect is ONLY for ONE ear – i.e. a monaural effect. It does _not_ exist if one frequency f_1 is input to one ear, and e.g. another/different/nearby frequency $f_2 \sim f_1$ is input into the other ear (this doesn't happen often in nature through…). The human brain is able/capable of processing binaural sound information to distinguish two (or more) closely-spaced frequencies, significantly better than monaural-only information of the same kind/type.

Center frequency (Hz)	Critical bandwidth (Hz)
100	90
200	90
500	110
1,000	150
2,000	280
5,000	700
10,000	1,200

In musical language, note that 1 whole note = 2 semitones; a 1/3 octave BW = 4 semitones = a major third. Note also that *e.g.* a 31-band audio spectrum analyzer covers the entire audio band (20 *Hz* – 20 *KHz*) in 1/3 octave per band.

Binaural Hearing and Sound Localization:

At frequencies below $f < 1000$ *Hz*, sound localization is primarily due to sensitivity to the inter-aural arrival time difference Δt (for sound pulses), or equivalently the relative phase difference $\Delta \varphi = \Delta t / \tau$ (for steady sounds) associated with sounds traveling paths L_1 vs. L_2:

- 13 -
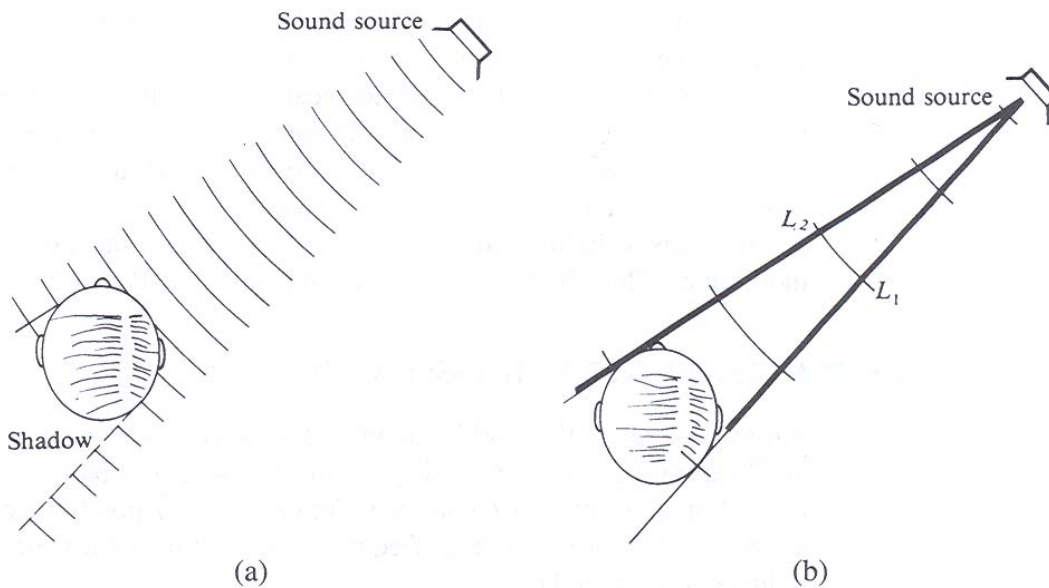
At frequencies above $f > 4000\ Hz$, sound localization is increasingly due to the perceived sound intensity level **_difference_** of both ears – the head casts a "shadow" on the away-side ear for increasingly high frequency sounds. At low frequencies, this effect disappears due to **_diffraction_** of the sound wave around the head… At frequencies of $f \sim 1000\ Hz$, the sound intensity level is only $\sim 8\ dB$ greater for the ear **_nearest_** the source, whereas at frequencies of f $\sim 10\ KHz$, the sound intensity level difference can often be $\sim 30\ dB$.

The human ears are separated by a typical distance of $d_{ears} \sim 6$" ($= 1/2$ foot, or $\sim 15\ cm = 0.15\ m$).

When the ear-ear separation distance is comparable to the wavelength of a sound, the corresponding (maximum) arrival time difference is $\Delta t = d_{ears}/v \sim \lambda/v = \tau_{osc} = 1/f$.

Thus, for $\lambda \sim d_{ears} \sim 0.15\ m$ and $v = 343\ m/s$ then $\Delta t \sim 0.15/343 = 0.44\ ms$ or: $f = 1/\tau_{osc} = 1/\Delta t \simeq 2300\ Hz$.

For frequencies higher than this, it becomes increasingly difficult for us to localize sound sources… The folds/creases in the pinna of our outer ears are there to aid/enhance localization of sounds in this higher frequency region!

It is also true that when $\lambda \gg d_{ear}$ (*i.e.* very low frequencies) we also have difficulties in localizing sounds – again due to **_diffraction_** of low frequency sound waves around our heads!

Practically, studies (we and others have carried out) show that we humans can accurately localize sounds in the frequency range of $100\ Hz \le f \le 1000\ Hz$.

Compare these results to our (very poor!) ability to localize sounds in **_water_**, where $v_{H_2O} \simeq 1500\ m/s$. The arrival time difference of sound waves (left – right) ears in water is

$\Delta t_{H_2O}^{L-R} \simeq (343/1500)\Delta t_{air}^{L-R} \sim 0.2\Delta t_{air}^{L-R} \Rightarrow$ 5× less in water! $\Rightarrow$ much harder for humans to localize sounds underwater than in air! However, many fish & other marine creatures (*e.g.* dolpins) can easily (and accurately) localize sounds underwater – because their hearing has been optimized for propagation of sound waves in water with speed $v_{H_2O} \simeq 1500\ m/s$!!!

Because we have lived in an air environment for millions of years, our human hearing has been specifically optimized for propagation of sound waves in air with speed $v_{air} \simeq 343\ m/s$!!!

Imagine how well we'd be able to localize sounds if we instead lived *e.g.* in a helium atmosphere, where $v_{He} \sim 970\ m/s \sim 3{\times}v_{air}$ or *e.g.* instead lived in an atmosphere of sulphur hexafluoride (SF_6), where $v_{SF6} \sim 150 \sim 0.44{\times}v_{air}\ m/s$!!!
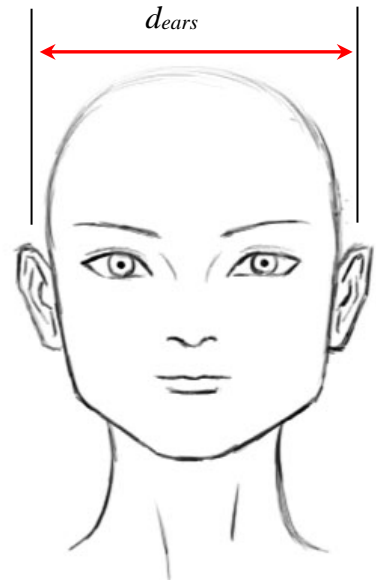
Sound Intensity & Human Hearing:

The human ear is exquisitely sensitive to detecting sounds (*n.b.* dogs hear ~ 100× better!) Due to this large dynamic range, our ears have an ~ *logarithmic* response to sound intensity.

Sound Intensity, I (*RMS Watts/m^2*) $\propto \xi^2_{rms}$ (*RMS particle displacement amplitude*)^2
  as well as: $I \propto u^2_{rms}$ (*RMS particle velocity amplitude*)^2
  as well as: $I \propto p^2_{rms}$ (*RMS over-pressure amplitude*)^2

The relationship of the {magnitude of the} *RMS* sound intensity at a given point $\vec{r}$ in a sound field associated *e.g.* with monochromatic/pure-tone sine-type travelling plane waves propagating in free air (*i.e.* the great outdoors) to the above quantities is:

$$\boxed{I_{rms}\left(\vec{r},t\right) = p^2_{rms}\left(\vec{r},t\right)\big/\rho_o c = \rho_o c \cdot u^2_{rms}\left(\vec{r},t\right) = \rho_o c \cdot \omega^2 \xi^2_{rms}\left(\vec{r},t\right) = cw^{ac}_{rms}\left(\vec{r},t\right)}$$

where:

$p_{rms}\left(\vec{r},t\right) =$ *RMS* over-pressure amplitude (*RMS Pascals*) at the point $\vec{r}$

$\rho_o \qquad =$ equilibrium mass density of dry air (@ *NTP*) = 1.204 *kg/m^3*

$c = v_{air} \quad =$ longitudinal propagation speed of sound in dry air = 343 *m/s* (@ *NTP*)

$u_{rms}\left(\vec{r},t\right) =$ so-called *RMS* longitudinal particle velocity amplitude (*RMS m/s*) at the point $\vec{r}$

$\omega = 2\pi f \;=$ angular frequency (*radians/sec*), $f =$ frequency (*Hz*)

$\xi_{rms}\left(\vec{r},t\right) =$ *RMS* longitudinal particle displacement amplitude (*RMS meters*) at the point $\vec{r}$.

The *RMS* acoustic energy *density* (*Joules/m^3*) at the point $\vec{r}$ is given by:

$$\boxed{w^{ac}_{rms}\left(\vec{r},t\right) = p^2_{rms}\left(\vec{r},t\right)\big/\rho_o c^2 = \rho_o u^2_{rms}\left(\vec{r},t\right) = \rho_o \omega^2 \xi^2_{rms}\left(\vec{r},t\right)}$$

The quantity $\boxed{z\left(\vec{r}\right) \equiv p_{rms}\left(\vec{r}\right)\big/u_{rms}\left(\vec{r}\right) \;(Pa\text{-}s/m)}$ is known as the *specific acoustic impedance* of the medium, which for monochromatic plane waves propagating in a free air sound field is a constant (*i.e.* independent of frequency): $\boxed{z^{free\,air}_{ac} \equiv z_o = \rho_o c \simeq 415\,kg/m^2\text{-}sec = 415\,\Omega_{ac}}$ (@ *NTP*).

Thus, we also see that for such plane waves propagating in a free-air sound field that:

$$\boxed{I_{rms}\left(\vec{r},t\right) = p_{rms}\left(\vec{r},t\right) \cdot u_{rms}\left(\vec{r},t\right) = p^2_{rms}\left(\vec{r},t\right)\big/z\left(\vec{r}\right) = u^2_{rms}\left(\vec{r},t\right) \cdot z\left(\vec{r}\right) = cw^{ac}_{rms}\left(\vec{r},t\right)}.$$

Note that *RMS* quantities are frequently/often used in acoustical (and other) physics because for a pure-tone/single-frequency sine wave-type sound, they conveniently correspond to the *time-averaged* value of that quantity:

$$\langle I \rangle = \frac{1}{\tau}\int_{\substack{one\\cycle}} I\left(t\right)dt = \frac{1}{2}I_{peak} = I_{rms} \propto \frac{1}{\tau}\int_{\substack{one\\cycle}} p^2\left(t\right)dt = \frac{1}{\tau}p^2_o\int_{\substack{one\\cycle}} \sin^2 \omega t\,dt = \frac{1}{\tau}\frac{1}{2}p^2_o = \frac{1}{\tau}p^2_{rms}$$

$p_{rms} = RMS$ (*R*oot-*M*ean-*S*quare) of sound (over-)pressure amplitude (*RMS Pascals = RMS Newtons/m^2*)

For a pure-tone (*i.e* single-frequency) sine wave: $p_{rms} \equiv \frac{1}{\sqrt{2}}p_o \; \{\Rightarrow p^2_{rms} \equiv \frac{1}{2}p^2_o = \langle p^2_o \rangle\}$

Famliar/Everyday Example: The 120 Volts/60 *Hz* AC line voltage in your house actually refers to the *RMS* voltage, *i.e.* the *RMS* voltage amplitude is $V_{rms} = 120$ Volts @ $f = 60$ *Hz*, hence the actual voltage amplitude (*aka* the peak amplitude) is $V_p = \sqrt{2} \cdot V_{rms} = 1.414 \cdot 120 = 169.7 \simeq 170\, Volts$.
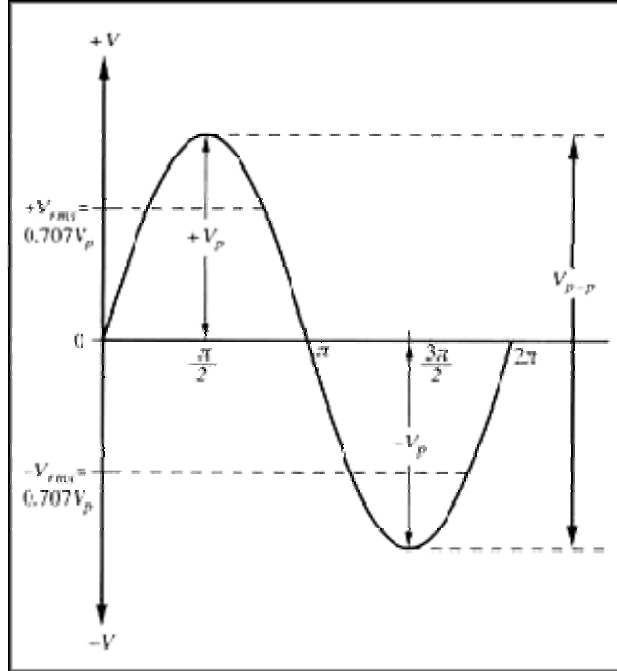
The time-averaged, or *RMS* sound intensity <u>threshold</u> of hearing (@ $f = 1$ KHz) is:
$<I_{thr}> \sim 2.5 \times 10^{-12}$ *RMS Watts*/$m^2 = 2.5$ *RMS pico-Watts*/m^2

Individual people may hear better/worse than the average person, and so threshold of hearing from one person to another can vary as much as 1/10 or 10× this!!! Since the human ear has an ~ logarithmic response to sound intensity, linear factors of ~ 2.5× are not really very significant, and thus for convenience' sake, we simply round this down to the so-called <u>reference</u> <u>standard</u> for the {average} sound intensity threshold of hearing, defined as:

$$\langle I_o \rangle = I_{o\,rms} = 10^{-12}\ RMS\ Watts \Big/ m^2 \quad \text{as the official \underline{Intensity} \underline{Threshold} \underline{of} \underline{Hearing}.}$$

Using $I_{rms} = p_{rms}^2 / \rho_o c$ with $\rho_o = 1.204\,kg/m^3$ and $c = 343\,m/s$ (@ NTP), we find that $\langle I_o \rangle = I_{o\,rms} = 10^{-12}\ RMS\ Watts \Big/ m^2$ corresponds to a *RMS* sound over-pressure threshold of

$$p_{o\,rms} = 2.0322 \times 10^{-5}\ RMS\ Newtons \Big/ m^2 \simeq 2 \times 10^{-5}\ RMS\ Pascals.$$

However, the sensitivity of human hearing is frequency dependent over the entire audio spectrum, and in fact the *RMS* reference intensity and pressure amplitudes $I_{o\,rms}$ and $p_{o\,rms}$ are specifically associated with pure tone/sine waves of frequency $f = 1$ *KHz*, because the human ear is most sensitive in the $f \sim 1$ – few *KHz* range, as shown in the figure below:
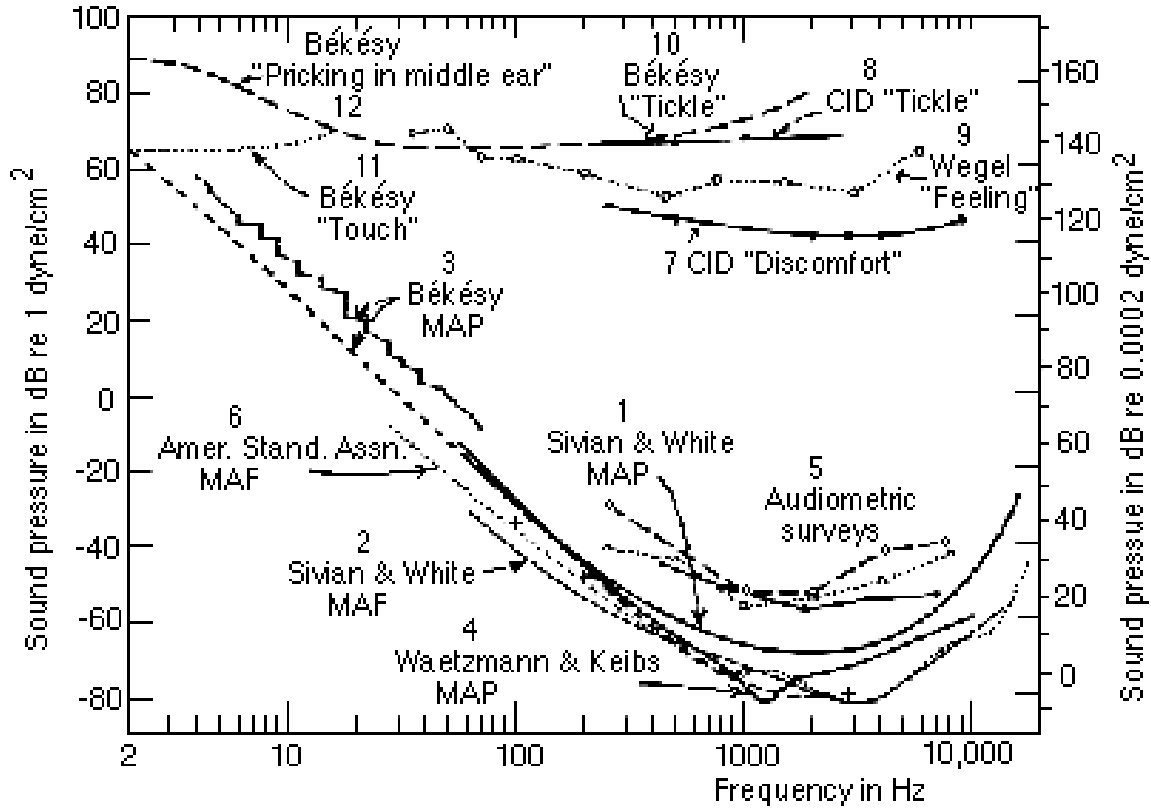
The so-called minimum audible pressure (MAP) (*aka* the reference pressure), is defined at $f = 1$ *KHz* and is:

$$\boxed{p_{o\ rms}\left(f = 1\ KHz\right) \simeq 2\times10^{-5}\ RMS\ Newtons\Big/m^2 = 2\times10^{-5}\ RMS\ Pascals}$$

In the above figure, note that the **dyne** is a *cgs* (*cm-gm-sec*) unit of force, hence the SI/*mks* $\Leftrightarrow$ *cgs* units conversion factor: 1 *Newton* = 10^5 *dynes*.

Recalling that $p_{atm} \simeq 10^5\ Pascals$, we see that humans are able to hear/detect pressure variations of order ~ 1 part in 10^{10} of atmospheric pressure!!!

The corresponding minimum audible longitudinal particle velocity $u_{o\ rms}$ (*aka* the reference particle velocity) and minimum audible longitudinal particle displacement $\xi_{o\ rms}$ (*aka* the reference particle displacement) at $f = 1$ *KHz* are:

$$\boxed{u_{o\ rms}\left(f = 1\ KHz\right) \simeq 4.8\times10^{-8}\ RMS\ m/s} \quad \text{and:} \quad \boxed{\xi_{o\ rms}\left(f = 1\ KHz\right) \simeq 7.7\times10^{-12}\ RMS\ m}$$

The latter should be compared *e.g.* with size of an atom, which is typically on the order of a few Angstroms, *i.e.* $d_{atom} \sim few\times10^{-10}\ m \gg \xi_{o\ rms} \sim 8\times10^{-12}\ m$, a factor of ~ 100× !!!

Sound *Intensity* Level:

The human ear responds ~ *logarithmically* to sound intensity:

Loudness = Sound *Intensity* Level, L_I = SIL: $$SIL = L_I \equiv 10 \, \log_{10}\left(\frac{I_{rms}}{I_{o \, rms}}\right) \; (dB)$$

   The unit of Loudness/Sound Intensity Level is the deci-Bel (*dB*) {*n.b.* deci = 10, *Bel* in honor of Alexander Graham Bell}.

Note that musicians have quantitatively defined <u>*six*</u> different loudness levels:

fortissimo	*fff*	$I = 10^{-2}$ RMS W/m^2	L_I (dB) = 100
	ff	10^{-3}	90
	f	10^{-4}	80
	p	10^{-6}	60
	pp	10^{-7}	50
pianissimo	*ppp*	10^{-8}	40

The Range of Human Hearing: Sound Intensity, Sound Intensity Level *vs.* Frequency:
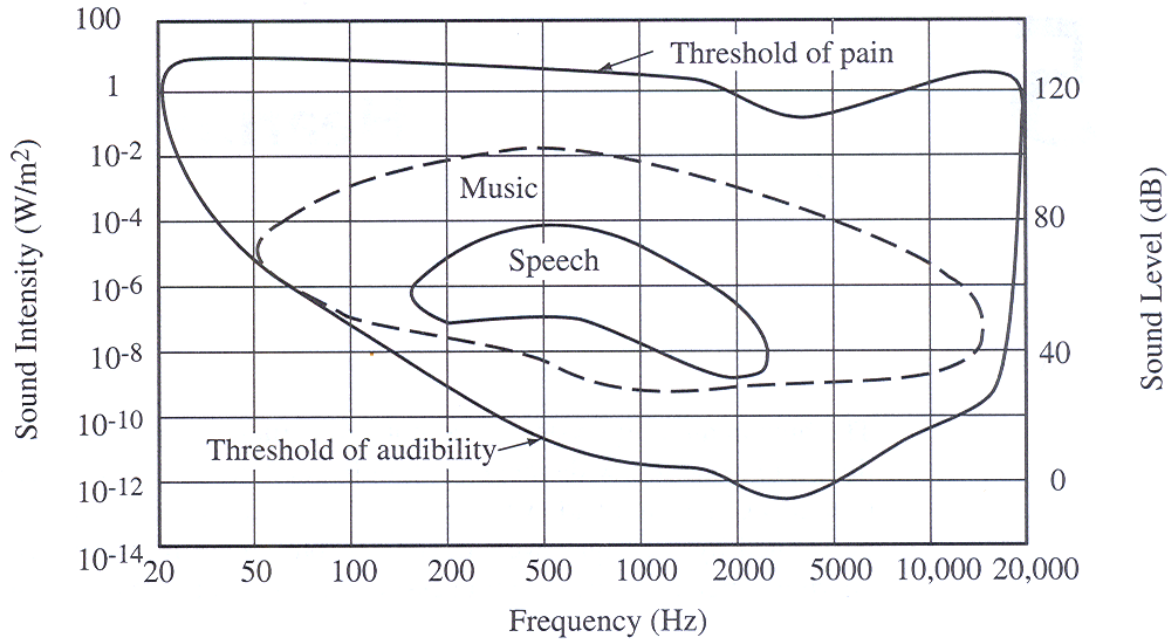
Sound *Pressure* Level: SPL = L_p = Sound Pressure Level (units = deci-Bels, *dB*)

$$SPL = L_p \equiv 10\log_{10}\left(\frac{p_{rms}^2}{p_{o \, rms}^2}\right) = 10\log_{10}\left(\frac{p_{rms}}{p_{o \, rms}}\right)^2 = 20\log_{10}\left(\frac{p_{rms}}{p_{o \, rms}}\right) \; (dB)$$

Note that for an acoustic plane wave propagating in **_free air_** (*e.g.* the great outdoors), $SPL = L_p$ and $SIL = L_I$ are essentially the same numerical values in *dB*, and are typically are within 0.1 *dB* of each other across the frequency spectrum of human hearing. However, *e.g.* inside an auditorium (or, more generally, in any **_confined_** space), due to sound reflection from the walls/ceiling/floor (creating multiple sound waves/standing waves), $SPL = L_p$ and $SIL = L_I$ will **_not_** necessarily be the same! {We will discuss this in more detail in subsequent lecture(s)}

Note also that most microphones – one (of many) kinds of sound transducers – are such that they are sensitive/respond to (over)-pressure. Hence, technically speaking, such microphones measure/determine the <u>Sound Pressure Level</u>, $SPL = L_p$ (not Sound Intensity Level, $SIL = L_I$).

Apparent <u>Loudness Level: Phons</u>

The perceived response of {average} human hearing to **_constant_** loudness levels (*aka* sound intensity levels) $SIL = L_I$ is **_not_** independent of frequency. The response of the human ear for very low (< 20 *Hz*) and very high frequencies (> 20 *KHz*) is increasingly poor. Note that the open-closed ¼-λ resonances associated with the ear canal affect our loudness level response.

Because human hearing is not flat with frequency, the perceived, or **_apparent_** loudness of a sound depends on frequency, and also on the actual intensity *I* (in *Watts*/m^2), or equivalently, the actual loudness L_I (in *dB*) {or sound pressure level L_P = SPL (in *dB*) of the sound.

In 1933, Fletcher and Munson obtained average values of the **_apparent_** loudness of sounds for human hearing as a function of these variables. The unit of **_apparent_** loudness L_{app} is the **Phon**, defined as the value of the **_SPL_** that has **_constant apparent loudness_** for (average) human hearing. The figure below shows the {ISO 226:2003 revised} Fletcher-Munson curves – contours of constant **_apparent_** loudness $L_{app}(f)$ *vs.* frequency, *f*.
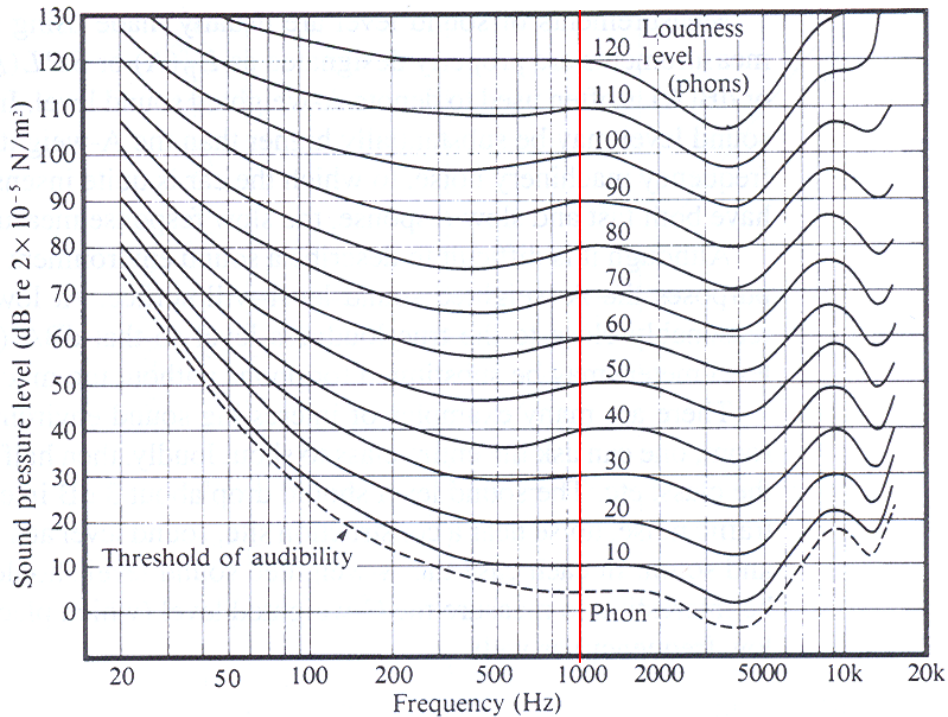
Note that at $f = 1000\ Hz$, L_{app} (Phons) $\equiv SPL$ (dB). At other frequencies, the graph clearly shows that L_{app} (Phons) $\neq SPL$ (dB) …

Sound pressure level (*SPL*) meters have 3 types of sound weighting networks:

A−weighting: the 40 Phon curve of above figure. Units: *dB-A SPL*
B−weighting: the 70 Phon curve of above figure. Units: *dB-B SPL*
*C*−weighting: flat, independent of frequency.   Units: *dB-C SPL*

   A device that measures *SPL* is known as a Sound Level Meter - the results of *SPL* measurement by this device can also be <u>weighted</u> by the average frequency response of the human ear. A *SPL* meter utilizes a flat-response pressure microphone, absolutely {NIST} calibrated in its sensitivity. See/show UIUC Physics 406 POM's Extech *SPL* meter…

   A *SPL* meter also often has different frequency-dependent weighting schemes, as shown in figure below. *C*-weighting has almost a flat frequency response, whereas *A* (*B*)-weighting has response similar to human ear response at low (high) sound pressure levels of 40 (70) *phons*, respectively.

A, *B* and *C*-Weighting Curves *vs*. Frequency:

20
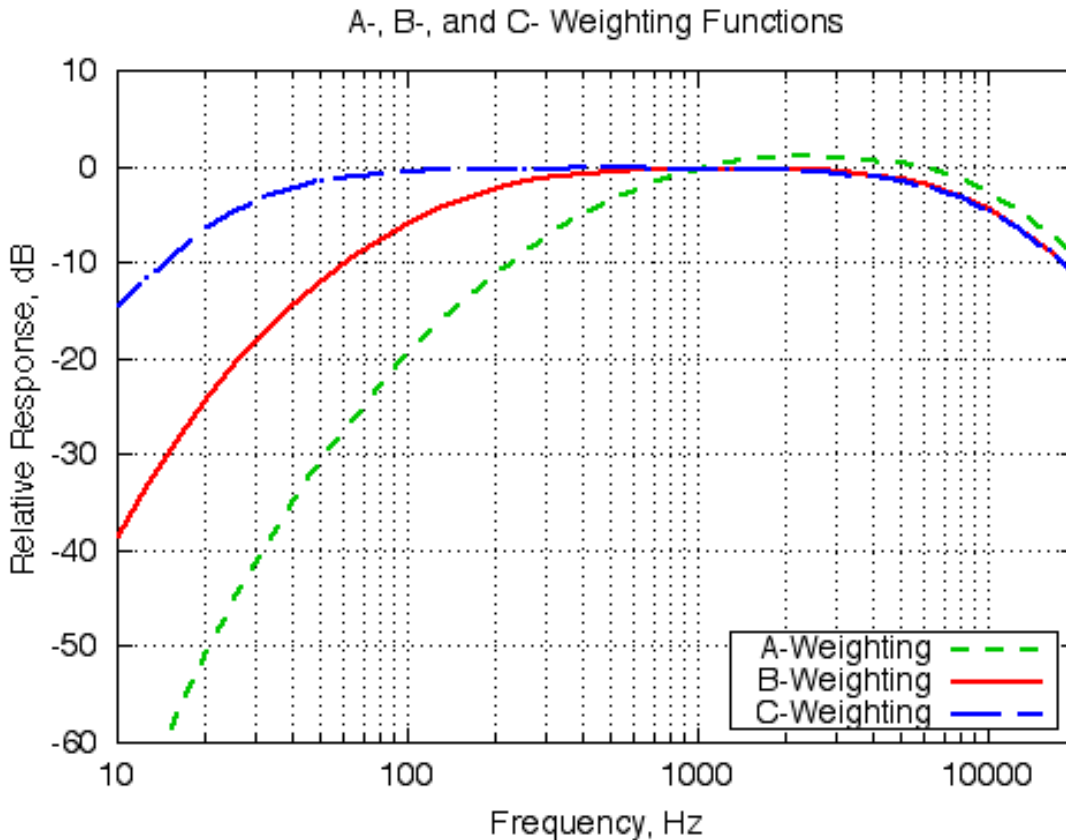
Relationship between Apparent/Perceived Loudness Level L_{app} (Phons units) and Apparent/Perceived Loudness N_{app} (Sones units) for *Pure* Tones

For ***apparent*** loudness levels of $L_{app} = 40$ phons or greater, for pure tones (and/or narrow bandwidth sounds) ***only***, there exists a straight line relation on <u>semi-log</u> plot (like $y = mx + b$) of:

$$\log_{10}\left(N_{app}(sones)\right) = m \, L_{app}(phons) + b$$

slope   intercept

where numerically:

slope $m = \log_{10}(2) = 0.30103$
intercept $b = -40.0 \log_{10}(2)$

Hence:

$$N_{app}(sones) = \left[10^{\frac{L_{app}(phons)-40}{10}}\right]^{0.30103}$$

or:

$$L_{app}(phons) = 40 + \frac{10\log_{10}N_{app}(sones)}{\log_{10}(2)}$$

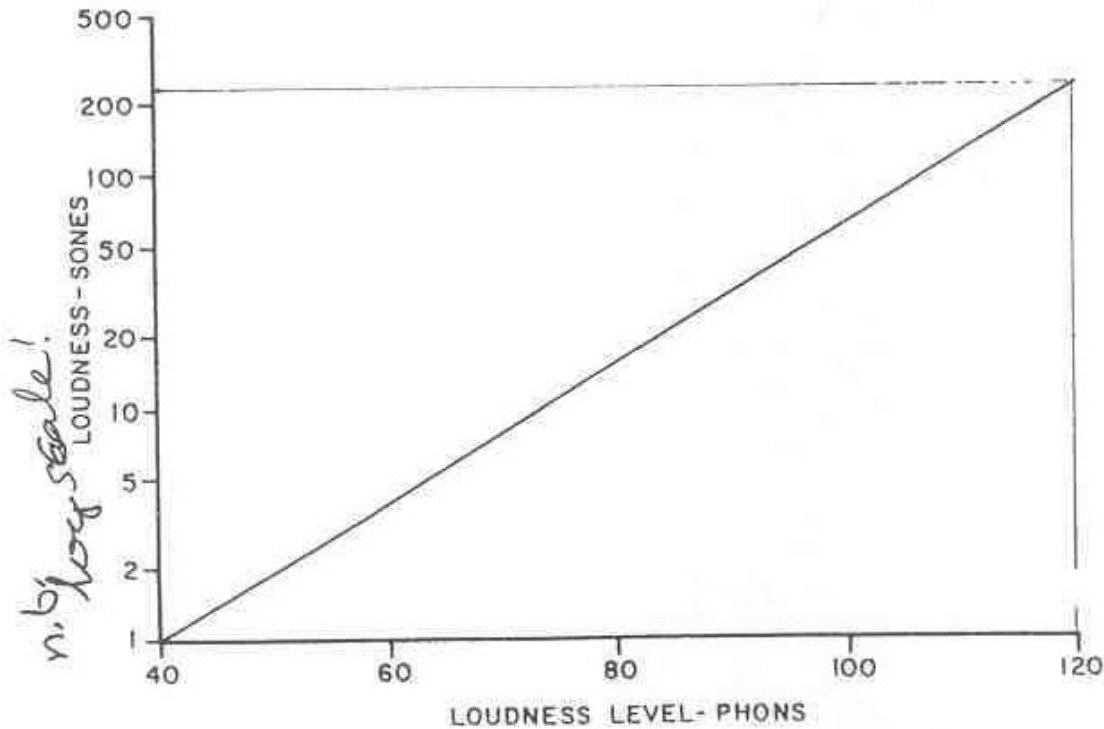

FIG. 4. Relation between loudness in sones and loudness level in phons.

$N_{app}(sones)$ is used primarily by psychologists in carrying out human psychoacoustics research.

The Just Noticeable Difference:

The ***Just Noticeable Difference*** (*JND*, in *dB*) in our human hearing is $JND \simeq \Delta L_p \sim 0.5 \ dB$.

However, the *JND* in our human hearing is <u>frequency</u> <u>dependent</u> **.and.** also sound pressure level/*SPL*-dependent, as shown in the figure below:
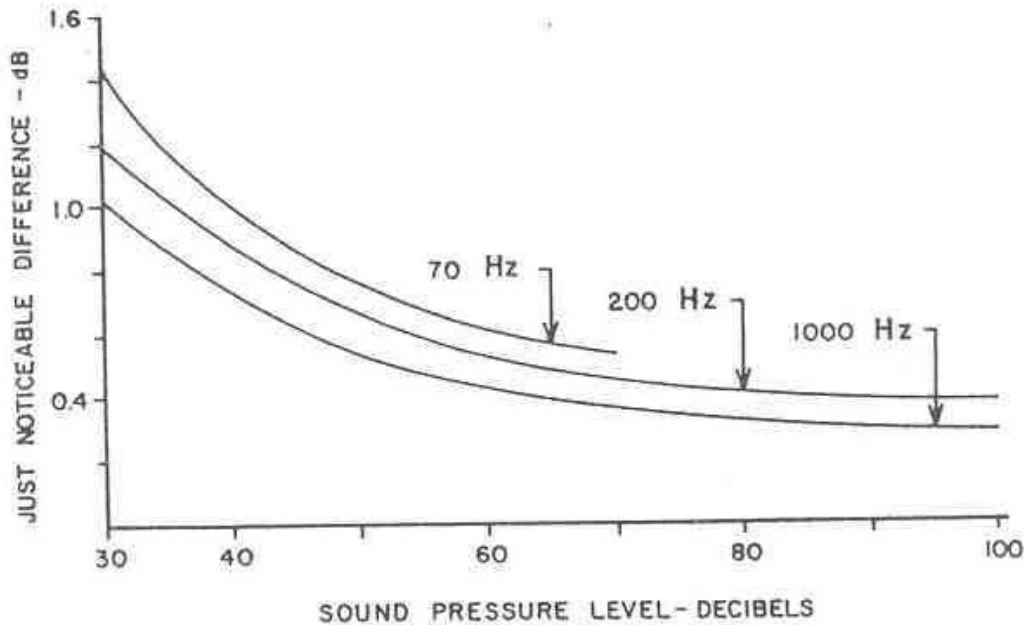

FIG. 2. Just noticeable difference in sound pressure level for three frequencies.

Question(s): Why do we humans hear in the frequency range that we do (20 *Hz* – 20 *KHz*)?
Why do we <u>no</u>t hear in the lower/higher frequency ranges (< 20 *Hz*, > 20 *KHz*)?

It is ***not*** at all an accident that we hear in the frequency range that we do! We humans, as ***social animals***, are primarily interested all-things human (as other social animals are primarily interested in ***their*** own species) – and hence we are primarily interested in hearing human-made sounds – as produced by our own voice(s). The frequency range of sounds produced by our own voice(s) – the totality of the physics associated with air as a medium + vibrating vocal chords in our larynx/voice box + hyoid bone + acoustic cavities of our lungs + throat + mouth + nasal passage/sinus cavity dictates what the acoustic power spectrum of the human voice can/cannot be. Hence over the millions of years of our evolution, our hearing co-evolved with the sounds that our voices make.

It is also not at all an accident that our ears are tuned to be especially sensitive *e.g.* to the sounds/cries produced by our infants and our young in the ~ 1 → few *KHz* range.

It is also ***no*** accident/a ***good*** thing that we do ***not*** hear too well in the infra-sound (*f* < 20 *Hz*) region – because it would have been/would be significantly detrimental to us if our hearing was constantly being "masked" by hearing draft/wind noises as we were walking and/or running!

The Difference Between Two *Uncorrelated* Loudnesses/Sound Intensity Levels:

Using the fact that: $\boxed{\log_{10} A - \log_{10} B = \log_{10}\left(A/B\right)}$

$\triangle L = L_2 - L_1$ = Difference in two Loudnesses (= Difference in two Sound Intensity *Levels*).

where:

$$L_1 = 10\log_{10}\left(I_1/I_o\right) \text{ and } L_2 = 10\log_{10}\left(I_2/I_o\right), \text{ then:}$$
$$\triangle L = L_2 - L_1 = 10\log_{10}\left(I_2/I_o\right) - 10\log_{10}\left(I_1/I_o\right)$$
$$= 10\left[\log_{10}\left(I_2/I_o\right) - \log_{10}\left(I_1/I_o\right)\right]$$
$$= 10\left[\left(\log_{10} I_2 - \log_{10} I_o\right) - \left(\log_{10} I_1 - \log_{10} I_o\right)\right]$$
$$= 10\left[\log_{10} I_2 - \log_{10} I_1\right] = 10\log_{10}\left(I_2/I_1\right)$$

If *e.g.* $I_2 = 2I_1$ then: $\triangle L = L_2 - L_1 = 10\log_{10}\left(2\right) = 10*0.301 = 3.01 \simeq 3\,dB$

i.e. there is only a $\approx 3\,dB$ difference in loudness/intensity levels for 2 (uncorrelated) sounds which differ by a factor of 2× in intensity, $I_2 = 2I_1$.

Adding *Uncorrelated* Sounds:

Two uncorrelated sounds with intensity levels L_1 and L_2 (*e.g.* at the same frequency)

$L_1 = 70\,dB$ and $L_2 = 80\,dB$ (@ $f = 1000$ Hz). Note that $L_2 = 80\,dB$ corresponds to $I_2 = 10\,I_1$:

$$L_1 = 70 = 10\log_{10}\left(I_1/I_o\right) \qquad L_2 = 80 = 10\log_{10}\left(I_2/I_o\right)$$
$$7 = \log_{10}\left(I_1/I_o\right) \qquad\qquad 8 = \log_{10}\left(I_2/I_o\right)$$
$$10^7 = \left(I_1/I_o\right) \qquad\qquad 10^8 = \left(I_2/I_o\right)$$
$$I_1 = 10^7 I_o = 10^7 * 10^{-12} \qquad I_2 = 10^8 I_o = 10^8 * 10^{-12}$$
$$= 10^{-5}\,W/m^2 \qquad\qquad = 10^{-4}\,W/m^2$$
$$\underline{\text{Thus: } I_2 = 10I_1 \text{ or: } I_1 = 0.1I_2}$$

Rule: Must add <u>Intensities</u>, <u>NOT</u> Loudnesses if sounds are *not* correlated
   – *i.e.* if sounds have no *phase coherence*

Then: $I_{sum} = I_{TOTAL} = I_1 + I_2$ {If sounds *are* correlated, then *must* add
        $= I_1 + 10\,I_1 = 11\,I_1$ *amplitudes* via phasor diagram $\rightarrow$ interference effect(s)!!!}

Thus: $L_{sum} = 10\,\log_{10}\left(11I_1/I_o\right) = 10\,\log_{10}\left(I_1/I_o\right) + 10\,\log_{10}\left(11\right) = 70\,dB + 10.4\,dB$
        $= 80.4\,dB \Rightarrow$ only slightly louder than $80\,dB$!!!

Example: Adding *N Uncorrelated* Equal Strength Sounds:

If N uncorrelated sound sources are superposed/added together, each with the same individual sound intensity I_1, the resulting loudness level is:

$$\boxed{L_{sum} = 10 \, \log_{10}(N \, I_1/I_o) = 10 \, \log_{10}(I_1/I_o) + 10 \log_{10}(N)} \quad \Rightarrow \quad \boxed{\Delta L = 10 \log_{10}(N)}$$
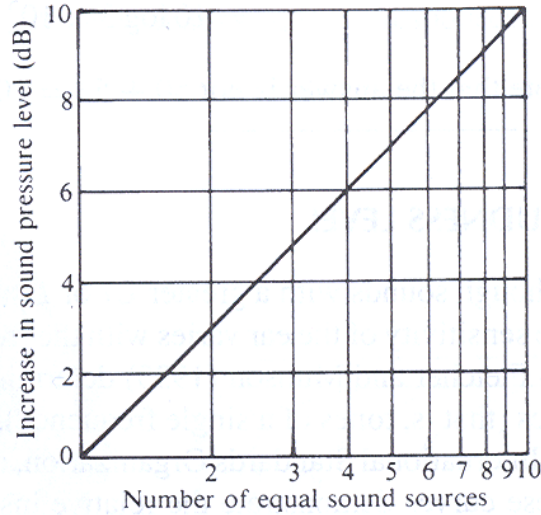
Adding *Correlated* Sounds:

If sounds are correlated (*i.e.* have a stable phase relation to each other) then <u>must</u> add sounds together at the <u>amplitude</u> level
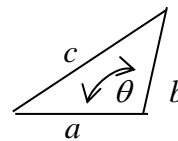
Suppose have two {*RMS*} over-pressure amplitudes p_1 and p_2

"$p_{ToT} = p_1 + p_2$" but p_{ToT} actually depends on the <u>phase relation</u> between $p_1 + p_2$

Use ***phasor diagram*** to calculate p_{ToT}:
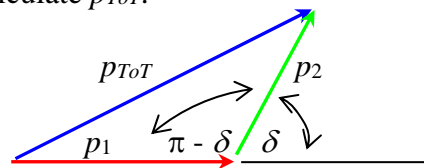
δ = relative phase angle between p_1 and p_2

Use the law of cosines:

$$c^2 = a^2 + b^2 - 2ab \cos\theta \qquad\qquad \theta = (\pi - \delta)$$

$$p_{ToT}^2 = p_1^2 + p_2^2 - 2p_1p_2 \cos(\pi - \delta) \text{ but: } \cos(\pi - \delta) = \underbrace{\cos\pi}_{=-1}\cos\delta + \underbrace{\sin\pi}_{=0}\sin\delta = -\cos\delta$$

$$\therefore \; p_{ToT}^2 = p_1^2 + p_2^2 + 2p_1p_2 \cos(\delta) \; \propto I_{ToT}$$

$$\text{or: } \; I_{ToT} = I_1^2 + I_2^2 + 2\sqrt{I_1}\sqrt{I_2}\cos\delta$$

$$p_{ToT} = \sqrt{p_1^2 + p_2^2 + 2p_1p_2\cos\delta}$$

Phase difference: $\delta \equiv \varphi_{p_1} - \varphi_{p_2}$	$\cos \delta$	Resultant Amplitude: p_{ToT}
0°	$+1$	$p_1 + p_2$
45°	$+\dfrac{1}{\sqrt{2}}$	$\sqrt{p_1^2 + p_2^2 + \sqrt{2}\, p_1 p_2}$
90°	0	$\sqrt{p_1^2 + p_2^2}$
135°	$-\dfrac{1}{\sqrt{2}}$	$\sqrt{p_1^2 + p_2^2 - \sqrt{2}\, p_1 p_2}$
180°	-1	$p_1 - p_2$
225°	$-\dfrac{1}{\sqrt{2}}$	$\sqrt{p_1^2 + p_2^2 - \sqrt{2}\, p_1 p_2}$
270°	0	$\sqrt{p_1^2 + p_2^2}$
315°	$+\dfrac{1}{\sqrt{2}}$	$\sqrt{p_1^2 + p_2^2 + \sqrt{2}\, p_1 p_2}$
360°	$+1$	$p_1 + p_2$

If have *N* sources of ***correlated*** sounds, then simply have to work out the phasor diagram for *N* individual phase-related amplitude components to obtain resultant/overall total amplitude.

The sound intensity $\left| I_{ToT} \right| \propto \left| p_{ToT} \right|^2$ The (magnitude) of the sound intensity, *I* is proportional to the <u>square</u> of the (magnitude) of the total/net *RMS* over-pressure amplitude, $|p|^2$.

For a "free-field" sound situation, the corresponding SPL and/or SIL for each entry in the above table can be computed using:

$$SPL = L_p \equiv 10\log_{10}\left(\frac{p_{tot}^2}{p_o^2}\right) = 10\log_{10}\left(\frac{p_{tot}}{p_o}\right)^2 = 20\log_{10}\left(\frac{p_{tot}}{p_o}\right) \ (dB)$$

$$\simeq SIL = L_I \equiv 10\log_{10}\left(\frac{I_{tot}}{I_o}\right)$$

Loudness of Complex (Poly-Frequency) Tones & The Critical Band:

If two pure tones of frequency f_1 and f_2 each individually have the same *SPL*, but are within the critical band Δf_{crit} of each other, we perceive the overall sound as not as loud as when the two frequencies are well-separated from each other, *i.e.* <u>outside</u> the critical band. The following figure shows this effect:
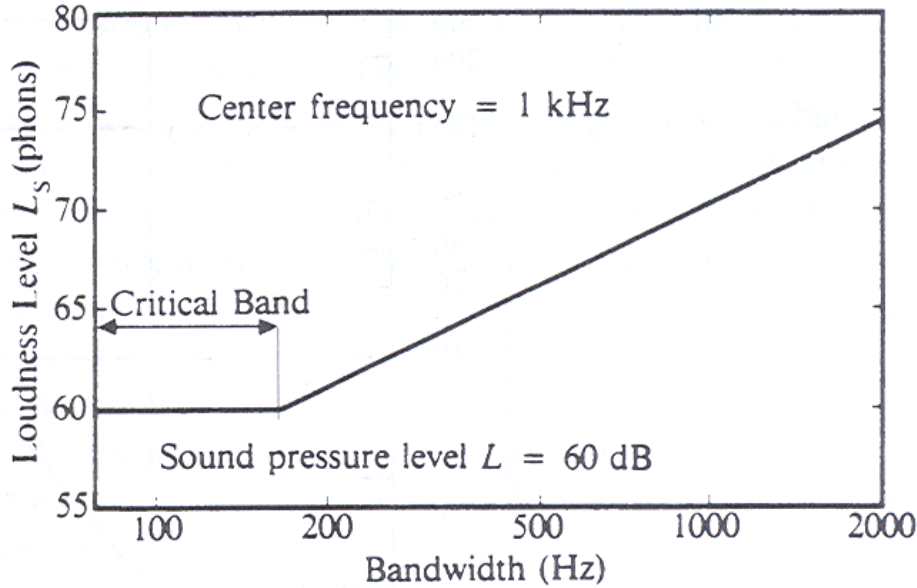
If white noise (= all frequencies in the audio band, of equal amplitude) is used to compare the overall apparent loudness of white noise e.g. with that of a pure tone sine wave type signal at $f = 1\ KHz$, as a function of *SPL* (in *dB*), the result shown in the graph on the right is obtained:
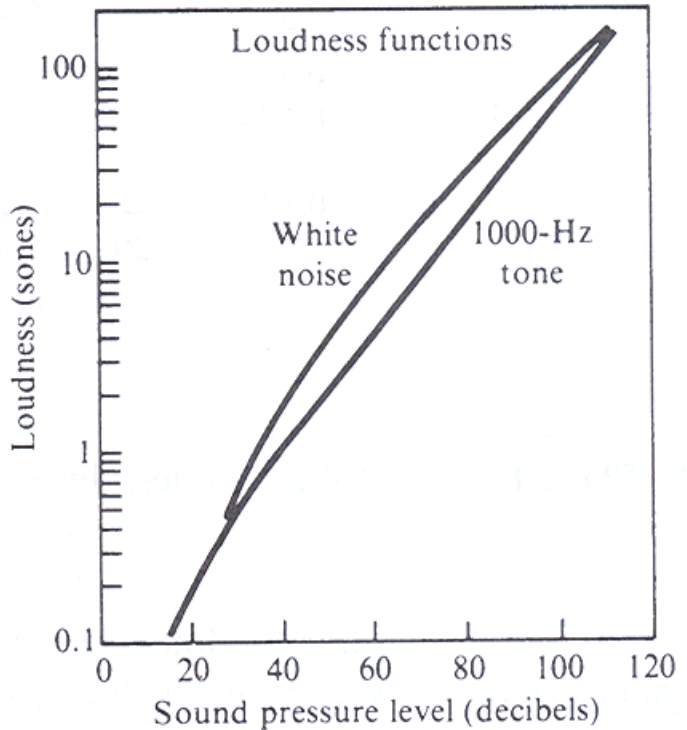
Masking:

   Superposition of two (or more) sound signals can make it difficult to "decode" one of them, *e.g*. listening to a friend talk to you in a crowded/noisy room – known as ***masking***.

Masking problem increases with age and with hearing damage (exposure to loud noises).

e.g. a $L = 60$ *dB*, $f = 1200$ *Hz* masking tone.

   See curve below for the *JND vs*. frequency for 1200 *Hz* masking tone *vs*. intensity level of masking tone. One can see the effect of the critical band and also see effect of 2nd harmonic (difference frequency, $\Delta f = f - f_{mask} = 1200$ *Hz*) – due to quadratic non-linear response term(s) in our ear (and/or brain)! {See UIUC P406POM lecture notes on distortion for details}
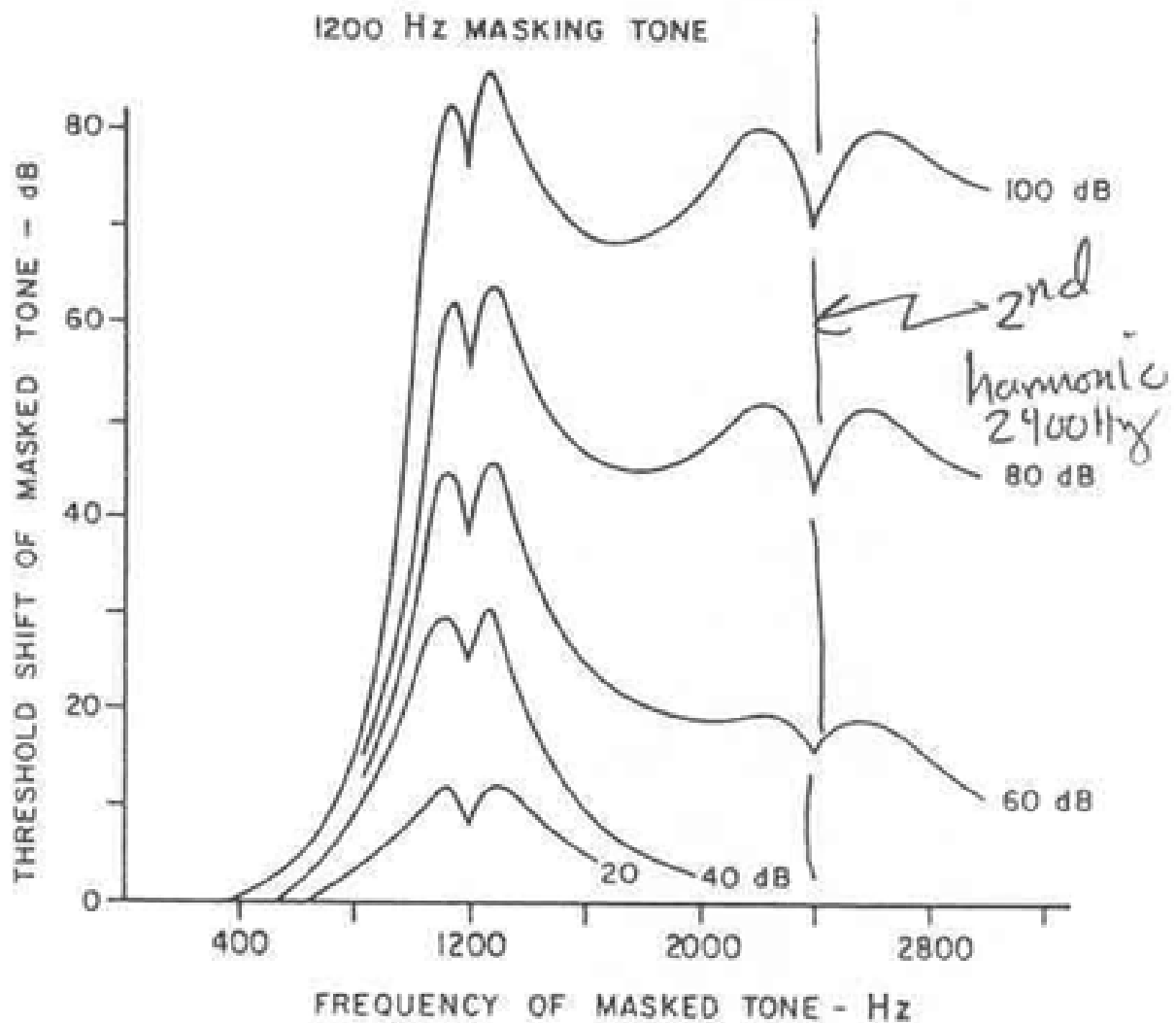

Fig. 5. Masking curves for a masking tone of 1200 hertz.

Hearing Loss/Disorders:

At birth, humans can hear over the frequency range 20 *Hz* – 20 *KHz*.

As we grow older, we experience "natural" hearing loss (Presbycusis), particularly in the higher frequency range. (See age-related hearing plots below…)

Very <u>loud</u> sounds can temporarily and/or permanently damage sensitive hearing nerves in the cochlea. Repeated acoustic trauma can cause permanent (and profound) hearing loss or deafness.

If you have ever experienced a temporary hearing loss due to loud sounds – you have had a warning! *n.b.* the <u>stereocilia</u> do regenerate (daily), but if the hair cells are damaged/die, there is <u>no</u> regrowth of hair cells!

* tinitis – ringing in the ears (can be due to more than one cause)
* ear infections can also lead to hearing loss, especially in young children & infants.
* loud explosions (artillery shells – military)
* determining factor of damage to hearing is <u>product</u> of exposure time $\times$ loudness level
* Hearing loss due to <u>over</u>-<u>stimulation</u> of hair cells – causes excito-toxicity – too much Ca^{2+} poisons neurons in the auditory nerve…

Extreme Acoustic Trauma - Guinea Pig Stereocilia Damage to 120 *dB* Sound Pressure Levels:
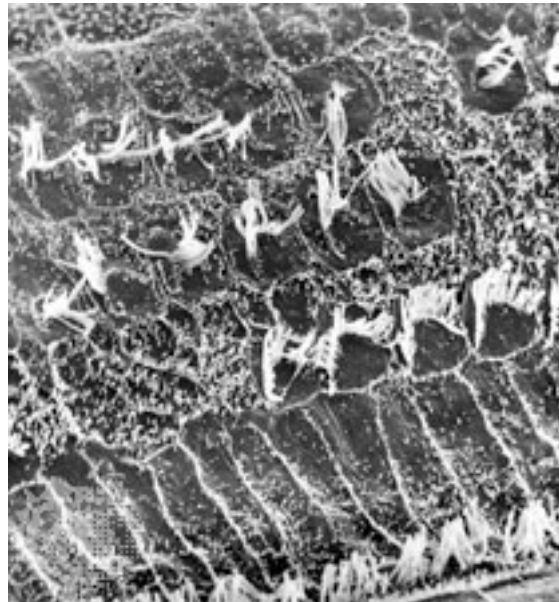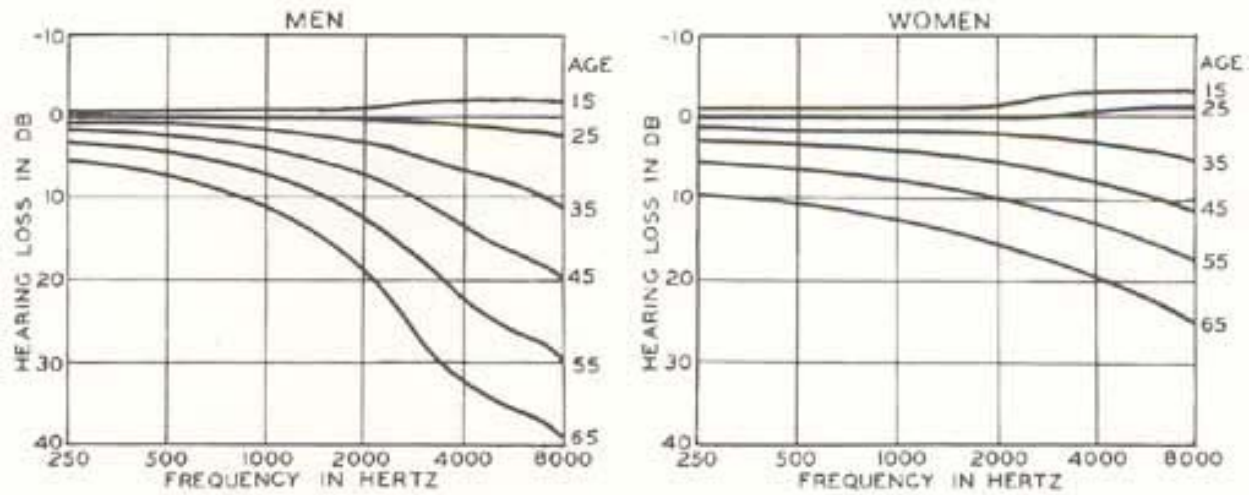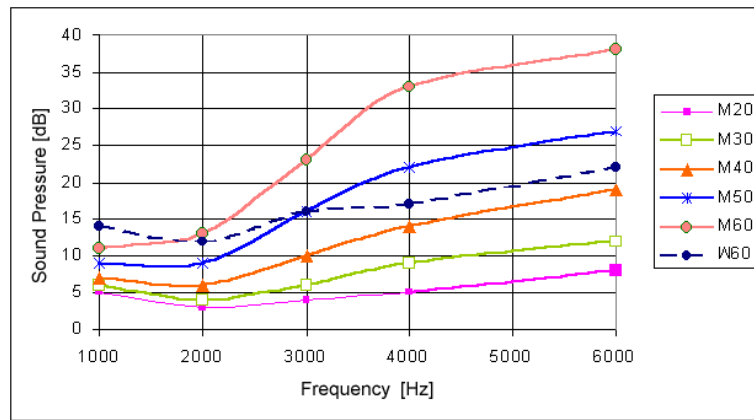

Before Exposure   After Exposure

Protect your hearing – it's all you've got!!!

Presbycusis: Hearing Loss *vs*. Age – Different for Men *vs*. Women:

Age-Related Hearing Sensitivity: Men (M_age) *vs*. Women (W_age) *vs*. Frequency:

Comparison of Normal *vs*. Age-Related Hearing Loss *vs*. Exposure to Gun Fire:
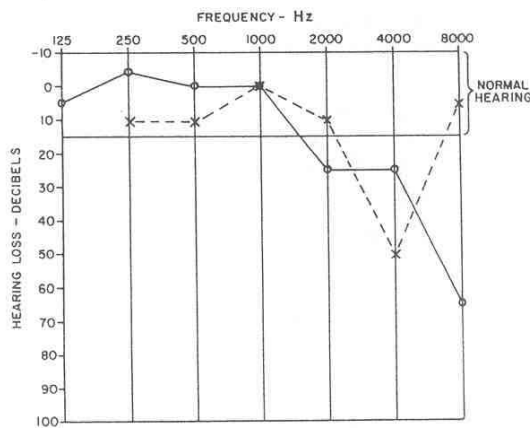

FIG. 6. Audiograms. Solid curve: presbycusis. Dotted curve: hearing loss produced by exposure to gunfire.

Median Noise-Induced Hearing Loss *vs*. Frequency – Long-Term Exposure:
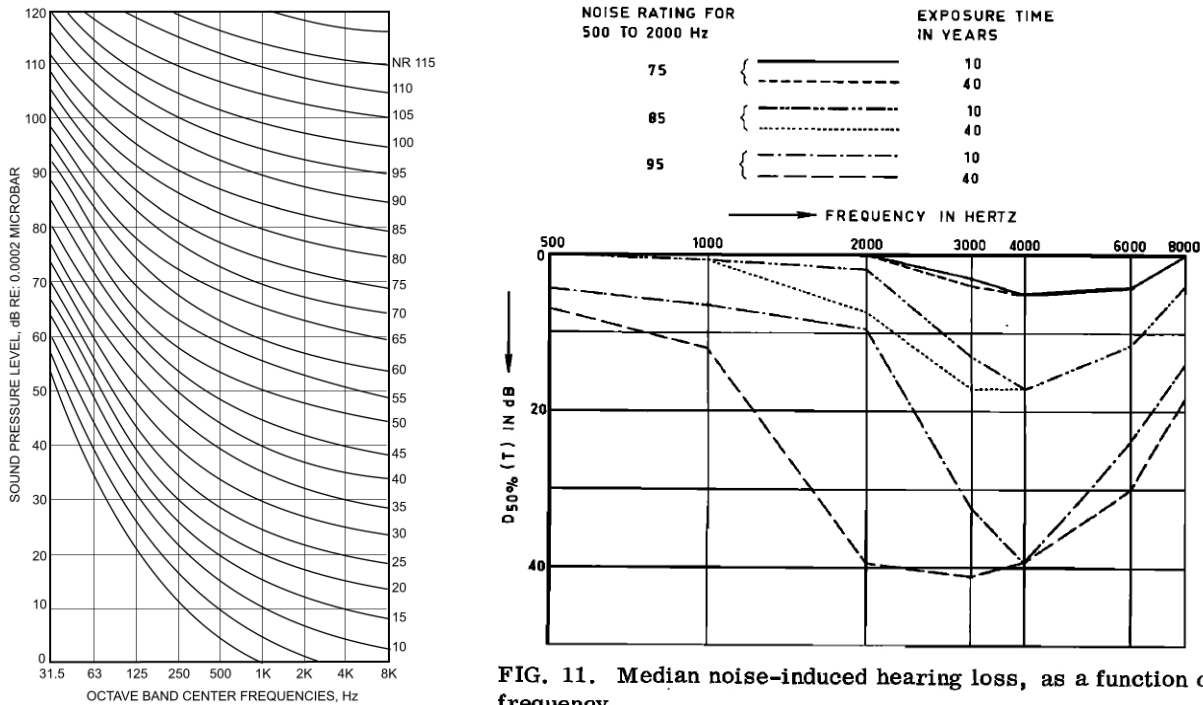
Noise Rating Curves:

FIG. 11. Median noise-induced hearing loss, as a function of frequency.

W. Passchier-Vermeer, "Hearing Loss Due to Continuous Exposure to Broad-Band Noise", Jour. Acoust. Soc. Am., Vol. 56, No. 5, *p*. 1585-1593, November 1974

OSHA Maximum Permissible Daily Exposure Limits – Industrial Noise:

Since 1970, there exist legal time limits for exposure to noise in the workplace – for industries doing business with the U.S. federal government, as shown in the figures and left-hand table below*, assuming an 8 hour workday, 5 days/week. These limits were obtained from extensive analyses of workplace-related hearing loss – permanent threshold shifts in human hearing, expressed in *dB* units. Note that the legal limits provide protection only for frequencies necessary for the understanding of speech. No allowance was made for exposure to noise outside of the work place. These legal noise level limits were determined so as to protect 85% of the exposed population, while assuming that financial compensation would be provided for the remaining 15%, who were assumed to be more susceptible to hearing loss due to noise exposure.

If daily industrial/work noise exposure is composed of two or more periods of noise exposure at different levels, their combined effect can be taken into account via the daily exposure requirement that $\sum t_i/T_i < 1$, where t_i is the ith exposure time at SPL_i and T_i is the OSHA exposure time limit at that SPL_i.

Recommended non-occupational daily noise exposure time limits are also shown figures in the right-hand table below**, for comparison.
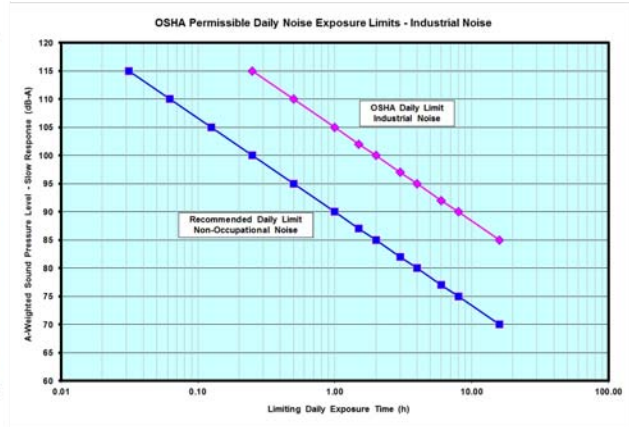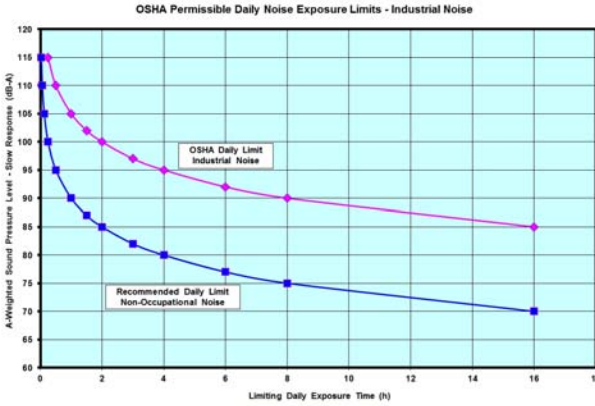
OSHA* Daily Exposure Time Limit – Industrial Noise (hrs)	SPL (dB-A)
16	85
8	90
6	92
4	95
3	97
2	100
1.5	102
1	105
½	110
¼	115

Recommended** Daily Exposure Time Limit – Non-Occupational Noise (hrs)	SPL (dB-A)
16	70
8	75
6	77
4	80
3	82
2	85
1.5	87
1	90
½	95
¼	100
1/8 hr ~ 12 min	105
1/16 hr ~ 6 min	110
1/32 hr ~ 3 min	115

  * More info available on online at the OSHA website:
http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=standards&p_id=9735

** A. Cohen, J. Anticaglia and H.H. Jones, " 'Sociocusis'- Hearing Loss From Non-Occupational Noise Exposure", *Sound and Vibration* Vol. 4, *p*. 12-20, November 1970.

   Note from the above tables that the daily exposure time limit(s) decrease by a factor of 2× for each $\Delta SPL = 5$ *dB-A* increase, which is also reflected in the above RHS semilog-*x* plot of *SPL vs.* $\log_{10}$(Daily Exposure Time), *i.e.* a straight-line $y(x) = mx + b$ relationship, where $y(x) = SPL$, $x = \log_{10}$(Daily Exposure Time), intercept $b = 105$ *dB* {90 *dB*} for the OSHA {Recommended} curve(s), respectively [since $\log_{10}(1.0) = 0$], and slope $m = -5$ *dB*/ $\log_{10}(2)$.

   Since $SPL = 10 \log_{10}(I/I_o)$ (*dB*), the above value of the slope *m* tells us that two different values of acoustic intensity limits I_1, I_2 and their associated maximum Daily Exposure Times $\Delta t_1^{exp}, \Delta t_2^{exp}$ are related to each other by:

$$\boxed{I_1 \sqrt{\Delta t_1^{exp}} = I_2 \sqrt{\Delta t_2^{exp}} = \text{constant}}$$

i.e. Damage to our hearing is proportional to the ***square-root*** of the exposure time $\sqrt{\Delta t^{exp}}$, as opposed to varying ***linearly*** with the exposure time Δt^{exp}, since $E = I \cdot \Delta t \cdot A_\perp \left(Joules\right)$.

Perceptual Hysteresis *vs*. Sound Pressure Level Effect:

   Many details/phenomena associated with human hearing need to be carefully taken into account in order to eliminate/account for bias effects – and thus reliably/accurately determine curves shown in the above figures – *e.g.* age & sex (*i.e.* men *vs*. women) of the human subjects participating in these statistical studies, to properly account for frequency-dependent, sound intensity fatigue and hysteresis effects, *etc*. Some examples are shown in the figures below:
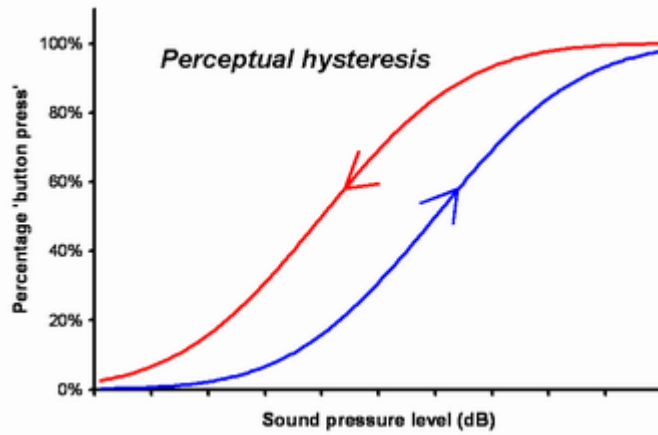
## Threshold Determination *vs*. Sound Pressure Level Trial:
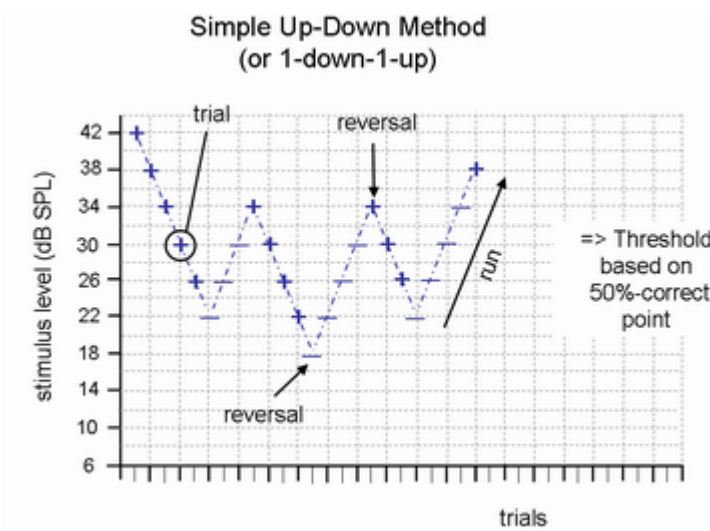
Brain Processing of Sound Signals:

Signals from basilar membrane in cochlea sent to <u>various</u> areas of the brain – at the frequenc(ies) of the acoustical signals incident on the ears (!!!). From recent technological advances – development of functional Magnetic Resonance Imaging (fMRI) – researchers have learned that a many areas of the brain simultaneously process these sound signals – underscoring the importance hearing (& music) to human beings. If interested in learning more about this, the recently-published book "*This is Your Brain on Music – the Science of Human Obsession*" by Daniel J. Levitin is *highly* recommended – see his website: http://www.yourbrainonmusic.com/

Auditory signals transmitted to the brain via auditory nerves undergo much additional processing in the brain.

* Speech sounds: mostly processed in *left* hemisphere of brain.
* Music sounds: mostly processed in *right* hemisphere of brain.
* <u>Separate</u> processing centers for consonance (human-like sounds) & dissonance
  (not human-like sounds)!
* Musical/sound/voice memories stored in ***several*** areas of the brain – explains robust retentivity
 / longevity/stability of acoustical/sound-type memories!
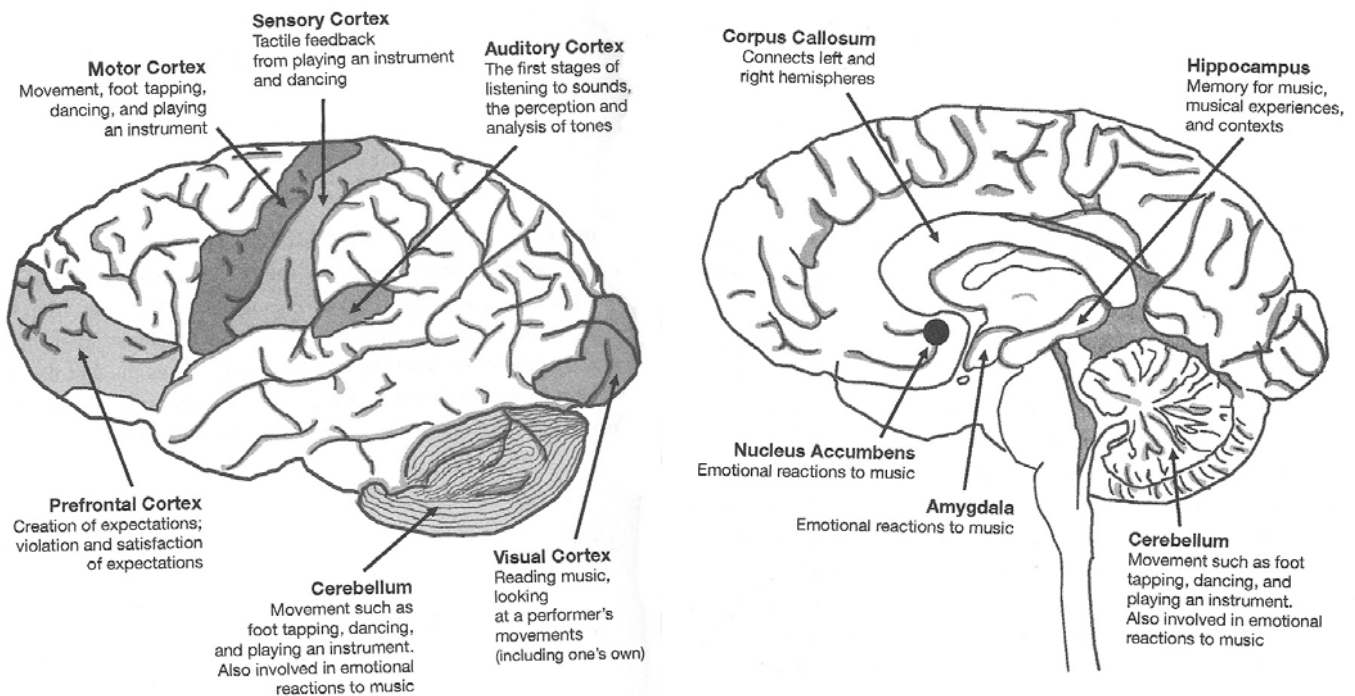
Because many areas of the brain process sound signals, there are also many possible ways for brain to malfunction. Again, researchers have learned a great deal on this via fMRI studies over the past decade. If interested in this subject, the recently published book "*Musicophilia – Tales of Music and the Brain*" by Oliver Sacks, MD is *highly* recommended – see his website: http://musicophilia.com/

Legal Disclaimer and Copyright Notice:

Legal Disclaimer:

The author specifically disclaims legal responsibility for any loss of profit, or any consequential, incidental, and/or other damages resulting from the mis-use of information contained in this document. The author has made every effort possible to ensure that the information contained in this document is factually and technically accurate and correct.

Copyright Notice: