COMP 546

Lecture 20

Head and Ear

Thurs. March 29, 2018

Impulse function at $t = 0$.

$$I(X, Y, Z, t) = \delta(X - X_0,\ Y - Y_0,\ Z - Z_0,\ t)$$

To define an impulse function properly in a continuous space requires more math.
Let's not spend our time doing that, since we just want qualitative behavior here.

Sound obeys the wave equation.

So, how is this function defined $t \neq 0$?

Impulse becomes expanding sphere

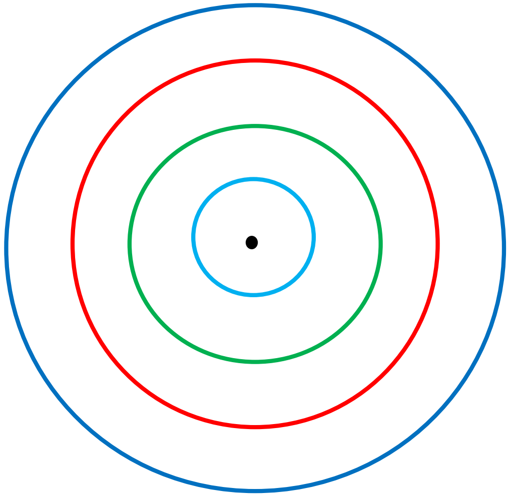
One can show that this follows from the wave equation.

$t = 4\,\Delta\,t$

$t = 3\,\Delta\,t$

$t = 2\,\Delta\,t$

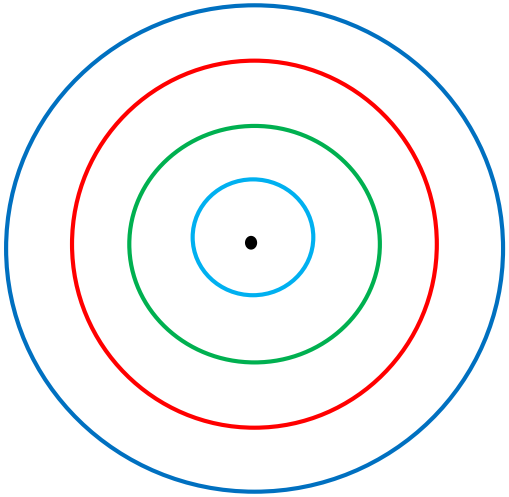
$t = \,\Delta\,t$



$r = v\,t$

Impulse sound energy is spread over a thin sphere of *fixed thickness*

and of area $4\pi\, r^2$ where $r^2 = (X - X_0)^2 + (Y - Y_0)^2 + (Z - Z_0)^2$.



$$r = v\, t$$

$$I^2 \sim \frac{1}{r^2}$$

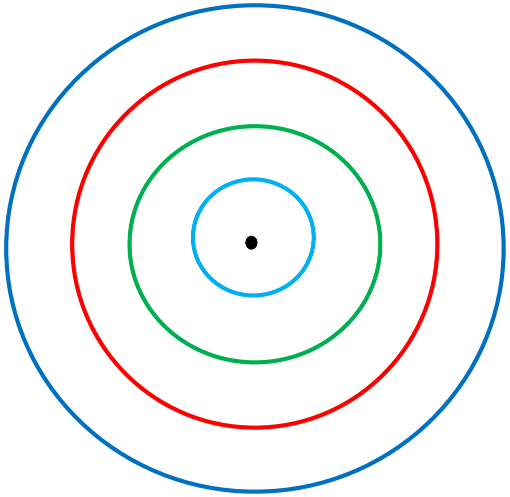
So, SPL $I \sim \dfrac{1}{r}$

$$I(X, Y, Z, t)$$

$$= \begin{cases} I_{src} \;\; \delta(X - X_0,\; Y - Y_0,\; Z - Z_0), & \text{when } t = 0 \\[2em] \dfrac{I_{src}}{r} \, \delta(r - v\, t), & \text{when } t > 0 \text{ and} \end{cases}$$

$$r = (X - X_0)^2 + (Y - Y_0)^2 + (Z - Z_0)^2$$

I_{src} is constant (~energy in impulse)



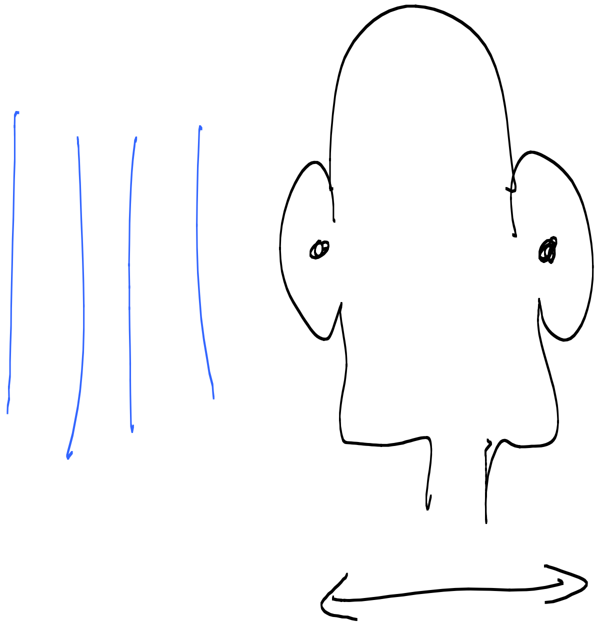
We can write a general sound source a sum of impulse functions:

$$I_{src}(t) = \sum_{t'=0}^{T-1} \delta(t - t')\, I_{src}(t')$$

Far from the source, where r is large, the wavefront is approximately locally planar.

Binaural hearing
(preview of next lecture)

If the sound arrives from the left (assuming planar wavefronts), what is the interaural delay?

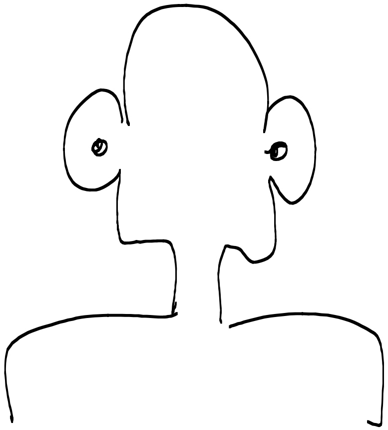


d = 17 cm

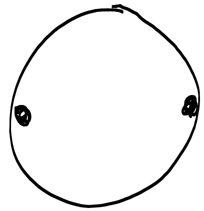
$$t = \frac{d}{v} = \frac{.17}{340}$$

$$\approx .5 \; ms$$

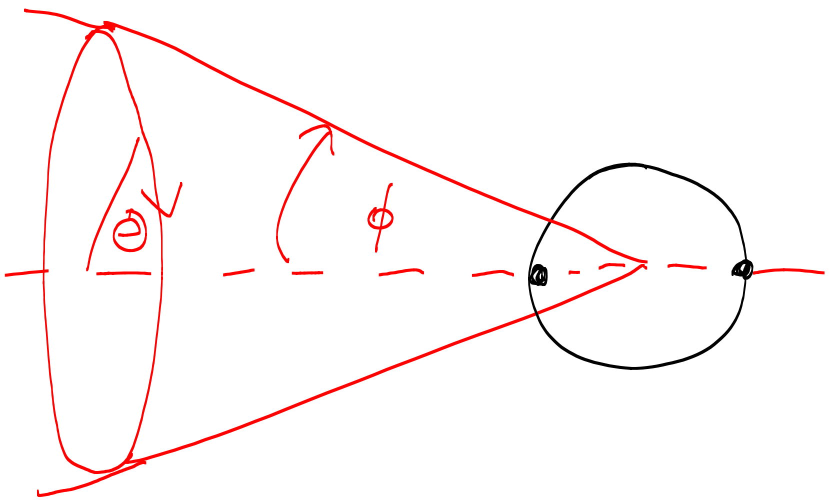
Naïve model: cone of confusion



Model head, shoulders, ears as a sphere.



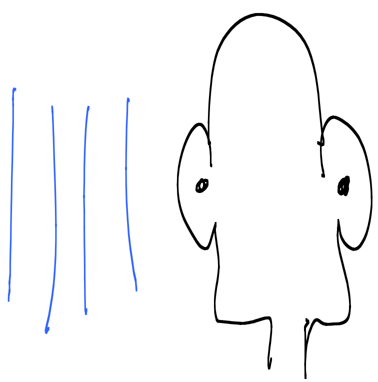
All incoming directions on a cone define the same delay & shadow effect.



Exercise: use time delay τ to estimate cone angle ϕ

Interaural differences

How can the auditory system estimate the delay and shadowing ?   Here is a simple model:



$$I_l\,(t) = \alpha \;\; I_r(t - \tau) + n(t)$$

shadow (attenuation) | delay | noise

Maximum likelihood: find the α and τ that minimize

$$\sum_{t=1}^{T} \{\, I_l\,(t) - \alpha \;\; I_r(t - \tau)\;\; \}^2$$

where $\tau < 0.5\, ms$.

To find the α and τ that minimize

$$\sum_{t=1}^{T}\{I_l\,(t)^2 - \alpha\, I_l\,(t)\, I_r(t-\tau) + I_r(t-\tau)^2\}$$

we first find the τ that maximizes

$$\sum_{t} I_l\,(t)\, I_r(t-\tau)\,.$$

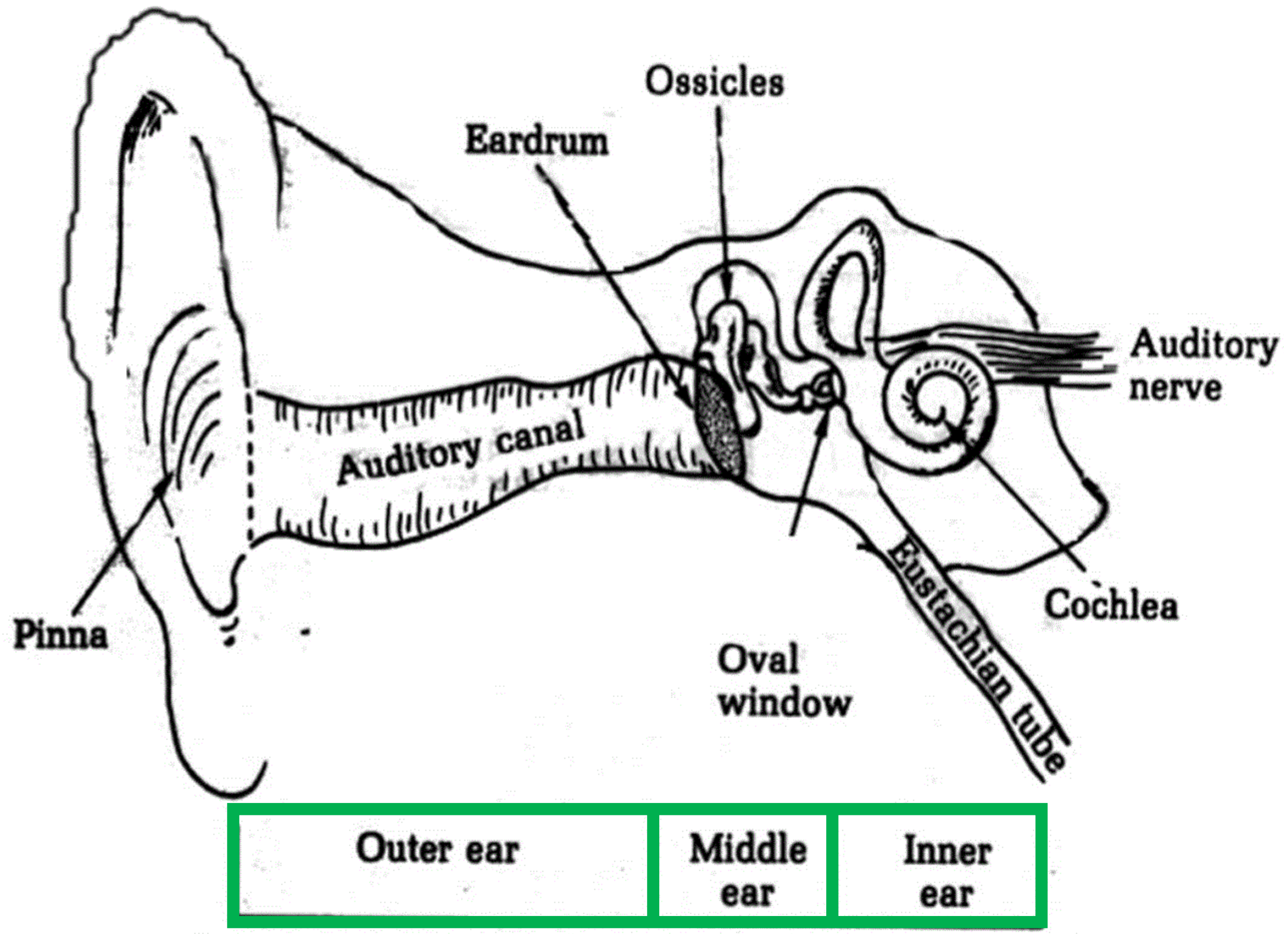
This ignores the small dependence of the 3rd term above on τ.

Then estimate α (shadowing):

$$\alpha^2 = \frac{\sum_{t=1}^{T} I_l(t)^2}{\sum_{t=1}^{T} I_r(t-\tau)^2}$$

Note that this gives two cues which we can combine.

The Human Ear

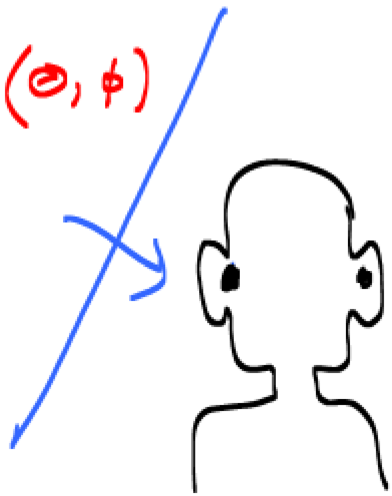


Outer Ear

Next ten slides:

How do head and outer ear transform the sound that arrives at the ear from various directions ?

Head related impulse response (HRIR)



Suppose sound is from direction (ϕ, θ).

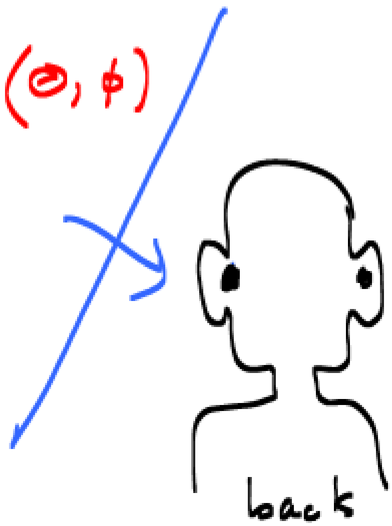
The wave is planar when it arrives at the head.

If the source is an impulse then sound measured at the ear drum of ear i is:

$$I(t) = h_i\,(t;\ \phi, \theta) * \delta(r - vt)$$

left or right

Sound source $I_{src}(t;\ \phi,\theta)$ transformed



Suppose sound is from direction (ϕ,θ) and emits $I_{src}(t;\ \phi,\theta)$.

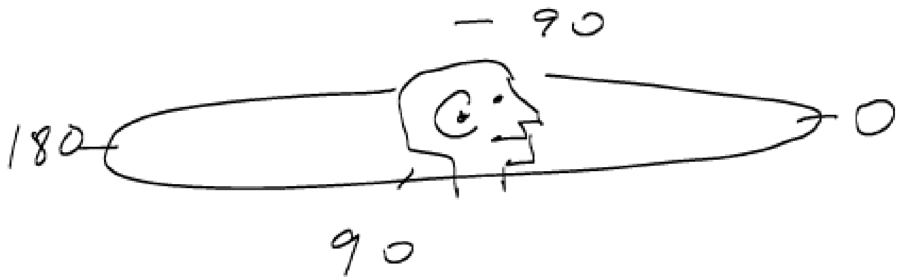
Then the sound measured at the ear drum of ear i is:

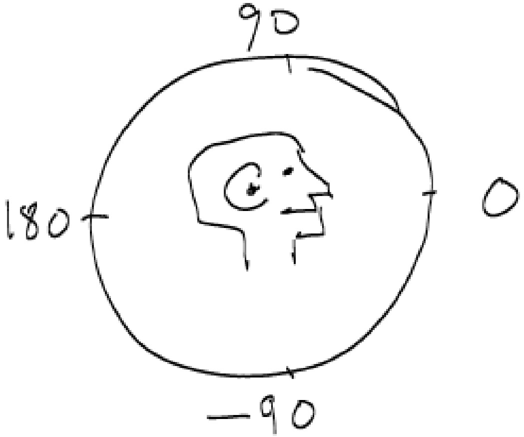
$$I(t) = h_i\,(t;\ \phi,\theta) * I_{src}(t;\ \phi,\theta)$$

(Ignoring time delay from source to ear.)

KEMAR mannequin

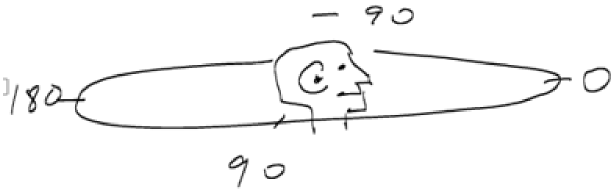
In following slides, I will show HRIR measurements $h_i\ (t;\ \ \phi, \theta)$.



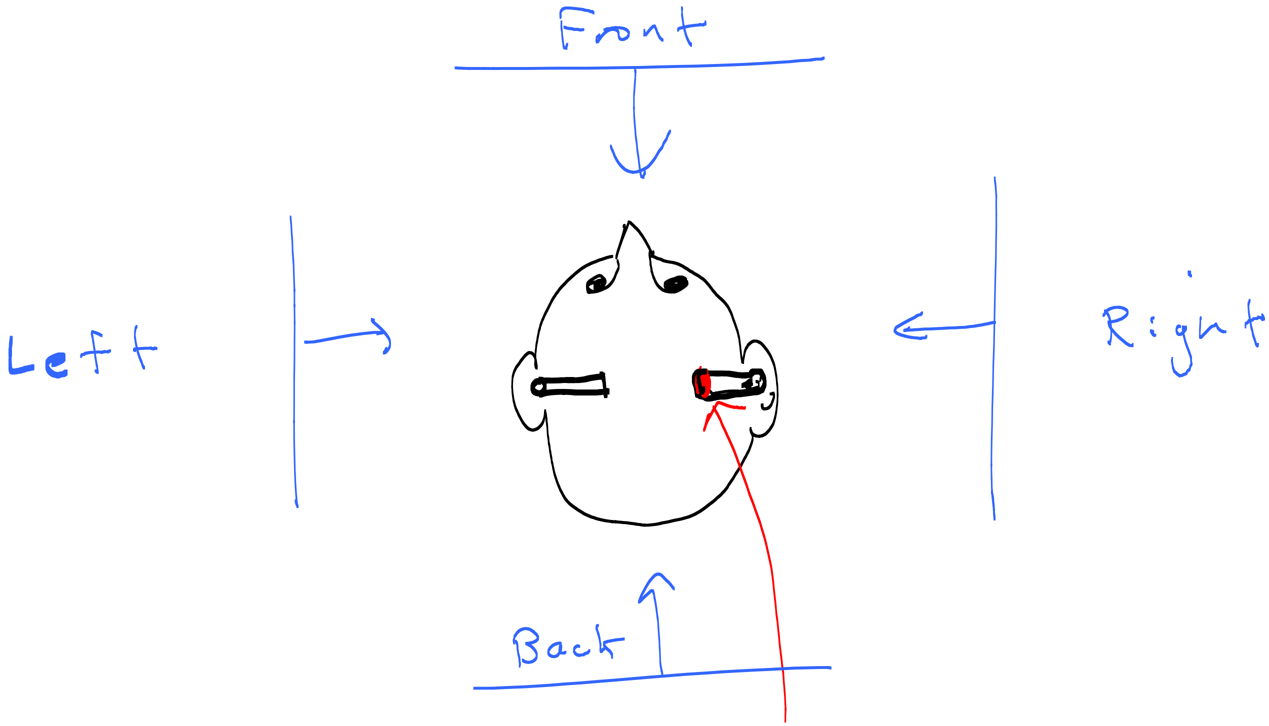


azimuth θ

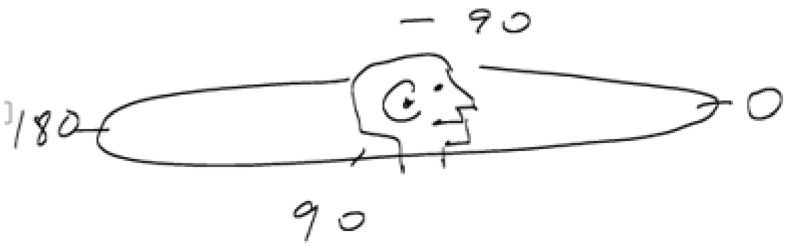
elevation ϕ



Azimuth θ (Elevation $\phi = 0$)

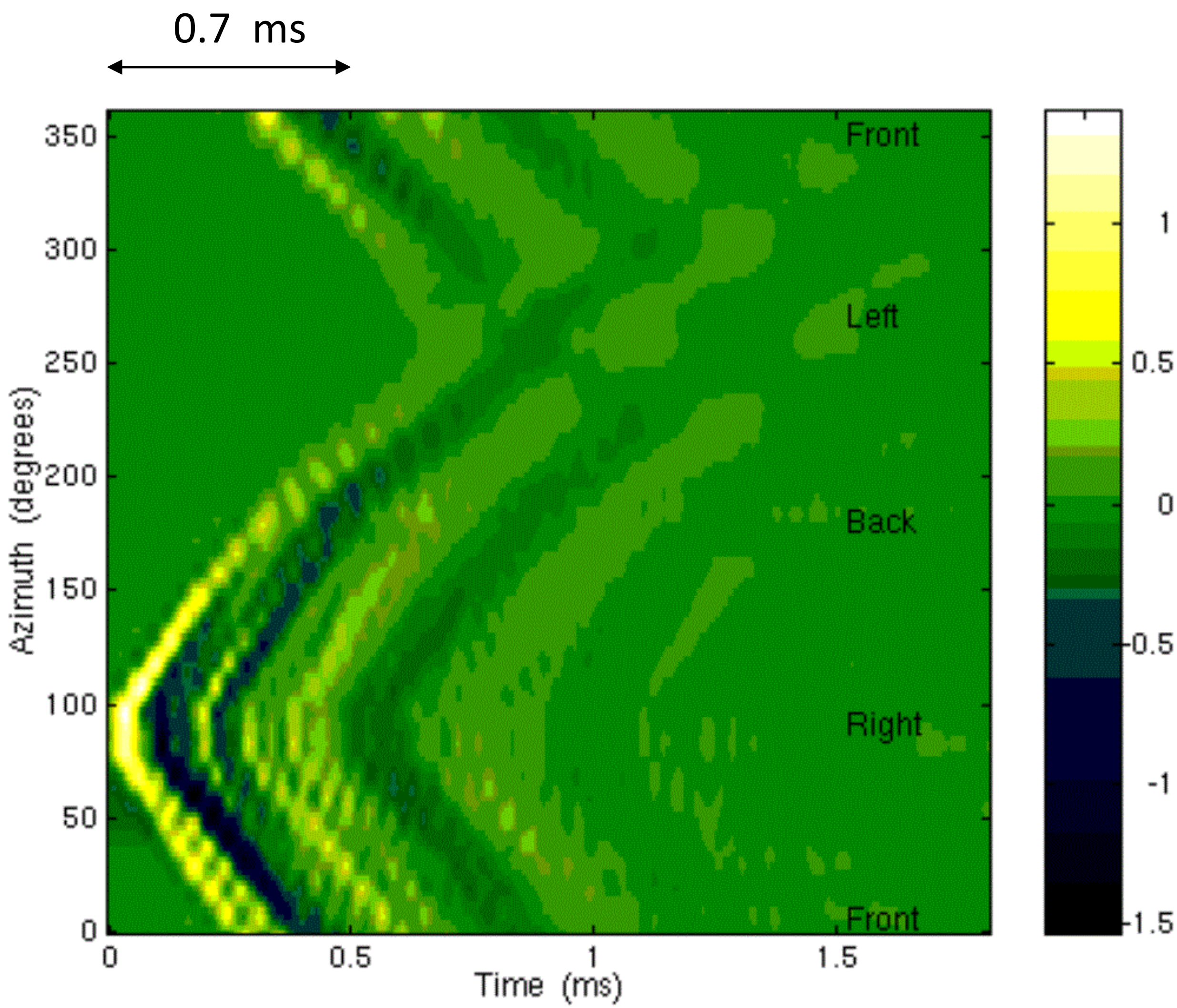


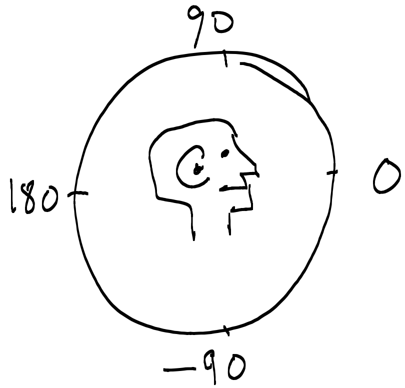
Suppose sound is measured at right ear drum.



HRIR

Source direction (azimuth)

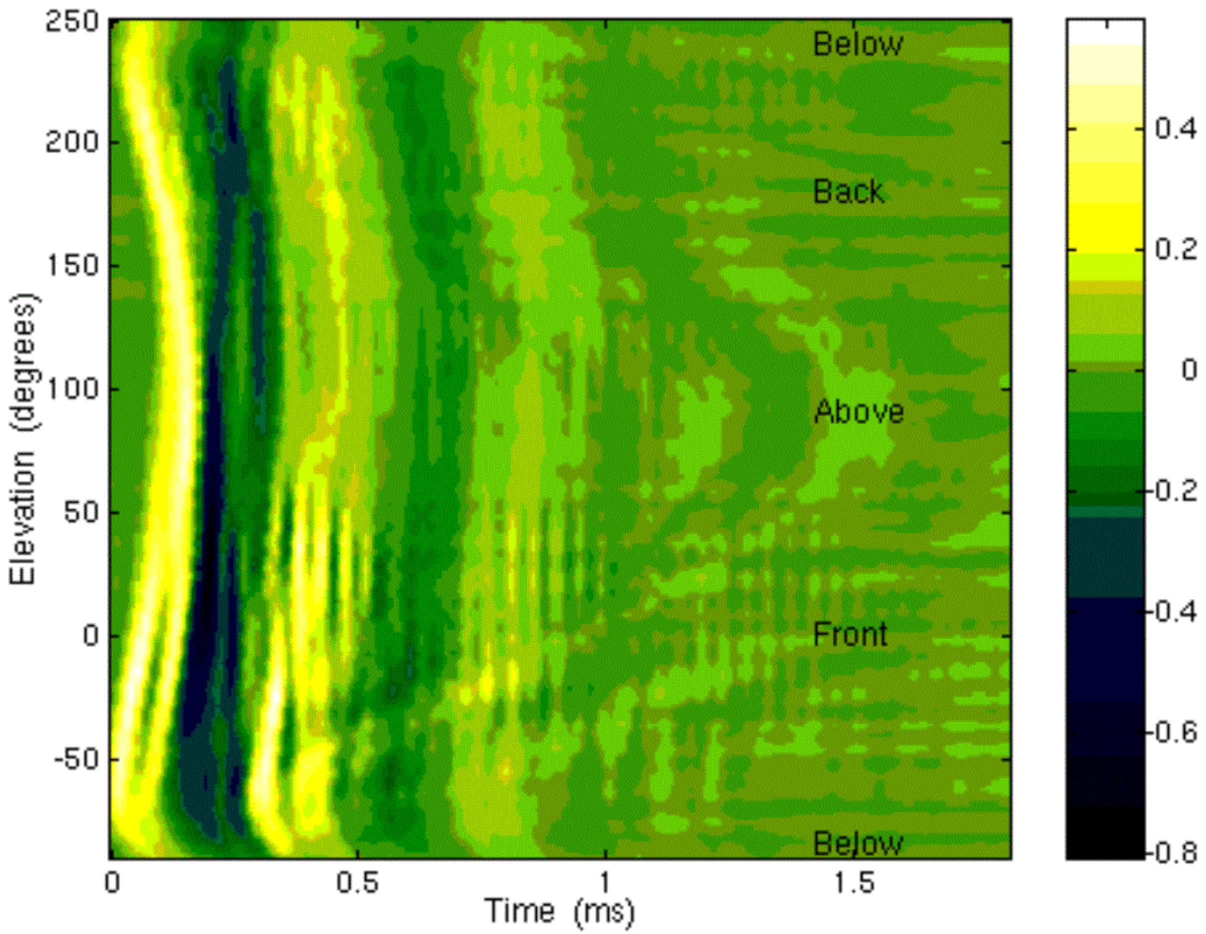




Arrival time differences are not as significant when azimuth = 0 and elevation is varied.

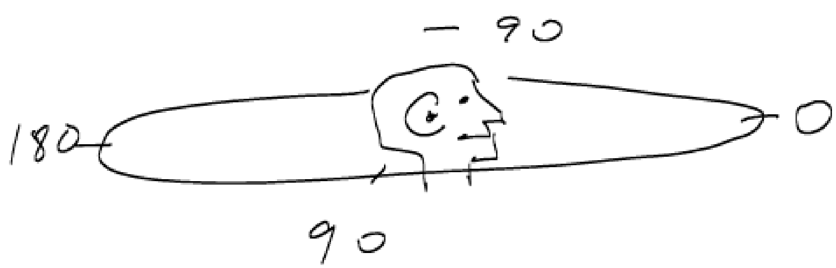
HRIR

Source direction (elevation)

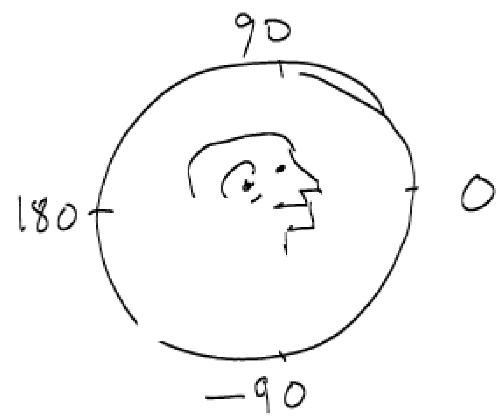


If head is symmetric about the medial plane (left/right), then :

$$h_{left}\,(t;\ \ \phi, \theta)\ \ =\ \ h_{right}\,(t;\ \ \phi, -\theta)$$



azimuth θ



elevation ϕ

$$I_{right}(t;\phi,\theta) \;\; = \underbrace{h_{right}\,(t;\;\; \phi,\theta)}_{\text{HRIR}} * I_{src}(t;\;\; \phi,\theta)$$

For each incoming sound direction (ϕ,θ), what is the Fourier transform with respect to variable t ?

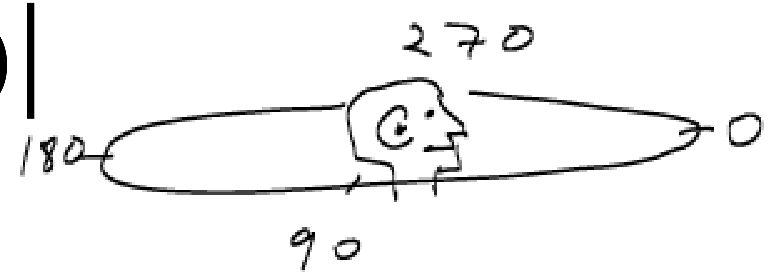
$$I_{right}(t;\phi,\theta) \;=\underbrace{h_{right}\,(t;\;\;\phi,\theta)}_{\text{HRIR}} * I_{src}(t;\;\;\phi,\theta)$$

For each incoming sound direction (ϕ,θ), what is the Fourier transform with respect to t ?

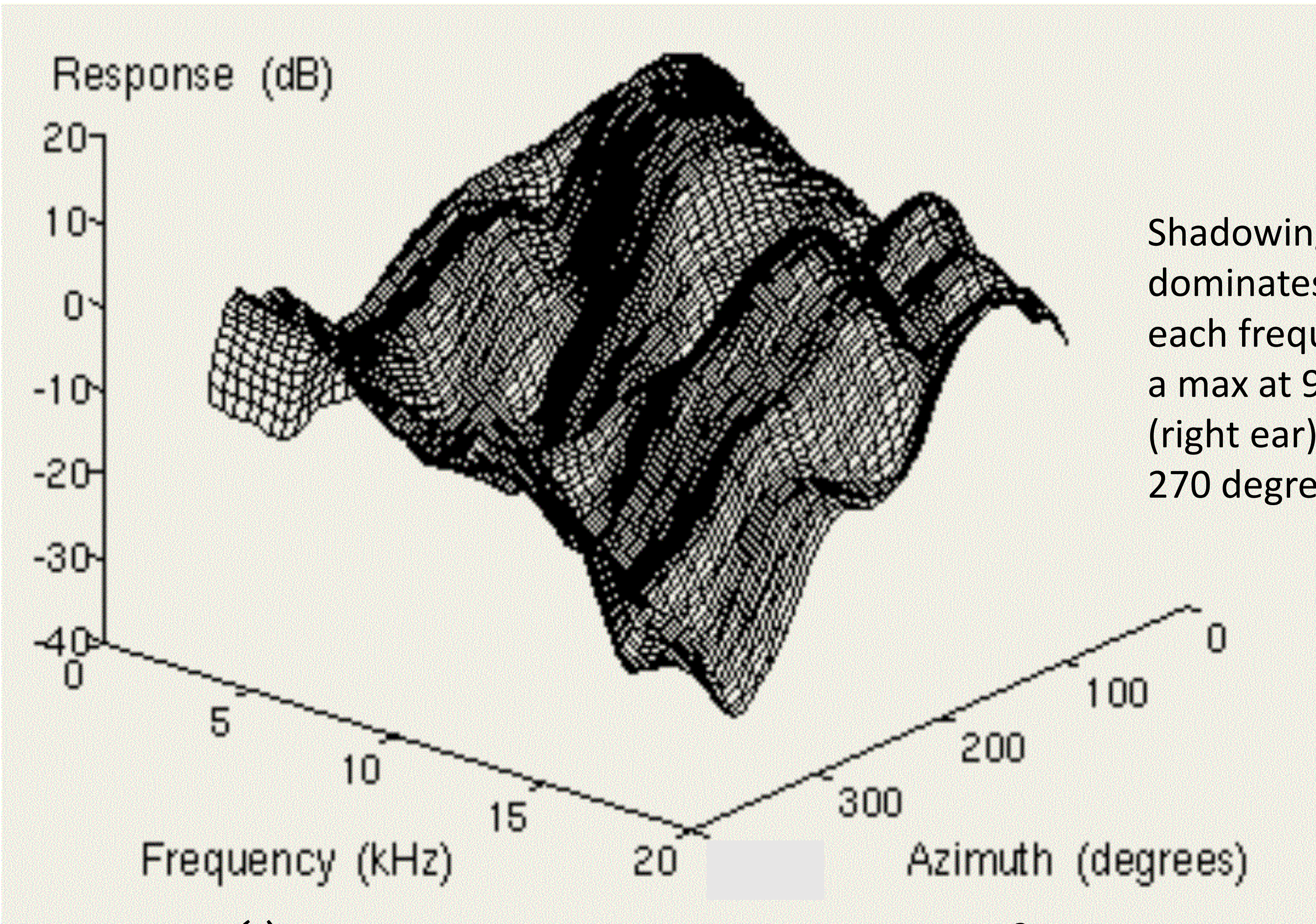
$$\hat{I}_{right}(\omega;\phi,\theta) \;=\underbrace{\hat{h}_{right}\,(\omega;\;\;\phi,\theta)}\quad \hat{I}_{src}(\omega;\;\;\phi,\theta)$$

Head Related "Transfer Function" **(HRTF)**

HRTF $\left|\hat{h}_{right}(\omega;\theta,\phi=0)\right|$



(plot for fixed elevation $\phi = 0$)

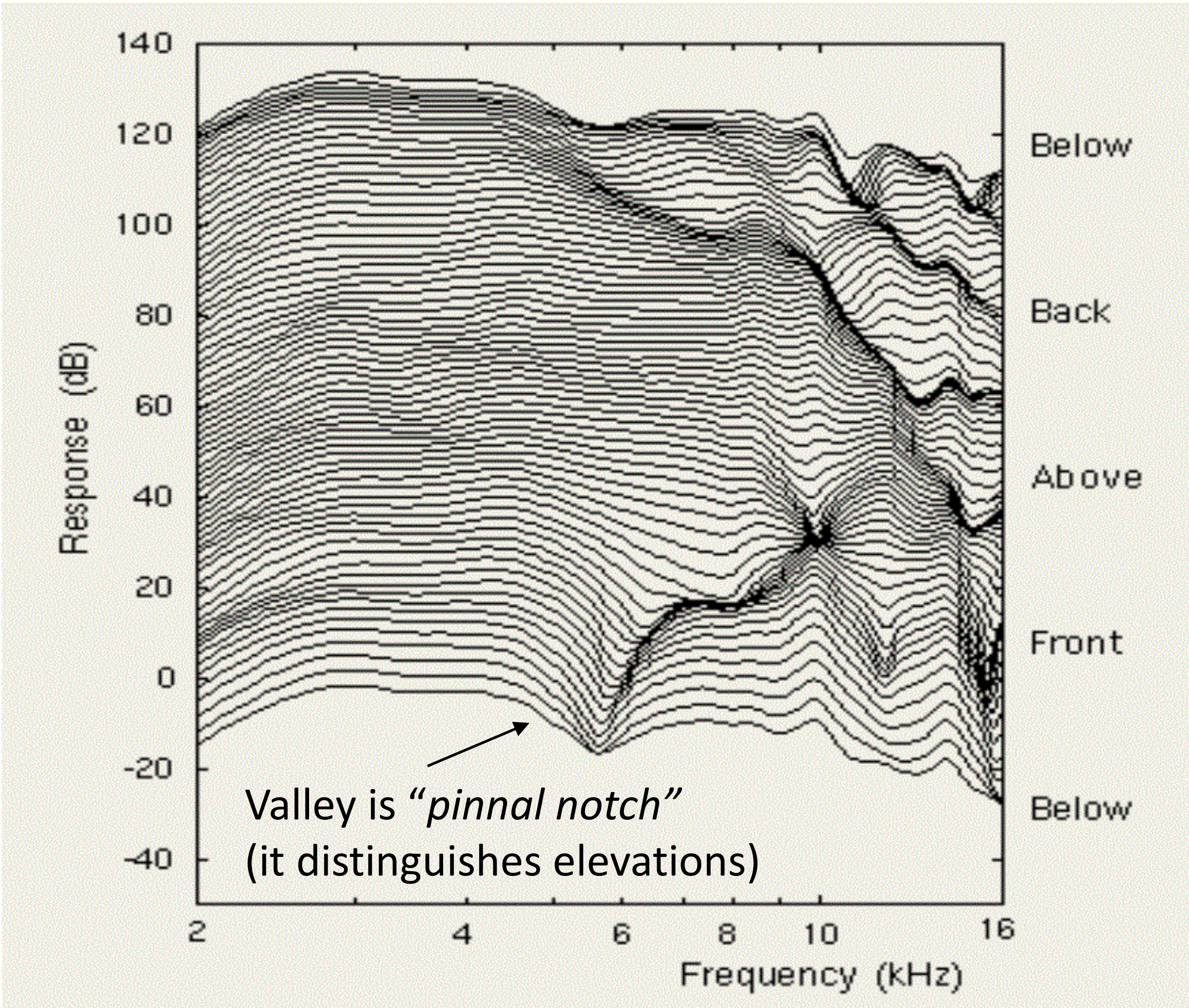


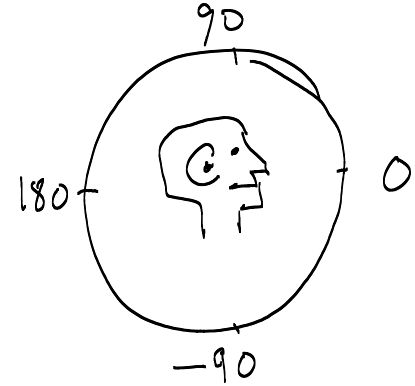
ω θ

Shadowing effect dominates: HRTF for each frequency ω has a max at 90 degrees (right ear) and min at 270 degrees (left ear).

HRTF $\left|\hat{h}_{right}\,(\omega; \theta = 0, \phi)\right|$

(plot for fixed azimuth $\theta = 0.$)

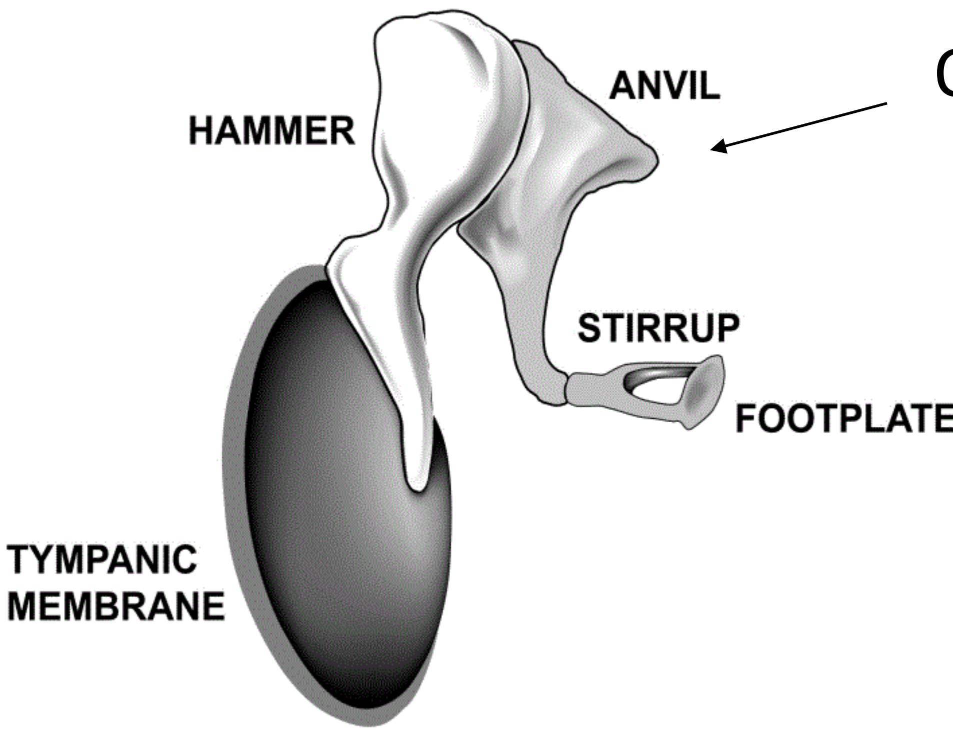




(medial plane)

Curves shifted for visualization

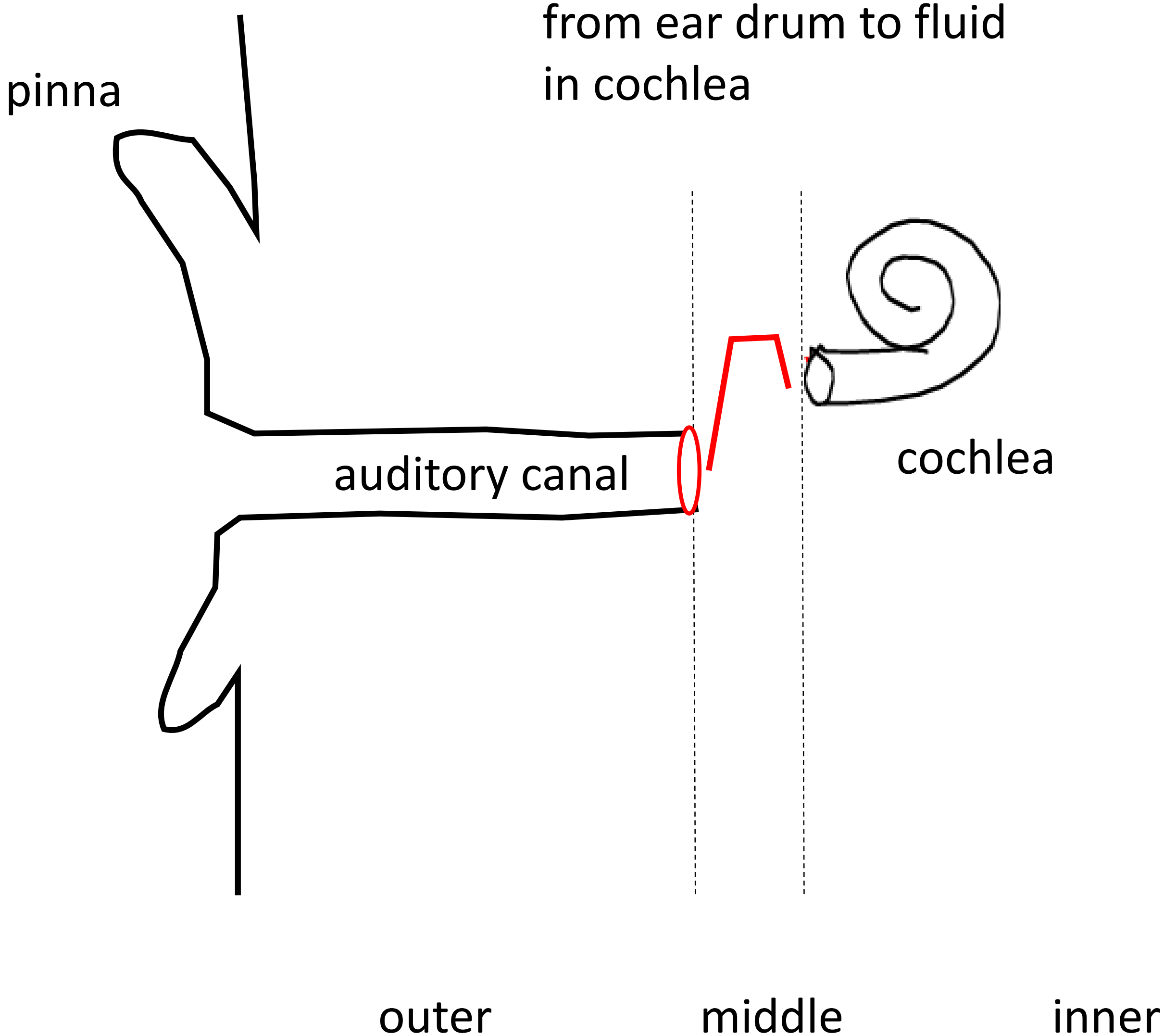
Middle Ear



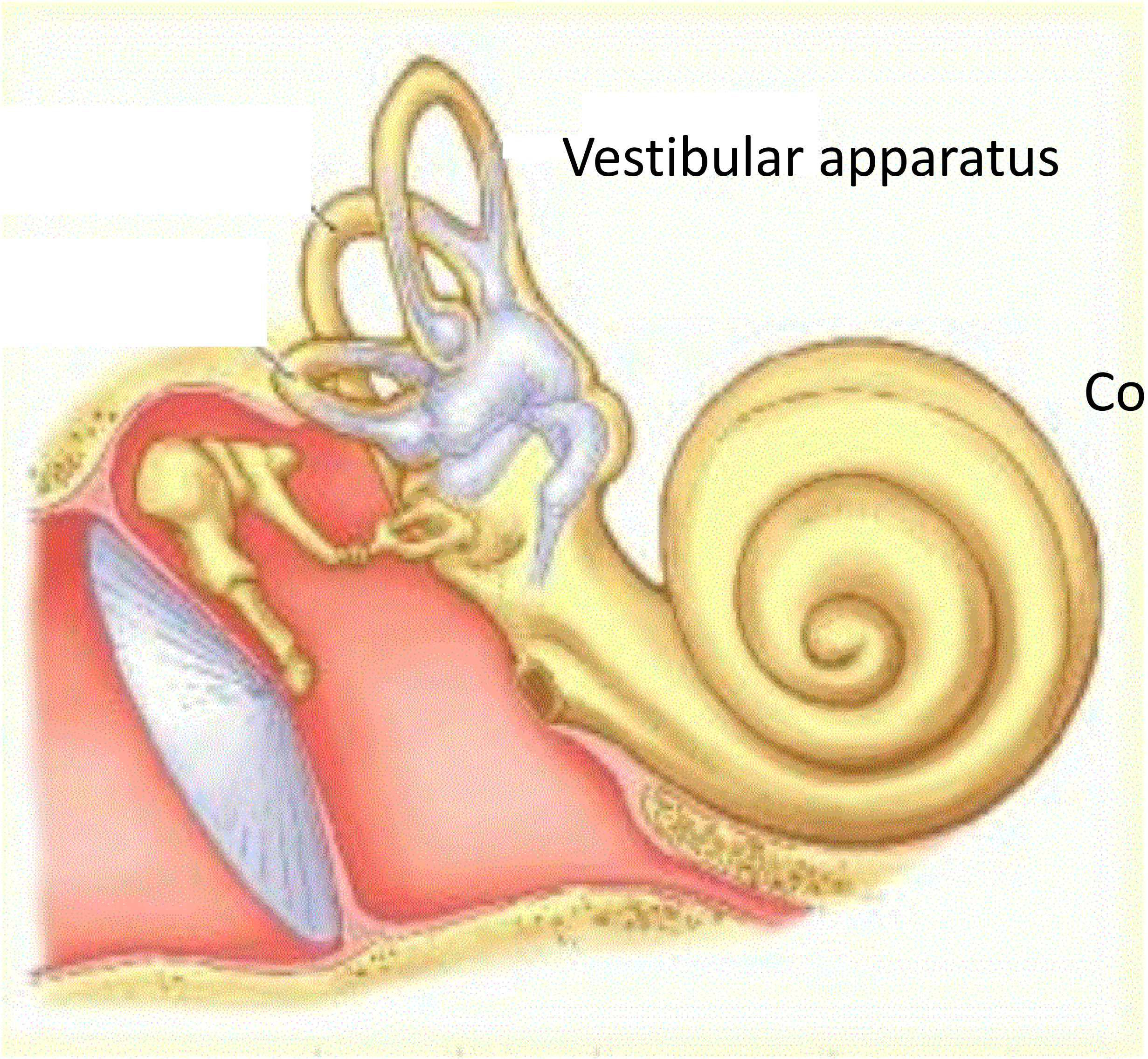
Ossicles (bones)

"Ear drum"

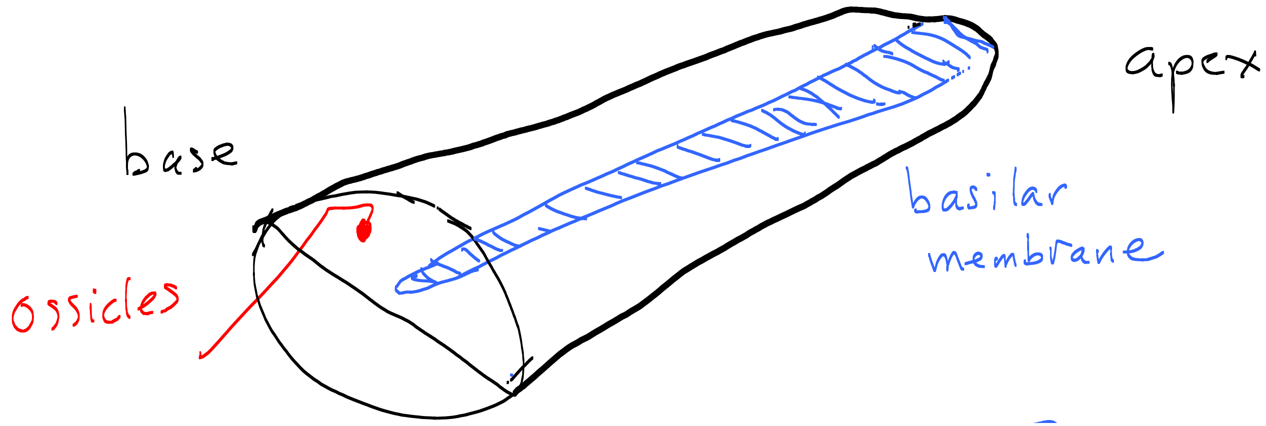
Ossicles act as a lever, transferring vibrations from ear drum to fluid in cochlea



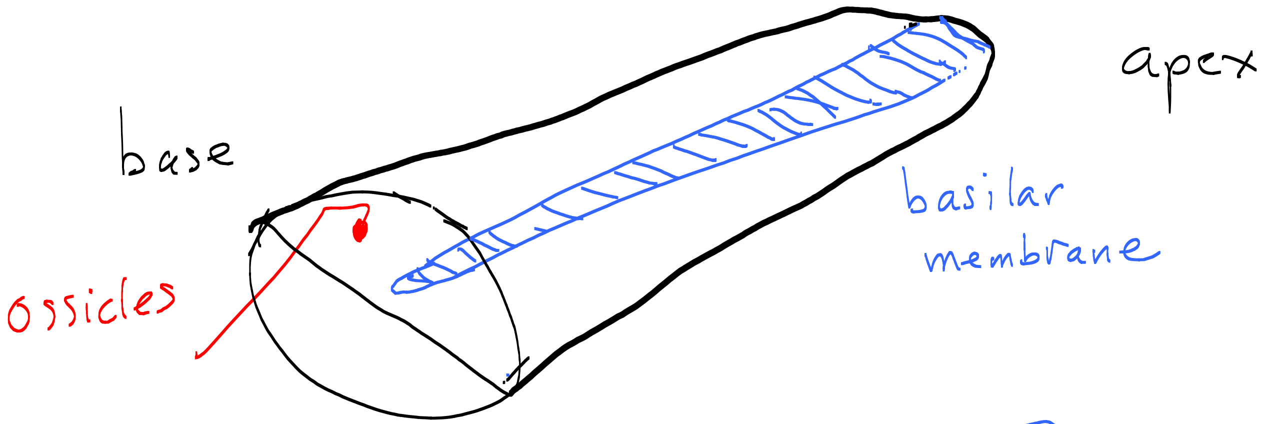
Inner ear



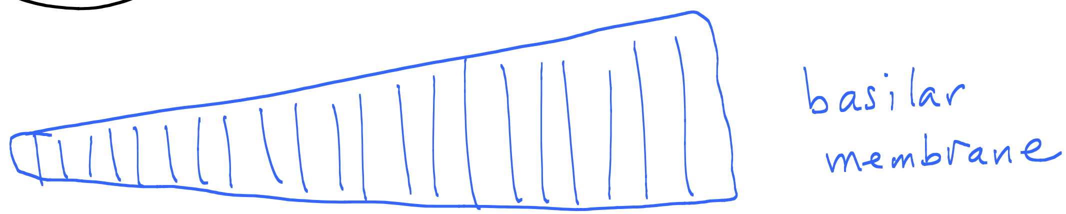
Cochlea (unrolled)



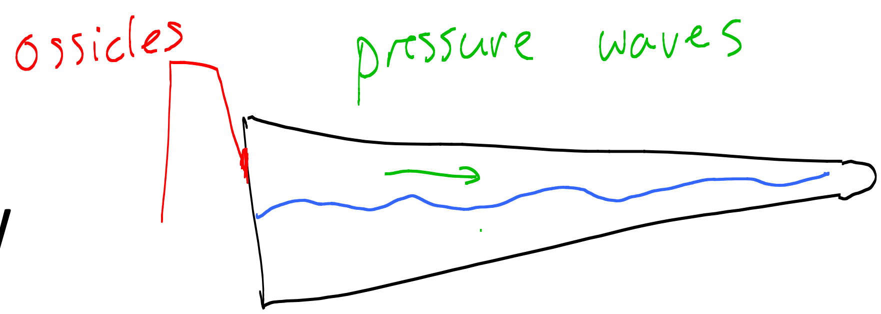
Cochlea (unrolled)



TOP VIEW

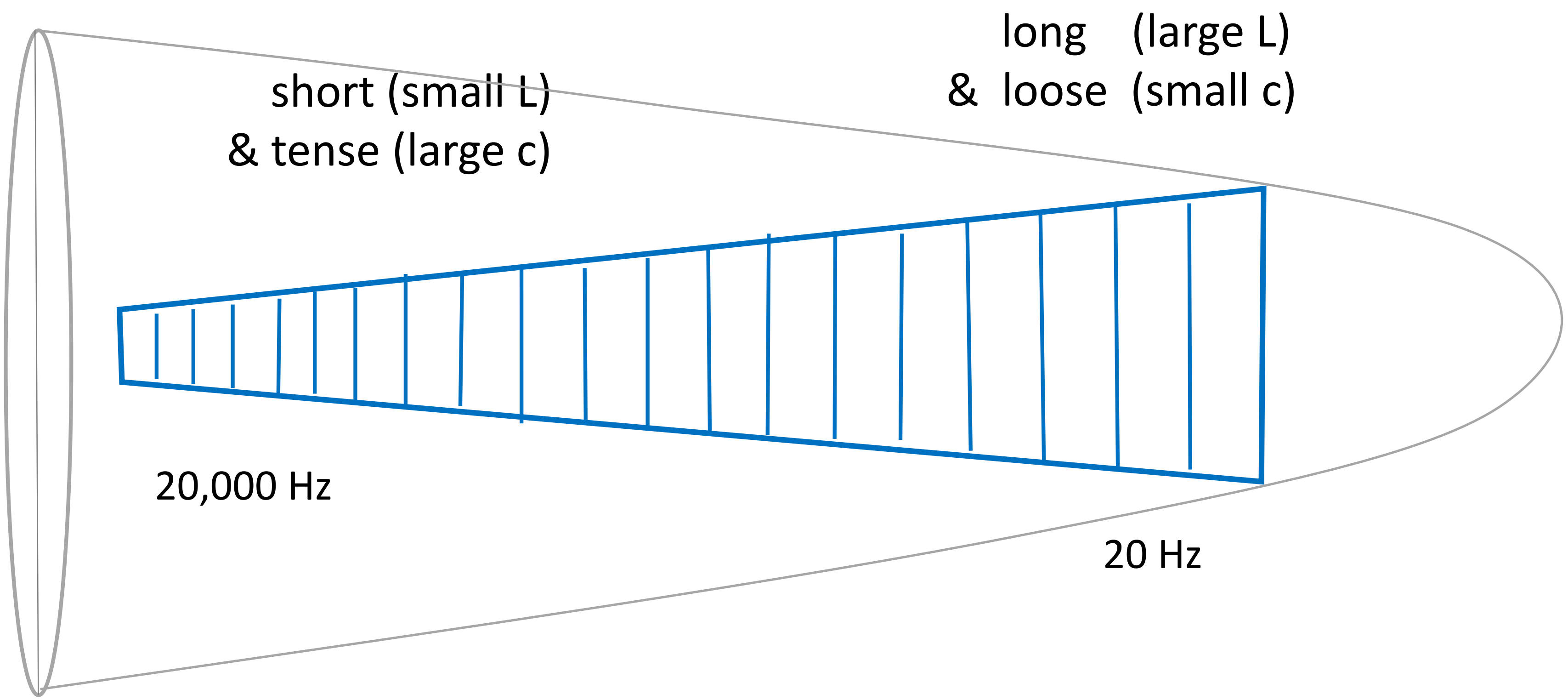


SIDE VIEW

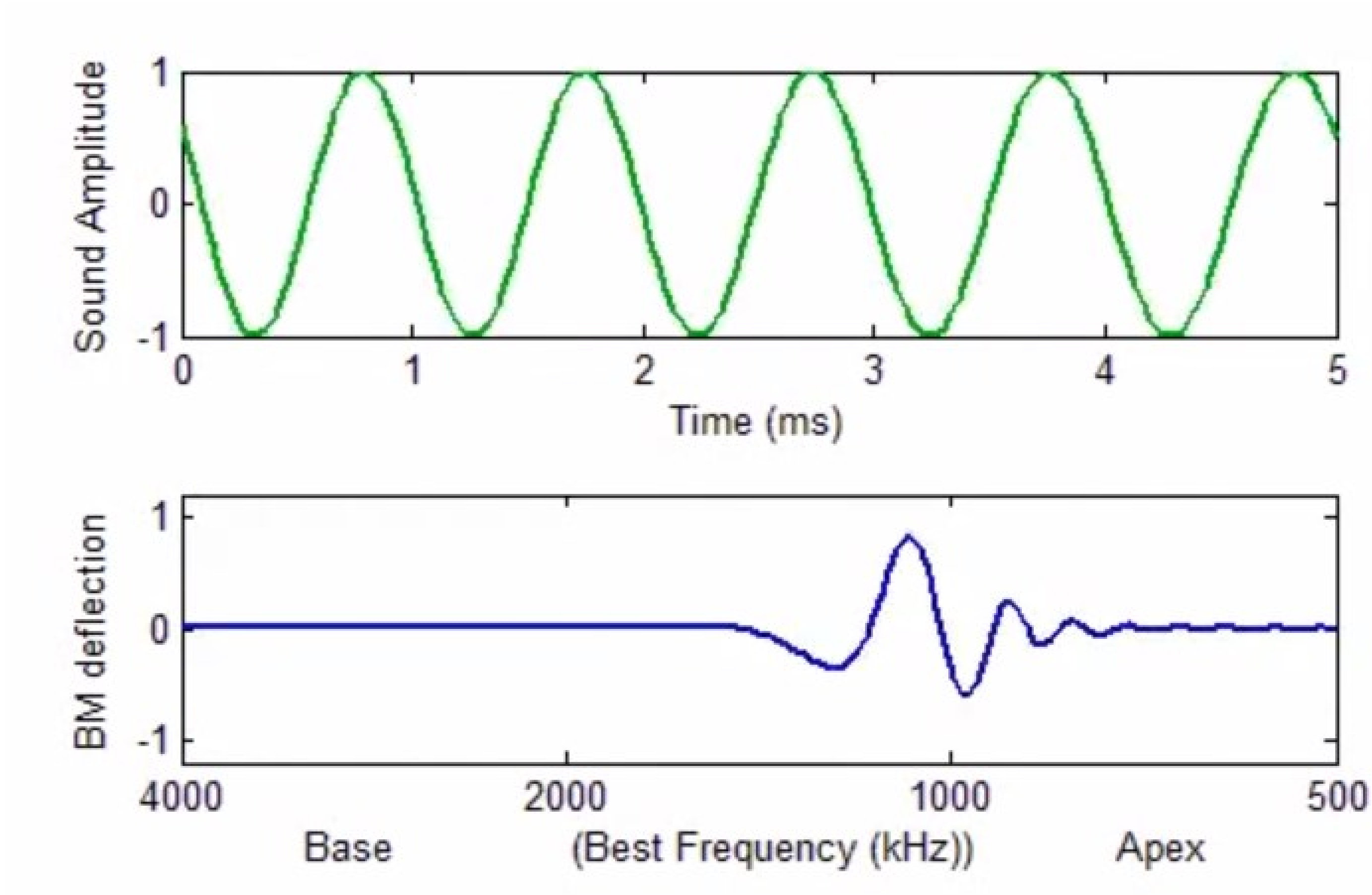


Recall vibrating string $\omega = \frac{c}{L}$

Both L and c vary on fibres on basilar membrane.



Basilar Membrane (BM)



http://auditoryneuroscience.com/topics/basilar-membrane-motion-0-frequency-modulated-tone

http://auditoryneuroscience.com/ear/bm_motion_2