Tampere University

Hearing

SGN 14007

Lecture 2

Annamaria Mesaros

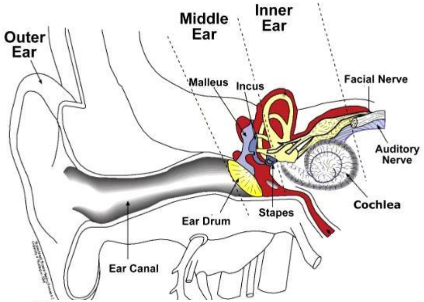
Auditory system

Two parts:

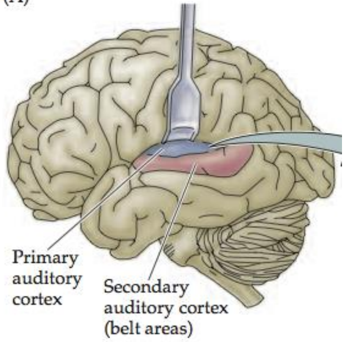
- Peripheral auditory system (outer, middle, inner ear)
- Auditory nervous system (in the brain)

Ear physiology: studies the peripheral system

Psychoacoustics: studies the entire process, relationships between physical sound stimuli and the subjective sensation



(A)



Auditory system

Dynamic range of hearing 120 dB

Audible spectrum of sound (freq.range) varies a lot between individuals:

- Only few can actually hear between 20 Hz and 20 kHz
- Sensitivity to low sounds (< 100Hz) is not very good
- Sensitivity to high sounds (> 12 kHz) decreases along with age

Selectivity of hearing:

- Can pick an instrument from among an orchestra
- Can follow a speaker in a noisy room (cocktail party effect)
- One can sleep in background noise but still wake up to an abnormal sound

Tampere University

Psychoacoustics

Perception involves information processing in the brain

- Information about the brain is limited

Psychoacoustics studies the relationships between sound stimuli and resulting sensations

- Attempt to model the process of perception

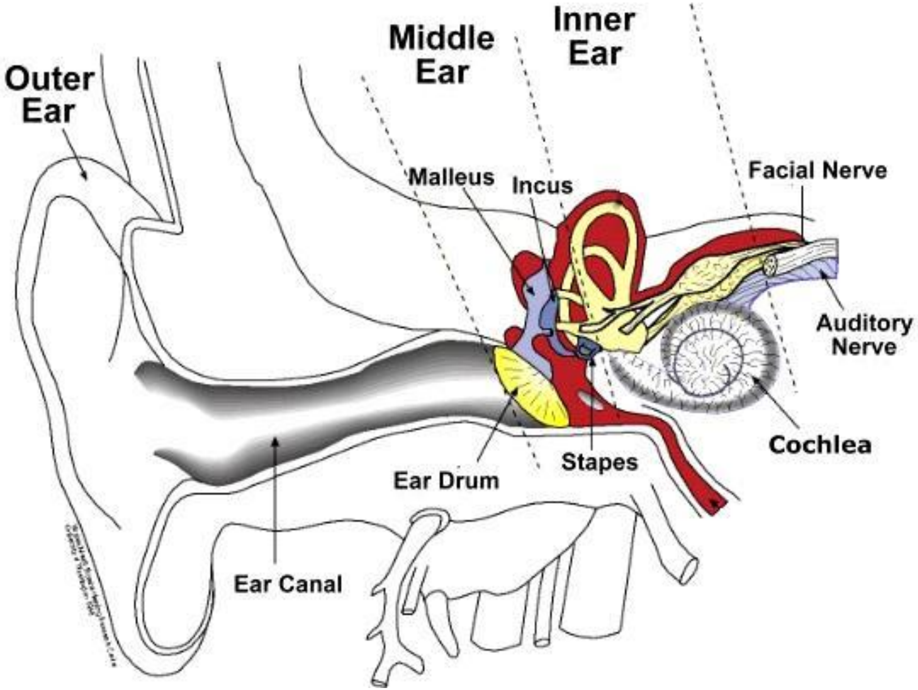
  For example trying to predict the perceived loudness / pitch / timbre from the acoustic properties of the sound signal

- Psychoacoustic listening test:

  Test subject listens to sound and asked to describe sensations

Ear physiology

The human ear consists of three main parts: outer ear, middle ear, and inner ear



Outer ear

Consists of pinna and auditory canal

Pinna

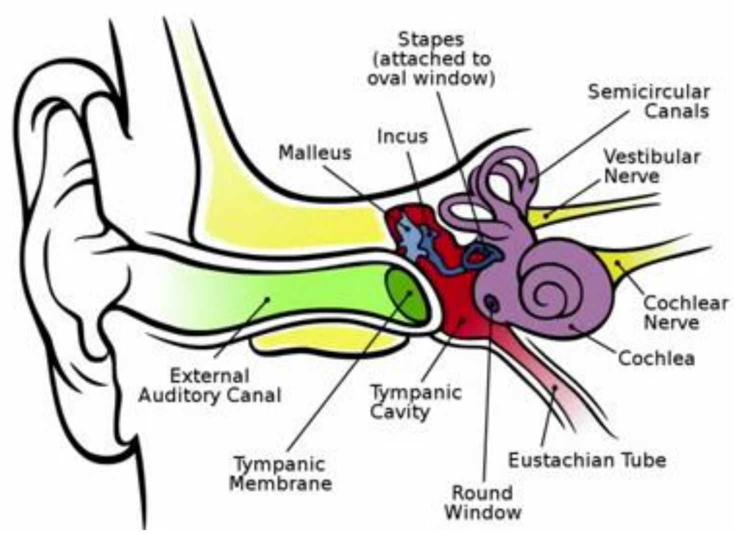
- Gathers sound; has direction-dependent response
- Vertically asymmetric shape, causes the outer ear to pass more high frequency content of an elevated source than an ear level source

Auditory canal (ear canal)

- Conveys sound to middle ear
- Acts as a resonator at 3-4 kHz, amplifies sound +10dB

Outer ear is passive and linear

- Its behavior can be completely described by laws of acoustic wave propagation



after Chittka, L. & Brockmann, A. , Perception Space —The Final Frontier (2005)

Middle ear

Consists of the eardrum and ossicles

Eardrum

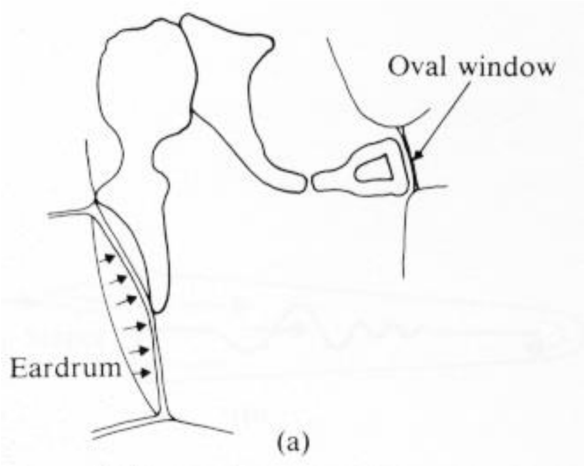
- Transforms sound waves into mechanical vibration

Ossicles

- Malleus (resting against the eardrum), Incus and Staples
- Transmit eardrum vibrations (in air) to the oval window of the inner ear (filled with fluid)

Acoustic reflex:

- When sound pressure level exceeds 50-60 dB, eardrum tension increases and staples is removed from the oval window
- Protects inner ear from damage

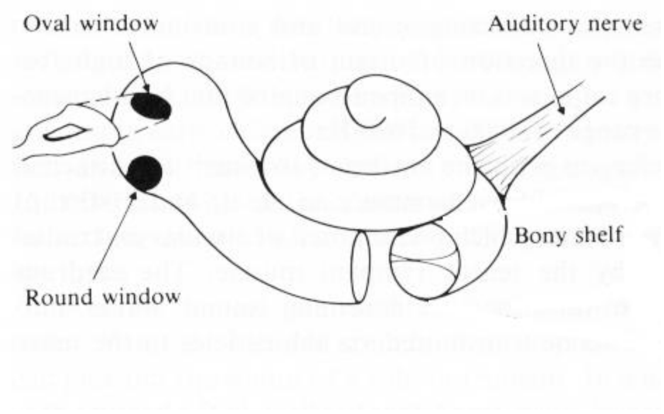


Tampere University

Inner ear

Contains the **cochlea** (greek: “snail shell”)

- Fluid-filled organ
- Spiral tube, approximately 35 millimeters long
- Vibrations on the cochlea’s oval window cause hydraulic pressure waves inside cochlea
- In the cochlea there is the **basilar membrane**
- On the basilar membrane there is the **organ of Corti** with **nerve cells** that are sensitive to vibration
- Nerve cells transform movement information into neural impulses in the **auditory nerve** which carries information to the brain



Basilar membrane

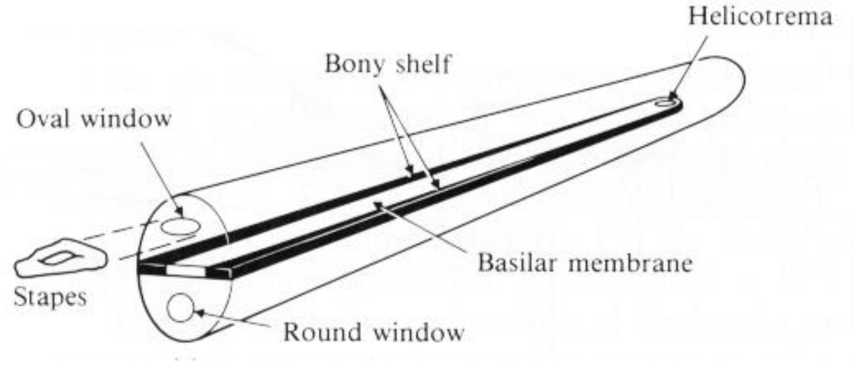
Sits on a flexible structure that divides the fluid into separate tunnels

- Hydraulic pressure waves travel along the cochlea

Varies gradually in tautness and shape along its length

- Low mass and stiffness at base (high resonance freq.),
- High mass and less stiffness at apex (low resonance freq.).

(cochlea, stretched out for illustration purposes)

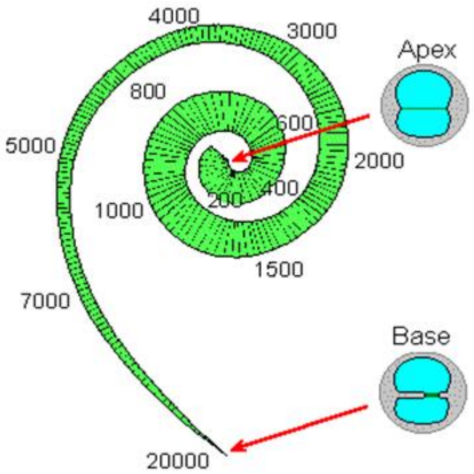


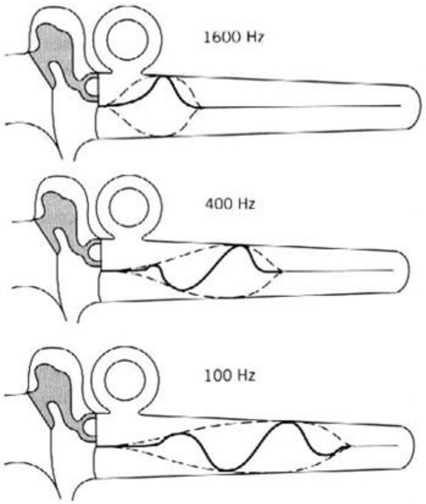
Basilar membrane

- Behaves as a vibrating spring-mass system
- Each location of BM has a **characteristic frequency** at which it vibrates maximally

  → mechanical spectrum analyser

- For a specific location the BM response is that of a bandpass filter





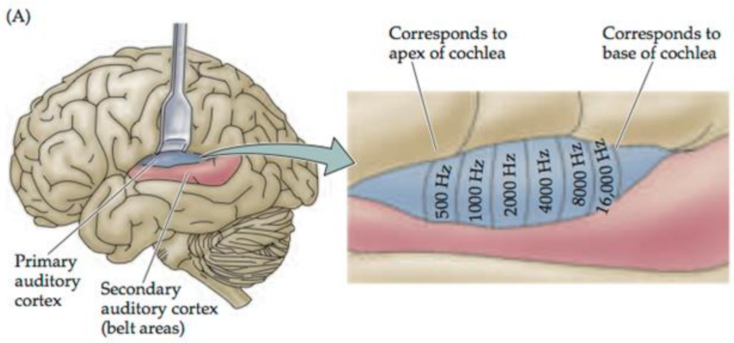
Sensory hair cells

- Distributed along the basilar membrane, transform its movement into neural impulses.
- Organized in rows running along the length of the basilar membrane, 3500 inner hair cells, 12000 outer hair cells
  - Anatomically and functionally distinct types
  - Outer hair cells mechanically amplify low-level sound that enters the cochlea
  - Inner hair cells transform the sound vibrations in the fluids of the cochlea into electrical signals; these are relayed via the auditory nerve to brain
  - Damage to these hair cells result in decreased hearing sensitivity (permanent)

- When an inner hair cell bends, it generates neural impulses into auditory nerve fibres. Impulse rate depends on vibration amplitude and frequency
  - Each hair cell is most sensitive to its characteristic frequency, which depends on the position on the BM

Auditory cortex

- The auditory pathway has several parallel paths from inner ear to the cortex
- Left and right ear information is combined early on and propagated to both sides of the brain.
- The primary auditory cortex (present on both sides) contains the precise frequency mapping of the basilar membrane
- The secondary auditory cortex receives more diffuse information
  - The processing that occurs in the auditory cortex is not well known



Tampere University

Thinking break (2 min)

Consider sound perception, and the attention mechanism that allows you to concentrate on different sounds in a mixture.

Psychoacoustics

Loudness and masking, pitch

Loudness

Loudness describes the subjective level of sound:

- The attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from quiet to loud
- Central part of psychoacoustics, consistent phenomenon
- Related to SPL, frequency content and duration of a sound

Loudness level (phon)

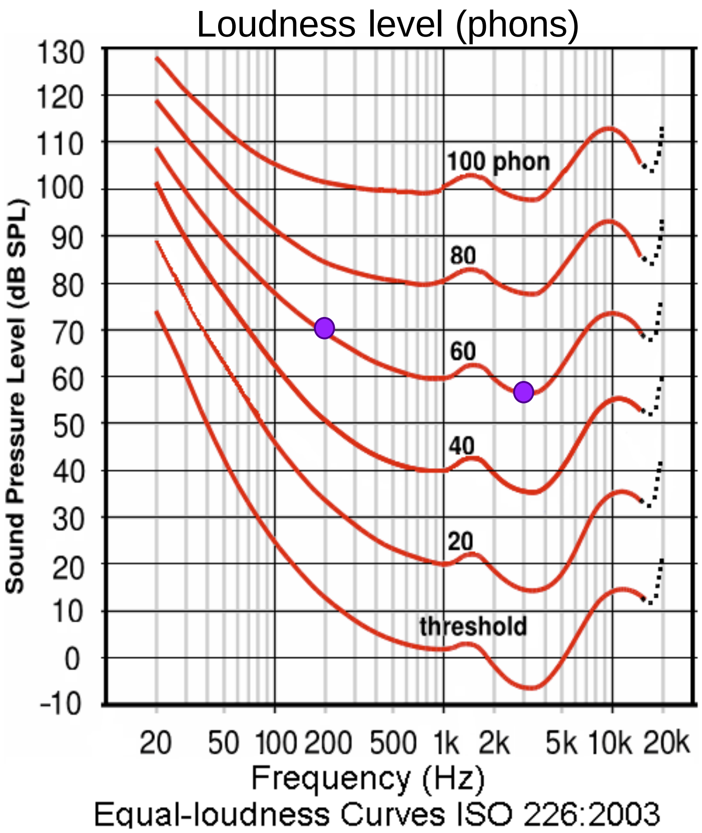
- Determined by comparison with standardized reference tone
- Sound pressure level (dB) of a 1000 Hz sinusoidal that has the same subjective loudness as the target sound
- e.g the heard sound is perceived as equally loud as 40 dB 1 kHz sinusoidal signal, its loudness level is 40 phons

Equal loudness curves

- **Equal loudness**: tones on the red curve as heard equally loud, even though their SPL is different!

  E.g. 70 dB at 200 Hz and 57 dB at 3000 Hz, 13 dB difference!

- Threshold of hearing
- Curves are based on several international studies



http://www.sengpielaudio.com/ISO226LoudnessCurves.gif

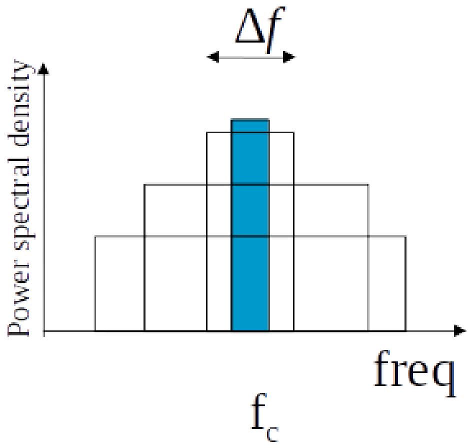
Critical bands

Frequency bandwidth of the “auditory filter” created by the cochlea

Linked to the inability of the auditory frequency analysis mechanism to resolve inputs that are close together in frequency

Implications on the perception of loudness:

- Ear analyzes sound at critical band resolution
- Each critical band contributes to the overall loudness level
- Comparing two band-limited noises with same center frequencies and SPL, if we increase the other’s bandwidth Δf, the **perceived loudness increases only after the critical bandwidth**

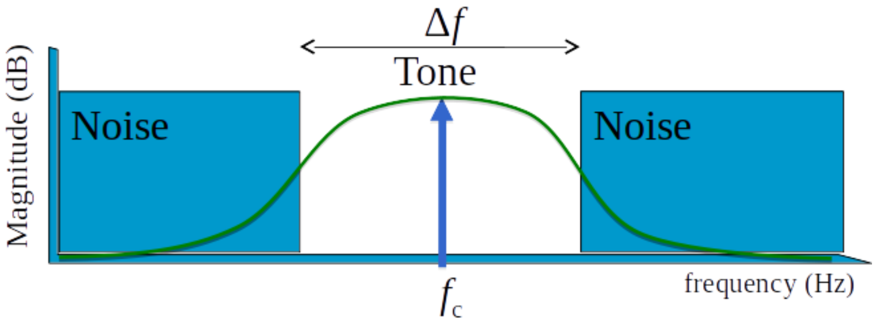


Critical bands

Basilar membrane: different frequencies resonate particularly strongly at different points along the membrane → auditory filters are **band-pass filters**

Equivalent rectangular bandwidth (ERB) - relationship between the auditory filter, frequency, and the critical bandwidth.

- Select equal amplitude of tone and noise
- Widen bandwidth Δf around tone until tone is detected
- Minimum bandwidth Δf ERB that allows tone detection is the auditory filter bandwidth
- Experimentally determined: $\Delta f_{ERB} = 24.7 + 0.108\, f_c$



Loudness of a complex sound

A broadband sound is perceived louder than a narrowband sound of equal SPL

- Loudness of complex sounds is calculated by using specific loudness of critical bands
- Specific loudness is roughly proportional to the log-power of the signal at the band
- Overall loudness is obtained by summing up specific loudness values over critical bands

Loudness varies with frequency!

Frequency sensitivity along basilar membrane is not linear

→ bandwidth of critical bands is different depending on frequency

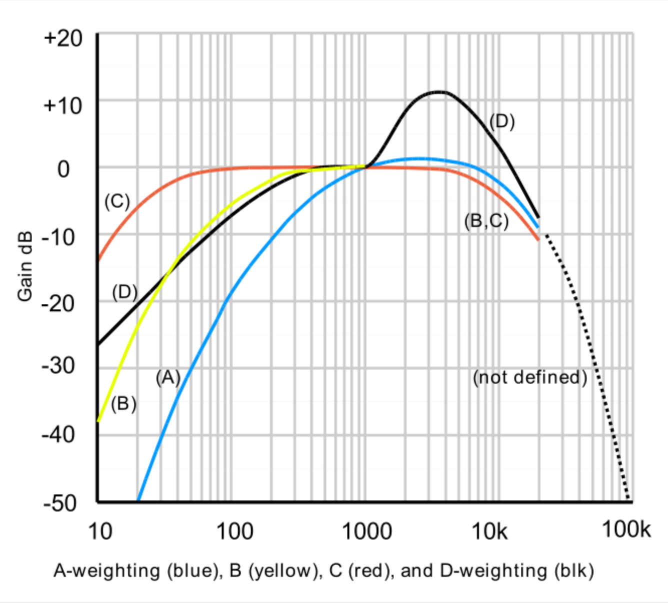
Measuring sound level

SPL is not good metric as perceptual measure of loudness

Weighing the SPL using curves related to frequency sensitivity of hearing

A-weighing is standard

- Roughly resembles the inverse of the hearing threshold curve
- Sensitivity of human ear for pure tones, relatively quiet level
- Others developed for different levels of SPL and more representative to noise rather than pure tones (D - curve peak)



Masking

Phenomenon when **perception of one sound is affected by the presence of another sound** (maskee and masker)

Masking depends on:

- Spectral structure of the sound
- Temporal structure

Exploited in audio compression (mp3)

Masking in frequency domain

- Also called **simultaneous masking**: a sound is made inaudible by a noise or unwanted sound that happens at the same time

- Ability to hear frequencies separately: frequency resolution of the ear

  → Sounds within the **same critical band** mask each other easily!

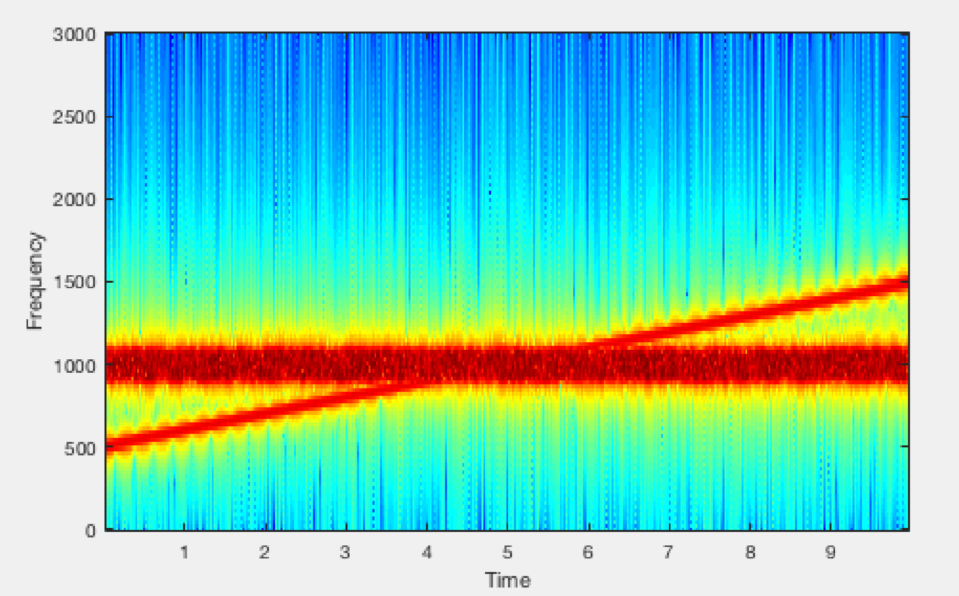
- Narrowband noise masks a tone (sinusoidal) easier than a tone masks noise

- **Masked threshold**: the threshold of audibility is raised because of the masker
  - Sounds with a level below the masked threshold are inaudible
  - In quiet, masked threshold=threshold of hearing

Masking in frequency domain

Narrow band noise at 1 kHz, bandwidth 160 Hz

Tone with frequency 0.5-1.5 Hz



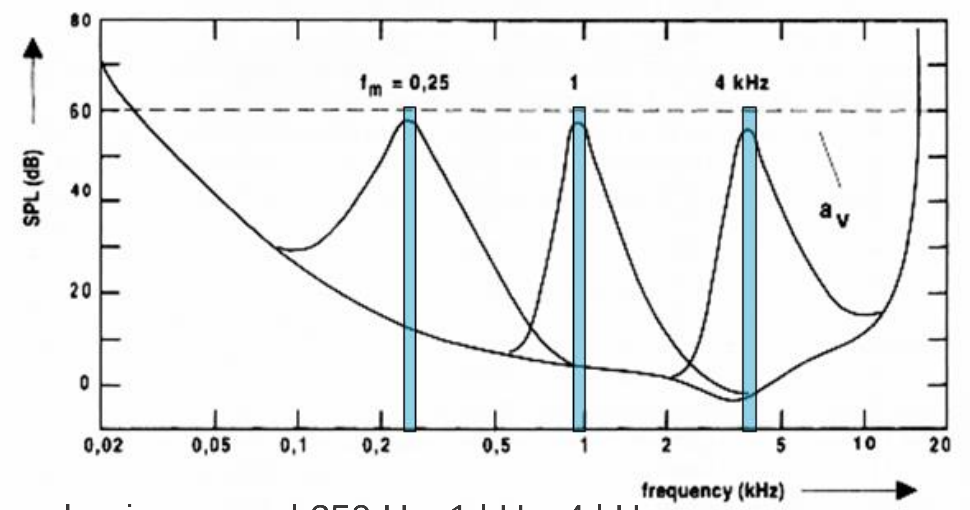
sin

noise

mixture

Masked thresholds

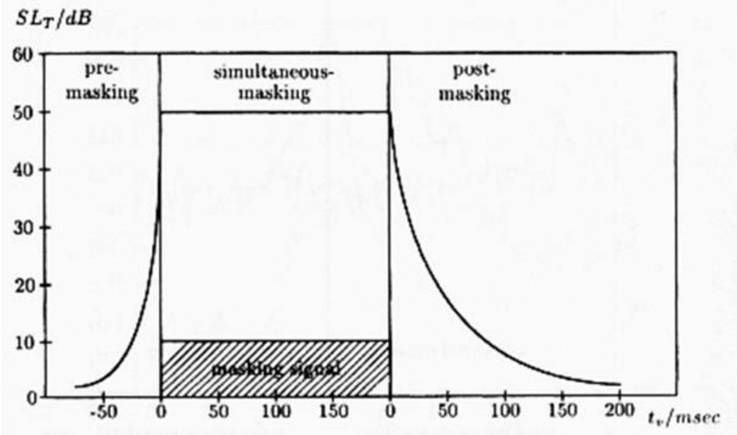
- Spreading function: The effect of masking extends to the spectral vicinity of the masker (spreads more towards high frequencies)
- Additivity of masking: joint masked threshold is approximately (but slightly more than) the sum of the components



Maskers: narrowband noise around 250 Hz, 1 kHz, 4 kHz

Masking in time domain

- Also called **non-simultaneous masking**: a sudden sound makes inaudible other sounds which are present immediately before or after it
- Does not extend far in time
- **Forward masking**: masking effect extends to times after the masker is switched off (can be up to 200ms)
- **Backward masking**: masking extends to times before the masker is switched on (relatively short, up to 50 ms)



Thinking break (2 min)

Think of your everyday life experience and what kind of auditory masking examples you encountered.

Pitch

Pitch:

- Subjective attribute of sounds that enables us to arrange them on a frequency-related scale ranging from low to high
- Sound has a certain pitch is human listeners can consistently match the frequency of a sinusoidal tone to the pitch of the sound
- Measured in Hz

Pitch vs fundamental frequency

- Fundamental frequency is a **physical attribute** while pitch is a **subjective attribute**
- In practice, perceived pitch ≅fundamental frequency
- “Perfect pitch” or “absolute pitch” - the ability to recognize the pitch of a musical note without any reference (minority of population)

Mel scale

Frequency scale created using psychoacoustic tests to quantify pitch

"Adjust the pitch of a test tone **two times higher** than the reference tone."

- In music this is called octave (ratio of fundamental frequencies is 2:1)
- Reference point: **1000 mels** assigned to 1000 Hz tone

Result:

- Below 500-1000 Hz subjects double the frequency
- Over 1000 Hz subjects use considerably larger steps

One commonly used mel scale approximation:

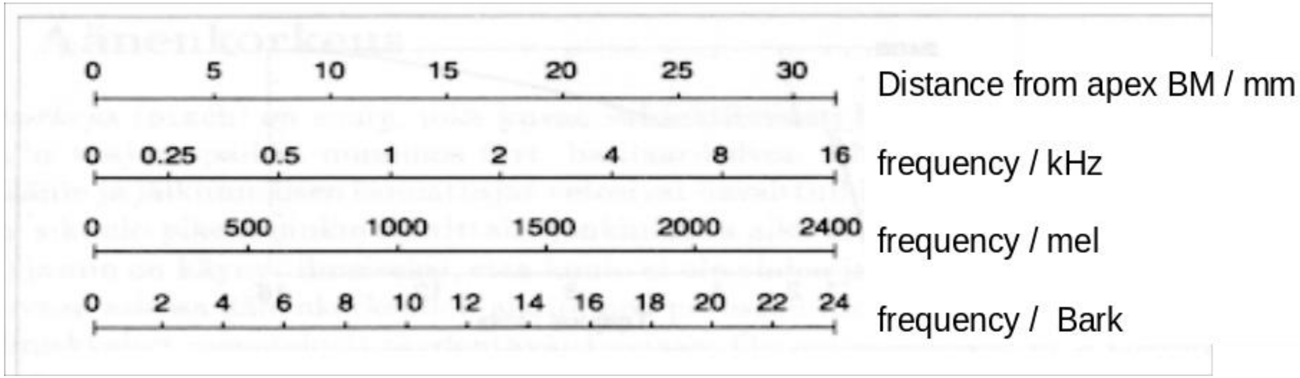
$$m = 2595 \log_{10}(1+f/700)$$

m is mel value

f is frequency (Hz)

Perceptually-motivated frequency scales

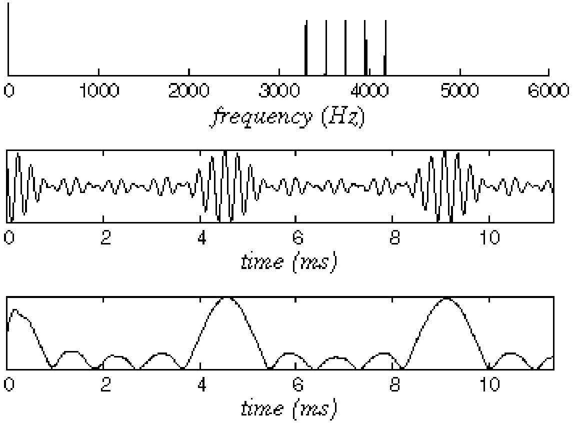
- Critical bands aldo determine a pitch scale (similar to mel scale if placed in succession)
- Bark, ERB, mel scales have approximately linear relation to the position of resonance on the basilar membrane
  - ERB bandwidth is approximately 0.9 mm in basilar membrane
  - 1 Bark = 100 mels



Pitch perception

Was explained by two competing theories:

- Place theory: “Peak activity along the basilar membrane determines pitch” (fails to explain missing fundamental)
- Periodicity theory: “Pitch depends on rate, not place, of response.” Neurons fire in sync with signals
- Current assumption: combination of the above
  - Sound is subdivided into subbands (critical bands)
  - Periodicity of the amplitude envelope is analyzed within bands
  - Results are combined across bands



Missing fundamental

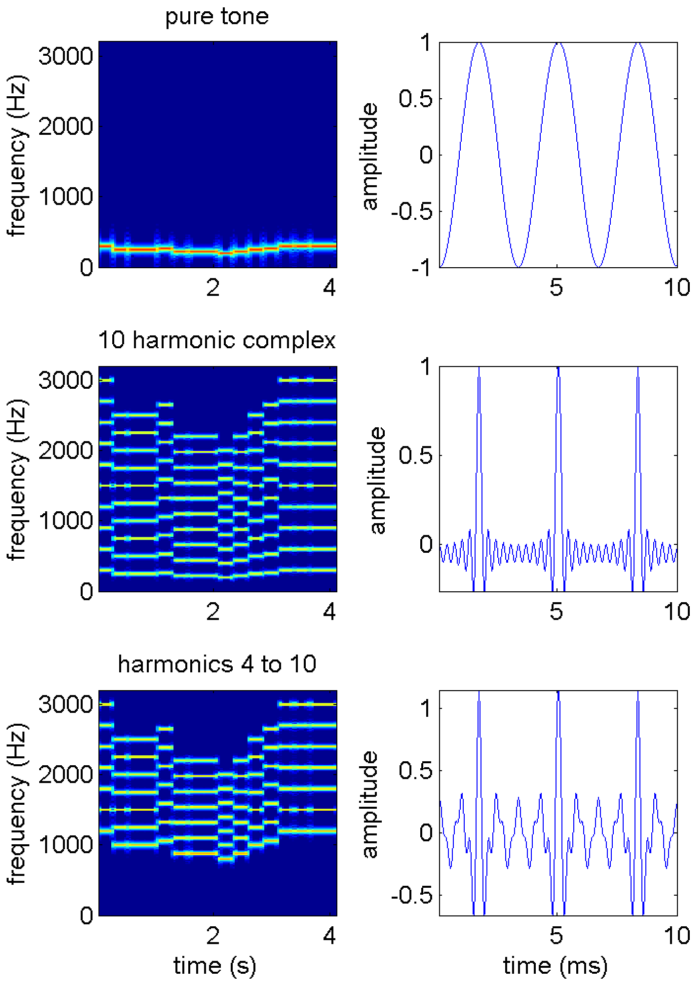
Demo:

http://auditoryneuroscience.com/pitch/missing-fundamentals

Pure tones

Harmonics 1-10

Harmonics 4-10



Spatial hearing

Spatial hearing

Most important auditory cues for localizing a sound source in space are:

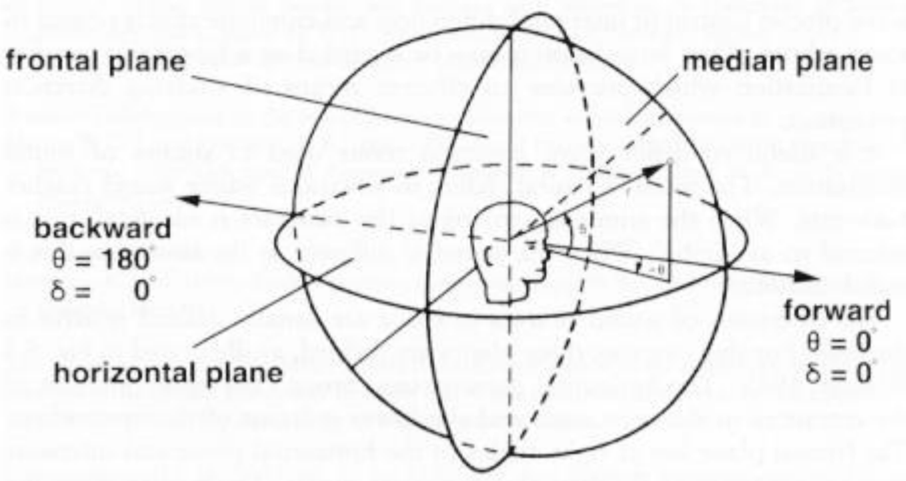
- Interaural **time difference (ITD)** - relatively constant with frequency
- Interaural **intensity difference (ILD)** - increases with frequency
- Direction-dependent filtering of the **sound spectrum** by head and pinna

Terms used for spatial hearing:

- Monaural : with one ear
- Binaural : with two ears
- Interaural : between the ears (interaural time difference etc)
- Lateralization : localizing a source within the head

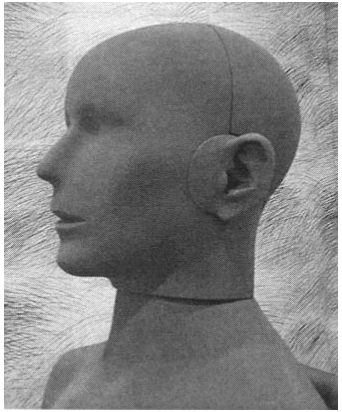
Monaural source localization

- Directional hearing works to some extent even with one ear
- Head and pinna form a **direction-dependent filter**
  - Direction-dependent changes in the spectrum of the sound arriving in the ear can be described with HRTFs (head-related transfer function)
- HRTFs are crucial for localizing sources in the median plane (vertical localization)



HRTF and measurement

- HRTFs can be measured by recording
  - Sound emitted by a source
  - Sounds arriving to the auditory canal or eardrum (transfer function of the auditory canal does not vary along with direction)
- Measurement in practice:
  - microphone in the ear of a test subject
  - head and torso simulator
- HRTFs are **individual.** (large differences!)
- Online databases exist

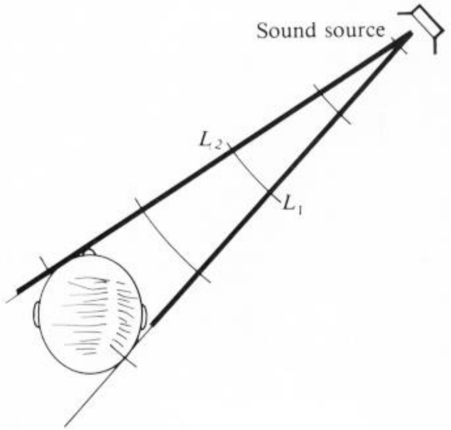


Localizing a tone

- Experimenting with sinusoidal tones helps to understand the localization of more complex sounds

- Angle-of-arrival perception for sinusoids **below 750 Hz** is based mainly on interaural time difference (ITD)

  - Delay is due to distance between the ears *D* and limited speed of sound *c*
  - Delay bounds between ears $\pm D/c$, i.e. between ± 600 to 700 µs
  - Higher frequencies cause ambiguous ITD

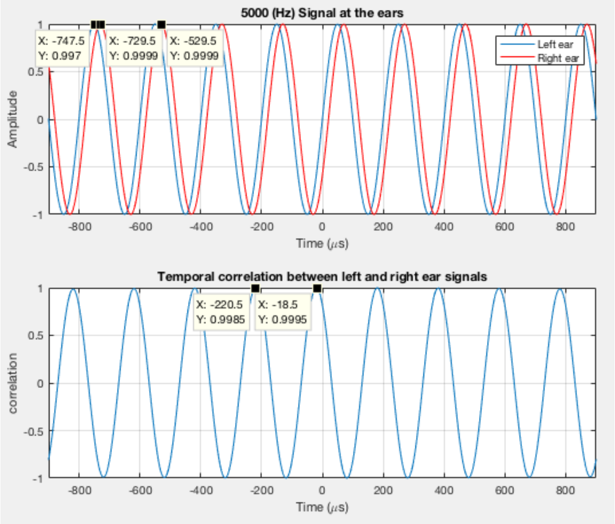


Phase ambiguity

Tone at frequency *f* arrives the left ear first; it has one possible ITD value if *f* < 750 Hz

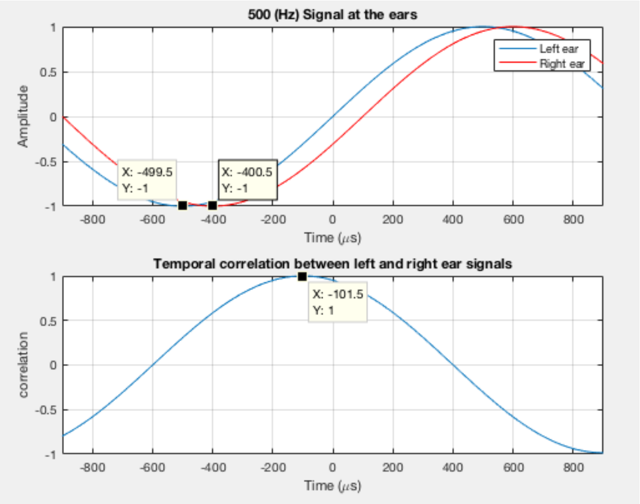
5 kHz tone arrives to ears:

multiple possible ITD values



500 Hz tone arrives to ears:

single ITD value

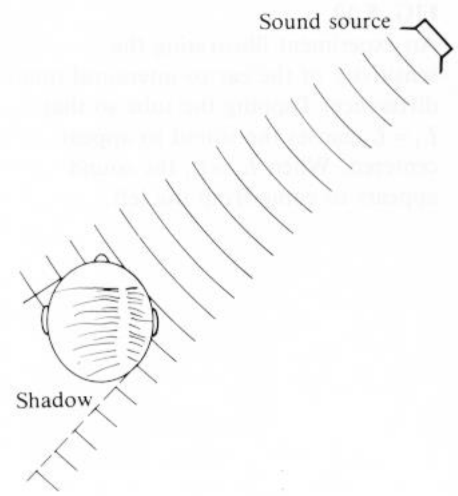


Here: -700µs source at right side, +700µs, source at left side, 0µs = front/back

Localizing a tone

Interaural time difference is useful only up to 750 Hz

- Above that, the time difference is ambiguous, since there are several wavelengths within the time difference
- Moving the head (or source movement) helps: can be done up to 1500 Hz
- At higher frequencies (> 750 Hz) the auditory system uses interaural level difference (ILD)
  - Head causes an acoustic ”shadow” (sound level is lower behind the head)
  - Works especially at high frequencies with shorter wavelenghts compared to head dimensions



Localizing complex sounds

Complex sounds refer to sounds that

- involve a number of different frequency components and
- vary over time

Localizing sound sources is typically a result of combining all the above-described mechanisms

- Interaural time difference (most important)
- Interaural intensity difference
- HRTFs

Wideband noise: directional hearing works well

Lateralization in headphone listening

When listening with headphones, the sounds are often localized **inside the head**, on the axis between the ears:

- Sound does not seem to come from outside the head because the diffraction caused by pinnae and head is missing
- Lateralization = localizing the sound within the head
- If the sounds are processed with HRTFs carefully, they move outside the head
  - Done by convolution of the measured HRIRs (head related impulse response, left and right ear of a sound source in desired direction) with the source signal

Precedence effect

The sound in indoors that reaches the listener contains the direct path, early reflections, and late reverberation.

Precedence effect: the brain suppresses early reflected sounds to aid in source direction perception

- Early reflections are not heard as separate sounds.
- Instead, they amplify the loudness of direct sound

Precedence effect works when:

- Early reflections arrive less than ~ 35 ms after direct sound.
- The spectrum of a reflection is similar to that of the direct sound
- Reflections are lower or at least not significantly louder than the direct sound
  - Even a sound with +10 dB SPL than the direct sound can not be heard

Tampere University

Summary

Basic function of the ear, how we hear sounds

Psychoacoustics

- Loudness
- Pitch
- Masking, critical bands

Binaural hearing

- How do we locate sounds?
- HRTF