GCT535: Sound Technology for Multimedia

# Fundamentals of Acoustics: Human Hearing

KAIST Graduate School of Culture Technology

**Juhan Nam**

# Introduction

- Human perception of sound
  - How sound is received and transformed in human ears
  - How perceptual attributes of sound is correlated with physical attributes of sound

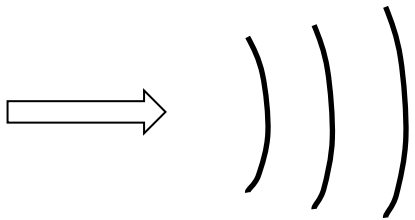
**Physical perspective (waves)**

**Psychological perspective**

**Generation**

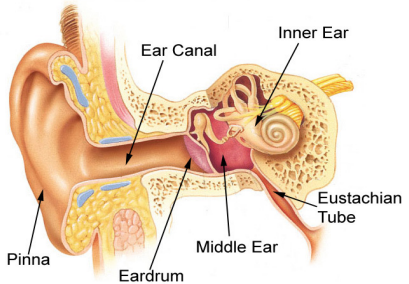
Vibration on an object

**Propagation**



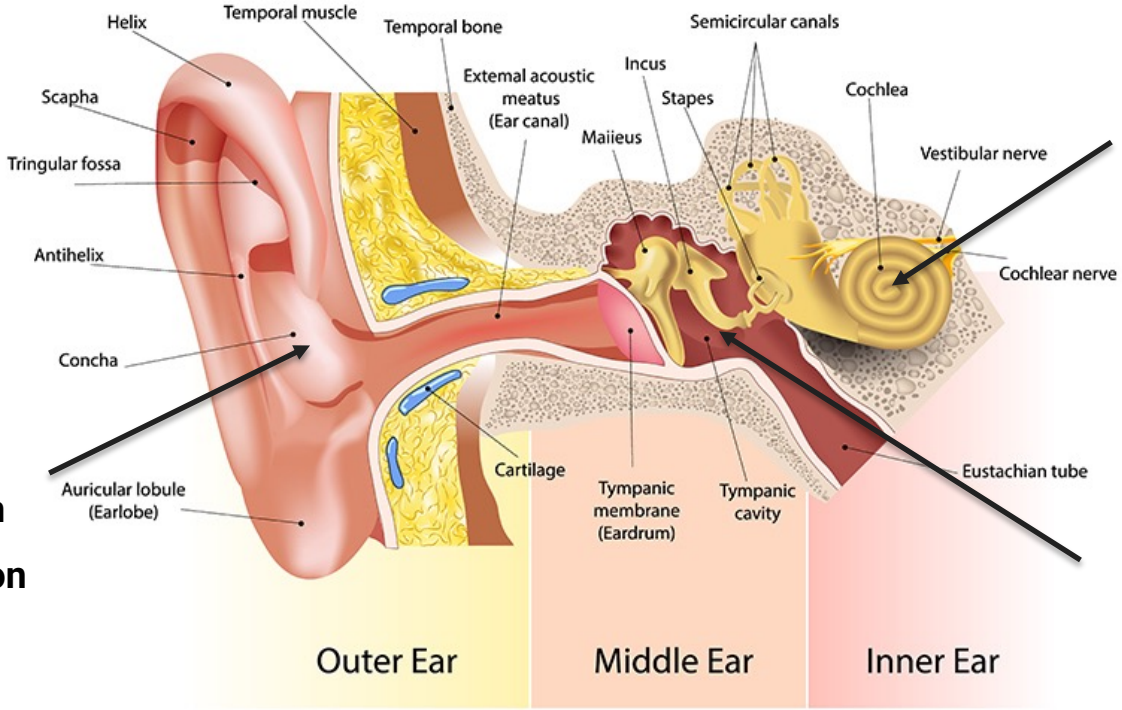
Traveling in air

**Perception**




Sensation of the air vibration through ears

# Human Ear

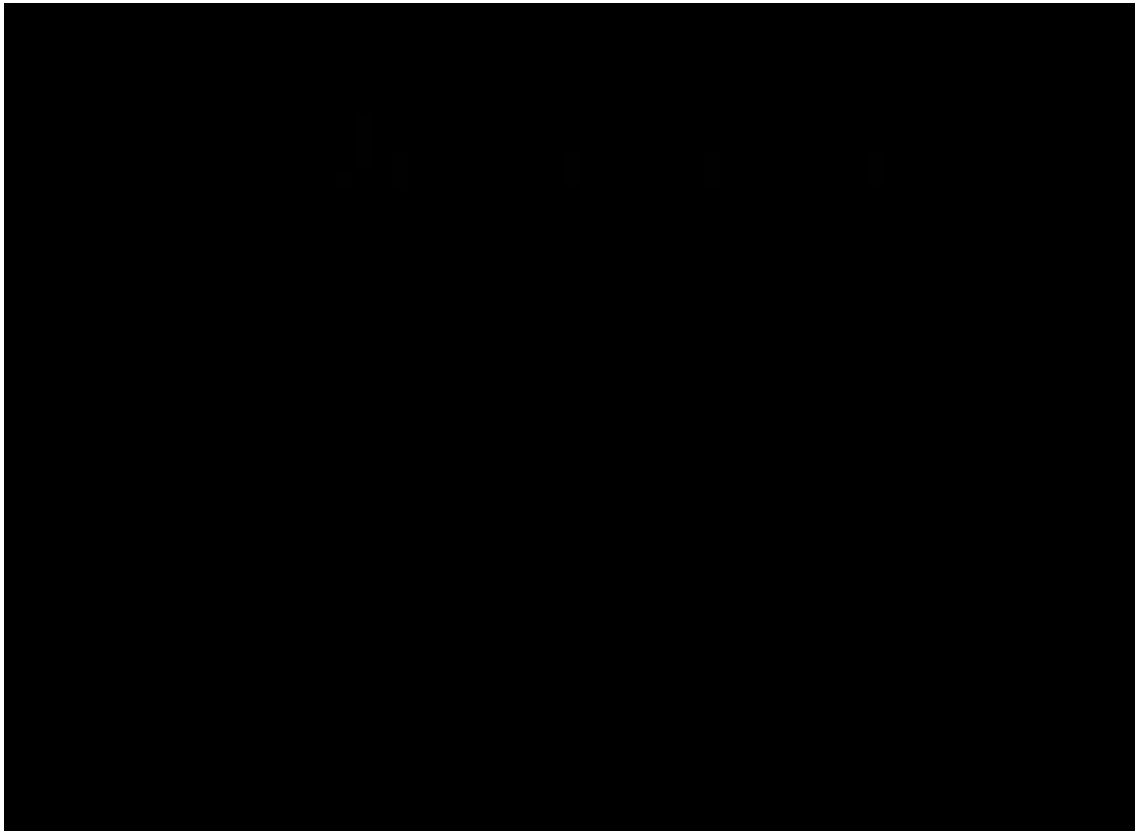



- ✓ **Protection**
- ✓ **Localization**
- ✓ **Amplification**

- ✓ **Band-pass Filtering**
- ✓ **Electric nerve firing**

- ✓ **Impedance Matching**
- ✓ **Amplification**

# Auditory Transduction in Human Ears



http://www.youtube.com/watch?v=PeTriGTENoc

# Auditory Transduction in Human Ears

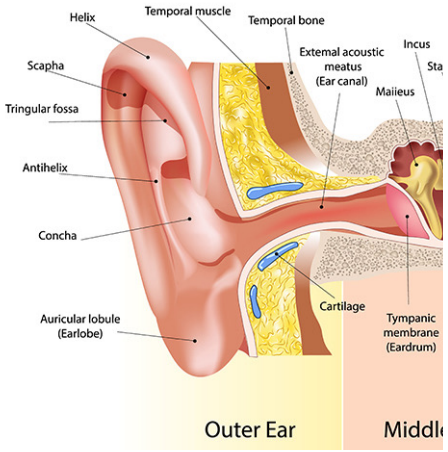
- Multiple steps of transduction
  - Air vibration (outer ear) → Mechanical vibration (middle ear) → Fluid vibration (inner) → Nerve firing (brain)

- High sensitivity by the amplification
  - Outer ear and middle ear

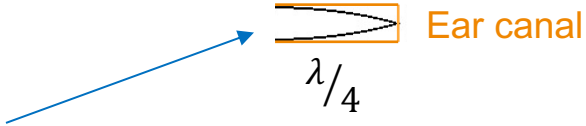
- Tonotopic organization
  - Inner ear can be regarded as a bandpass filterbank

# Outer Ear

- Pinnae
  - Collect and localize sounds:
    http://www.douglas-self.com/MUSEUM/COMMS/ear/ear.htm
  - **Localization**: related to recognize the direction of sound
    - head-related transfer function (HRTF)



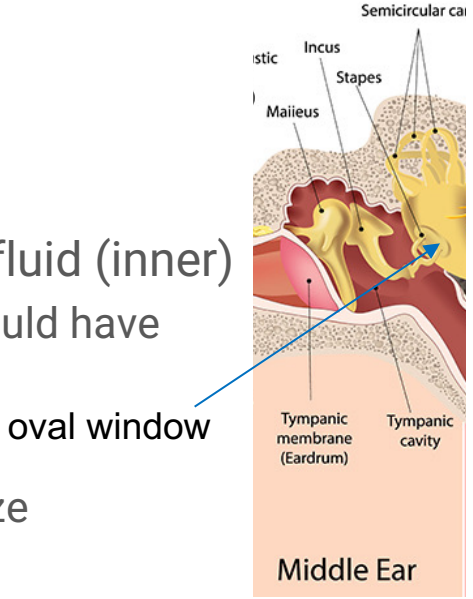

- Ear canal
  - Protect ear drums
  - Quarter-wave resonance:  boost the vibration around 3kHz by 15-20 dB




- Ear drum
  - Membrane that transduces air vibration to mechanical vibration
  - Malleus (hammer) is attached to it

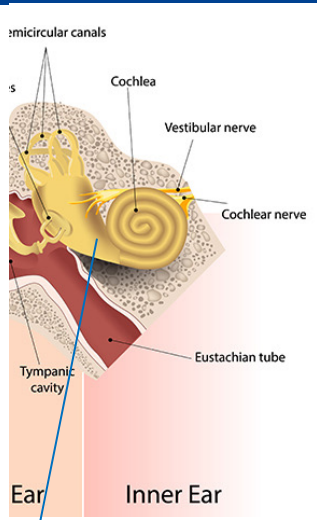
# Middle Ear

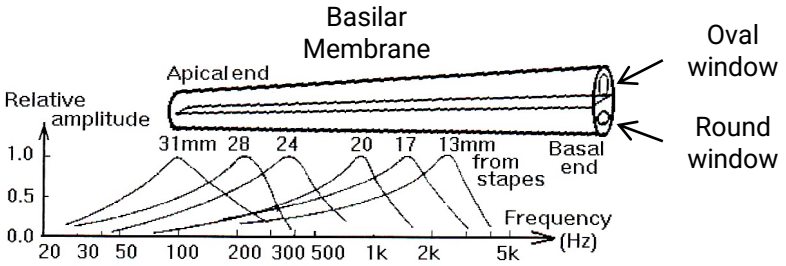
- Ossicles
  - malleus (hammer), incus (anvil) and stapes(stirrup)
  - The smallest bones in human body
  - **Impedance matching** between air pressure (outer) and fluid (inner)
    - Without ossicles, only about 1/30 of the sound energy would have been transferred to inner ear
  - **Amplification:** the three bones work as a lever
    - The vibration from ear drum is transferred to a smaller size membrane (oval window) in the inner ear
- Muscles
  - Reduce the sound transmission in response to loud sounds

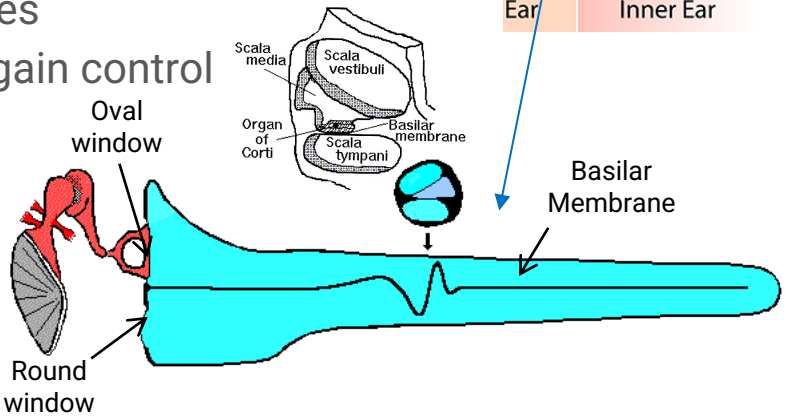


# Inner Ears

- Cochlea: transduces fluid vibration to nerve firing
  - Basilar membrane
    - Vibrate at a different position depending on the frequency
    - Similar to band-pass filters
    - The bandwidth becomes wider as the frequency goes up
  - Organ of Corti
    - One row of inner hair-cell fire neural spikes
    - Three rows of outer hair-cell: automatic gain control




Source: http://acousticslab.org/psychoacoustics/PMFiles/Module03a.htm

# Sound Transformation in Ear

- Outer ear
  - HRTF: sound location-dependent filters
  - Ear canal: resonance around 3kHz by 15-20 dB
- Middle ear
  - Amplification (three bones)
- Inner ear
  - Basilar membrane: sub-band decomposition
  - Inner hair-cell: non-linear processing

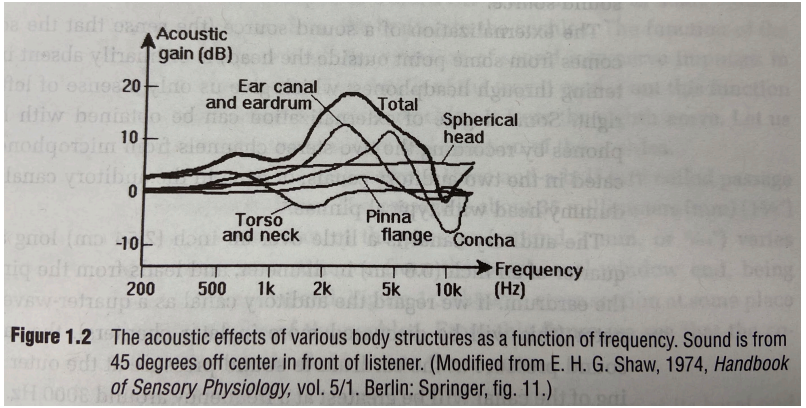
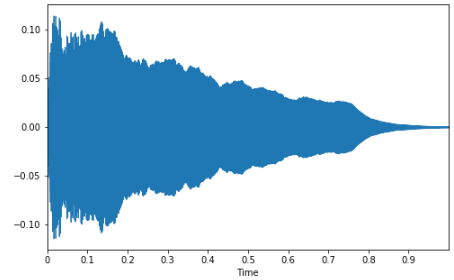
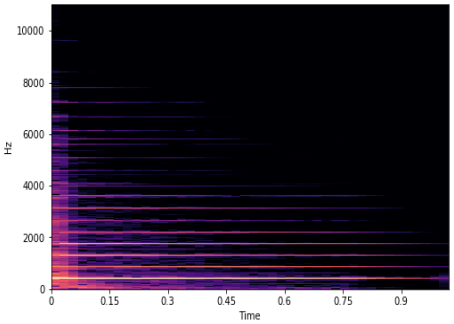



**Figure 1.2** The acoustic effects of various body structures as a function of frequency. Sound is from 45 degrees off center in front of listener. (Modified from E. H. G. Shaw, 1974, *Handbook of Sensory Physiology,* vol. 5/1. Berlin: Springer, fig. 11.)

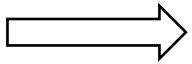
Source: P. R. Cook (Editor) Music, Cognition, and Computerized Sound (book)

# Perception of Sound

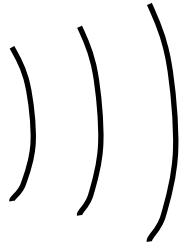


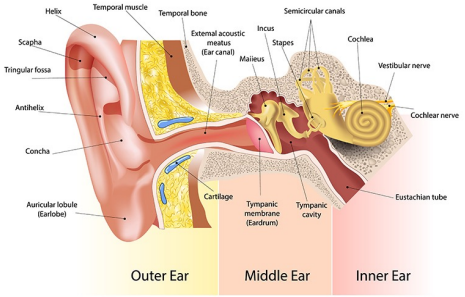




- Amplitude
- Frequency
- Envelope
- Duration
- Spectrum



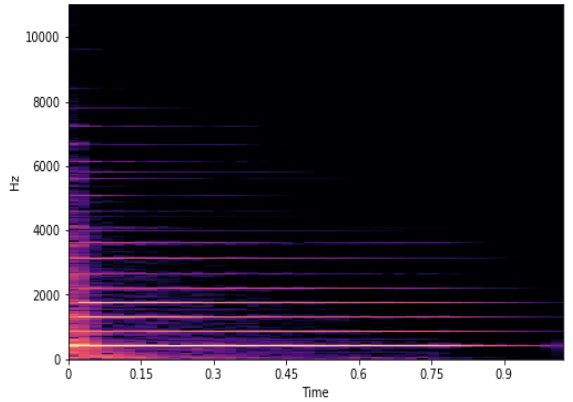
- **Loudness**
- **Pitch**
- **Timbre**





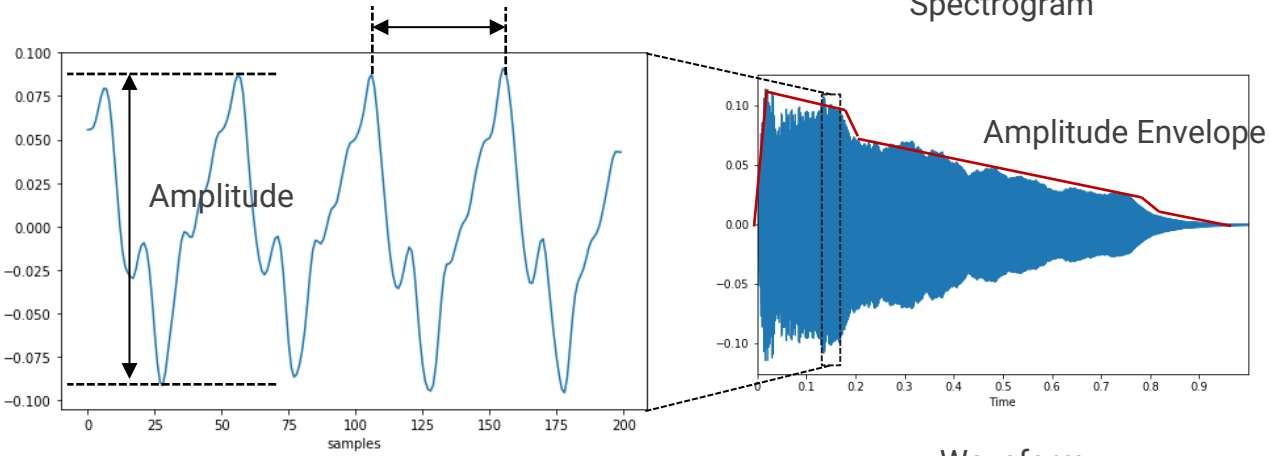
# Physical Attributes of Sound

- Amplitude
- Frequency (Hz)
- Envelope
- Duration
- Spectrum




Spectrogram

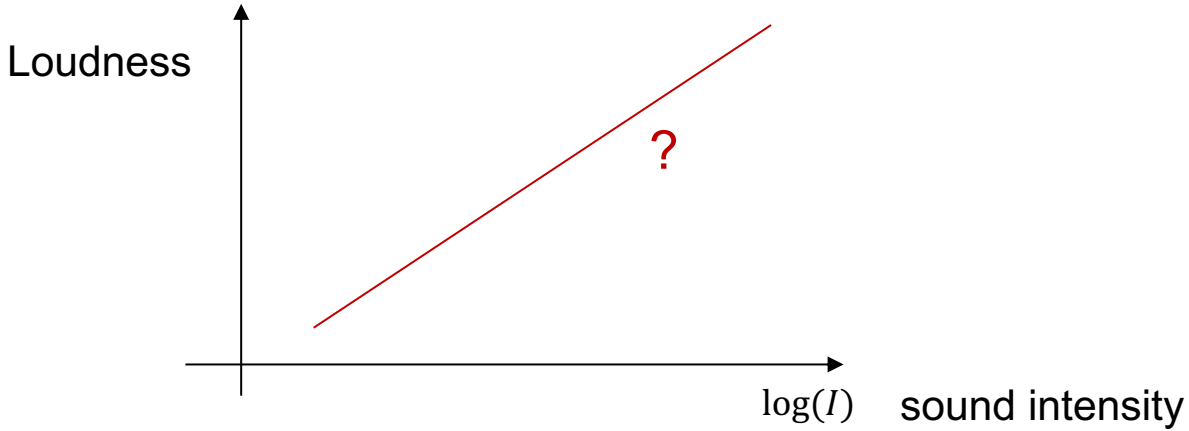
Period = 1/(fundamental frequency)




Waveform

# Loudness Perception

- Subjective perception of sound intensity (the squared amplitude)
- Weber's law: $\Delta I \propto I$
  - Just Noticeable Difference (JND) of intensity ($\Delta I$) is proportional to the intensity ($I$) at the moment
  - A log function satisfies this rule: but we needs some adjustment



# Loudness Perception

- Power law (Steven, 1936)
  - A simple approximation to loudness:

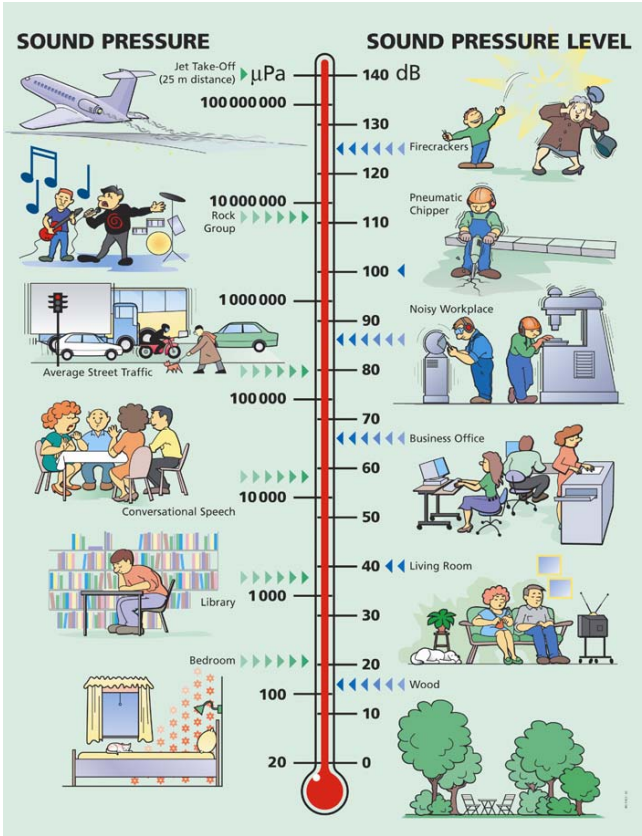
$$L = I^{0.3} \longrightarrow \log(L) = 0.3\log(I)$$

  - A doubling of perceived loudness corresponds to a factor of 10 increase in intensity ($10^{0.3} = 2$)
  - This perceptual unit of loudness is called “**Sone**”

# Loudness Perception

- Sound Pressure Level (SPL)
  - Physically measured sound intensity
  - Decibel scale: $20\log_{10}(P/P_0)$
  - The softest audible level: $P_0 = 2\times10^{-5}\ N/m^2$

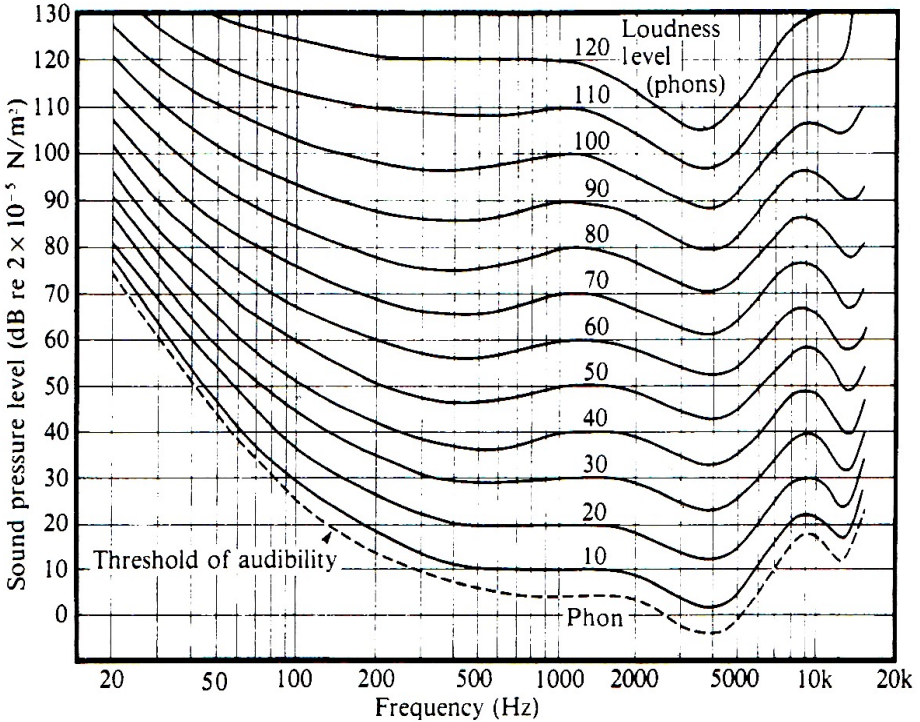

SPL meter




Source: http://www.audioholics.com/home-theater-connection/basic-home-theater-setup-guide/splmeter500x332.jpg/image_view_fullscreen
Source: http://www.bksv.com/doc/sve/7.%20Audiology%20(0210).pdf

# Equal-Loudness Curves

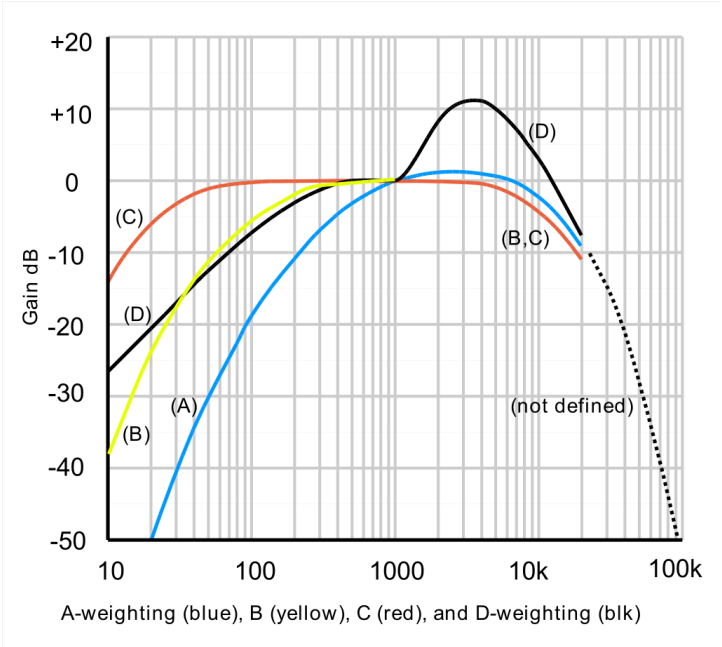
- Loudness depends on frequency (Fletcher and Munson, 1933)
  - 1kH is used as a reference
  - Most sensitive to 2-5 kHz tones
    - Resonance in the ear canal
  - See the threshold of audibility
    - Dynamic ranges are different




Do your own test: http://newt.phys.unsw.edu.au/jw/hearing.html

# Perceptual Weighting

- Weighting spectrum based on perceptual sensitivity of loudness
  - Inverse of the equal-loudness curve
  - A-weighting is a common choice



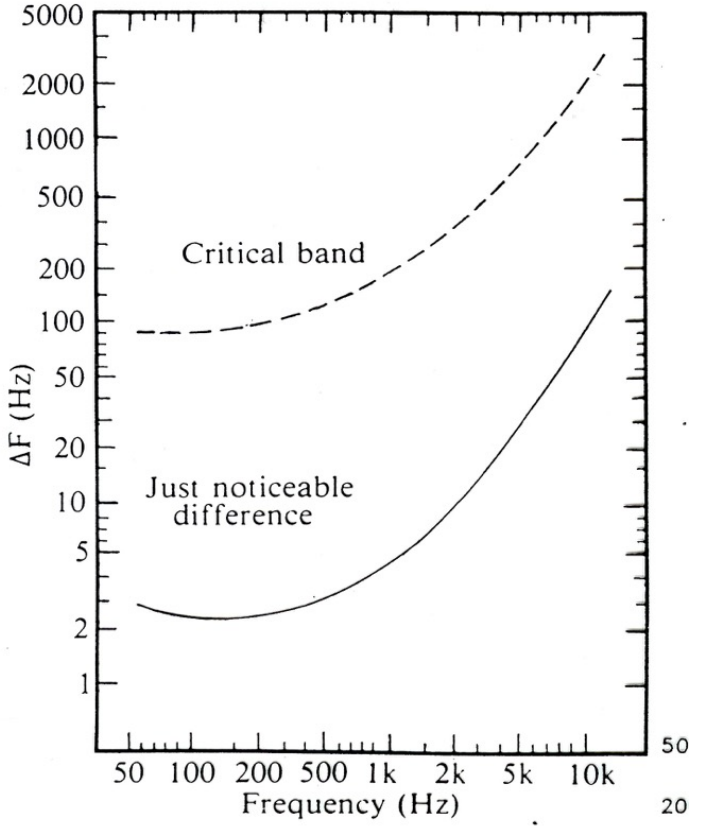
A-weighting (blue), B (yellow), C (red), and D-weighting (blk)

# Pitch Perception

- Perceptual correlate of fundamental frequency (F0)
  - Auditory attribute of sound according to which sounds can be ordered on a scale from low and high (ANSI, 1994)
  - Measured by subjective test
- Pitch is mainly determined by fundamental frequency (F0).
  - However, it also depends on pressure, spectrum, envelope and duration.
- Nonetheless, pitch and fundamental frequency are often exchangeable used

# Pitch Perception

- Audible pitch range
  - From 20Hz to 20kHz
  - Upper limits gradually decreases with age and also how much you are exposed to strong noises
- Pitch resolution
  - Just noticeable difference (JND) depends on not only the frequency but also the sound level and the duration of tone.
  - This is related to pitch scale

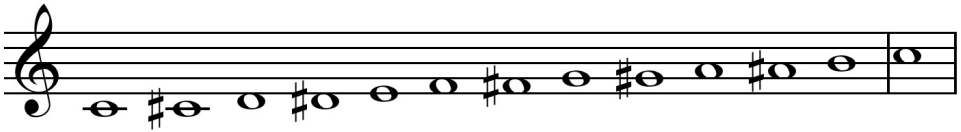


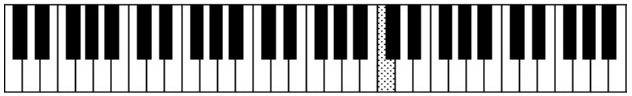
# Pitch Scale

- Human ears are sensitive to frequency changes in a log scale
  - Physical location of basilar membrane
- Approximated pitch scale by subjective tests
  - Mel scale $m = 2595 \log_{10}(1 + f / 700)$
    - Based on pitch ratio of tones
    - Most popularly used for audio analysis
  - Bark scale $Bark = 13 \arctan(0.00075 f) + 3.5 \arctan((f / 7500)^2)$
    - Critical band measurement by masking
    - Used in audio compression

# Pitch Scale in Music

- Musical temperament
  - Determining the frequencies of all 12 notes in the chromatic scale
  - Tuning musical instruments

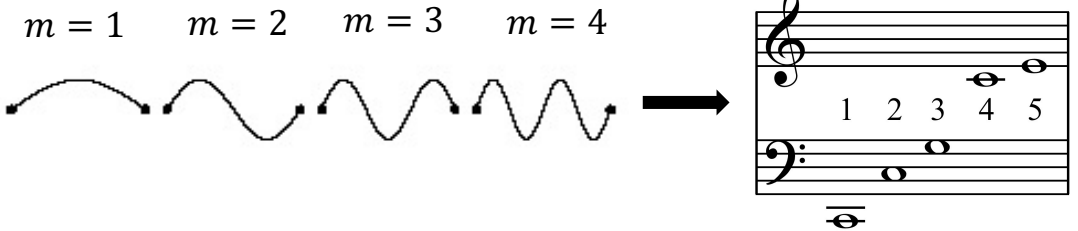


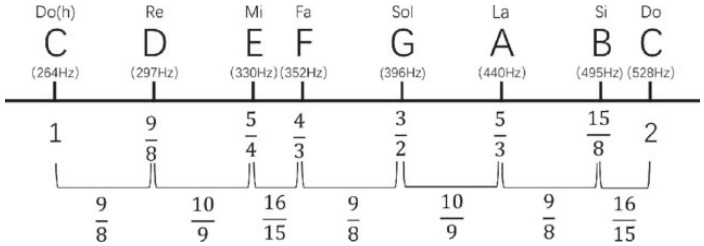


- Types
  - Pythagorean tuning
  - Just intonation
  - Equal temperament

# Pitch Scale in Music

- Pythagorean tuning
  - Based on the 3:2 ratio for perfect fifth interval
  - “Wolf fifth” issue: $(3/2)^{12} \neq 2^7$
    - C0→ G0→ D1→ A1→ E2→ B2→ F#3 → Db4→ Ab4 → Eb5 → Bb5 → F6 → C7
- Just intonation
  - Based on harmonic series of tone: two notes have a integer frequency ratio
  - Issue: key transpose changes the temperament



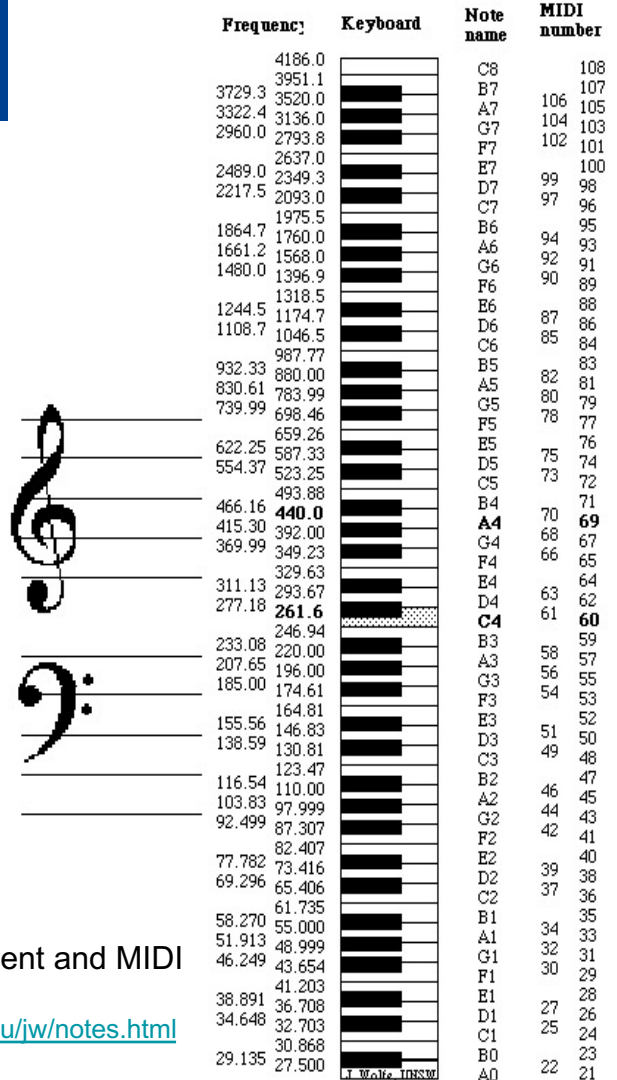



# Pitch Scale in Music

- Equal temperament
  - 1: $2^{1/12}$ ratio between two adjacent notes
  - Music note (*m*) and frequency (*f*) in Hz

$$m = 12\log_2\left(\frac{f}{440}\right) + 69, \quad f = 440 \cdot 2^{(m-69)/12}$$

  - Invariant to key transpose




Equal temperament and MIDI

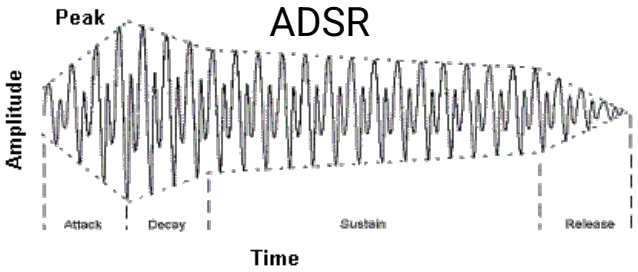
Source: https://newt.phys.unsw.edu.au/jw/notes.html

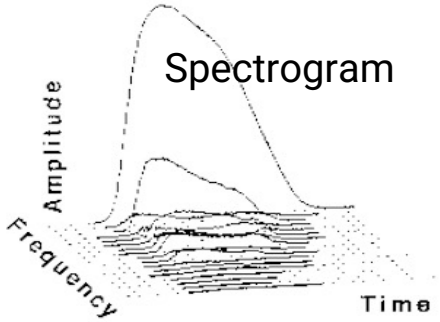
# Timbre Perception

- Attribute of sensation by which a listener can judge two sounds having the same loudness and pitch are dissimilar (ANSI)
  - Tone color or quality that defines a particular sound
- Associated with classifying or identifying sound sources
  - Class: speech, piano, guitar
  - Identity: human voice
- Timbre is a multi-dimensional concept
  - Multiple physical attributes are associated with timbre
  - Timbre space: measuring timbre dimension based on perceptual similarity of different sounds

# Physical Attributes in Timbre Perception

- Physical attributes that change timbre (Schouten, 1968)
  - Harmonicity: the range between tonal and noise-like character
  - Time envelope (ADSR)
  - Spectral envelope
  - Changes of spectral envelope and fundamental frequency
  - The onset of a sound differing notably from the sustained vibration
  - Inharmonicity




Timbre control knobs in synthesizer

# Timbre Space

- Perceptual multi-dimensional attributes based on measuring similarity
  - Ask human to listen a pair of sounds and judge the degree of similarity as a score
  - The similarity matrix is processed using multi-dimensional scaling, a dimensionality reduction algorithm which determines the timbre space
- Acoustic correlation with the three (reduced) dimensions
  - Spectral energy distribution
  - Attack and decay time
  - Amount of inharmonic sound in the attack

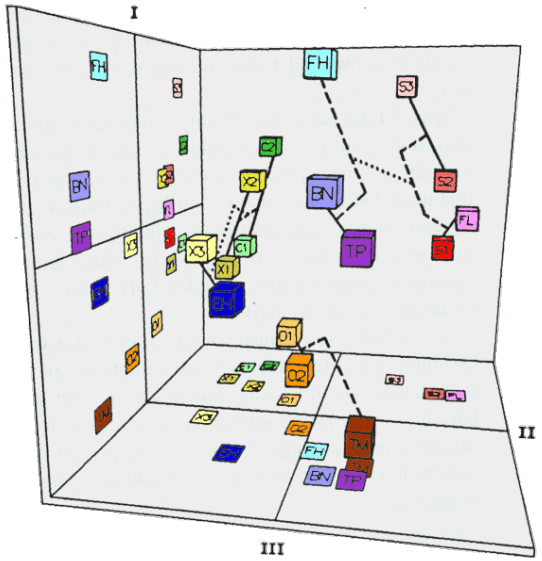



FIG. 1. Three-dimensional spatial solution for 35 similarity matrices generated by multidimensional scaling program INDSCAL (Carroll and Chang, 1970). Hierarchical clustering analysis (Johnson, 1967) is represented by connecting lines, in clustering strengths order: solid, dashed, dotted. Two-dimensional projections of the configuration appear on the wall and floor. Abbreviations for stimulus points: O1, O2 = oboes; C1, C2 = clarinets; X1, X2, X3 = saxophones; EH = English horn; FH = French horn; S1, S2, S3 = strings, TP = trumpet; TM = trombone; FL = flute; BN = bassoon.

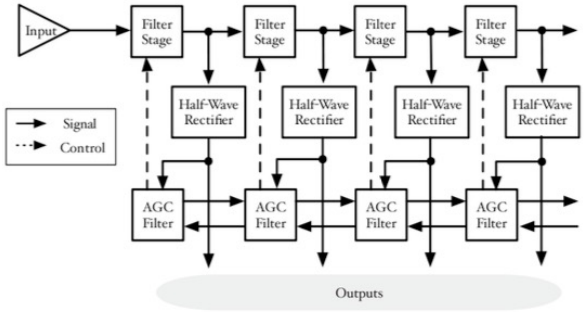
(Grey, 1977)

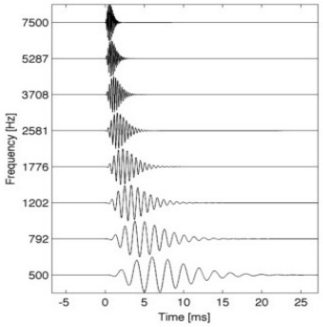
# Summary

| Physical Attributes | Perception Attributes | | |
|---|---|---|---|
| | Loudness | Pitch | Timbre |
| Amplitude | *** | * | * |
| Frequency | ** | *** | ** |
| Spectrum | * | * | *** |
| Envelope | * | * | ** |
| Duration | * | * | * |

# References

- Basic theory of human hearing
  - P. R. Cook (Editor), “Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics”
  - John R. Pierce, “The science of musical sound”
  - Brian C J Moore, “An introduction to the Psychology of Hearing”
- Computational model of human hearing
  - Dick Lyon, “Human and Machine Hearing: Extracting Meaning from Sound”




Gamma-tone Filterbank