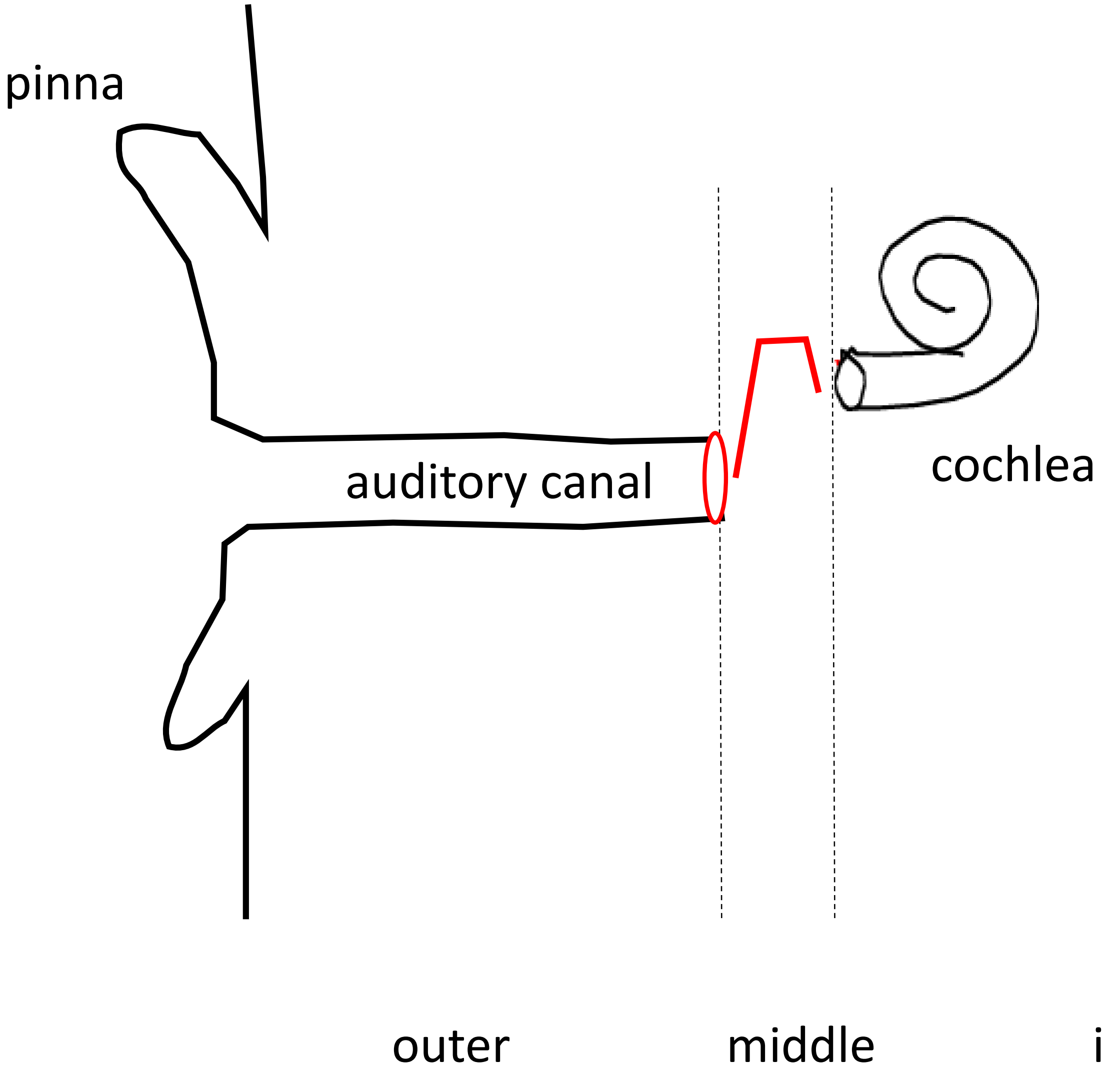
COMP 546

Lecture 21

Cochlea to brain, Source Localization

Tues. April 3, 2018

Ear



Eye	Ear
• Lens	• ?
• Retina	• ?
• Photoreceptors (light -> chemical)	• ?
• Ganglion cells (spikes)	• ?
• Optic nerve	• ?

Eye

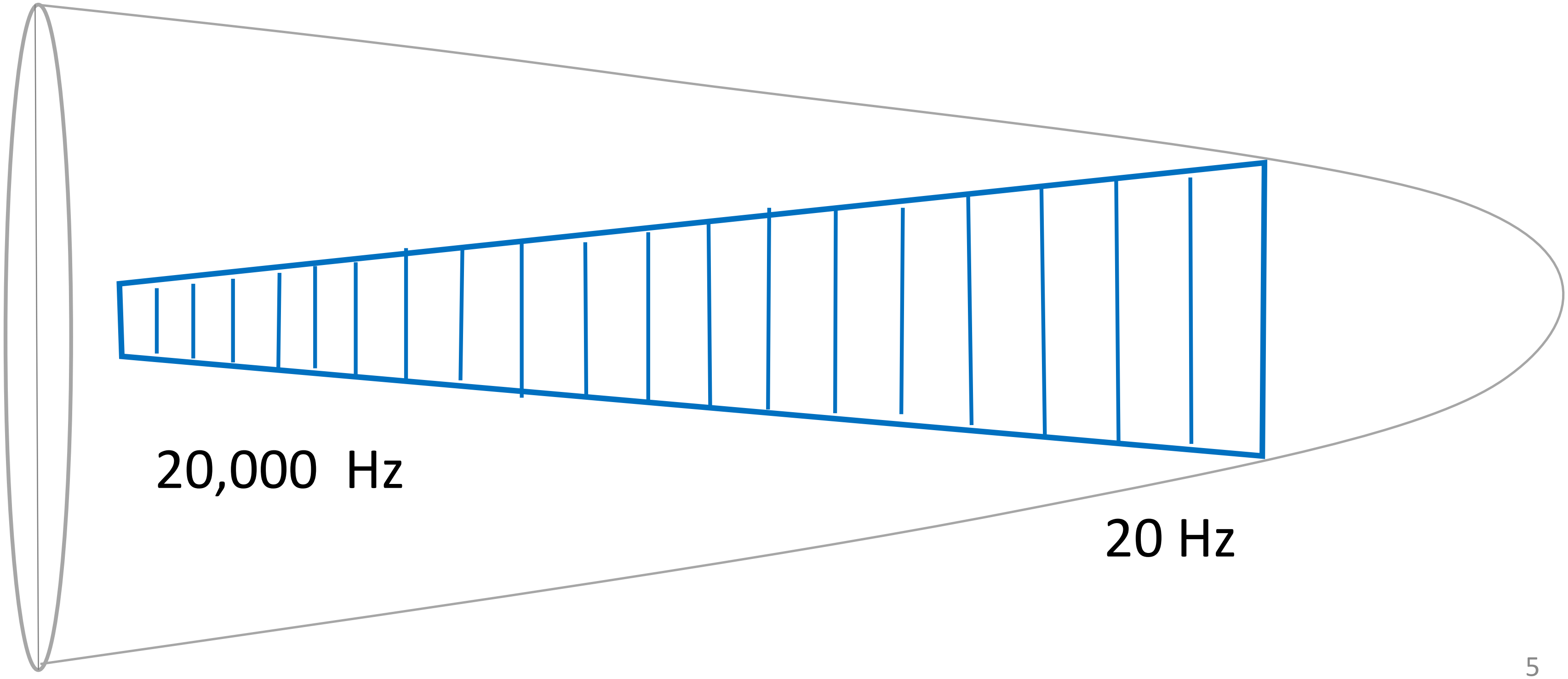
- Lens
- Retina
    - Photoreceptors
      (light -> chemical)
    - Ganglion cells (spikes)
- Optic nerve

Ear

- Outer ear
- Cochlea
    - hair cells
      (mechanical -> chemical)
    - Ganglion cells (spikes)
- VestibuloCochlear nerve

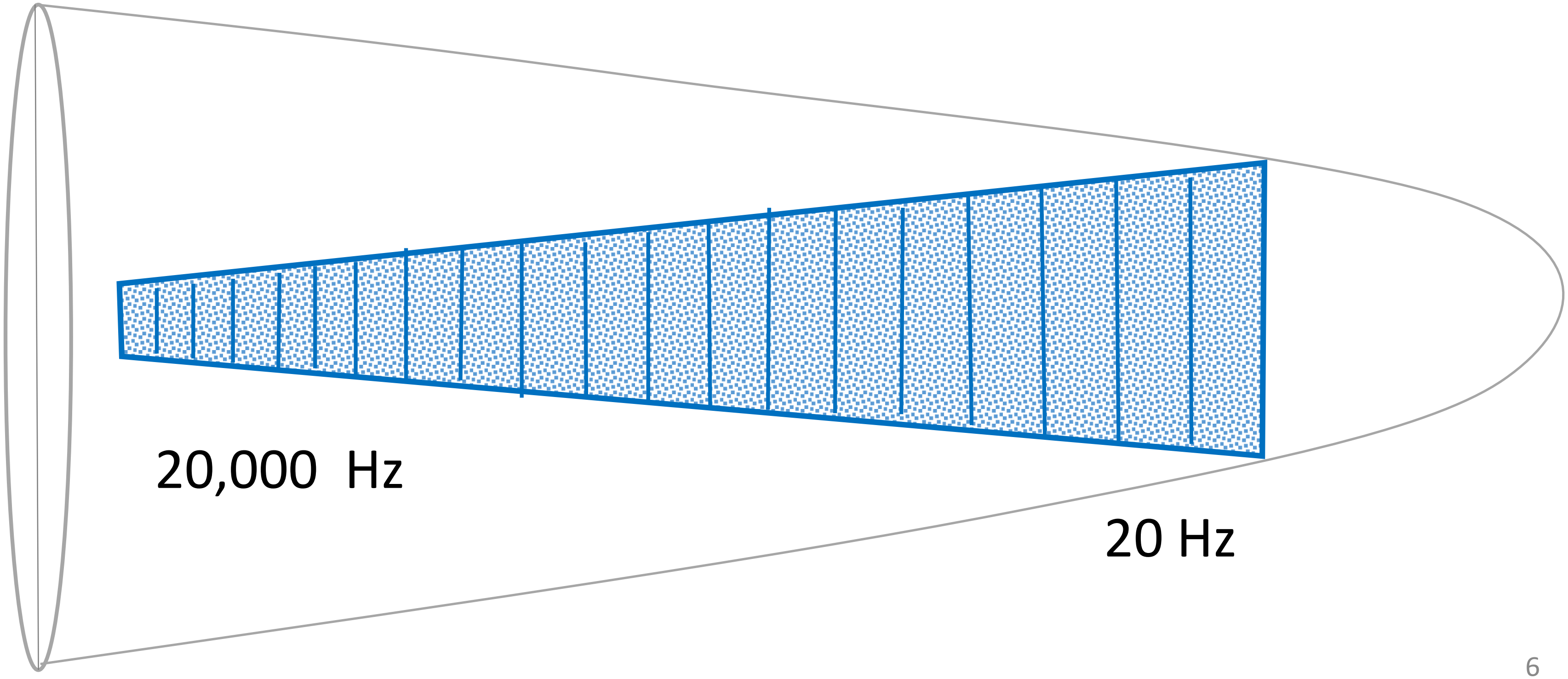
Basilar Membrane

BM fibres have bandpass frequency *mechanical* responses.

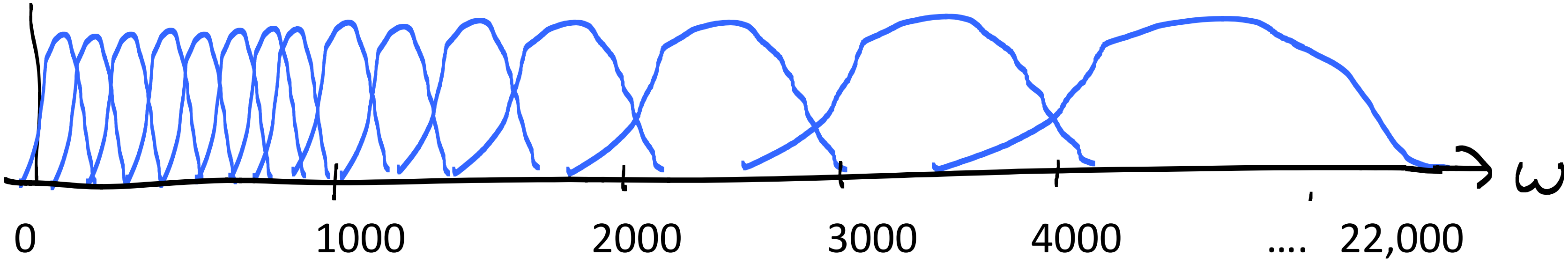


Basilar Membrane:
Place code ("tonotopic")

Nerve cells *(hair + ganglion)* are distributed along the BM.
They have similar bandpass frequency response functions.

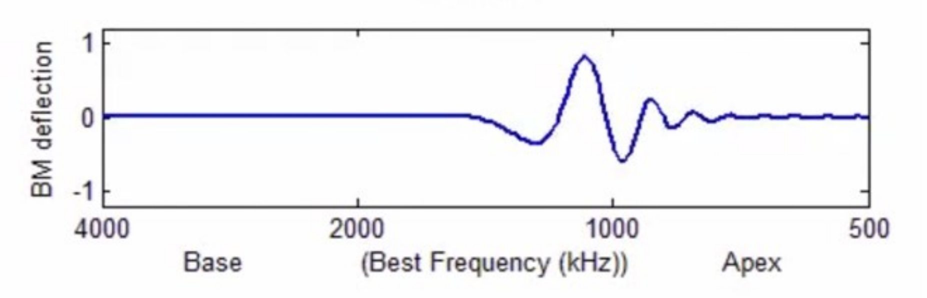


Bandpass responses
(more details next lecture)



Neural coding of sound in cochlea

- *Basilar membrane* responds by vibrating with sound.



- *Hair cells* at each BM location release neurotransmitter that signal BM amplitude at that location

- *Ganglion cells* respond to neurotransmitter signals by spiking

Louder sound within frequency band

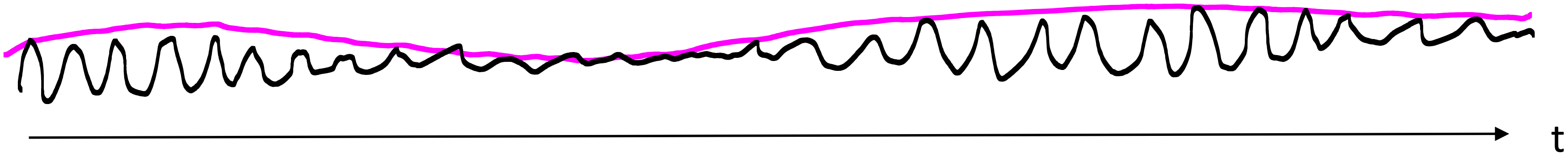
→ greater amplitude of BM vibration at that location

→ greater release of neurotransmitter by hair cell

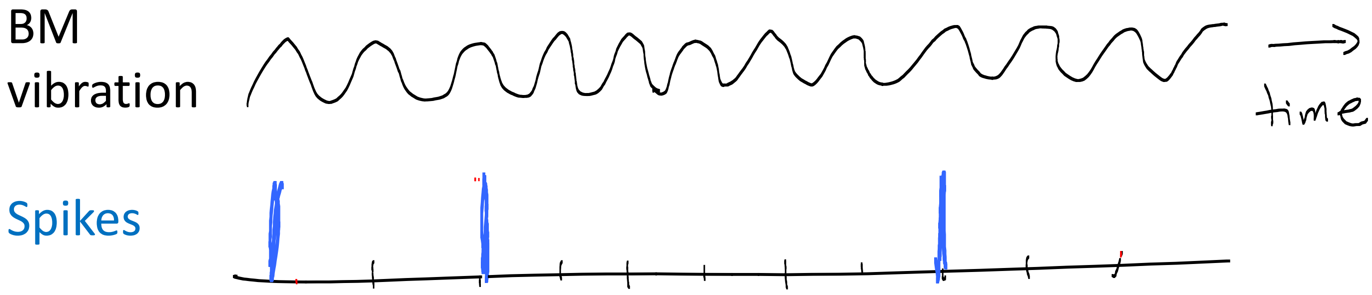
→ greater probability of spike at each peak of filtered wave

Hair cell neurotransmitter release can signal *exact timing* of BM *amplitude peaks* for frequencies up to ~2 kHz.

For higher frequencies, hair cells encode only the envelope of BM vibrations.



Timing of ganglion cell spikes: for frequencies up to 2 KHz ("phase locking")

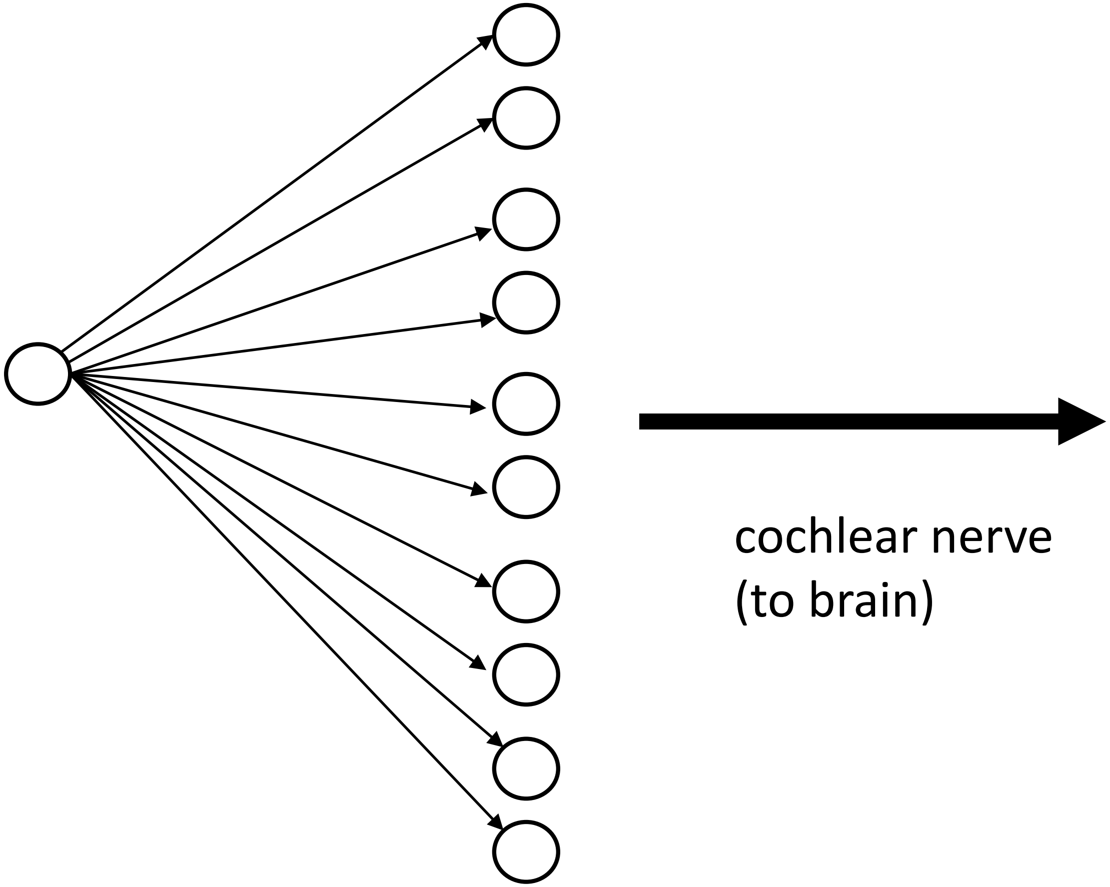


Hair cells release more neurotransmitter at BM amplitude peaks.

Ganglion cells respond to neurotransmitter peaks by spiking.

This allows exact timing of BM vibrations to be encoded by spikes.

Ganglion cells cannot spike faster than 500 times per second. So we need many ganglion cells for each hair cell.

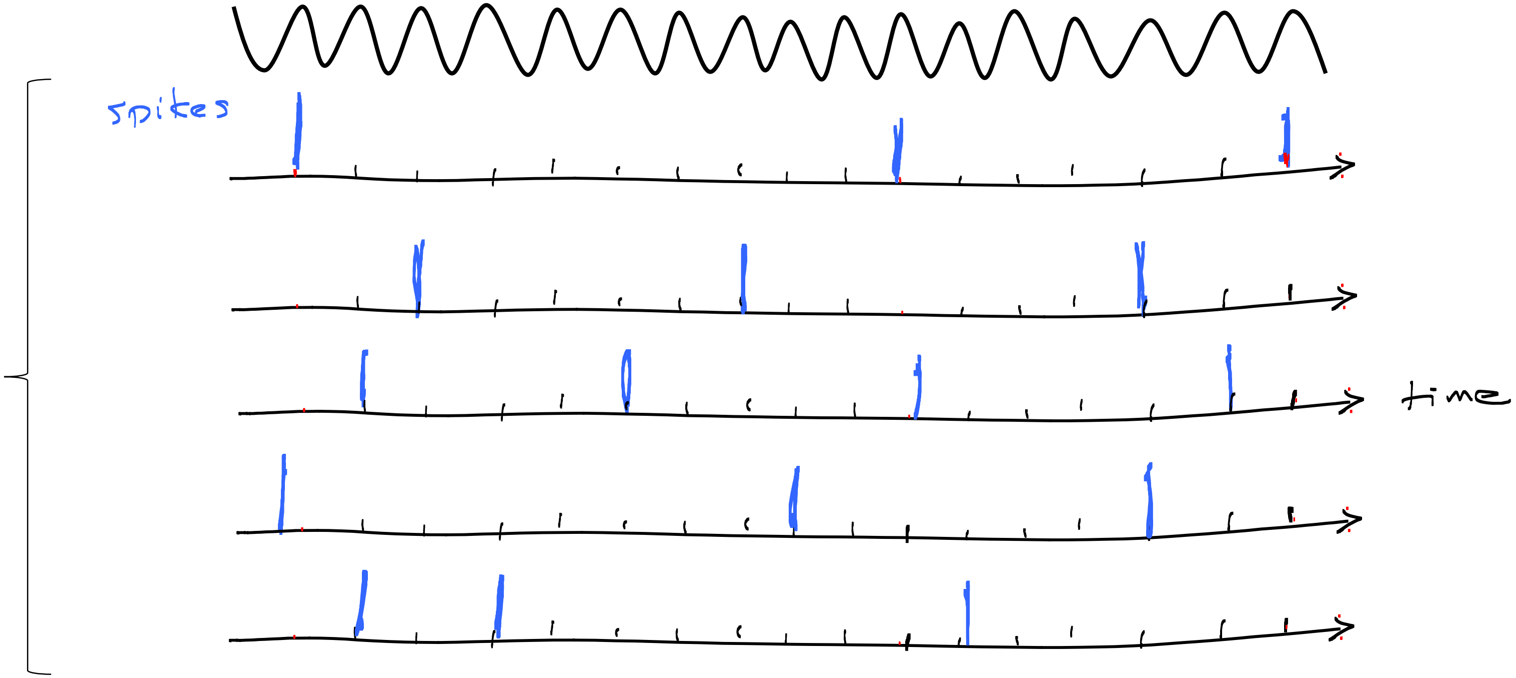


3,000 **hair cells**
in each cochlea
(left and right)

cochlear nerve
(to brain)

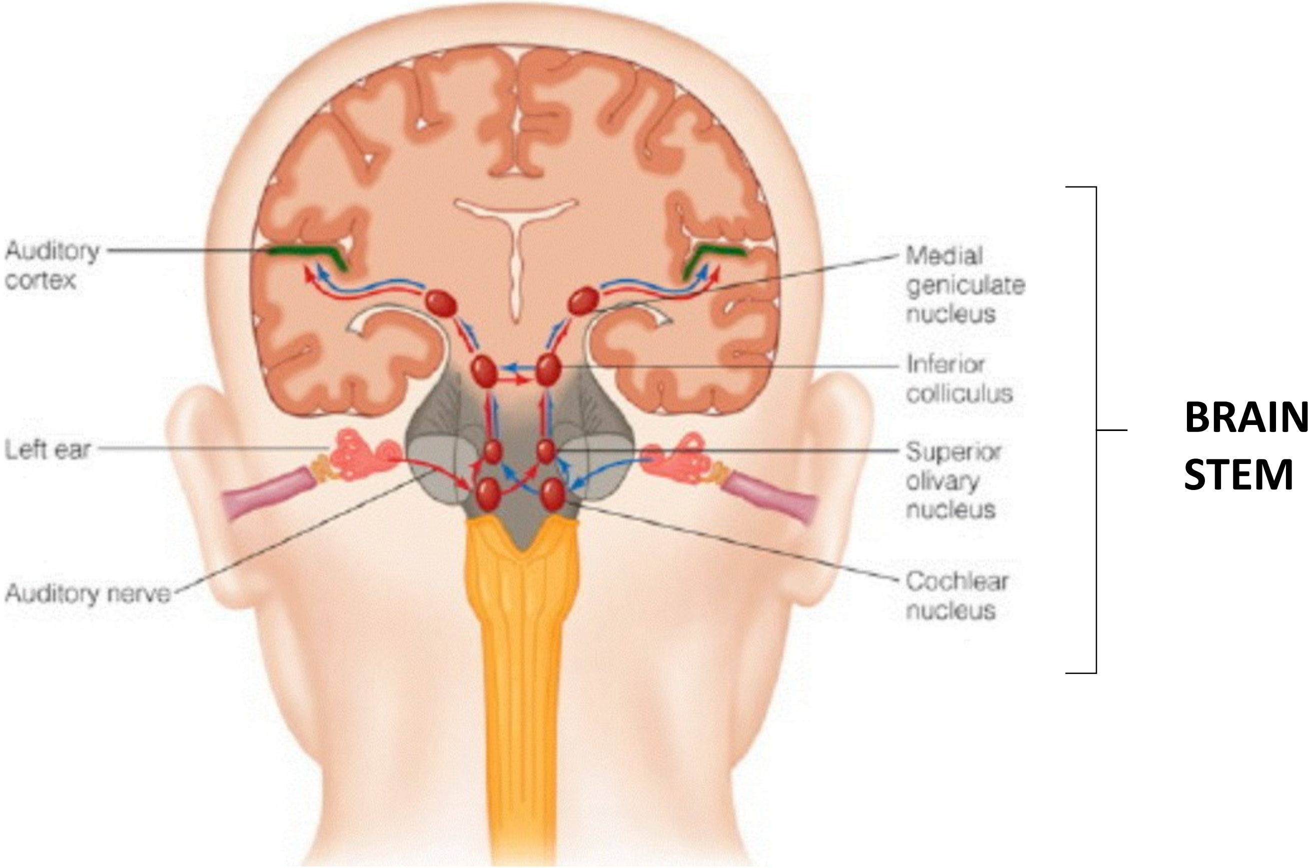
30,00 **ganglion cells**
in each cochlea

“Volley” code



Different ganglion cells at same spatial position on BM

From cochlea to brain stem

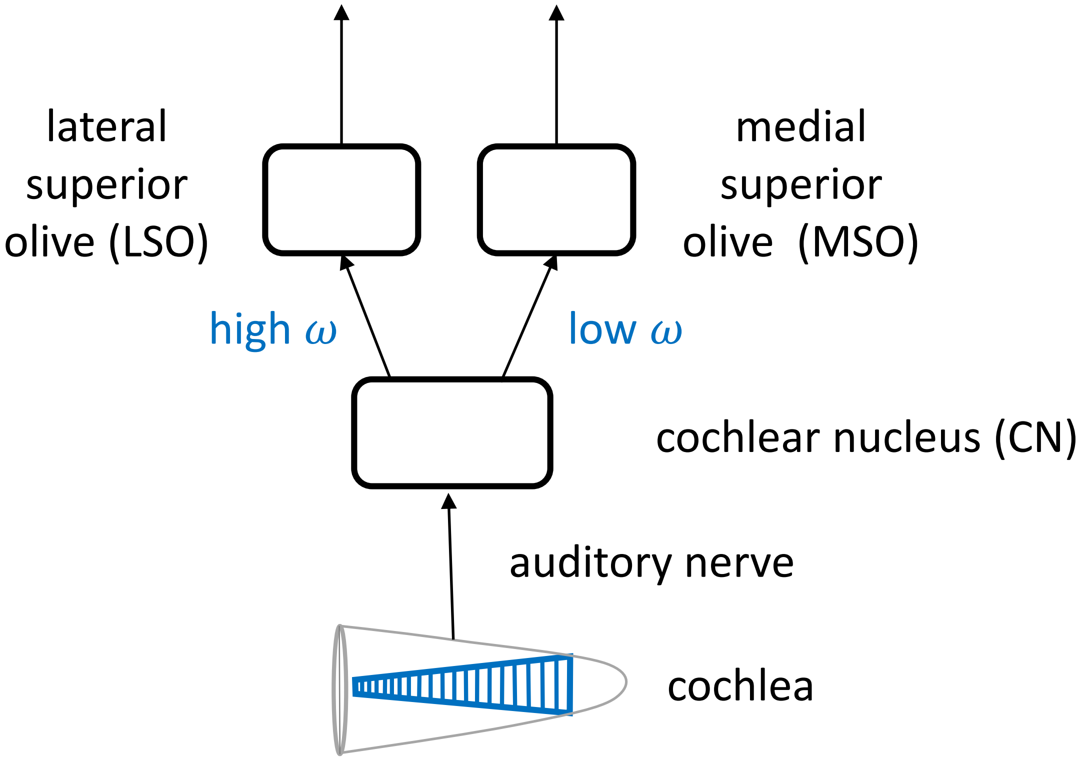


cochlea → cochlear nucleus

→ lateral and medial superior olive (LSO, MSO)

… → auditory cortex

Tonotopic maps



lateral superior olive (LSO)

medial superior olive (MSO)

high ω

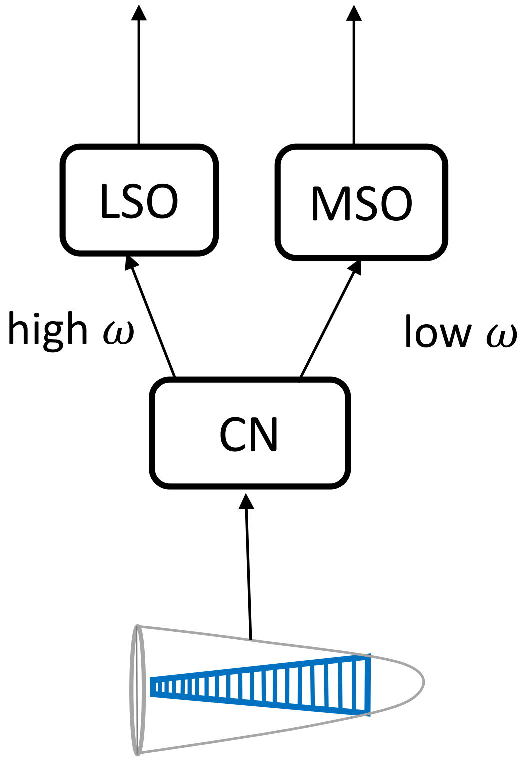
low ω

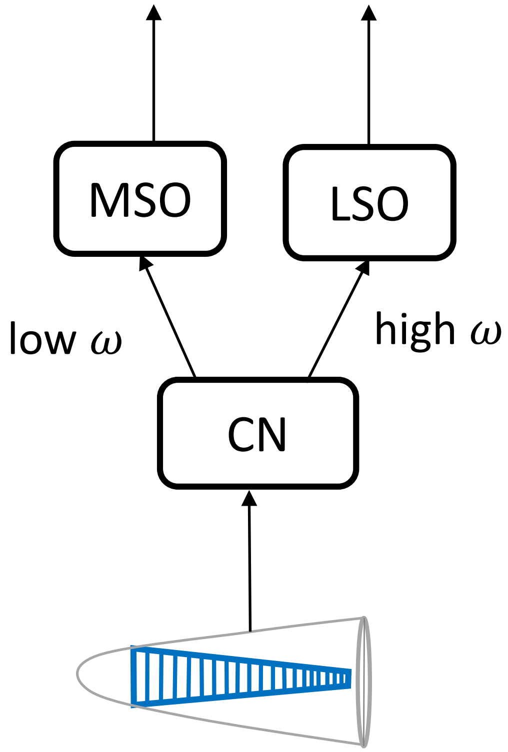
cochlear nucleus (CN)

auditory nerve

cochlea

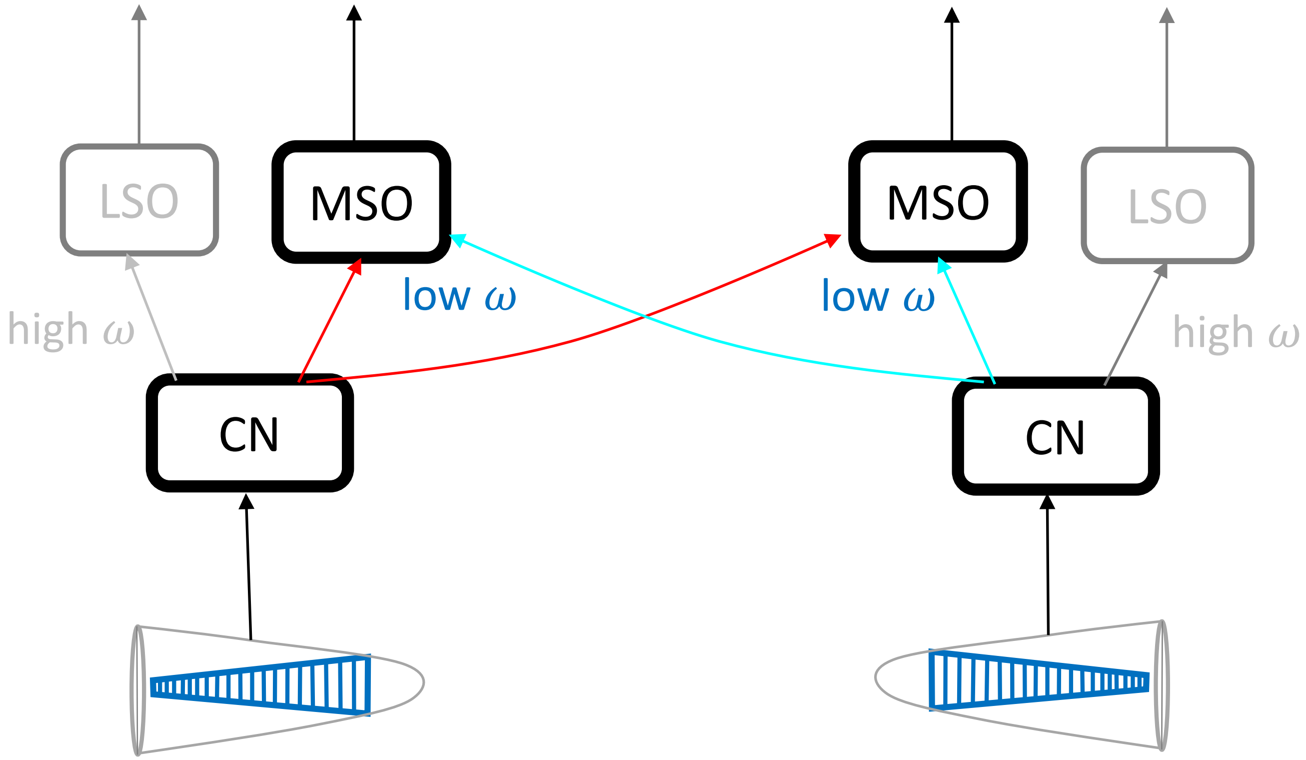
Binaural Hearing





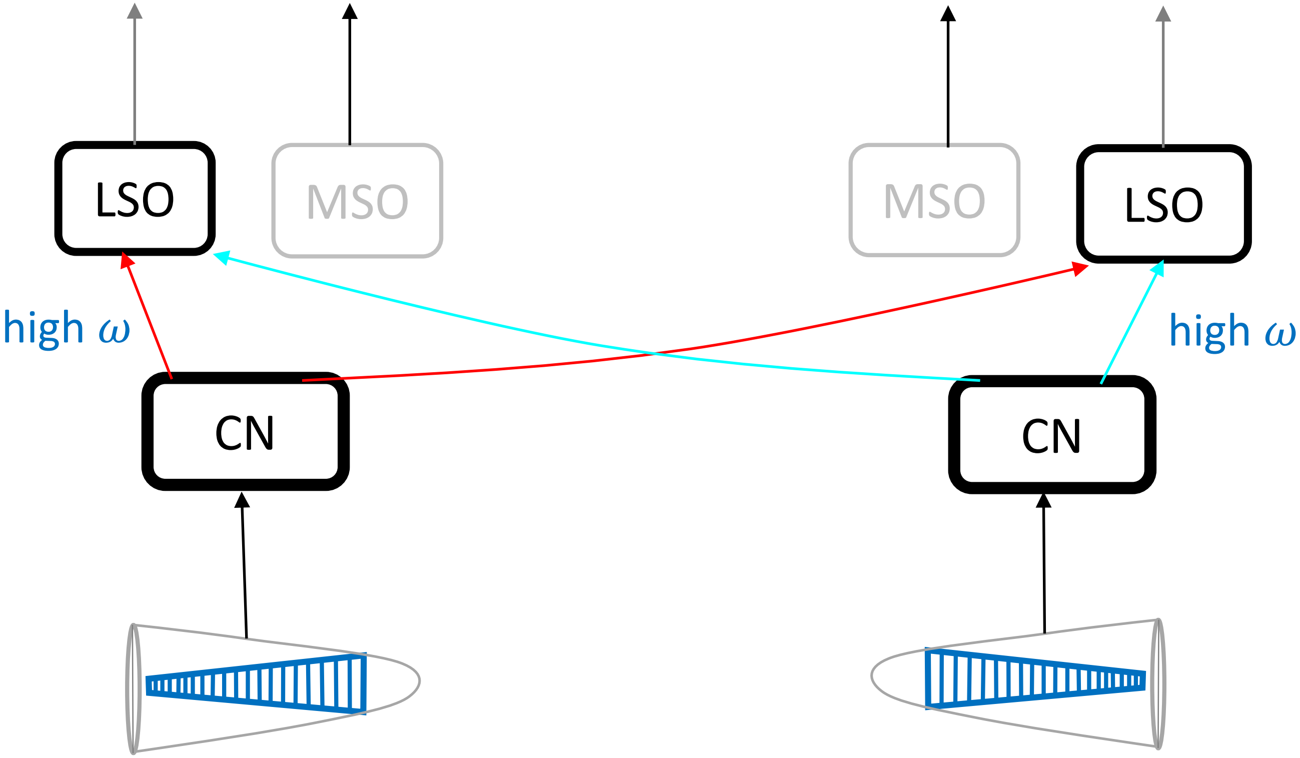
Binaural Hearing

MSO combines low frequency signals.



Binaural Hearing

LSO combines high frequency signals.



Levels of Analysis

high

- what is the task ? what problem is being solved?
- brain areas and pathways
- neural coding
- neural mechanisms

low

For high frequency bands,

- the head casts a shadow
- the timing of the peaks cannot be accurately coded by the spikes (only the rate of spikes is informative)

For low frequency bands,

- the head casts a weak shadow only
- the timing of the peaks *can* be encoded by spikes

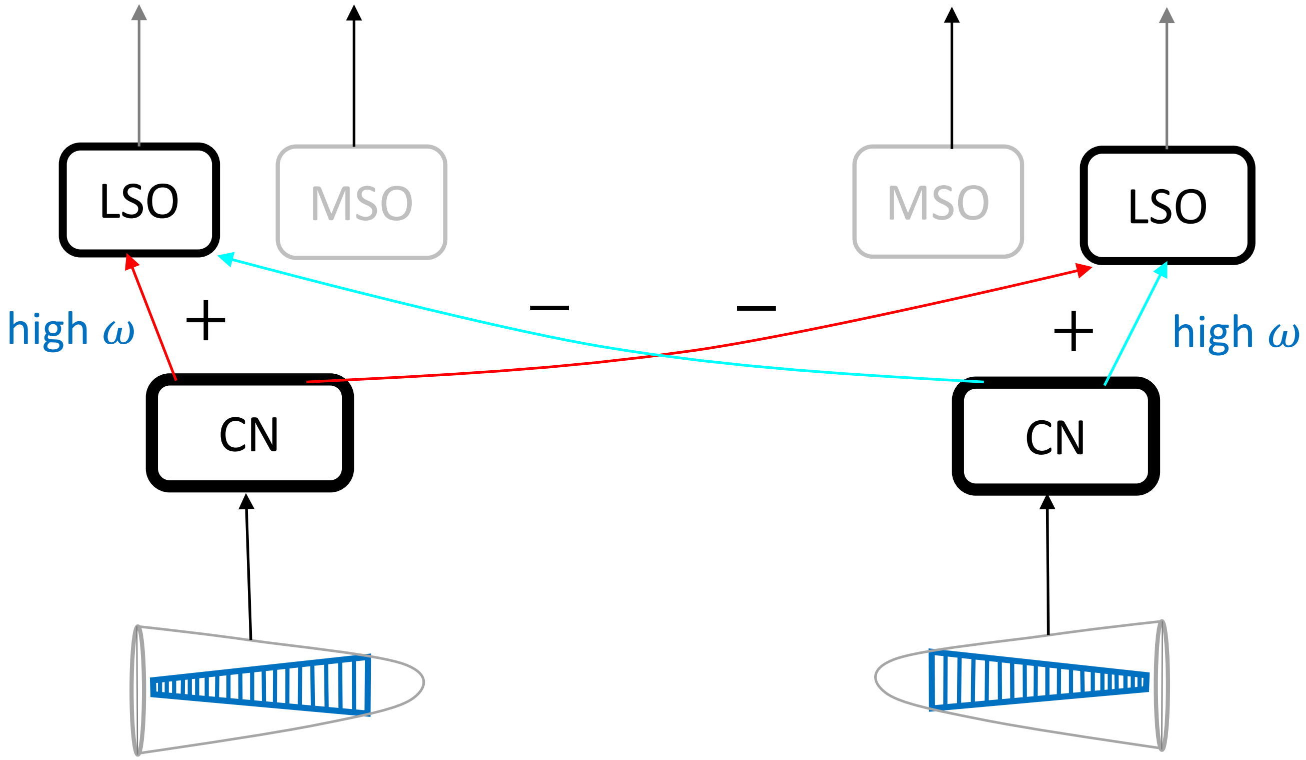
Duplex theory of binaural hearing

(Rayleigh, 1907)

- level differences computed for higher frequencies (ILD -- interaural level differences)
- timing differences computed for lower frequencies (ITD - interaural timing differences)

Level differences (high frequencies)

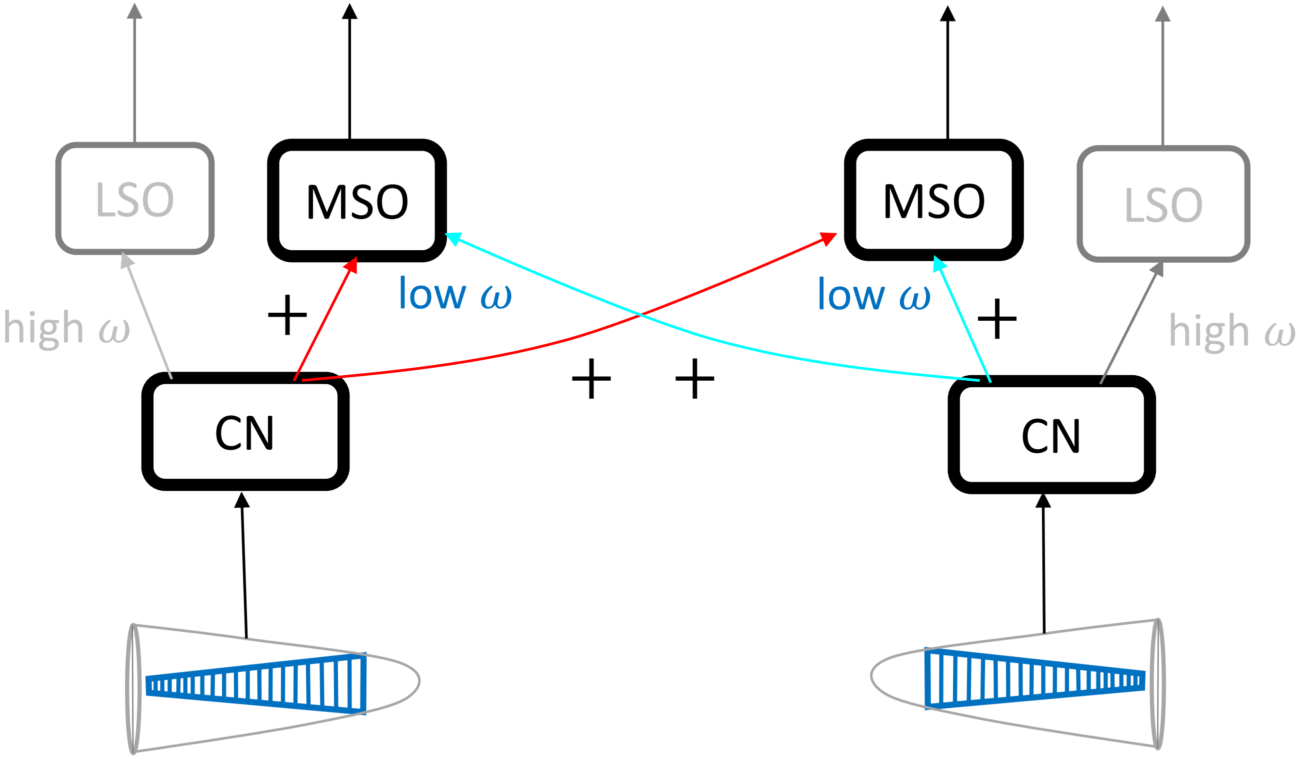
Excitatory input comes from the ear on the same side. Inhibitory input comes from ear on the opposite side.



Timing differences (low frequencies)

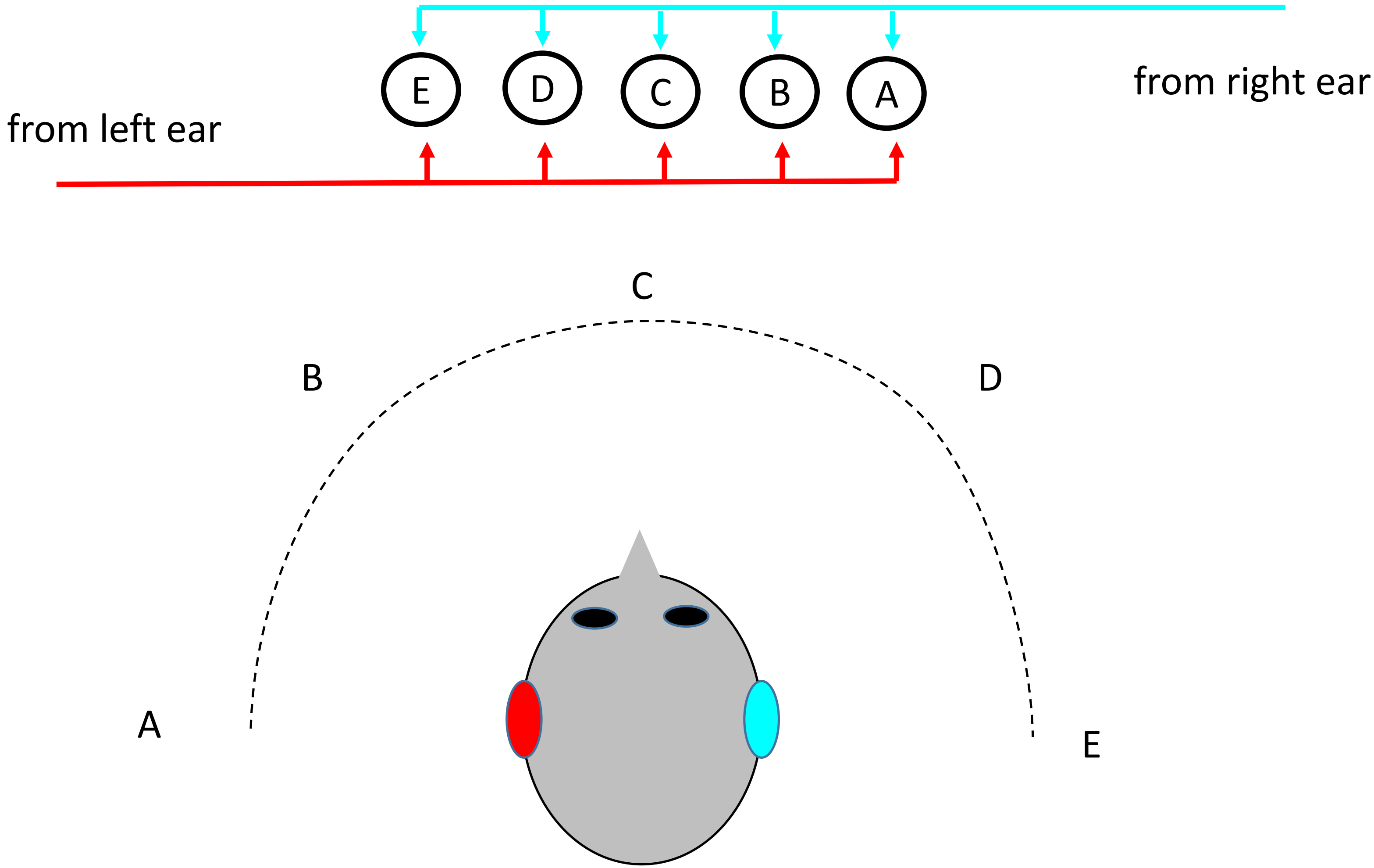
Sum excitatory input from both sides.

Reminiscent of binocular complex cells in V1 ?

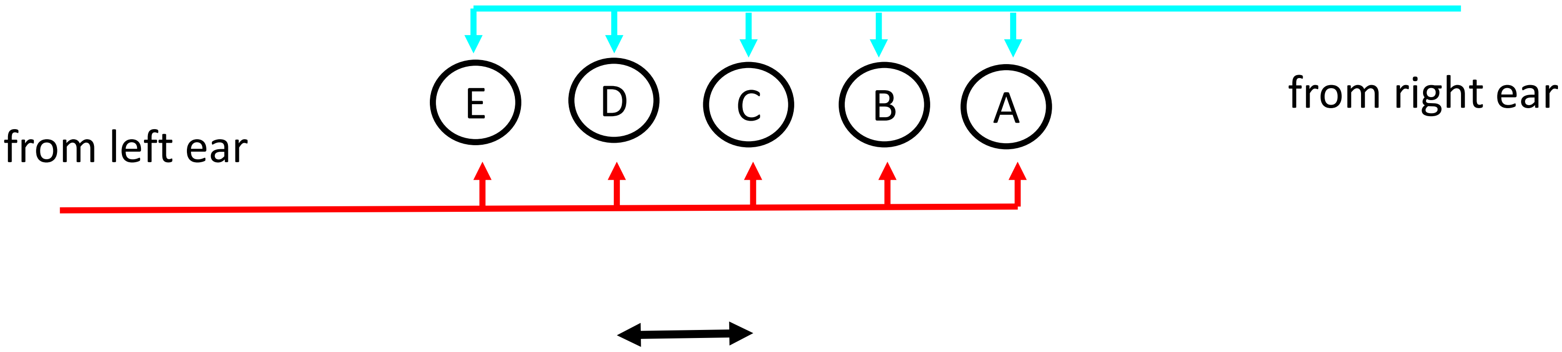


Jeffress Model (1948) for timing differences

http://auditoryneuroscience.com/topics/jeffress-model-animation



Spike Timing precision required for Jeffress Model ?



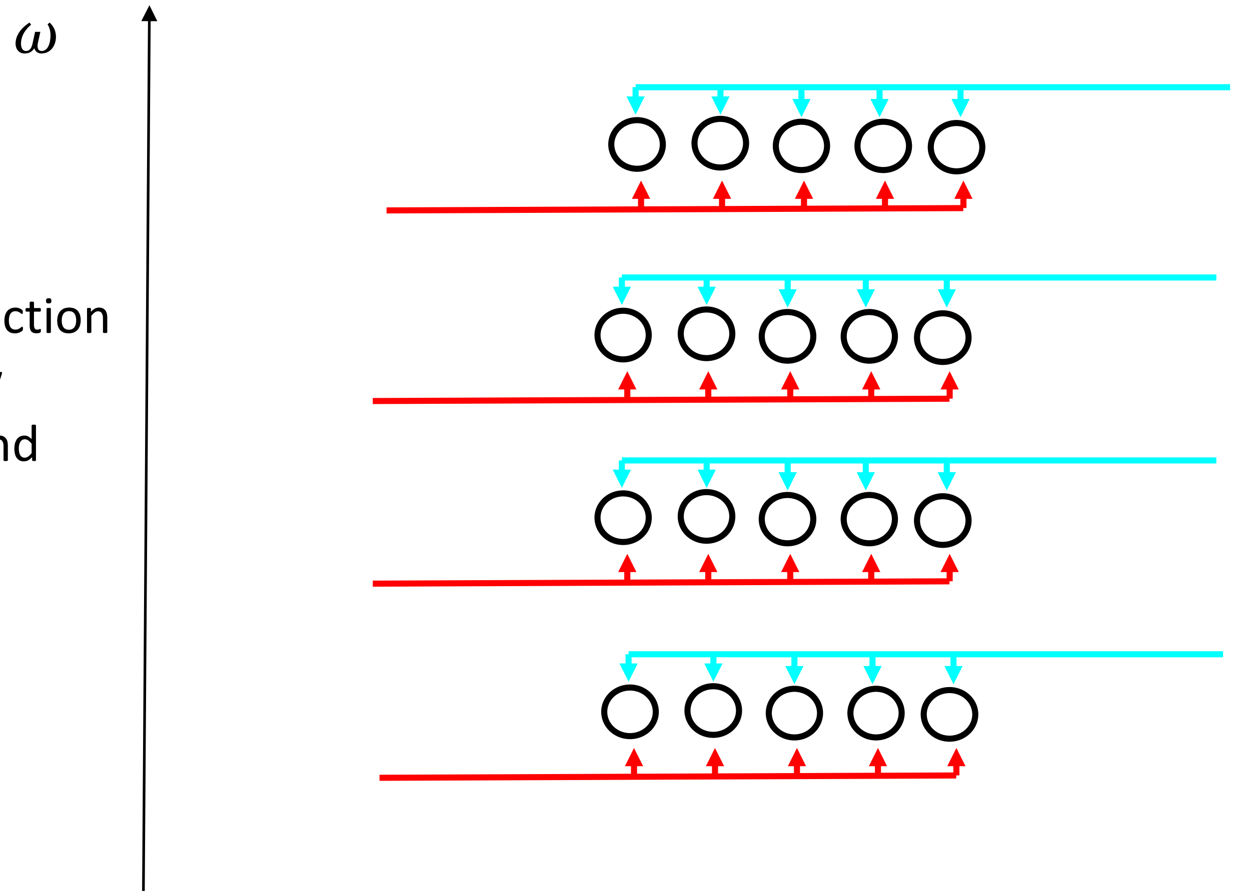
distance = $\frac{1}{10}\ millimetres$

speed of spike = $10\ metres\ \ second^{-1}$

$\Longrightarrow$ Δ time $= \frac{distance}{speed} = \frac{1}{100}\ millisecond$

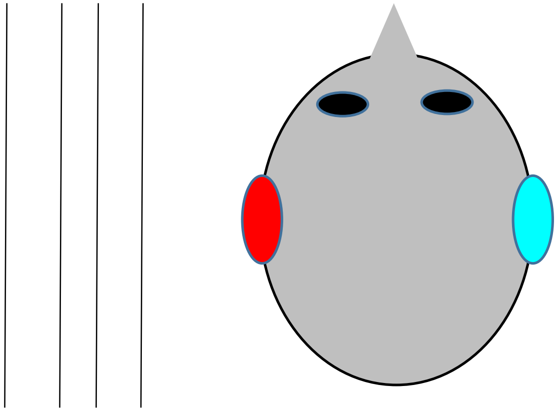
See Exercises 19 Q2c

Jeffress model remains controversial. It is not known exactly how "coincidence detection" occurs in MSO.



Coincidence detection for each low frequency band

Naïve *Computational Model* of Source Localization (Recall lecture 20)



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

$\uparrow$ shadow

$\uparrow$ delay

$\uparrow$ model error

Find the α and τ that minimize

$$\sum_{t=1}^{T} \{\, I_l\,(t) - \alpha\;\; I_r(t-\tau)\;\; \}^2 \qquad \text{where } \tau < 0.5\, ms.$$

Timing difference: find the τ that maximizes

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

Level difference:

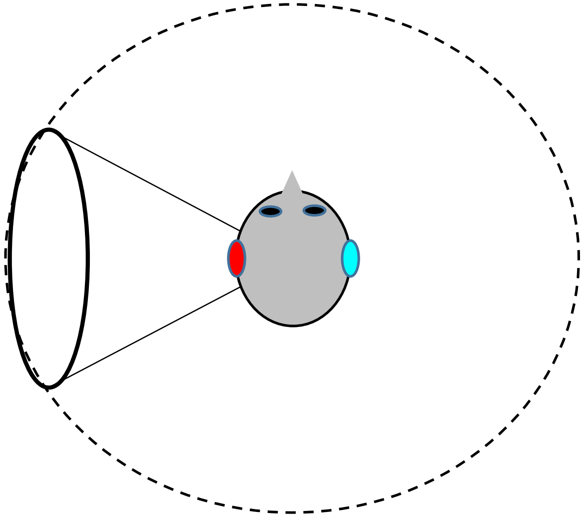
$$10\; log_{10}\, \frac{\sum_{t=1}^{T}\; I_l\,(t)^2}{\sum_{t=1}^{T}\; I_r\,(t)^2}$$

For each low frequency band j, find the τ that maximizes

$$\sum_{t} I_{left}{}^{j}(t) \;\; I_{right}{}^{j}(t-\tau) \; .$$

(or use summation model similar to binocular cells or Jeffress model)

An estimated value of delay τ in frequency band j is consistent with various possible source directions (ϕ, θ).



Similar to cone of confusion, but more general because of frequency dependence

For each high frequency band j, compute interaural level difference (ILD) :

$$ILD_j = 10\, log_{10} \frac{\sum_{t=1}^{T} I_{left}{}^{j}\,(t)^2}{\sum_{t=1}^{T} I_{right}{}^{j}\,(t)^2}$$

What does each ILD_j tell us ?

Recall head related impulse response function (HRIR) from last lecture..

If the source direction is (θ, ϕ), and $g^{j}(t)$ is the filter for band j.

then…

$$I_{left}^{j}(t;\ \phi,\theta)\ = g^{j}(t) * h_{left}(t;\ \phi,\theta)\ * I_{src}(t;\ \phi,\theta)$$

$$I_{right}^{j}(t;\ \phi,\theta) = g^{j}(t) *\ h_{right}(t;\ \phi,\theta)\ *\ I_{src}(t;\ \phi,\theta)$$

Take the Fourier transform and apply convolution theorem :

$$\widehat{I}^{\,j}_{left}(\omega;\ \phi,\theta)\ =\ \hat{g}^{j}(\omega)\ \hat{h}_{left}(\omega;\ \phi,\theta)\ \hat{I}_{src}(\omega;\ \phi,\theta)$$

$$\widehat{I}^{\,j}_{right}(\omega;\ \phi,\theta) = \hat{g}^{j}(\omega)\ \hat{h}_{right}(\omega;\ \phi,\theta)\ \ \hat{I}_{src}(\omega;\ \phi,\theta)$$

Take the Fourier transform and apply convolution theorem :

$$\widehat{I}^{\,j}_{left}(\omega;\ \phi,\theta) \ = \ \hat{g}^{j}(\omega)\ \hat{h}_{left}(\omega;\ \phi,\theta)\ \hat{I}_{src}(\omega;\ \phi,\theta)$$

$$\widehat{I}^{\,j}_{right}(\omega;\ \phi,\theta) = \ \hat{g}^{j}(\omega)\ \hat{h}_{right}(\omega;\ \phi,\theta)\ \ \hat{I}_{src}(\omega;\ \phi,\theta)$$

If there is just one source direction (ϕ,θ), then for each frequency ω within band j :

$$\frac{\widehat{I}^{\,j}_{left}(\omega)}{\widehat{I}^{\,j}_{right}(\omega)} \approx \frac{\hat{h}_{left}(\omega;\ \phi,\theta)}{\hat{h}_{right}(\omega;\ \phi,\theta)}$$

One can show using Parseval's theorem of Fourier transforms that if $h_{left}(\omega;\ \phi,\theta)$ and $h_{right}(\omega;\ \phi,\theta)$ are approximately constant within band j, then:

$$\frac{\sum_{t=1}^{T}\ {I_{left}}^{j}\ (t)^{2}}{\sum_{t=1}^{T}\ {I_{right}}^{j}\ (t)^{2}} \approx \frac{|\ \hat{h}_{left}^{j}(\ \phi,\theta)\ \ |^{2}}{|\ \hat{h}_{right}^{j}(\ \phi,\theta)\ \ |^{2}}$$

One can show using Parseval's theorem of Fourier transforms that if $h_{left}(\omega;\ \phi,\theta)$ and $h_{right}(\omega;\ \phi,\theta)$ are approximately constant within band j, then:

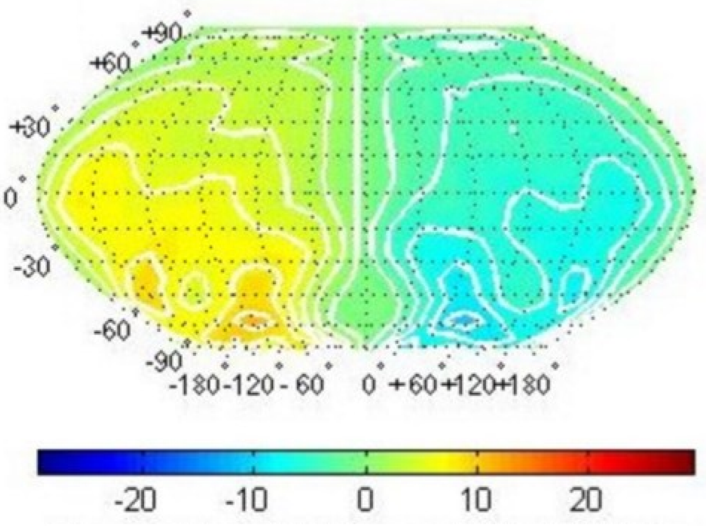
$$\frac{\sum_{t=1}^{T}\ I_{left}{}^{j}\,(t)^{2}}{\sum_{t=1}^{T}\ I_{right}{}^{j}\,(t)^{2}} \approx \frac{|\ \hat{h}^{j}_{left}(\,\phi,\theta)\ \ |^{2}}{|\ \hat{h}^{j}_{right}(\,\phi,\theta)\ \ |^{2}}$$

The ear can measure this…

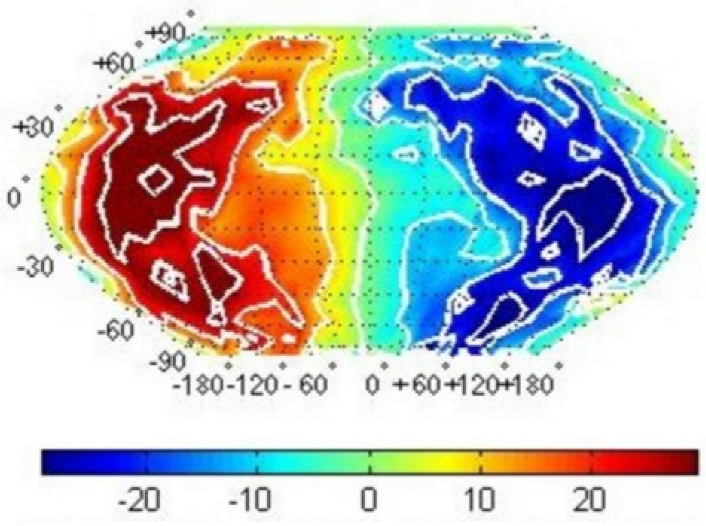
and can infer source directions (ϕ,θ) that are consistent with it.

https://auditoryneuroscience.com/topics/acoustic-cues-sound-location

Interaural Level Difference (dB) as a function of (ϕ, θ) for two fixed ω.



700 Hz



11,000 Hz

Each iso-contour in each frequency band
is consistent with a measured level difference (dB).

Monaural spectral cues
(Spatial localization with one ear?)

$$I^{j}(t;\ \phi,\theta)\ = g^{j}(t) * h(t;\ \phi,\theta)\ * I_{src}(t;\ \phi,\theta)$$

$$\hat{I}^{j}(\omega;\ \phi,\theta)\ =\ \hat{g}^{j}(\omega)\ \hat{h}^{j}(\omega;\ \phi,\theta)\ \hat{I}_{src}(\omega;\ \phi,\theta)$$

Pattern of peaks and notches across bands will be due to HRTF, not to the source.

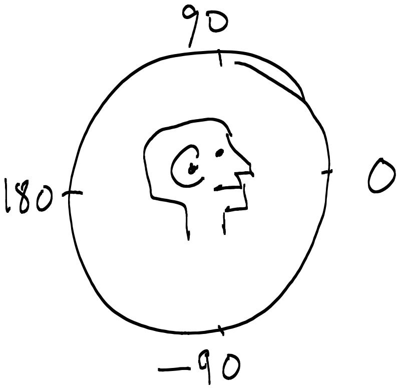
If the source is noise, then all frequencies make the same contribution on average.

"Pinnal notch" frequency varies with elevation of source
e.g. in the medial plane.

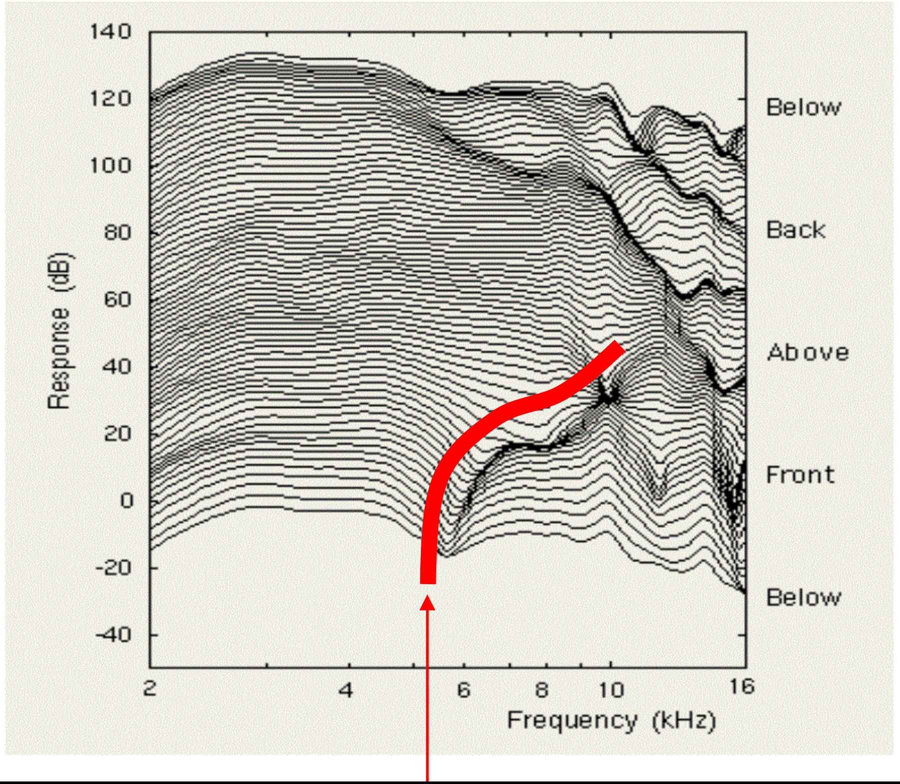
$$\hat{I}^j(\omega; \cancelto{0}{\phi},\theta) = \hat{g}^j(\omega)\ \hat{h}^j\cancelto{0}{(\omega}; \ \phi,\theta)\ \hat{I}_{src}(\omega; \ \phi,\theta)$$

HRTF from last lecture

Azimuth $\theta = 0$



e.g. medial plane



Levels of Analysis

high

- what is the task ? what problem is being solved?
  Source localization using level and timing differences within frequency channels.

- brain areas and pathways
  (cochlea to CN to MSO and LSO in the brainstem)

- neural coding
  (gave sketch only)

- neural mechanisms

low