CMSC 498M: Chapter 10
Sound in Games

Source:
- Lecture notes from Web and Prof. Ramani Duraiswami at UMCP

Overview:
- Nature of sound and human hearing
- Audio design in games

Overview

- Nature of sound and human hearing
- Audio design in games

The Nature of Sound

What is Sound?:

- **Vibratory energy** caused by movement of **physical objects**
- As mechanical wave energy, it requires a **medium** such as air or water in which to propagate

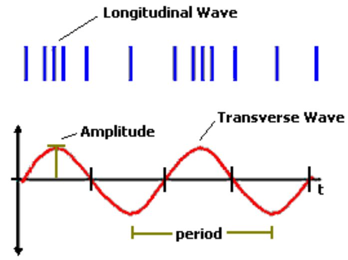
Properties:

Frequency: **Rate** of vibration

- Experience as **pitch** (high or low)

Amplitude: **Intensity** of vibration.

- Experience as **loudness**
- Measured in **decibels** (dB) (too loud/too long ⇒ hearing loss)



The Nature of Sound

Frequency:

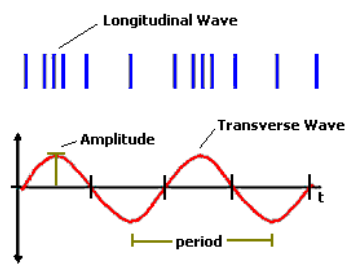
- Number of sound pulses that travel past a fixed point in a second
- Measured in **Hertz** (Hz), that is, **cycles per second**
- **Humans**: 20-20,000 Hz; **Bats**: 100,000 Hz

Speed of Sound:

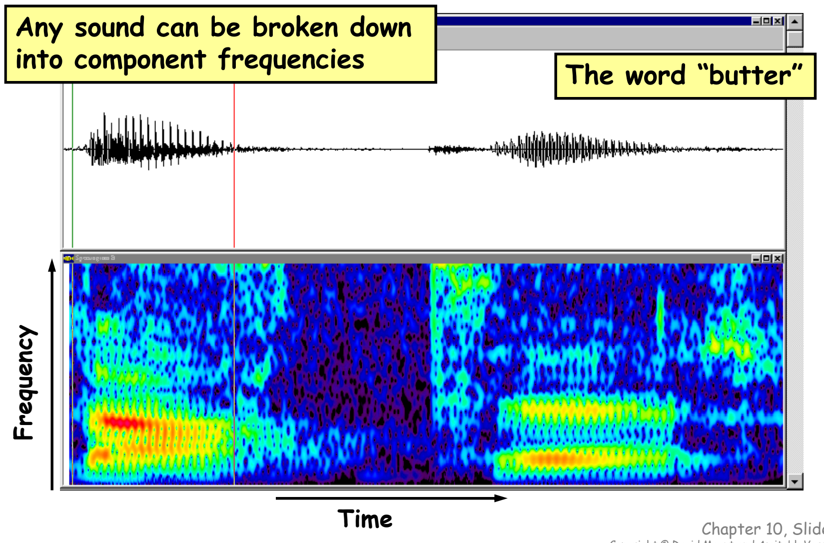
- How fast sound pulse travels
- Sound travels at **same speed** in a given medium, irrespective of pitch
- **344 meters/sec** in air

Pitch:

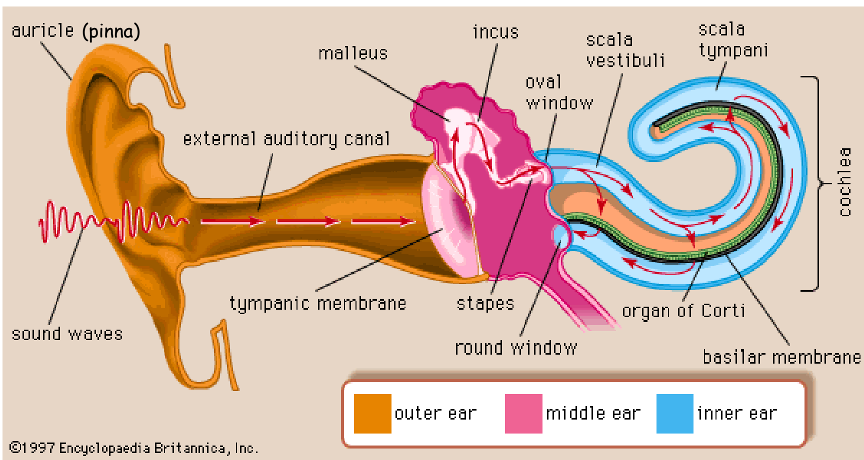
- As with color in vision, this is **perception**
- A **high frequency** sound is heard as a **high pitch**.



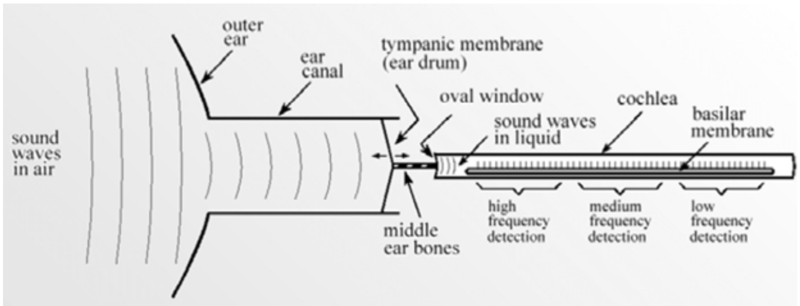
The Nature of Sound



Human Hearing



Human Hearing



- **Outer ear** collects sound waves from the environment
- **Ear drum** → **middle ear bones** → **fluid-filled cochlea**
- 12,000 nerve cells form the **cochlear nerve**
- Each nerve cell only responds to **narrow range** of audio frequencies
- The ear is a **frequency spectrum analyzer**

Volume

Volume:

- Perception of volume **logarithmic** in **amplitude**
- A **decibel** is proportional to the log of the ratio of two intensities: $10 \cdot \log_{10}(I_1/I_2)$
- A **3 decibel** (dB) increase is roughly equivalent to amplitude **doubling**
- As **absolute measure**, it's in comparison to **threshold of audibility**
  - 0 dB: **inaudible**
  - 60 dB: **normal speech**
  - 80 dB: a **shout**, city street **traffic noise**
  - 100 dB: **motocycle/snowmobile**
  - 120-130 dB: front row at a **heavy metal rock** concert (AC/DC, Led Zepplin, Deep Purple)
  - 140 dB: **gun blast**

Pitch

Pitch:
- Perception of pitch is **logarithmic** in **frequency**
- **Higher frequencies** are perceived as **higher pitches**
- Measured in **hertz** Hz (cycles per second) and **kilohertz** kHz.
- **Humans audible range**: roughly 5 to 20,000 Hz
- **Examples**:
  - 8 Hz: Lowest frequency produced by **whales**
  - 32 Hz: Lowest note on **piano**
  - 250 Hz: Typical **speaking voice**
  - 440 Hz: Baby's **cry**
  - 4096 Hz: Highest note on a **piano**
  - 200,000 Hz: Highest frequency produced by **dolphins**

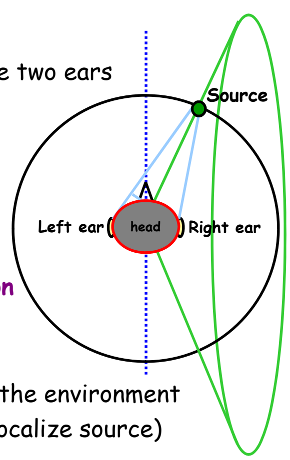
How do we perceive sound location?

Initial idea:
- Measure attributes of received sound at the two ears

Compare sound received at two ears:
- Interaural **Level** Differences (ILD)
- Interaural **Time** Differences (ITD)

Surfaces of constant Time Delay:
- $|x - x_L| - |x - x_R| = c \cdot \delta \cdot t$
- Delays same for points on **cone-of-confusion**
- Level differences are **small**

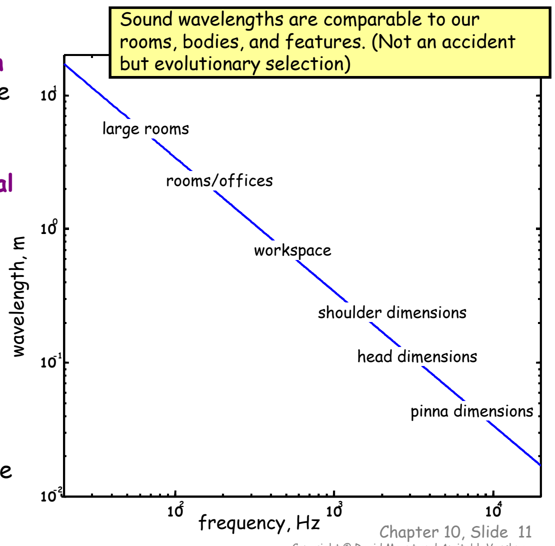
Other mechanisms necessary:
- **Scattering** of sound off our **bodies** and off the environment
- **Purposive motion** (e.g. moving left-right to localize source)



Sound and Human Spaces

Sound Wavelengths:
- Comparable to **human dimensions** and of the spaces we live in
- Wavelength λ is **inversely proportional** to frequency f

Cases: a = object size
- $\lambda \gg a$: the wave is **not affected** by object
- $\lambda \sim a$: scattered wave is complex and **diffraction effects are important**
- $\lambda \ll a$: wave behaves like a **ray**

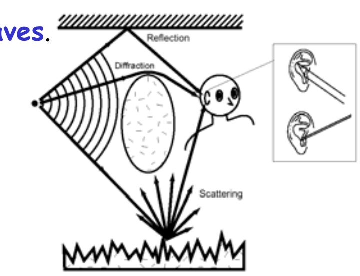


Modeling Sound Scattering

Interactions change received sound waves.

Scattering of body and ears:
- **Bodies** ~ 50 cm
- **Heads** ~ 25 cm
- **Ears** ~ 4 cm
- Not much multiple scattering

Scattering off surroundings:
- **Rooms** ~ 2m – 10m
- More multiple scattering
- Larger sizes → lower frequencies



Pinna (Outer Ear) Shape and Hearing

Pinna (or **Auricle**):

- The **outer ear** with its complex **contours**

Low frequency sounds:

- Pinna behaves similarly to a **reflector dish**, thus **amplifying** the sound

High frequency sounds and Pinna Notch:

- Some sounds enter ear **directly** while others **reflect** off the contours of the pinna
- **Reflections** enter the ear canal at a very slight **delay**
- This delay results in **phase cancellation**, where the frequency component whose wave period is twice the delay period is virtually **eliminated**, called **pinna notch**
- This allows us to perceive **subtle variations** in the **locations** of sounds
- Each ear is different, so the effect **varies** for each **individual**

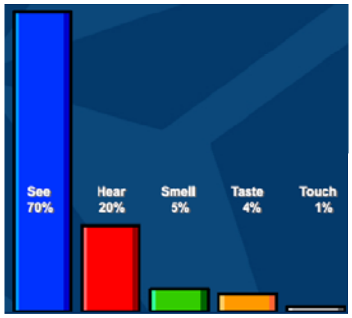
Overview

- Nature of sound and human hearing
- Audio design in games

Sound in Games

Role of Sound in Games:

- "[Sound] serves the **story**, creates a **mood**, … and can be the key to bringing the **visuals to life**" [LoBrutto,1994]
- **20% of perception is acoustic** [Dobbler et. al. 2002]



Sound Effects:

- Indicates **what is going** on around the player
- May alert player to **significant events**
- May include **background sounds** and **musical score**

Narration:

- **Reinforces action**
- Explains **non-obvious events** and story

Sound Effects

Background (or ambient) sound effects:

- Do **not** explicitly synchronize with the picture—**indicate setting**
- **Examples**: Wind, rain, traffic, crowd noises (called "**walla**")

Hard sound effects:

- **Common sounds** that appear **on screen**
- **Examples**: Door slams, weapons firing, and cars driving by

Foley sound effects:

- Sounds that **synchronize precisely** with action on screen
- Require the **expertise of a Foley artist** to record properly
- **Examples**: Footsteps, punching, rustling of cloth

Design sound effects and background score:

- Sounds that **do not occur in nature** or impossible to record
- Commonly used in **science fiction and fantasy** genres
- Use of **music** to create an emotional mood

Foley Artists

Ever wondered?
- Why footsteps sound louder in movies?
- Why they echo so much?
- Why the clanging of swords is so ringing and metallic?

Answer:
- **Foley artists.**
- Named after Jack Foley, a famous 1930's sound engineer.

Web sites:
- http://www.marblehead.net/foley/index.html
- http://www2.digidesign.com/digizine/-dz_main.cfm?edition_id=142&navid=1053

Foley Artists

Effect:	How it might be made:
– Galloping horses	Banging empty coconut shells together
– Kissing	Kissing back of hand
– Punching someone	Thumping watermelons
– High heels	High heels on wooden platform
– Bone-breaking blow	Breaking celery or bamboo or twisting a head of lettuce
– Footsteps in snow	Squeezing a box of corn starch
– Star Trek sliding doors	Pulling a piece of paper from an envelope
– Star Wars sliding doors	Flare gun plus sneakers squeak
– Bird flapping its wings	Flapping a pair of gloves
– Grass or leaves crunching	Balling up audio tape

Spoken Dialog

Two approaches:

- Synthetic **text to speech** (analogous to procedural animation)
- **Sound capture** (analogous to motion capture)

Text to speech:

- Sound is **artificially synthesized**
- May sound **artificial** or **computer-like**
- **Significant progress** over the last decade in generating natural sounding dialog from text
- It is still **hard to convey emotion** and patterns of intonation realistically

Examples:

- http://www.research.att.com/~ttsweb/tts/demo.php
- http://www.neospeech.com/

Spoken Dialog

Sound Capture:

- Hire **actors** to read a script
- Approach that is mostly followed in **expensive games** and **animated movies**
- **Good vocals** can make up for **weak animation**
- Pixar and Disney movies feature **top talent** for voices

Summary

Summary:
- Nature of sound and human hearing
- Audio design in games