Audio: Physics, Physiology and Perception

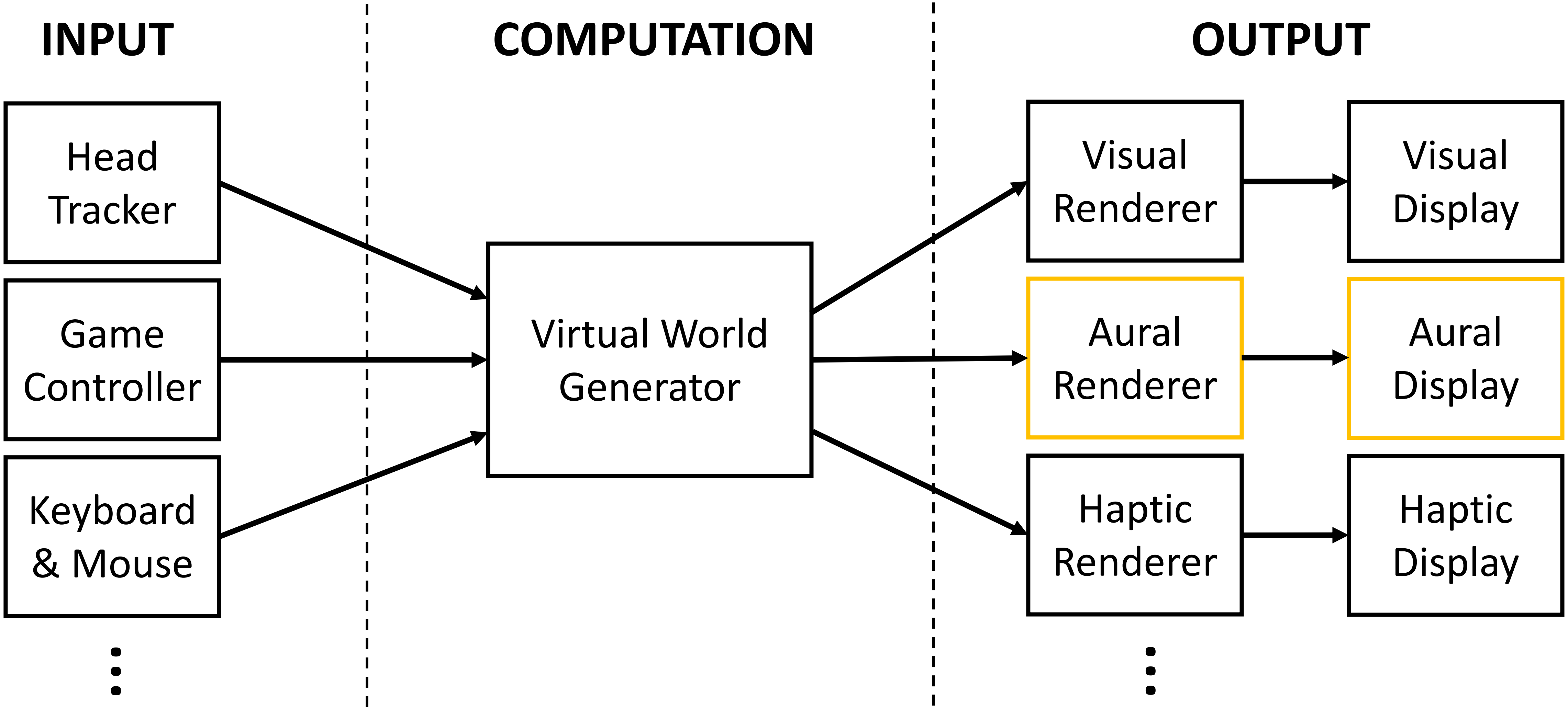
CS 6334 Virtual Reality

Professor Yapeng Tian

The University of Texas at Dallas

A lot of slides of course lectures borrowed from Professor Yu Xiang's VR class

Review of VR Systems



The Physics of Sound

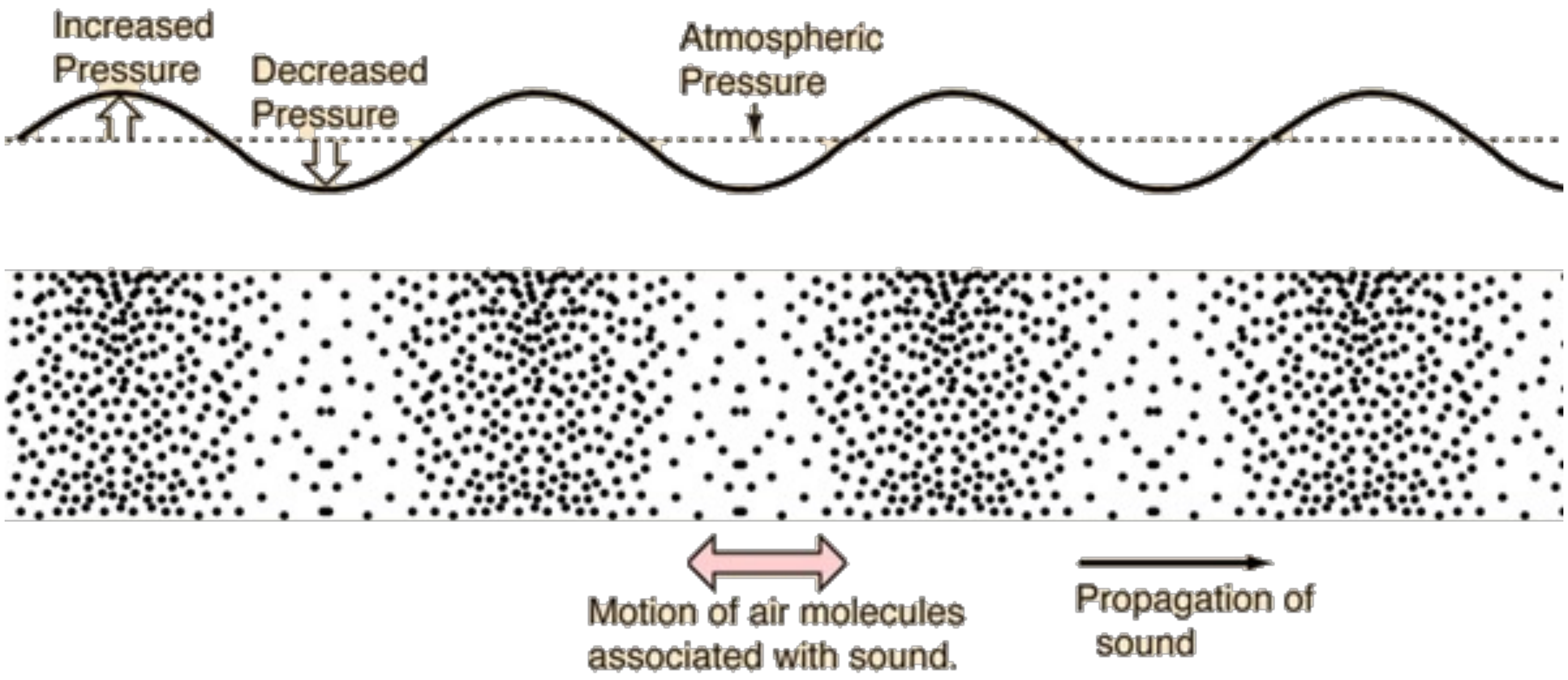
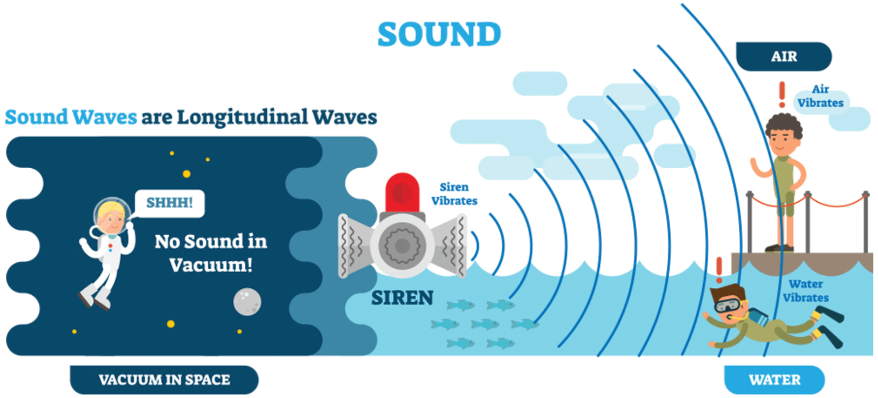


Figure by Dale Pond

The Physics of Sound



- Sound corresponds to vibration in a medium
  - Air, water, solids
  - No sound in vacuum
- Pressure variation
  - Compression extreme to a decompressed, rarefaction extreme
- Sound waves are typically produced by vibrating solid materials
  - Striking a large bell
  - Air flow: flute
  - Humans: lungs to force air through the vocal cords

The Physics of Sound

- Sound pressure level is reported in decibels (dB)

$$N_{db} = 20 * \log_{10}(p_e/p_r)$$

Pressure level of the peak compression

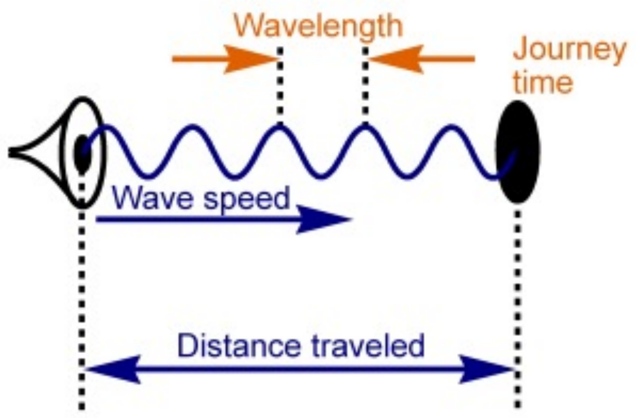
Reference pressure level
2×10^{-7} newtons / square meter

Breathing: 10 dB
Conversion in restaurants: 60 dB
Motorcycle at ft: 90 dB
Jet take-off at 25 meters: 150 dB

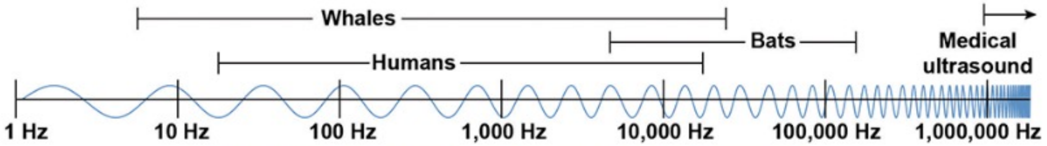
The Physics of Sound

- Attenuation
  - Sound intensity decreases by a constant factor (fixed percentage) for every unit distance from the source (exponential decay)

- Propagation speed
  - 343.2 meters per second through air at 20° C (68 ° F)
  - Light is about 874,000 times faster



The Physics of Sound

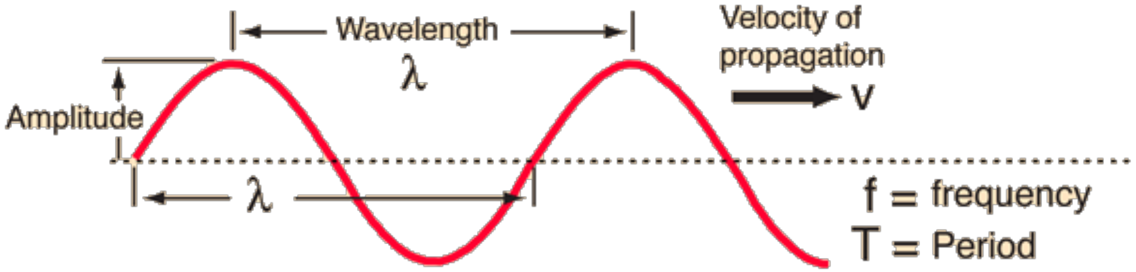


Frequencies of sound and average range of hearing

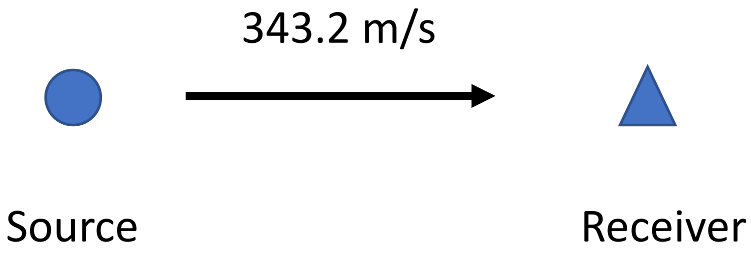
- Frequency
  - The number of compressions per second (called pitch)
  - 20 Hz to 20,000 Hz for human hearing
  - Ultrasound: above 20,000 Hz
  - Infrasound: below 20 Hz

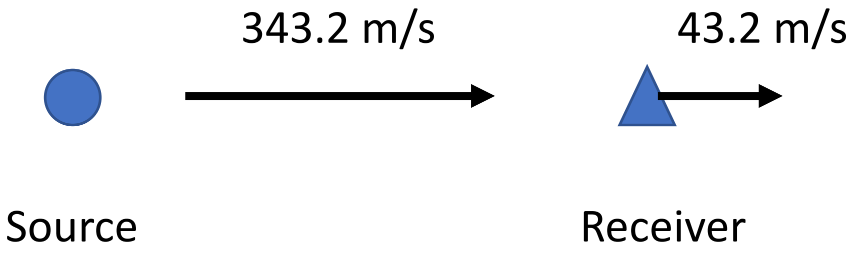
$$f = \frac{v}{\lambda}$$

- Wavelength
  - At 20 Hz: $\lambda = 343.2/20 = 17.1\mathrm{m}$
  - At 20,000 Hz: $\lambda = 17.1\mathrm{mm}$



Doppler Effect





$$f = \frac{v}{\lambda}$$

343.2 m/s

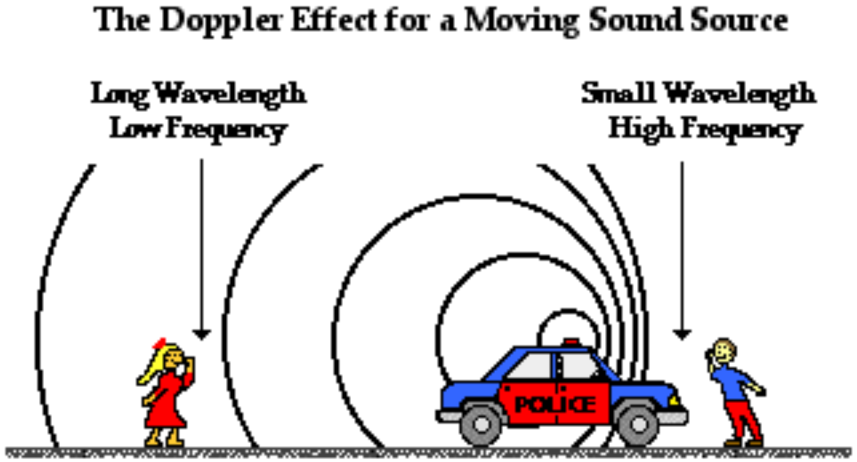
$$f = \frac{v}{\lambda}$$

300 m/s

Doppler Effect

- The received frequency shifts due to the relative motion between the source and the receiver

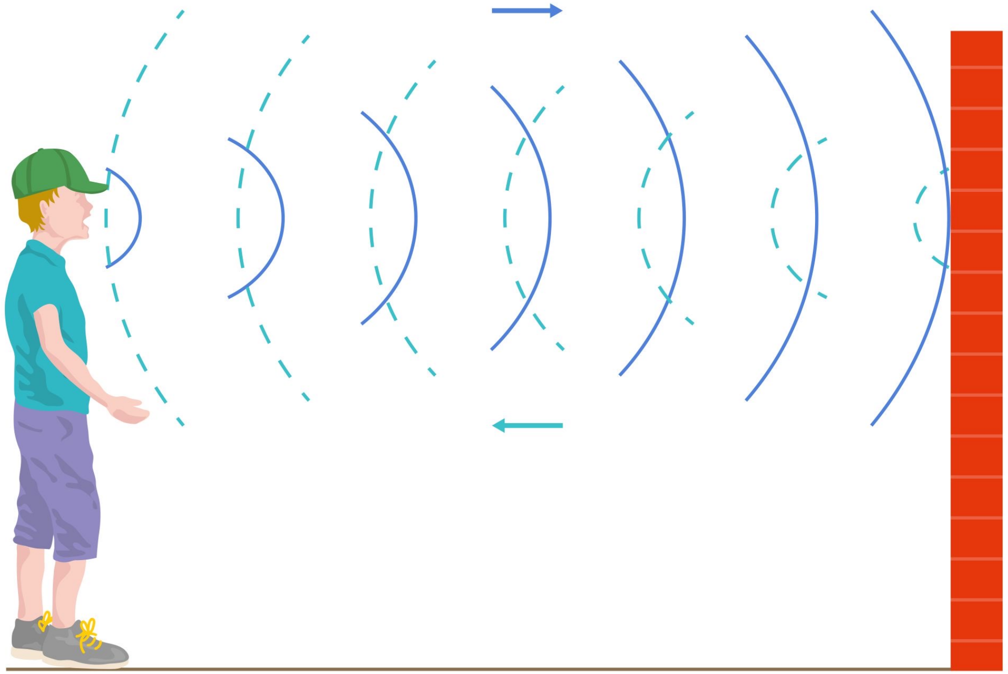
$$f_r = \left(\frac{s + v_r}{s + v_s}\right) f_s$$



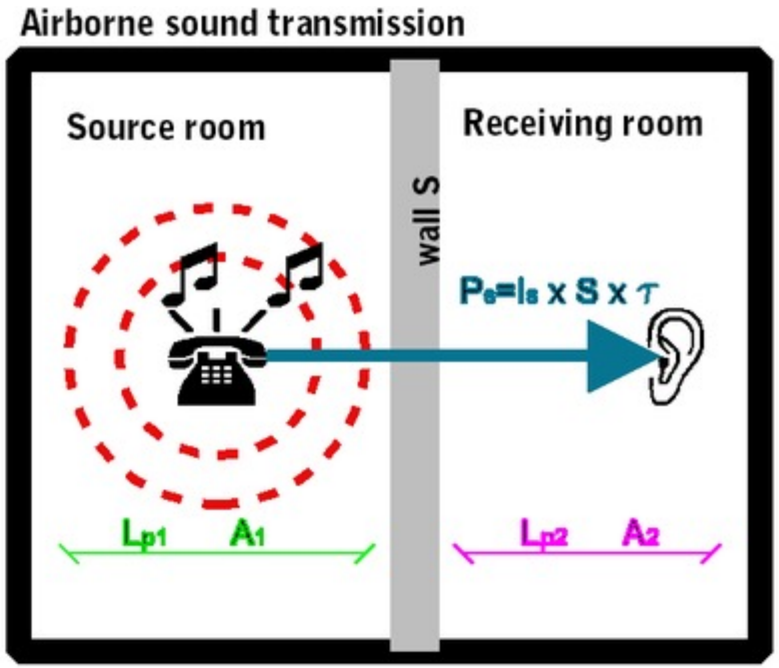
- s is the propagation speed in the medium
- vr is the velocity of the receiver
- vs is the velocity of the source

Siren seems to change pitch as a policy car passed by

Reflection and Transmission



Reflection



Transmission

Diffraction

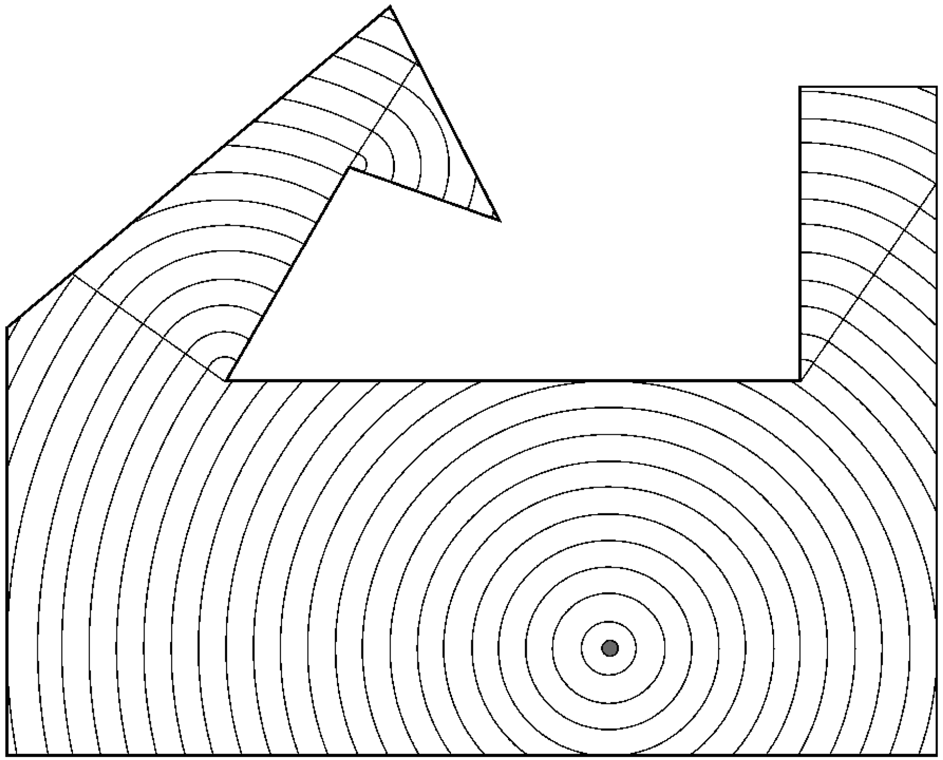
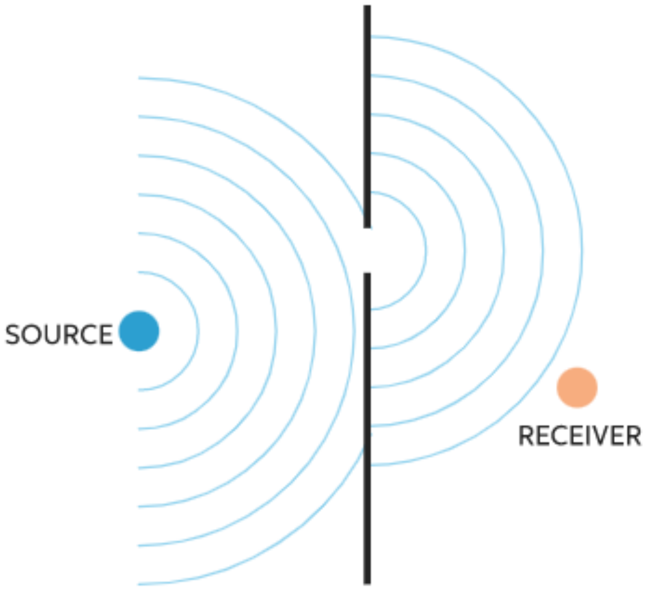


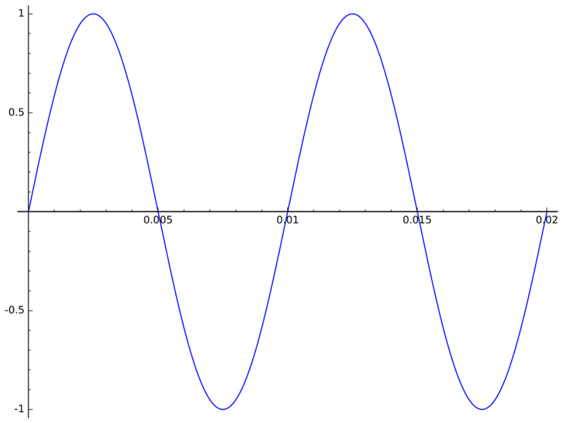
Figure 11.2: Waves can even bend around corners, due to *diffraction*. A top-down view of a room is shown. At each of the three interior corners, the propagating wavefront expands around it.



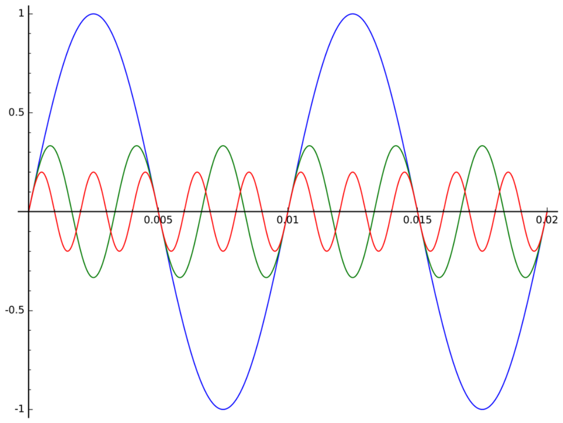
More diffraction occurs for longer wavelengths

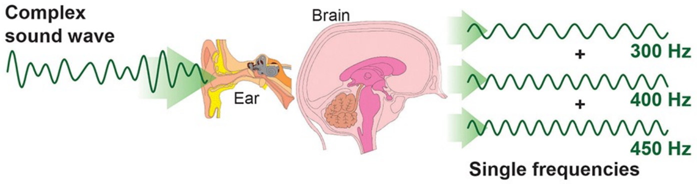
Spectral Decomposition

A pure tone (sinusoid)

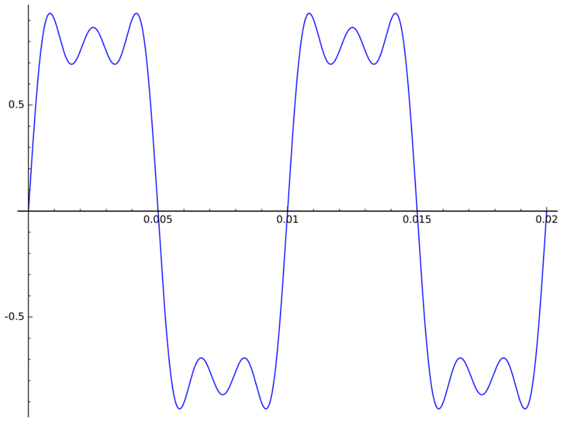


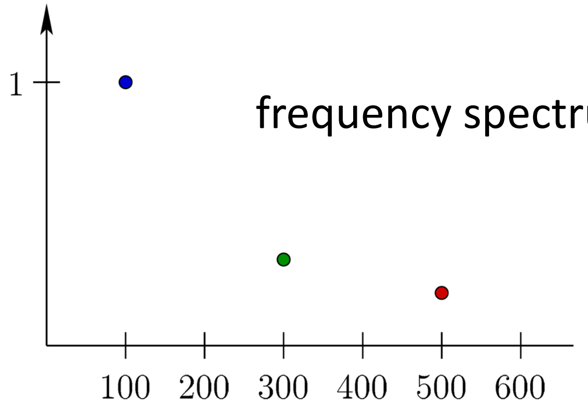
Three pure tones





Fourier analysis: any periodic function can be decomposed into sinusoids

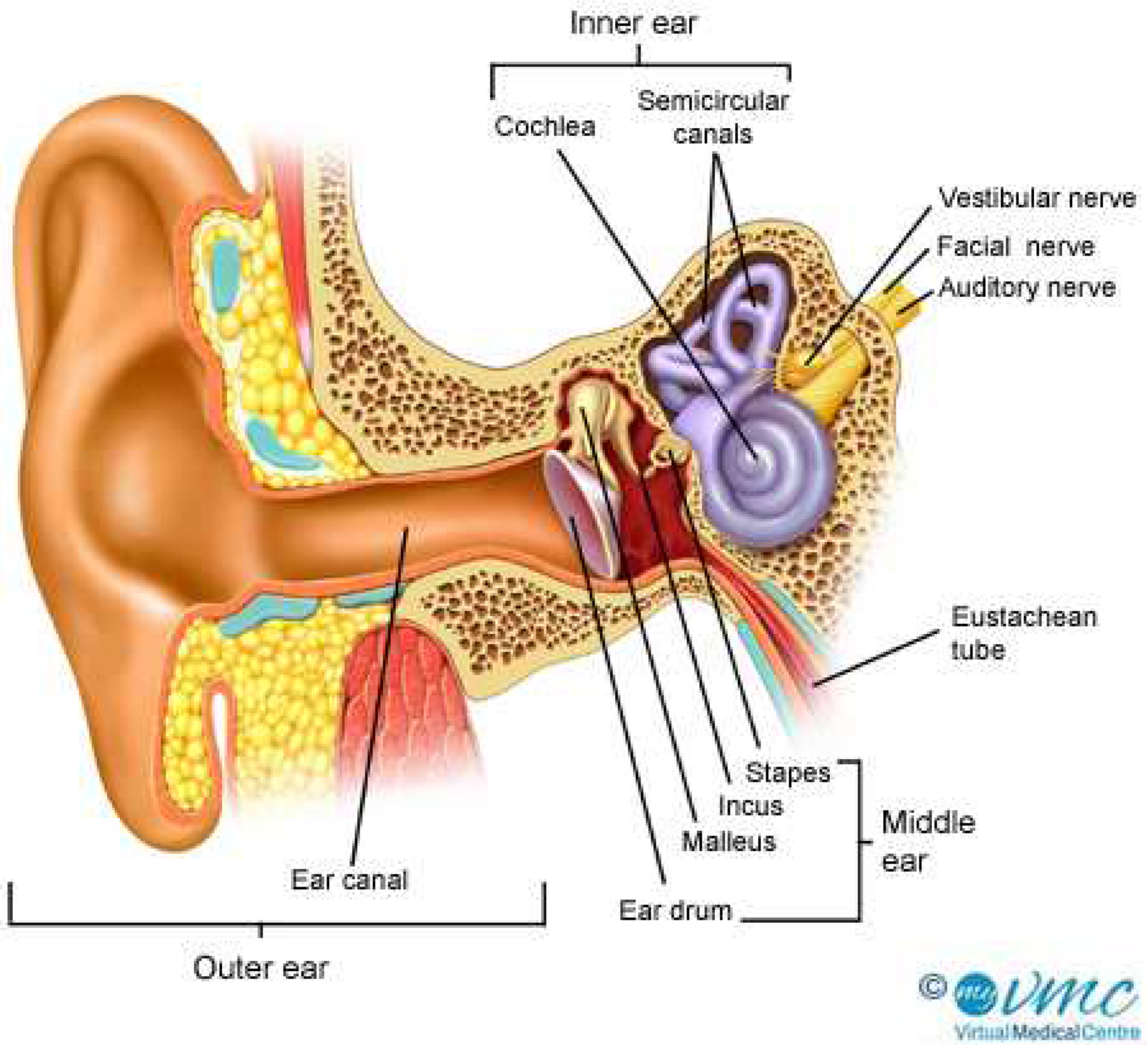




frequency spectrum

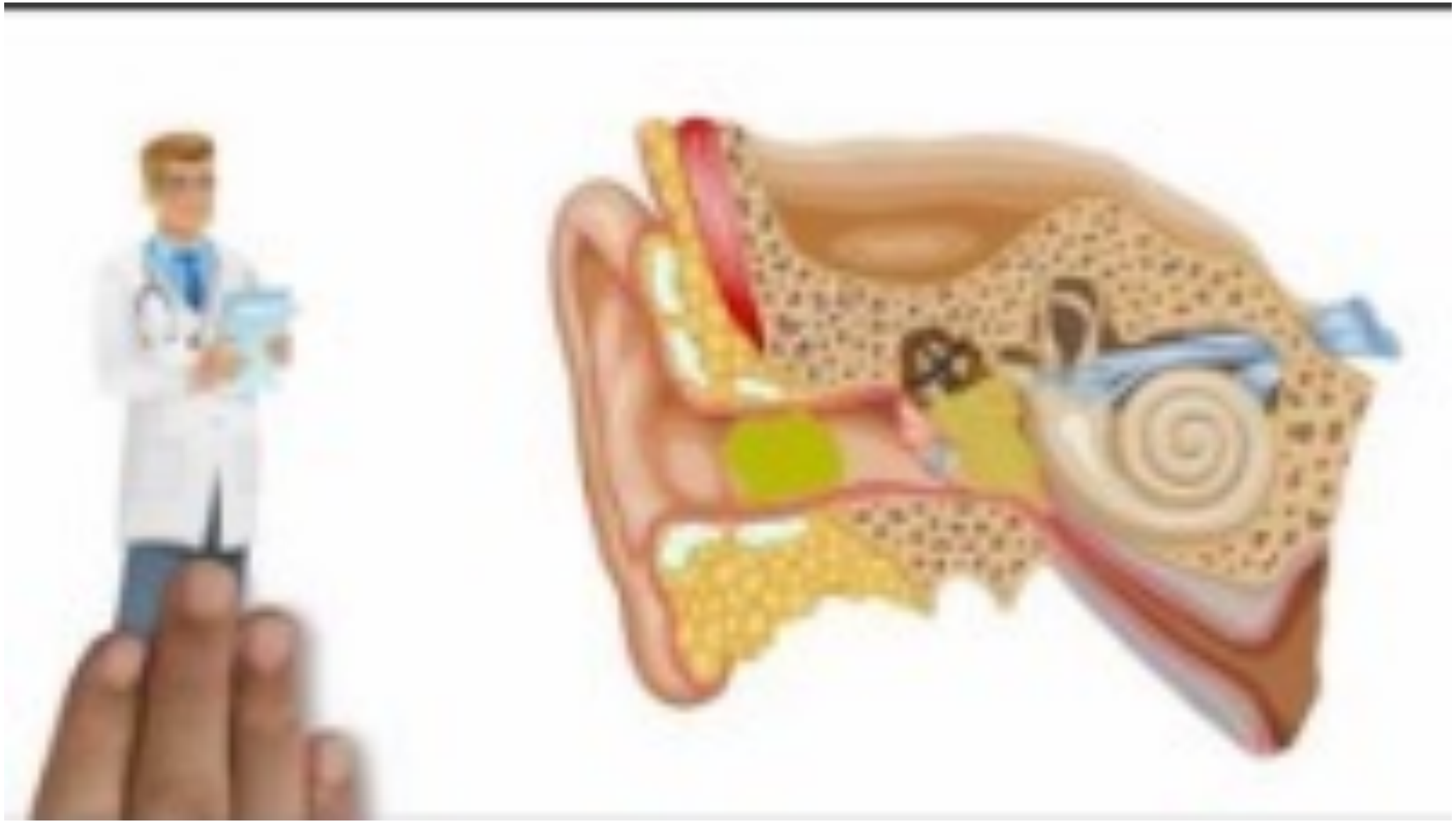
Directly adding the three pure tones

The Human Ear



- Ear drum, vibrate when receiving sound
- Middle ear (3 bones), converting vibrating air molecules in the outer ear into vibrating liquid in the inner ear
- Inner ear, the vestibular organs and the cochlea

Auditory system



https://www.youtube.com/watch?v=SU_aecxckRg

Auditory Perception

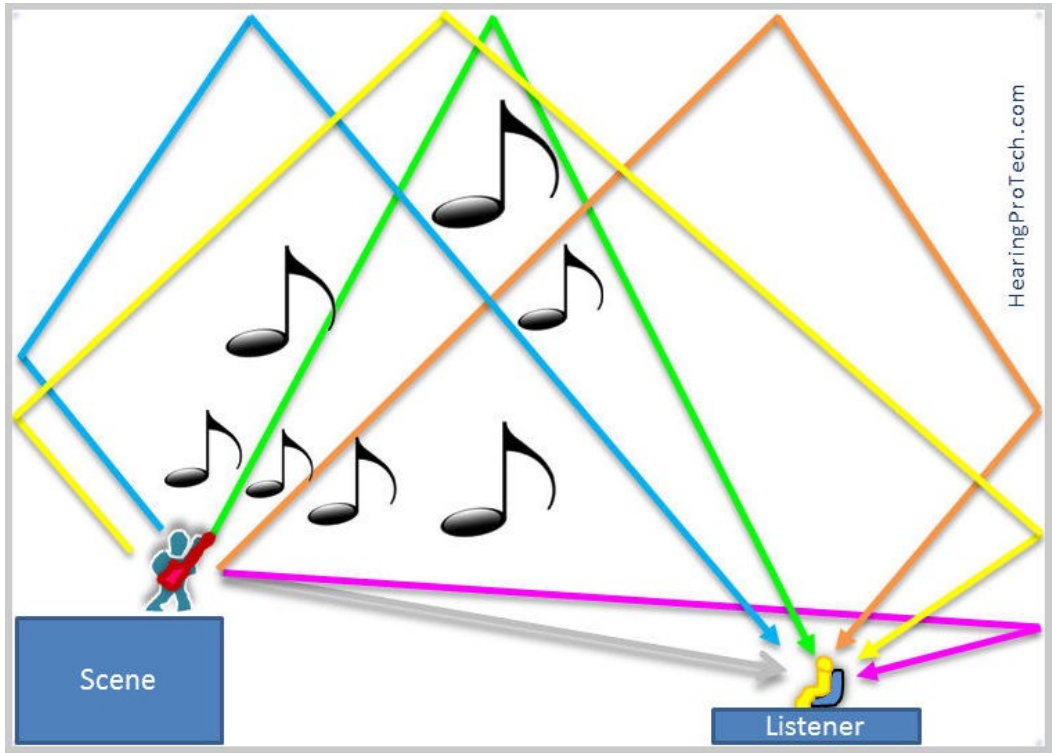
- Precedence effect
  - only one sound is perceived if two nearby identical sounds arrive at slightly different times

Echoes and reverberations

https://www.youtube.com/watch?v=LR-AYApQbNE

Auditory Perception

- Precedence effect
  - Sounds often reflect from surfaces, causing reverberation, which is the delayed arrival at the ears of many “copies” of the sound



- Rather than hearing a jumble with echoes, people perceive a single sound
- Perception based on the first arrival, which usually has the largest amplitude

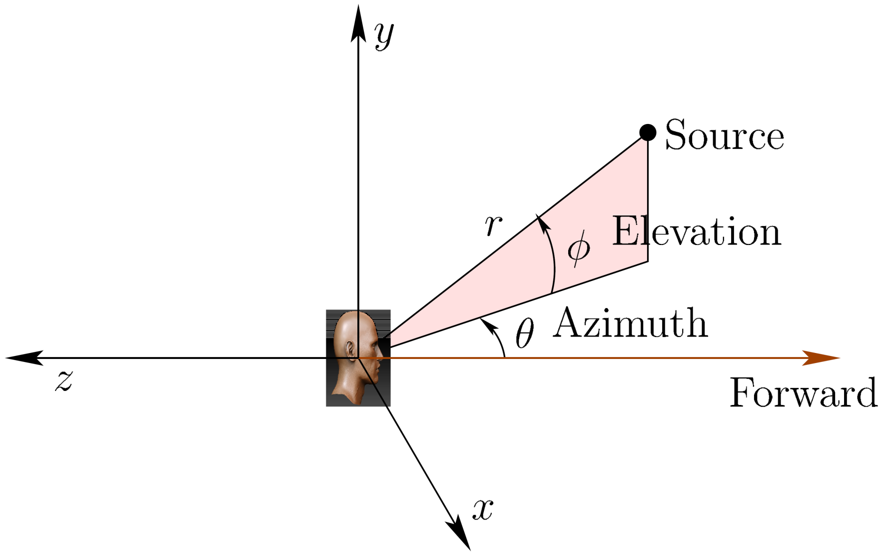
Pitch (Frequency) Perception

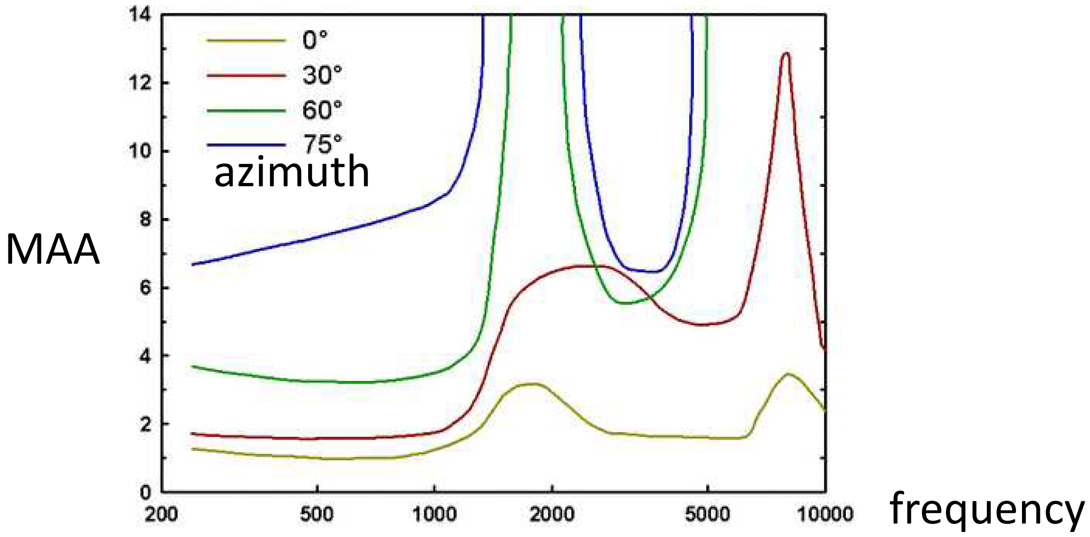
- Critical band masking
  - Block out waves that have frequencies outside of a particular range of interest

- Perception of differences in pitch
  - Just noticeable differences (JNDs)
  - < 1000Hz, JND 1Hz
  - 10,000Hz, JND 100Hz

Localization

- Estimating the location of a sound source by hearing it (crucial for VR)
- Minimum Audible Angle (MAA): minimum amount of angular variation that can be detected by a human listener



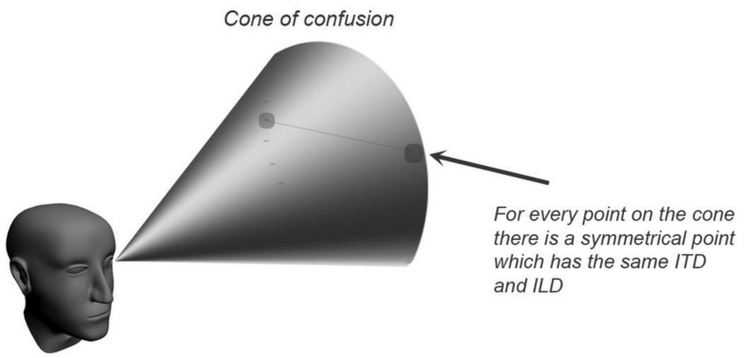


Monaural Cues for Localization

- The pinna is shaped asymmetrically so that incoming sound is distorted in a way that depends on the direction from which it arrives, especially the elevation.
- The amplitude of a sound decreases quadratically with distance.
- For distant sounds, a distortion of the frequency spectrum occurs because higher-frequency components attenuate more quickly than low-frequency components.
- The reverberations entering the ear as the sounds bounce around; this is especially strong in a room

Binaural Cues for Localization

- Interaural Level Difference (ILD), the difference in sound magnitude as heard by each ear
- Interaural Time Difference (ITD), the distance between the two ears is approximately 21.5cm, which results in different arrival times of the sound from a source



The cone of confusion is the set of locations where a point source might lie after using the ITD binaural cue.

Head Motion for Localization

- Auditory parallax, nearby audio sources change their azimuth and elevation faster than distant ones

- Integrating different cone of confusion for every head pose

- Doppler effect caused by the motion of a source relative to the receiver

Further Reading

- Section 11.1, 11.2 and 11.3, Virtual Reality, Steven LaValle