Human-machine interaction in virtual reality

Paul MacNeilage, Psychology

Eelke Folmer, Computer Science

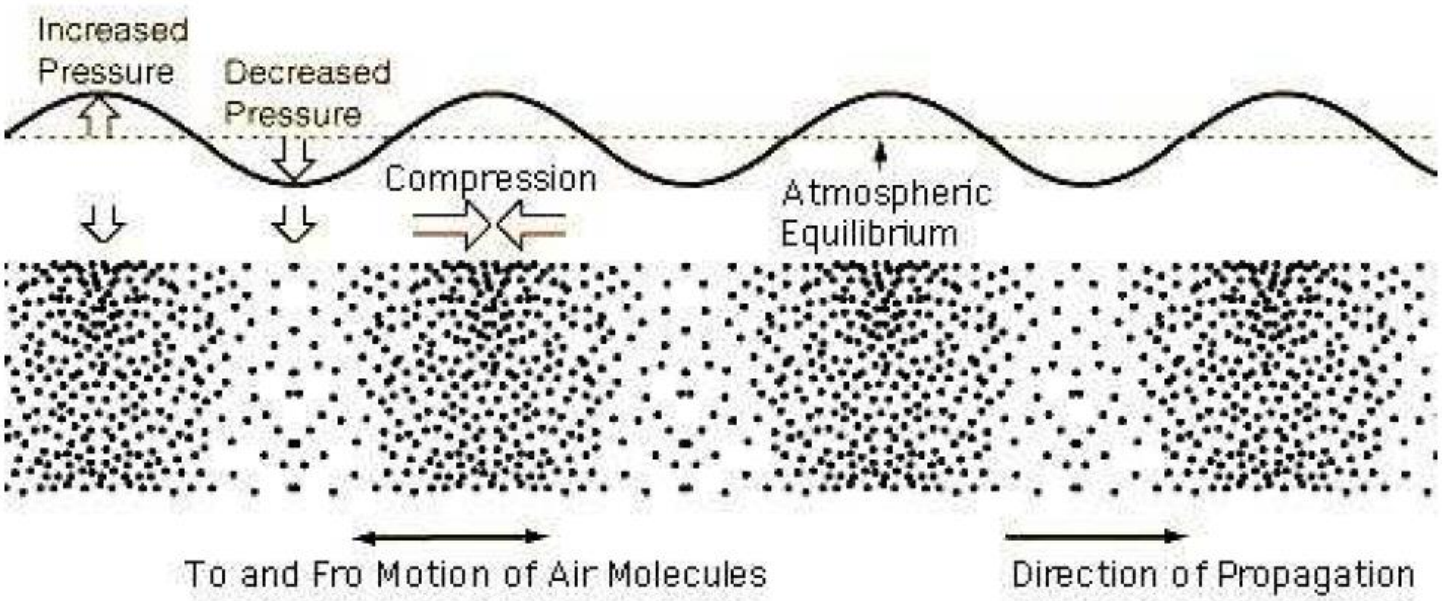
Auditory Stimuli in VR

- Physics of sound
- Physiology of hearing
- Auditory Perception
- Auditory Rendering

- Keep in mind comparison with light / vision

Physics of Sound

- Sound is vibration of molecules in a medium



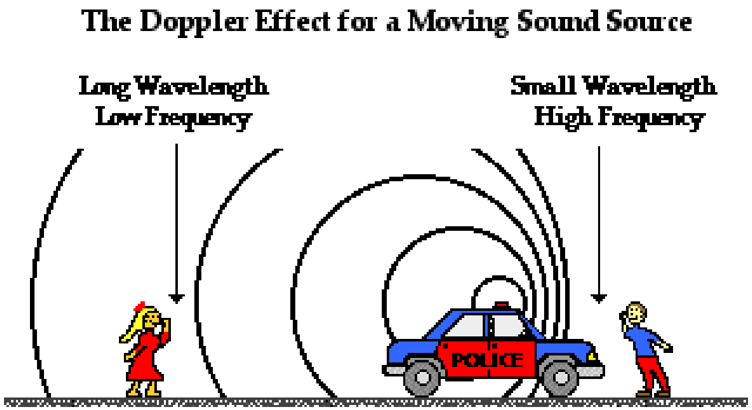
- Waves, just like light!

Important properties of sound

- Attenuation – fixed percentage per distance
- Speed of sound - 343.2 m/s (767 m/h)
- Frequency of audible sound – 20 to 20,000 Hz
- Wavelength – 17 m to 17 mm

Doppler Effect

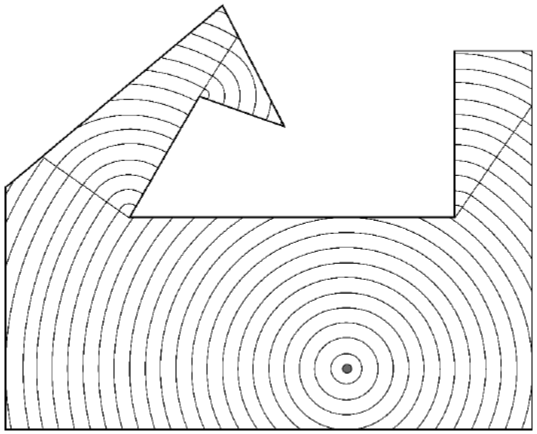
- When sounds or listeners are moving…



  - Waves compressed for sounds moving towards
  - Waves expanded for sounds moving away

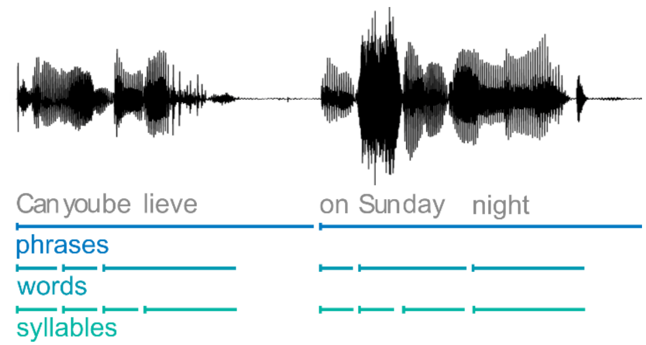
Reflection, Transmission, Diffraction

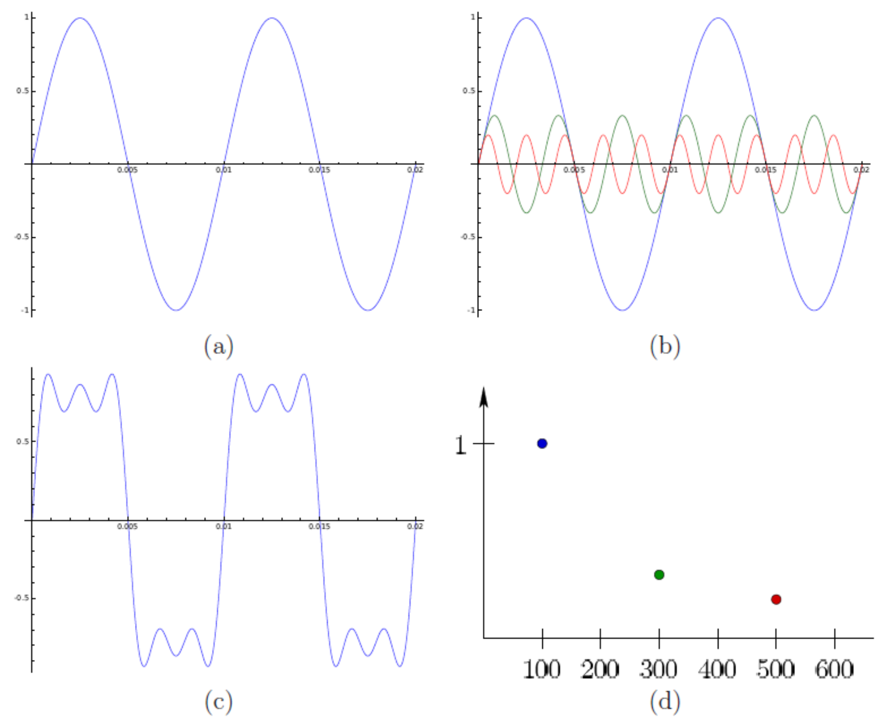
- Reflection – sounds bounce back
- Transmission – sounds pass through
- Diffraction – sounds bend around corners



Fourier Analysis

- Every sound can be represented as a waveform
- Any waveform can be broken down into a sum of sine waves
- Compressed representation, good for efficient computing





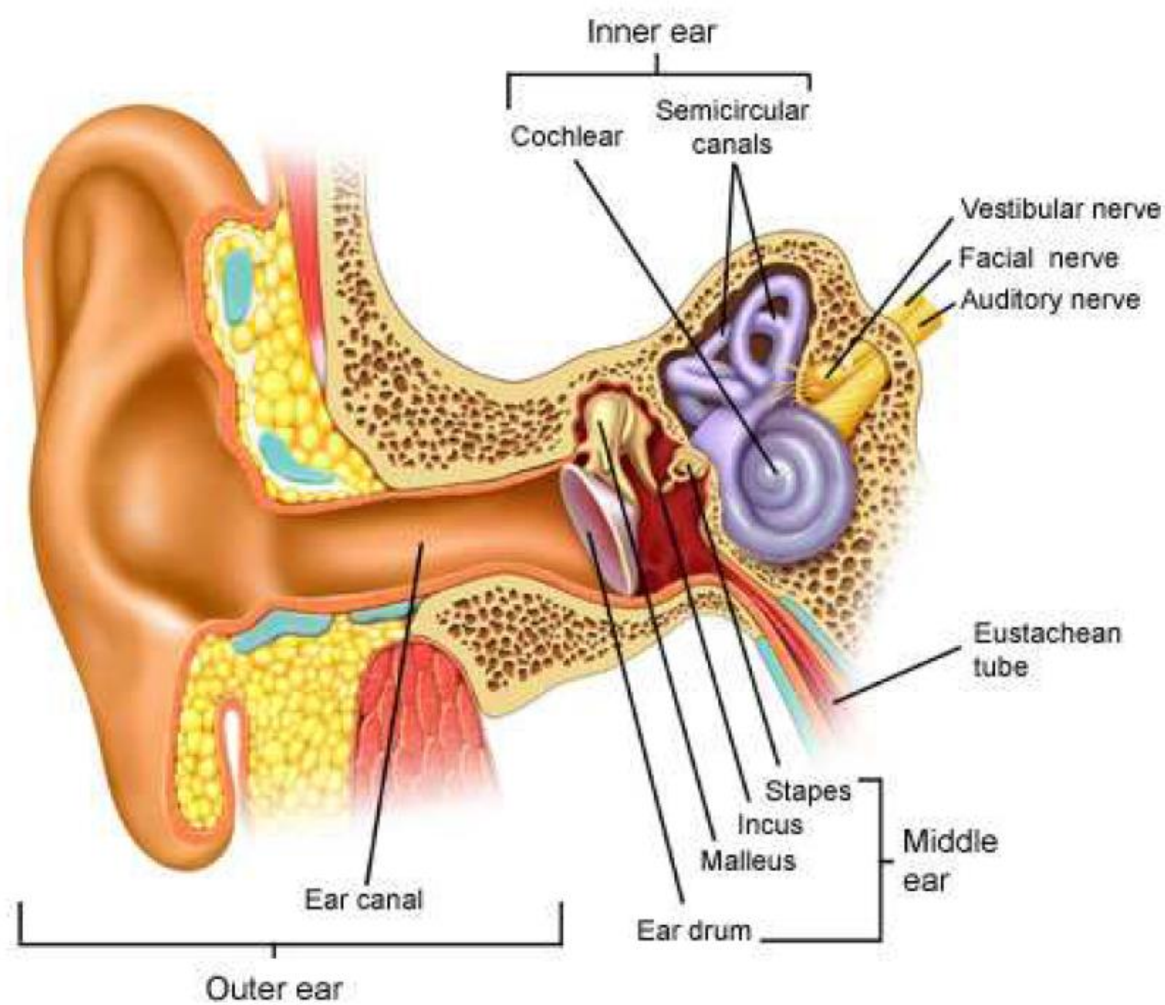
Auditory Stimuli in VR

- Physics of sound
- Physiology of hearing
- Auditory Perception
- Auditory Rendering

- Keep in mind comparison with light / vision

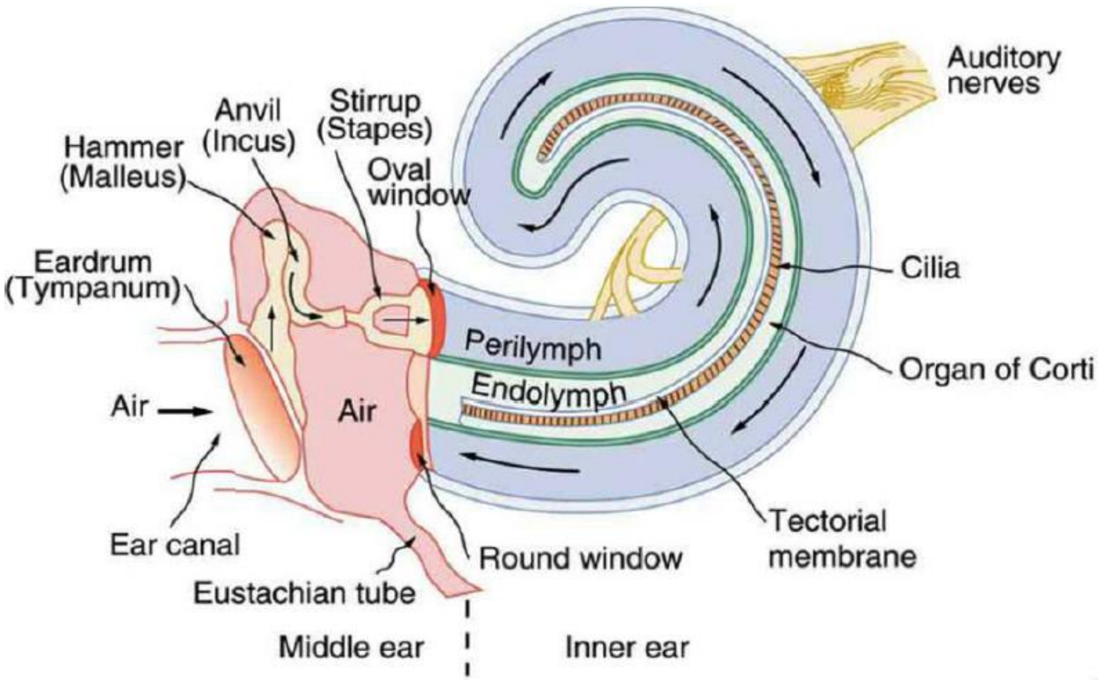
Anatomy of the Ear

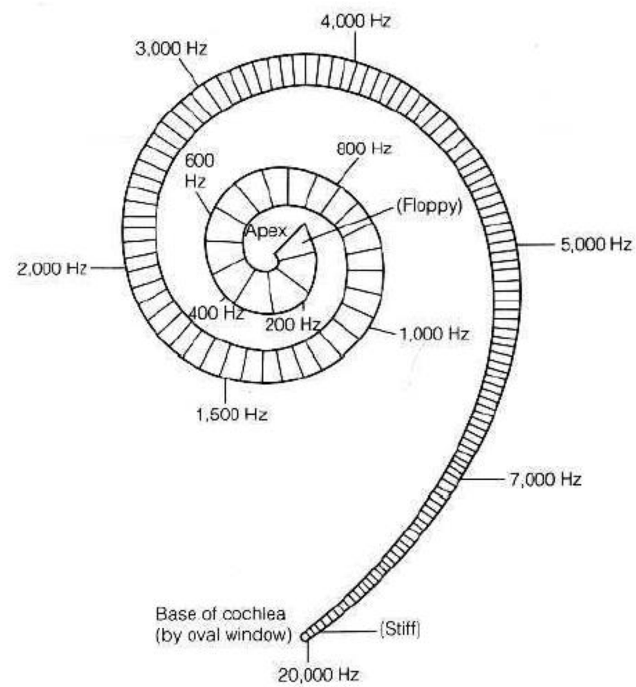
- Outer
  - 1500 to 7500Hz Amplified
- Middle
  - Hammer, anvil, stirrup
  - Transmission not reflectance
- Inner
  - Cochlea
  - Vibration -> spikes



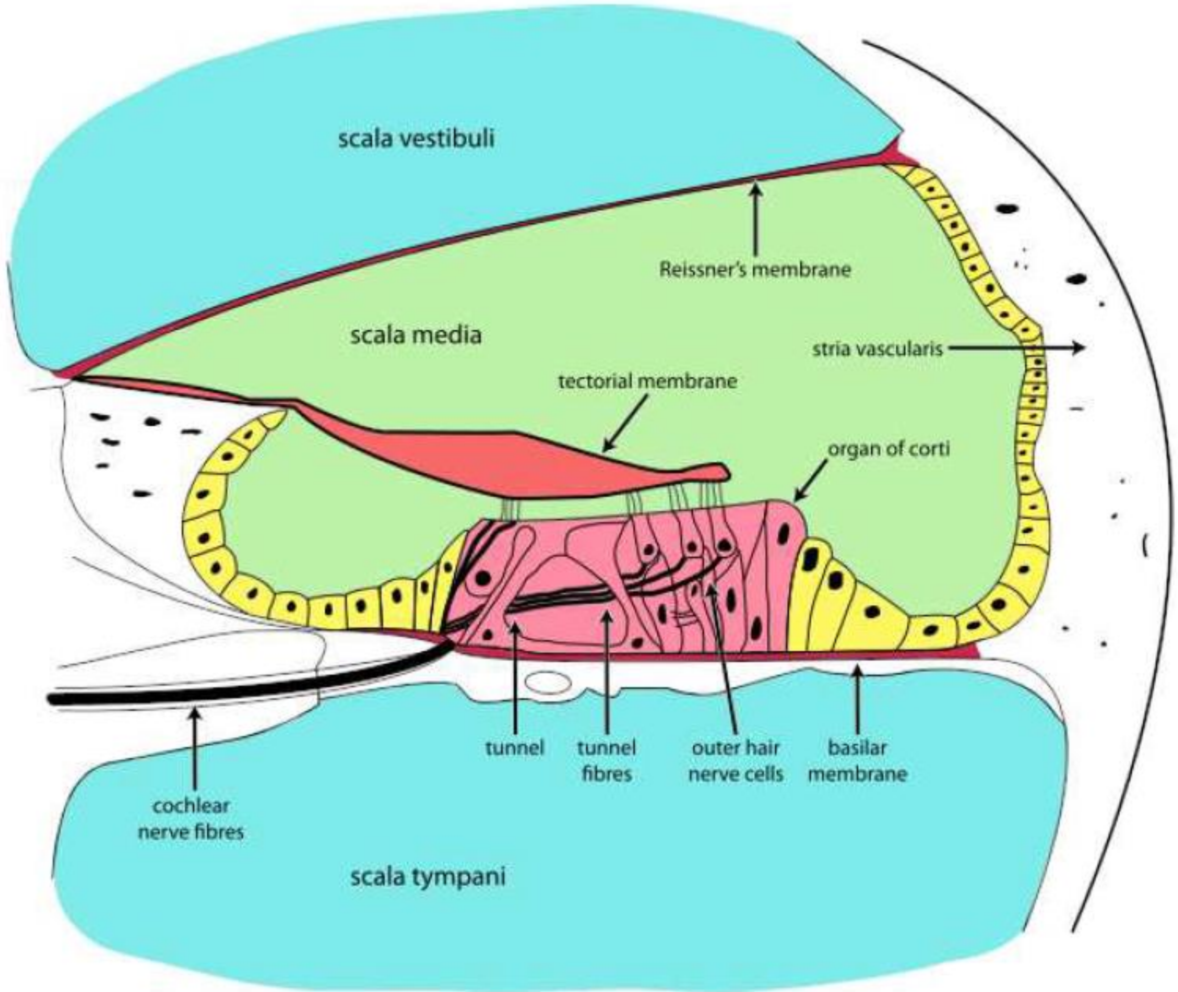
Vibrations to Spikes

- Cochlea decomposes sound into frequencies, amplitudes, and phases



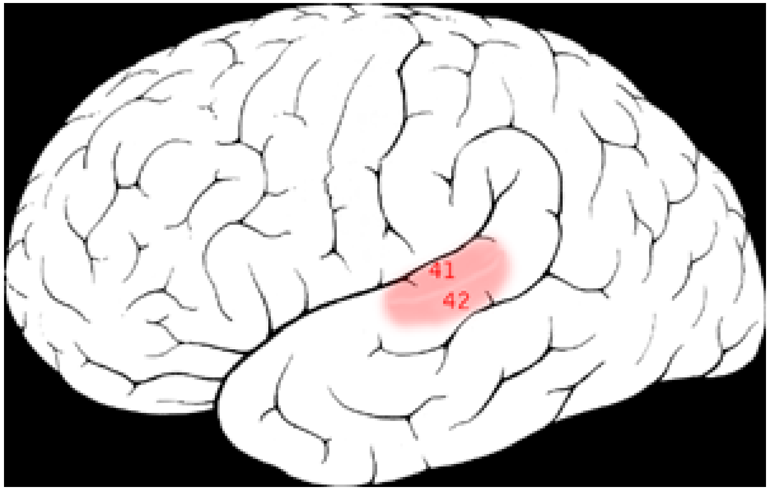


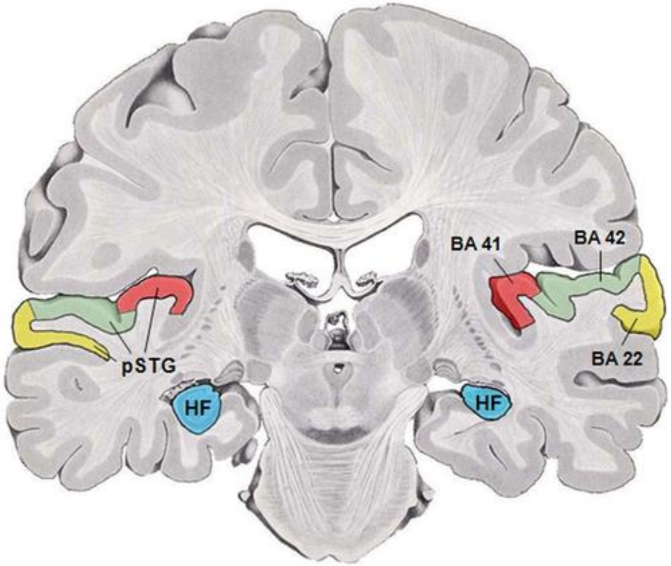
Vibrations to Spikes



Primary Auditory Cortex

- Brain region specialized for auditory processing
- Tonotopic organization





Auditory Stimuli in VR

- Physics of sound
- Physiology of hearing
- Auditory Perception
- Auditory Rendering

- Keep in mind comparison with light / vision

Auditory Perception

- Many parallels with visual perception
  - Psychophysics
    - Measuring perceptual response to physical stimulus
  - Illusions
    - Misperception provides clues to processing

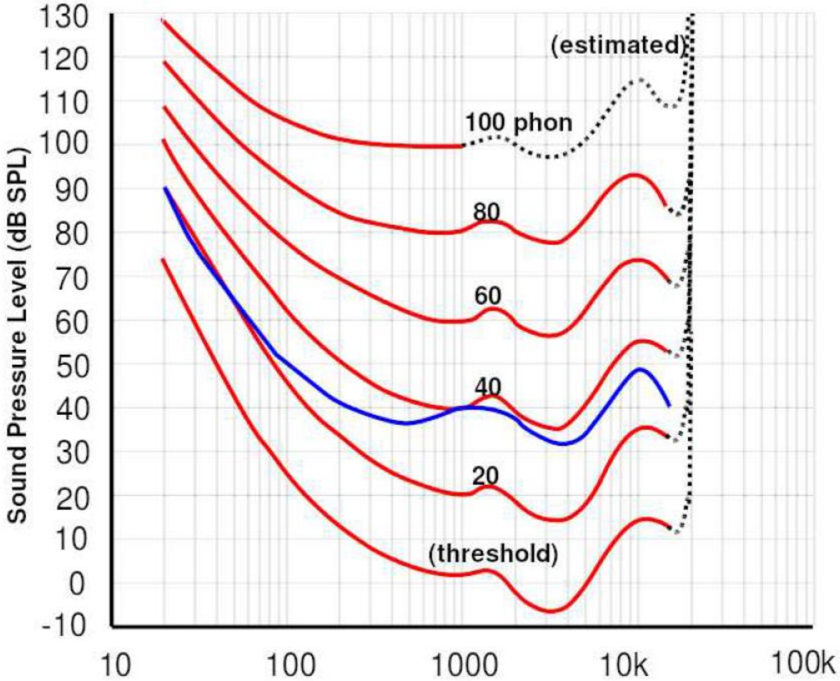
Precedence Effect

- Same sound, different times, only one sound perceived

- Suppress reverberations, echoes (below echo threshold, 3 to 61 ms)

Equal Loudness Contours

- Low-frequency sounds perceived quieter



Equal-loudness contours (red) (from ISO 226:2003 revision)
Original ISO standard shown (blue) for 40-phons

Pitch Perception

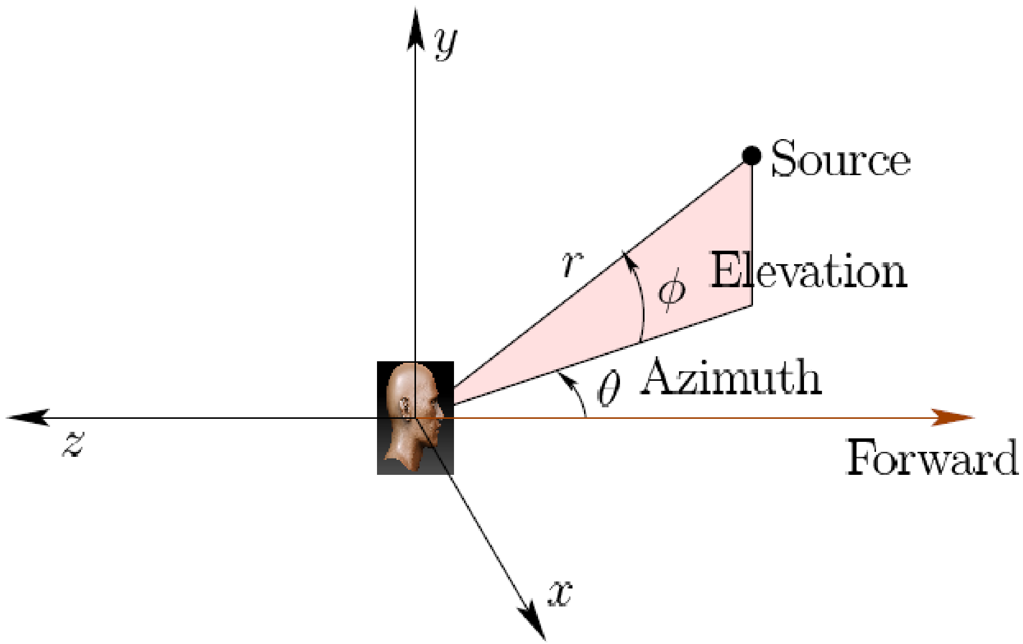
- Discrimination depends on frequency
  - 1000 Hz, JND of 1 Hz, 0.1%
  - 10,000 Hz, JND of 100 Hz, 1%

- Critical band masking – suppression of nearby frequencies when one frequency is played

- Missing fundamental – filling in; when higher harmonics are played, missing fundamental is heard

Localization

- Where is the sound coming from?

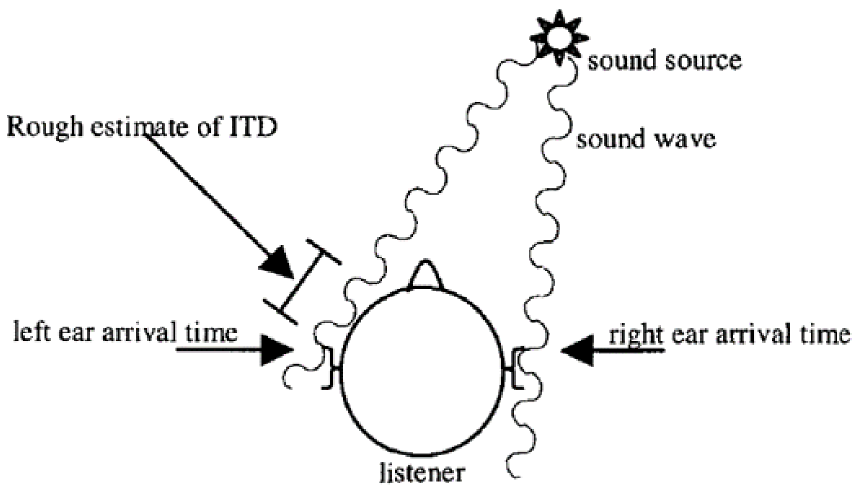


Binaural Cues

- Inter-aural time difference (ITD)
  - Sounds arrive at the two ears at different times

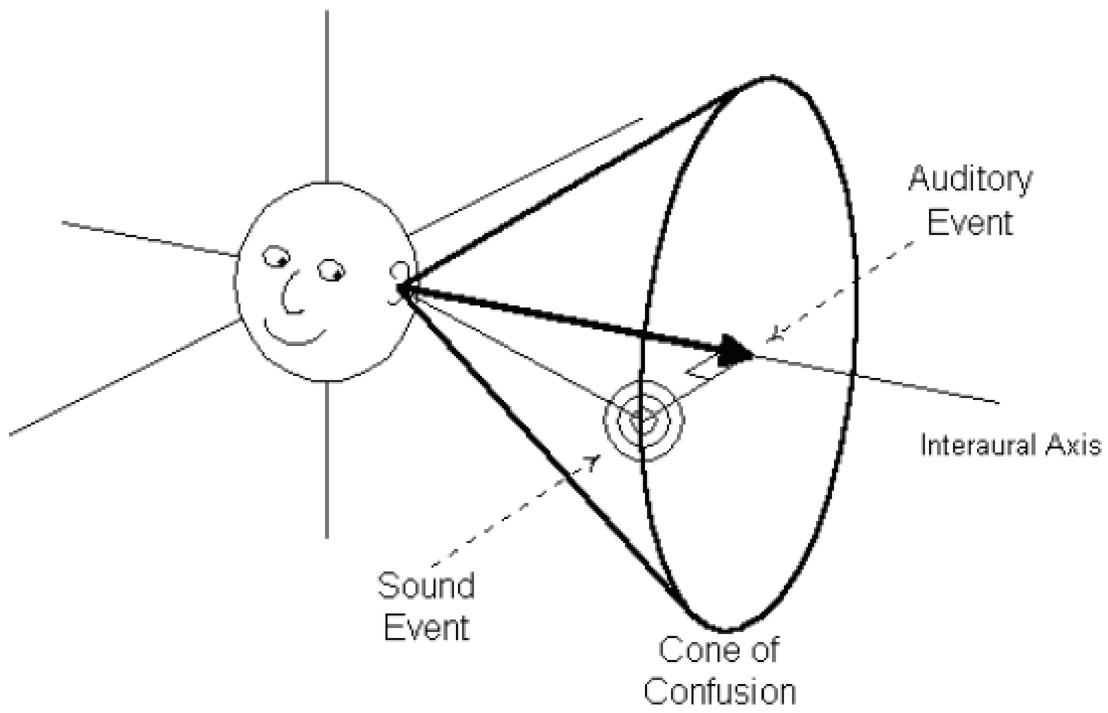
- Inter-aural loudness difference (ILD)
  - Loudness at the two ears is different (acoustic shadowing)

- Differences act as cues to azimuth



Cone of Confusion

- All locations consistent with a given ITD



Monoaural Cues

- Familiar loudness – cue to distance
- Frequency filtering by pinna - depends on elevation
- Frequency filtering by distance – high frequencies fade faster
- Reverberations – tell about environmental structure (e.g. echolocation)

Auditory Motion

- Provides additional information through multiple “views” of the sound source
- Similar to visual motion parallax – stimulus changes contingent on head motion
- Major contribution to “externalization” of perceived sound location

Published 2018

- First scientific evidence --

Psychophysical evidence for auditory motion parallax

Daria Genzel[a,b,1], Michael Schutte[a,1], W. Owen Brimijoin[c], Paul R. MacNeilage[b,d,2], and Lutz Wiegrebe[a,b,3]

[a]Department Biology II, Ludwig Maximilians University Munich, 82152 Planegg-Martinsried, Germany; [b]Bernstein Center for Computational Neuroscience Munich, 82152 Planegg-Martinsried, Germany; [c]Glasgow Royal Infirmary, Medical Research Council/Chief Scientist Office Institute of Hearing Research (Scottish Section), G31 2ER Glasgow, United Kingdom; and [d]Deutsches Schwindel- und Gleichgewichtszentrum, University Hospital of Munich, 81377 Munich, Germany

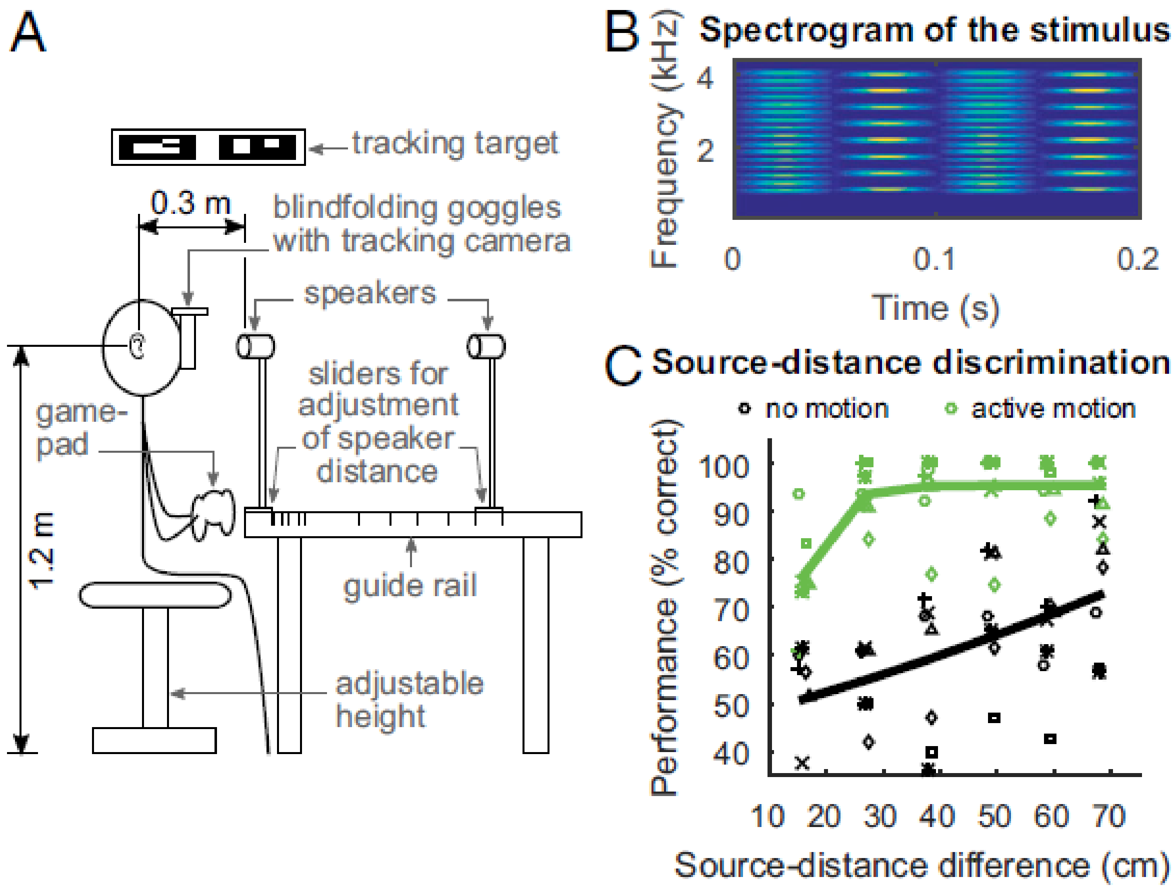
Edited by Wilson S. Geisler, University of Texas at Austin, Austin, TX, and approved February 16, 2018 (received for review July 6, 2017)

Distance is important: From an ecological perspective, knowledge about the distance to either prey or predator is vital. However, the distance of an unknown sound source is particularly difficult to assess, especially in anechoic environments. In vision, changes in perspective resulting from observer motion produce a reliable, consistent, and unambiguous impression of depth known as motion parallax. Here we demonstrate with formal psychophysics that humans can exploit auditory motion parallax, i.e., the change in the dynamic binaural cues elicited by self-motion, to assess the relative depths of two sound sources. Our data show that sensitivity to relative depth is best when subjects move actively; performance deteriorates when subjects are moved by a motion platform or when the sound sources themselves move. This is true even though the dynamic binaural cues elicited by these three types of motion are identical. Our data demonstrate a perceptual strategy to segregate intermittent sound sources in depth and highlight the tight interaction between self-motion and binaural processing that allows assessment of the spatial layout of complex acoustic scenes.

and spectral composition of the emitted sounds, distance estimation for humans is indeed impossible (10). This is not surprising, considering that an important visual distance cue (binocular disparity) is not available in audition, not least because humans cannot point each of their ears toward a sound source. Some visual depth cues have auditory counterparts, (e.g., blur is related to frequency-dependent atmospheric attenuation, and relative size to loudness), but many others are unavailable (e.g., occlusion, texture gradients, shading).

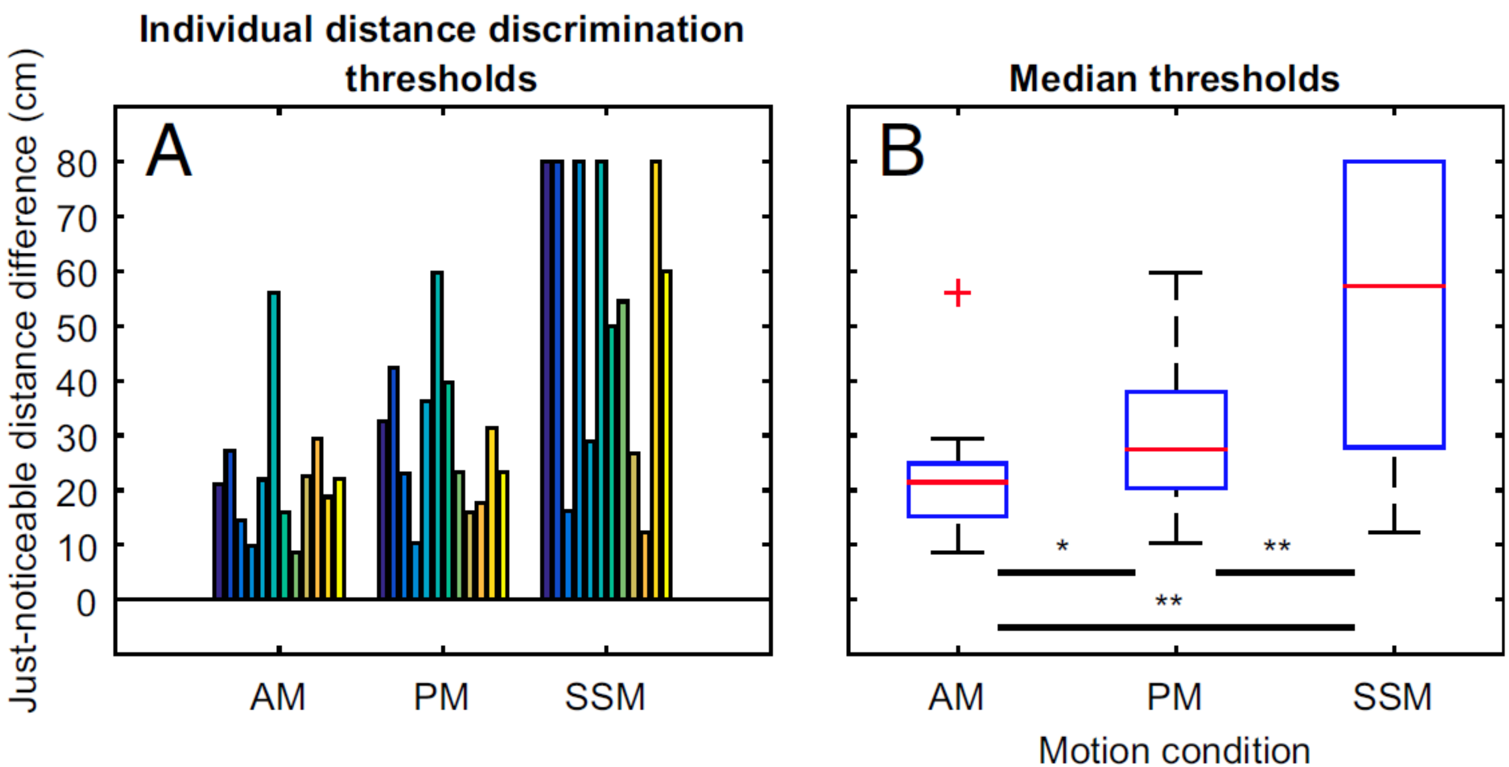
In reverberant rooms, the ratio of the sound energy in the first wave front relative to the energy reflected from the surfaces is a function of distance and allows the estimation of sound-source distance without motion (11–14). Recent theoretical work has pointed out that motion of the interaural axis (and specifically translational head motion) also allows fixing sound-source distance, through the analysis of auditory motion parallax (15). To date, however, it is unexplored to what extent auditory motion parallax may be exploited by human subjects to perceptually

Auditory Motion Parallax



Auditory Motion Parallax

- Better during active than passive



Auditory Stimuli in VR

- Physics of sound
- Physiology of hearing
- Auditory Perception
- Auditory Rendering

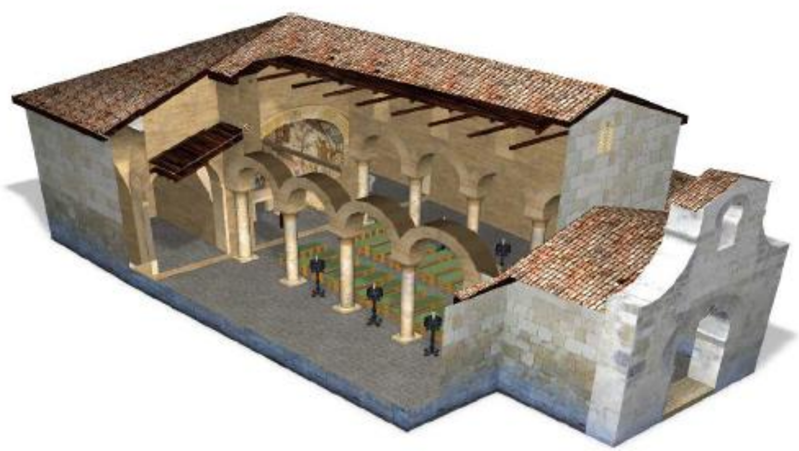
- Keep in mind comparison with light / vision

Auditory Rendering

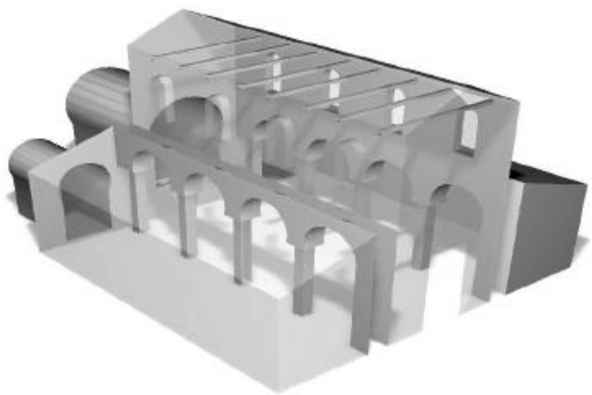
- Signal processing
- Auditory modeling
- Auralization

Auditory Modeling

- Similar to visual, but less spatial resolution required



(a)



(b)

Auditory Modeling

- Model of sound propagation in a simulated environment

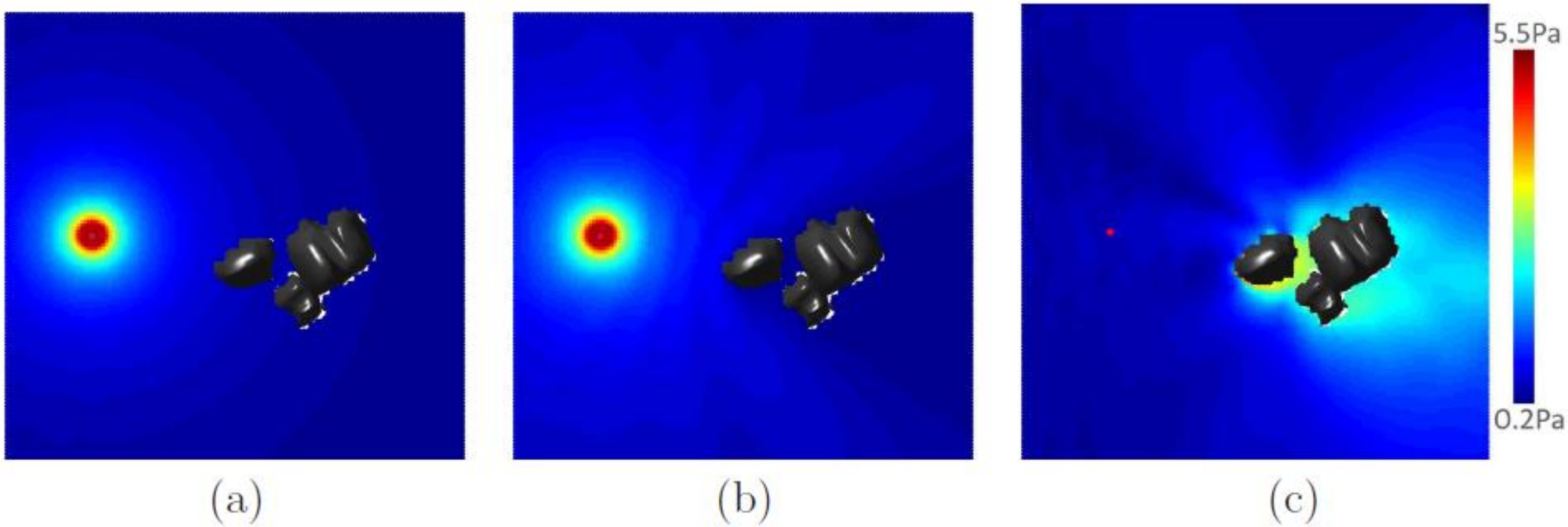


Figure 11.14: Computed results for sound propagation by numerically solving the Helmzoltz wave equation (taken from [178]): (a) The pressure magnitude before obstacle interaction is considered. (b) The pressure after taking into account scattering. (c) The scattering component, which is the pressure from (b) minus the pressure from (a).

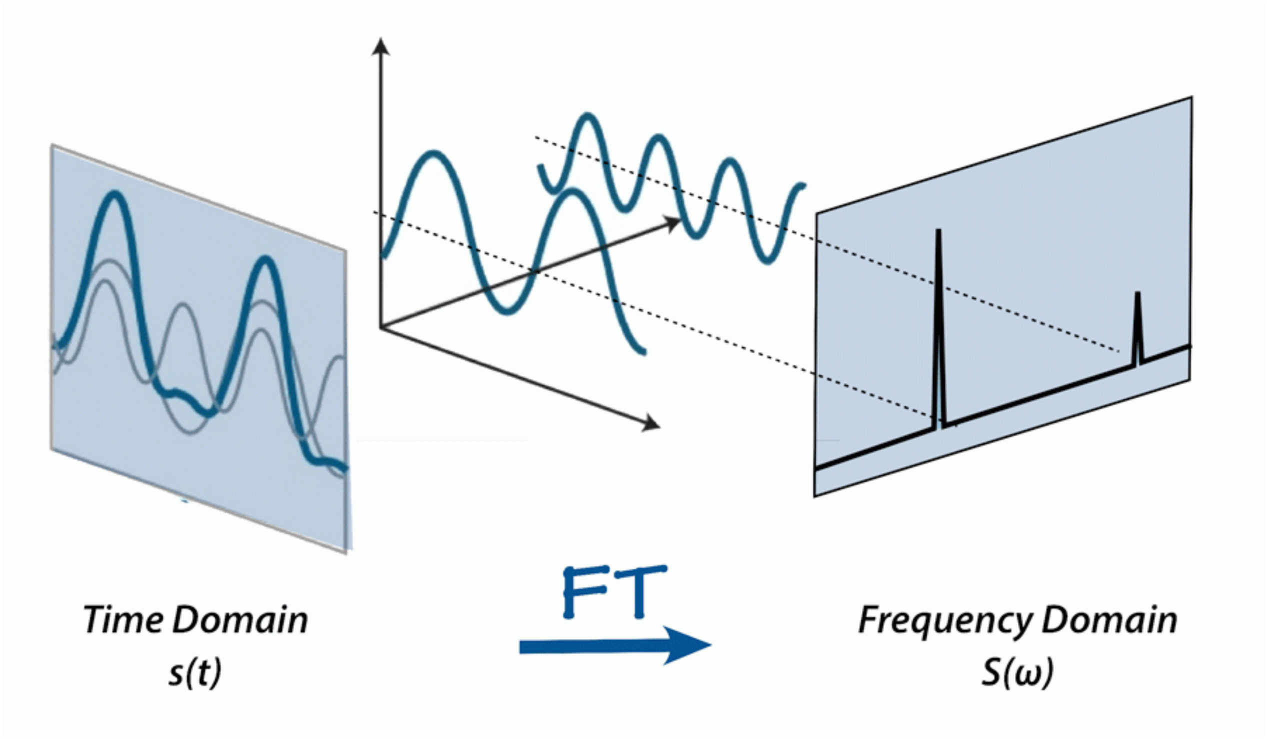
Auditory Rendering

- Audio Virtual World Generator (VWG)
- End-to-end system must be very fast!

Head Movement → Auditory VWG → Auralization

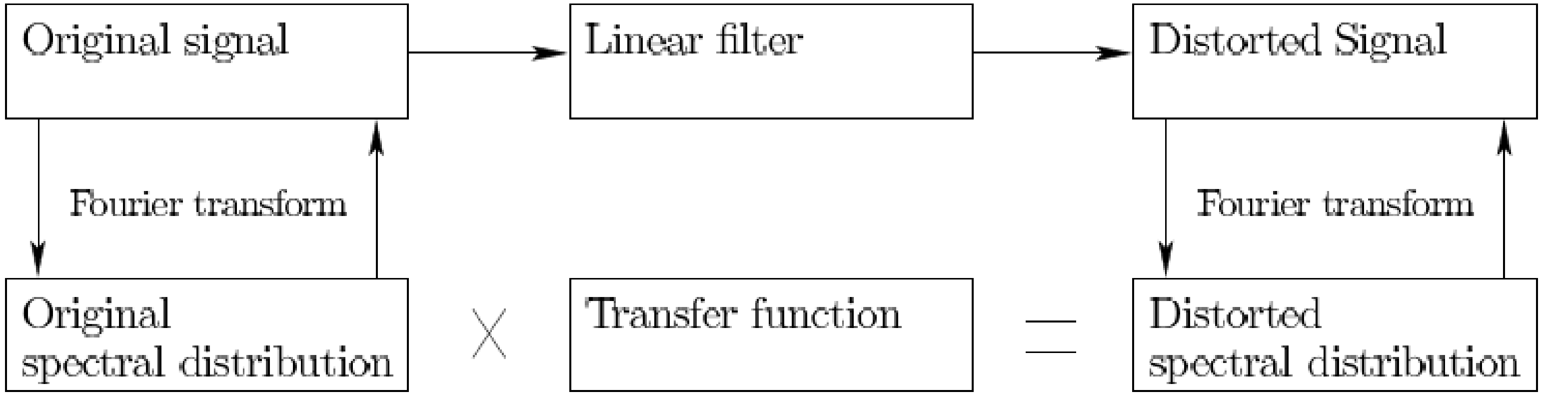
Signal Processing

- Time versus frequency domain
- Depends on Fourier transform



Signal Processing

- Computationally simpler processing



- Sampling rate needs to be twice highest frequency (e.g. CD and DVD rate over 44K)

Auditory Stimuli in VR

- Physics of sound
- Physiology of hearing
- Auditory Perception
- Auditory Rendering

- Keep in mind comparison with light / vision