CTP431: Fundamentals of Computer Music

Fundamentals of Musical Acoustics

KAIST Graduate School of Culture Technology

Juhan Nam

Goals

- Understand the physics of musical tone generation
  - Underlying mathematics
  - Properties of waves
  - Physical attributes of sound

- Human perception of sound
  - Physiology of human ears
  - Perceptual attributes of sound

What is Sound?

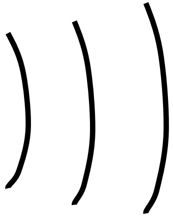
- Vibration of air pressure that we can hear through ears

Physical perspective (waves)

Generation

Vibration on an object

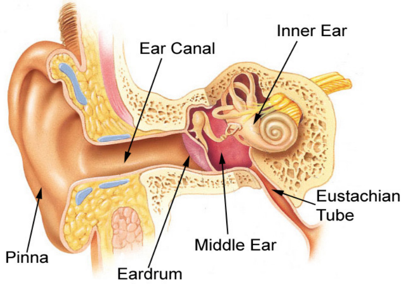
Propagation



Traveling in air

Psychological perspective

Perception



Sensation of the air vibration through ears

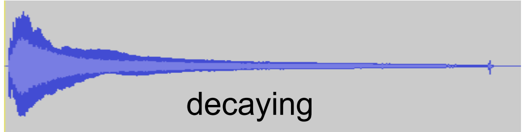
Introduction to Musical Tones

Sound Generation

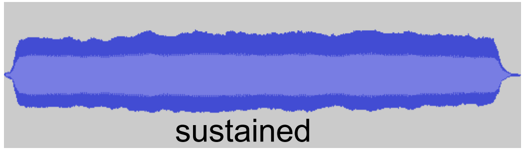
Sound Generation

Amplitude Envelope

Plucking, Striking
(momentary force)



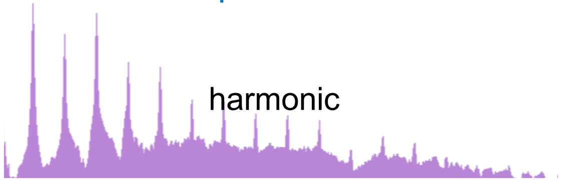
Bowing, Blowing
(continuous force)



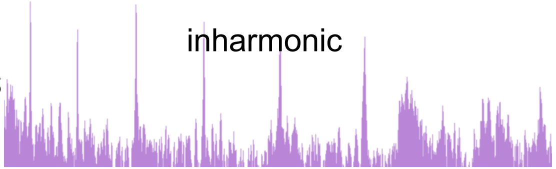
Playing → Force → **Musical Instrument** → **Sound**

Spectrum

< 1D vibration >
String: Piano, Guitar, Violin
Pipe: Clarinet, Flute



< 2D vibration >
Membrane: Drum, Percussions
Bar: Marimba, Xylophone



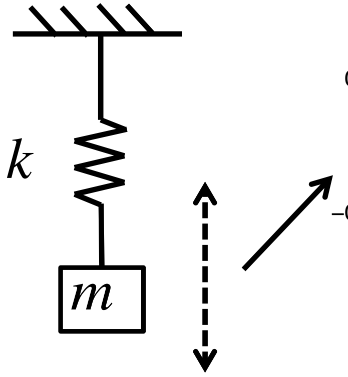
Oscillation: Simple Harmonic Motion

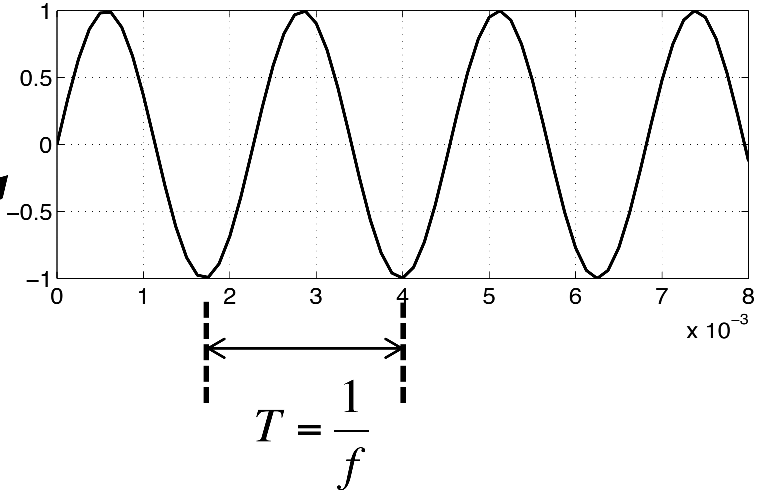
- A mass-spring model

Newton's second law

$$F = -kx = m\frac{d^2x}{dt^2}$$

Restoration force
by Hooke's law





- Generate a sinusoid oscillation
  - Pure tone: $x = A\sin(\omega t) = A\sin(2\pi f t)$

$\omega = \sqrt{k/m}$ angular frequency

$f = \omega / 2\pi$ frequency

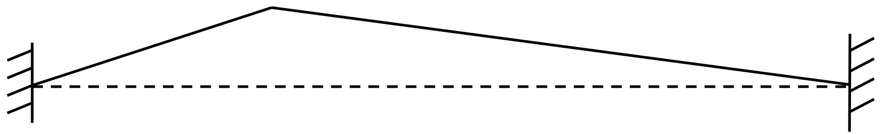
$T = 1/f$ period

Oscillation: Plucked String

https://www.youtube.com/watch?v=_X72on6CSL0

Oscillation: Plucked String

- One-dimensional ideal string vibration



Mathematics

Wave Equation $$K\frac{\partial^2 y}{\partial x^2} = \varepsilon \frac{\partial^2 y}{\partial t^2}$$ → Propagation

Boundary Condition Fixed ends → Reflection and Superposition

Initial Condition Plucking at a position → Modes and Standing Waves

Properties of Waves

Propagation

Reflection and Superposition

Modes and Standing Waves

Wave Equation

- For 1-D string (or other media) with displacement $y(x, t)$

$$\boldsymbol{F} = \boldsymbol{ma}$$

$$K \frac{\partial^2 y}{\partial x^2} = \varepsilon \frac{\partial^2 y}{\partial t^2}$$

Curvature (how much the string is bended)

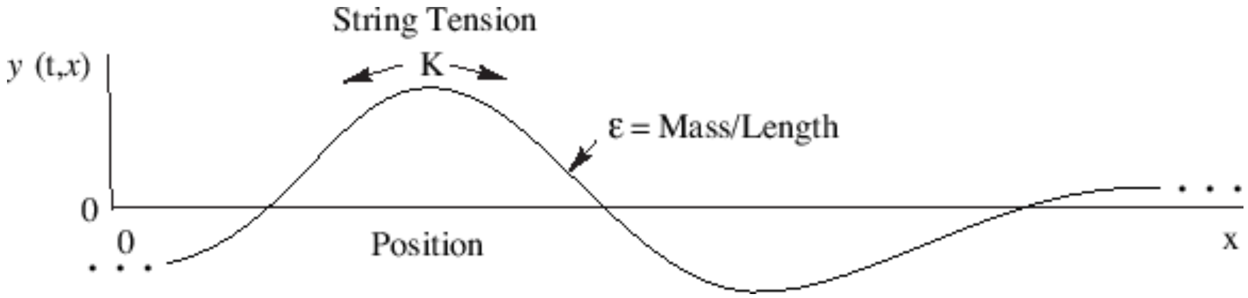
Stiffness (how stiff the string is)

Acceleration

Linear mass density

$$c = \sqrt{\frac{K}{\varepsilon}}$$

Speed of Wave



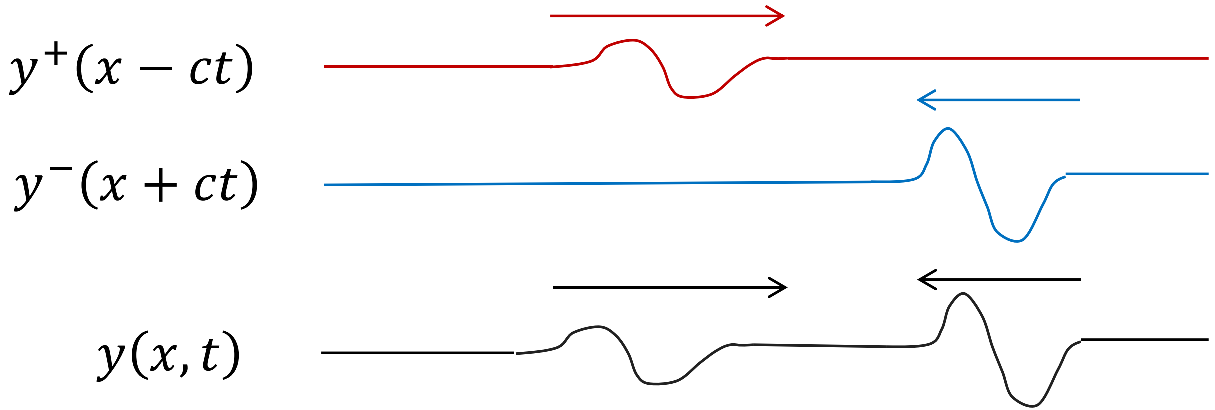
Source: https://ccrma.stanford.edu/~jos/pasp/Ideal_Vibrating_String.html

Wave Equation

- The general solution
  - The sum of **leftward-moving** y^- and **rightward-moving** y^+ traveling waves with a constant speed c

$$y(x,t) = y^+(x-ct) + y^-(x+ct)$$

  - The shape of waveform does not change: set by the **initial conditions**

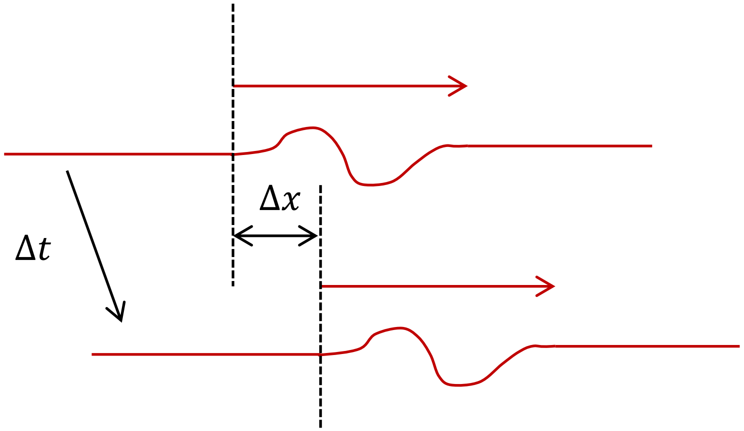


Wave Equation

- Why is it a traveling wave ?
  - Rightward-traveling wave case

$$y(x,t) = y^{+}(x - ct)$$

After Δt sec $\longrightarrow$ $y(x, t + \Delta t) = y^{+}(x - c(t + \Delta t))$



$$= y^{+}\,(x - ct - c\Delta t)$$

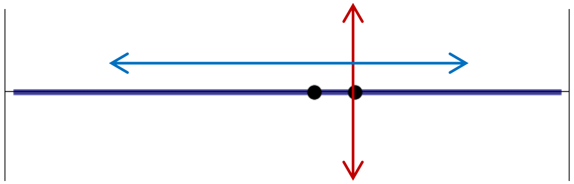
$$= y^{+}\,(x - ct - \Delta x) \quad \longleftarrow \quad c\Delta t = \Delta x$$

$$= y^{+}\,(x - \Delta x - ct)$$

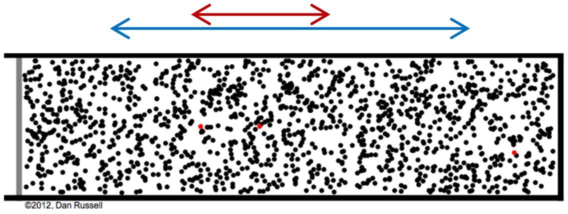
$$= y\,(x - \Delta x, t) \quad \longleftarrow$$ rightward-traveling by Δx

Wave Propagation

- The wave equation explains "wave propagation" in a medium
- Note that, while wave travels and transmits energy, the medium itself oscillates at the same position
  - Transverse wave
    - The wave travels through the medium in the perpendicular direction to the wave motion
    - Example: string, membrane, bar

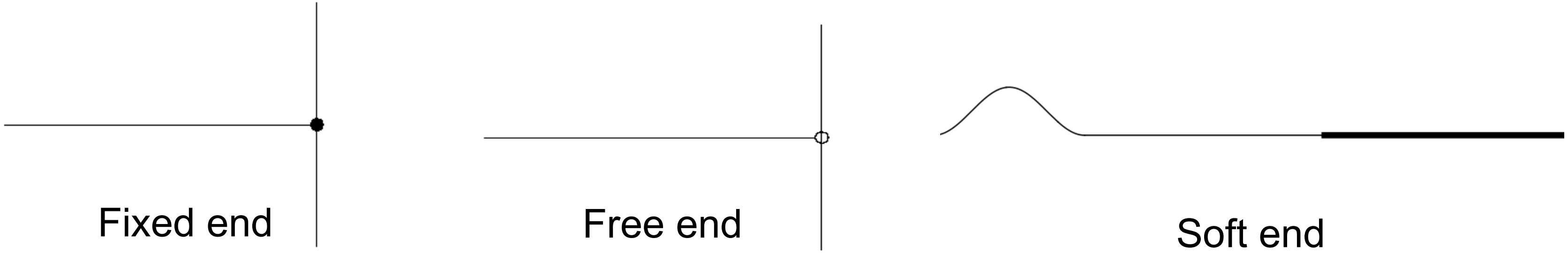


  - Longitudinal wave
    - The wave travels through the medium in the same direction as the wave motion
    - Example: pipe (air)



Wave Reflection

- Wave is reflected when the impedance (Z) of medium changes

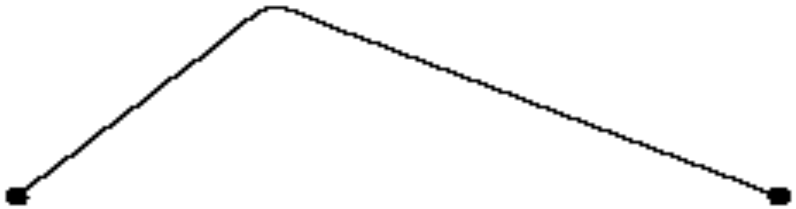


  - Related to the **boundary conditions**
  - Fixed end (Z is ∞): the incident wave is flipped
  - Open end (Z is 0): the incident wave is mirrored
  - Soft / hard end (Z is between 0 and ∞): the incident wave is partially flipped and transferred to the other side (e.g., hard end: bridge in string)

Source: http://www.acs.psu.edu/drussell/Demos/reflect/reflect.html

Wave Superposition

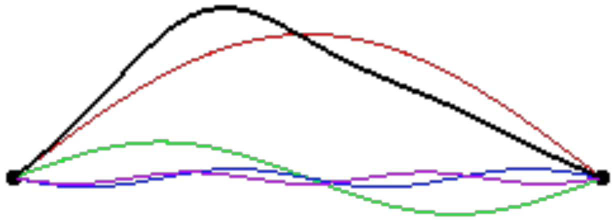
- Two traveling waves can form a single wave as a linear sum
- The superposition of the traveling wave and its reflections form this motion in the video



Source: https://www.acs.psu.edu/drussell/Demos/string/Fixed.html

Fourier Analysis

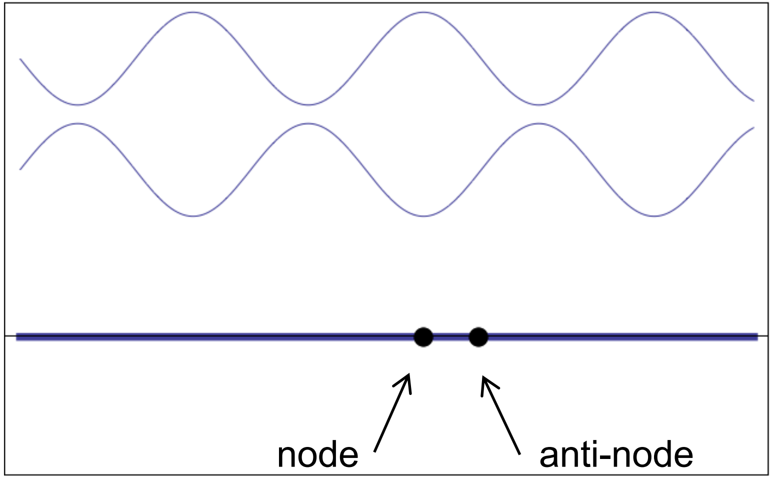
- Provides a way of representing the initial condition by the sum of sinusoids
  - The frequency (or period) of the sinusoids has an integer multiples relation
  - The amplitude and phase are determined by **Fourier series** as the waveform is a continuous-time and periodic signal



$$x(t) = \frac{1}{T}\left(A_0 + A_1 \cos\left(\frac{2\pi t}{T} + \phi_1\right) + A_2 \cos\left(\frac{2\pi 2t}{T} + \phi_2\right) + A_3 \cos\left(\frac{2\pi 3t}{T} + \phi_3\right) + \cdots\right.$$

Standing Wave

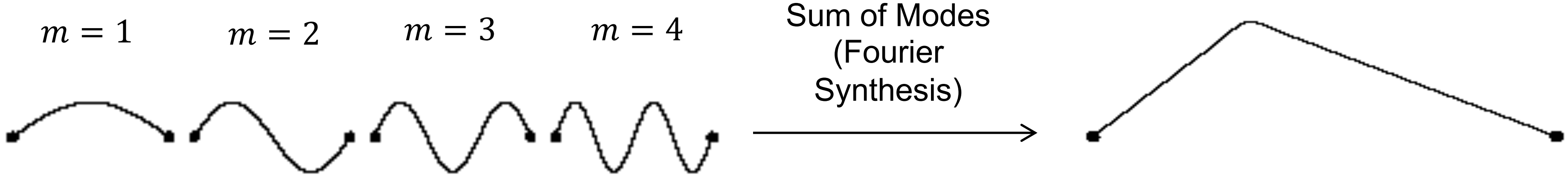
- Each of the sinusoidal shape travels to the left and right direction
- The sum of two travelling waves in opposite directions with the same frequency cancel or reinforce each other, creating a stationary oscillation



Source: http://www.acs.psu.edu/drussell/Demos/superposition/superposition.html

Mode

- Each of the standing waves form a mode with a sinusoidal oscillation
  - They are harmonic in frequency: **harmonic modes**
  - **Overtone**: higher modes other than the fundamental mode
  - **Partials**: indicates the frequency of each mode in the spectrum



$$\lambda = \frac{2L}{m}$$

Wave length

$$f = \frac{c}{\lambda} = \frac{m \cdot c}{2L}$$

Frequency

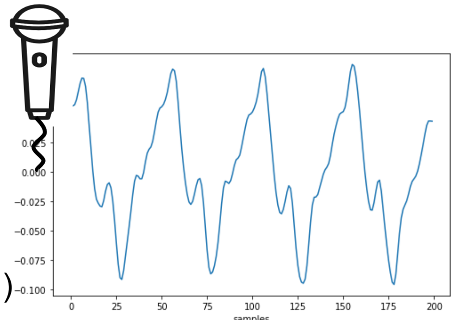
c : speed of wave

L : the length of string

m : mode number (1,2,3,…)

$$c = \sqrt{\frac{K}{\varepsilon}}$$

speed of wave
(determined by the tension and material)



Damping

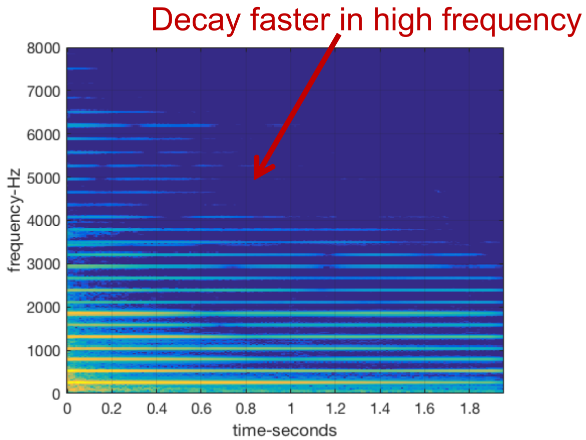
- A more practical setting of wave equation has a damping term

$$K\frac{\partial^2 y}{\partial x^2} = \varepsilon\frac{\partial^2 y}{\partial t^2} + \mu\frac{\partial y}{\partial t}$$

velocity

friction coefficient (resistance)

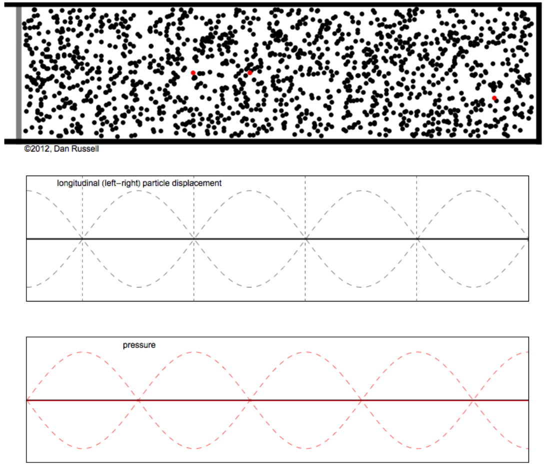
- The damping causes **exponential decaying** traveling wave
  - Higher frequency modes decay faster
  - The motion of plucked string becomes rounded over time



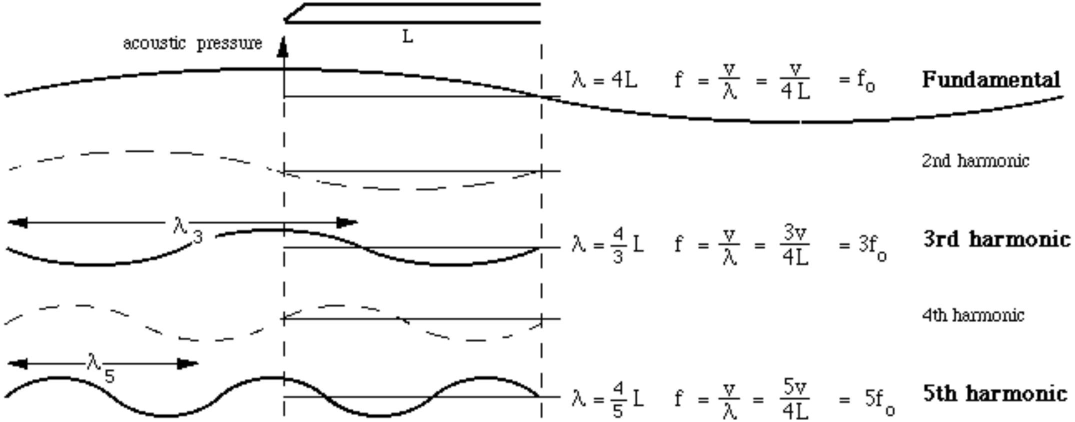
Spectrogram of a piano note

Pipes

- Longitudinal pressure wave to travel in the air column
  - Woodwind or brass instrument: flute, clarinet, trumpet
  - Blowing: continuous excitation
  - Clarinet: odd harmonics are generated



Closed-end  Open-end

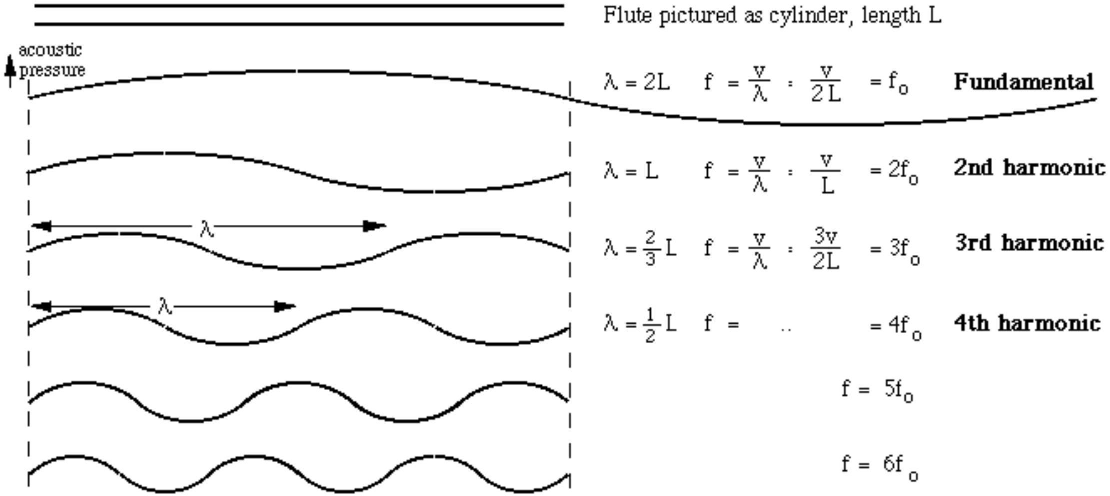


Source: https://www.acs.psu.edu/drussell/Demos/StandingWaves/StandingWaves.html https://newt.phys.unsw.edu.au/jw/clarinetacoustics.html

Pipes

- Longitudinal pressure wave to travel in the air column
  - Flute: All harmonic tones are generated

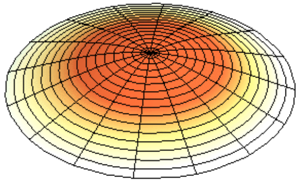
Open-end  Open-end



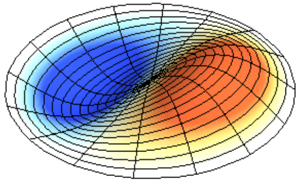
https://newt.phys.unsw.edu.au/jw/fluteacoustics.html

Membrane

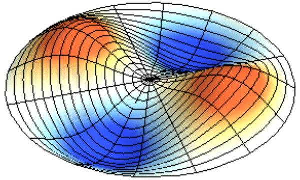
- 2D wave equation: $z(x, y, t)$
  - Drum, percussion
  - Boundary condition: determined by the shape of membrane
  - Circular harmonic oscillation → generate inharmonic tones



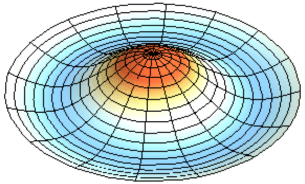
(0,1) Mode



(1,1) Mode



(2,1) Mode



(0,2) Mode

Source: https://www.acs.psu.edu/drussell/Demos/MembraneCircle/Circle.html

Chladni Plate

- Circular or rectangular vibrating plates that visualize 2D standing waves

https://www.youtube.com/watch?v=1yaqUI4b974

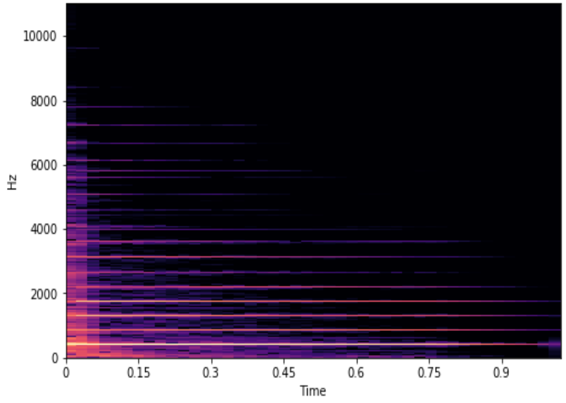
- Artistic examples
  - https://www.youtube.com/watch?v=Q3oItpVa9fs
  - https://www.youtube.com/watch?v=qxdJ-ebNi14

Summary

- Wave equation: travelling wave
  - Initial condition: by the human play on the musical instrument
  - Boundary condition: by the structure of the musical instrument
- Wave Properties
  - Propagation (traveling) by the wave equations
  - Reflection
  - Superposition
  - Standing waves
- Vibration modes
  - 1D: sinusoidal harmonic oscillations → harmonic tone
  - 2D: circular harmonic oscillations → complex inharmonic tone

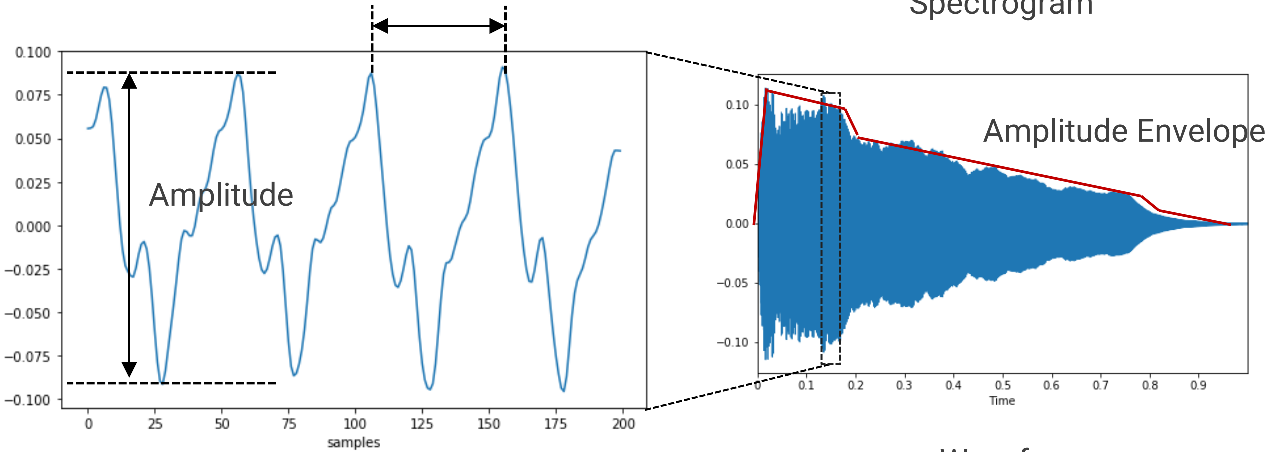
Physical Attributes of Sound

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



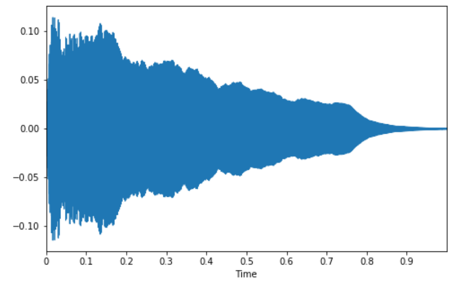
Spectrogram

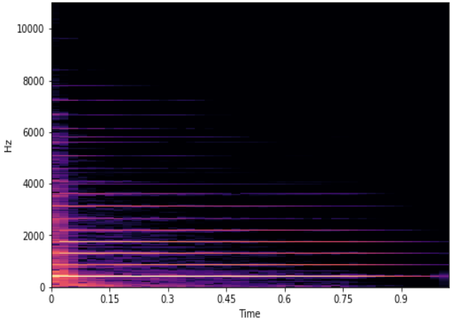
Period = 1/(fundamental frequency)



Waveform

Perceptual Attributes of Sound

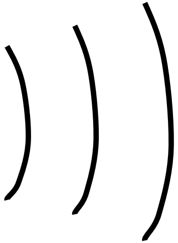


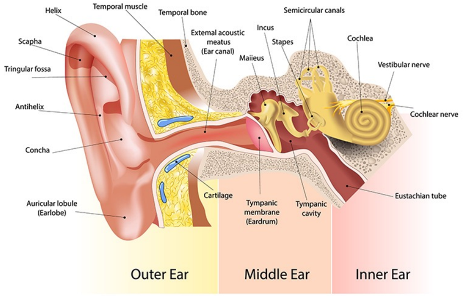


- Amplitude
- Frequency
- Envelope
- Duration
- Spectrum

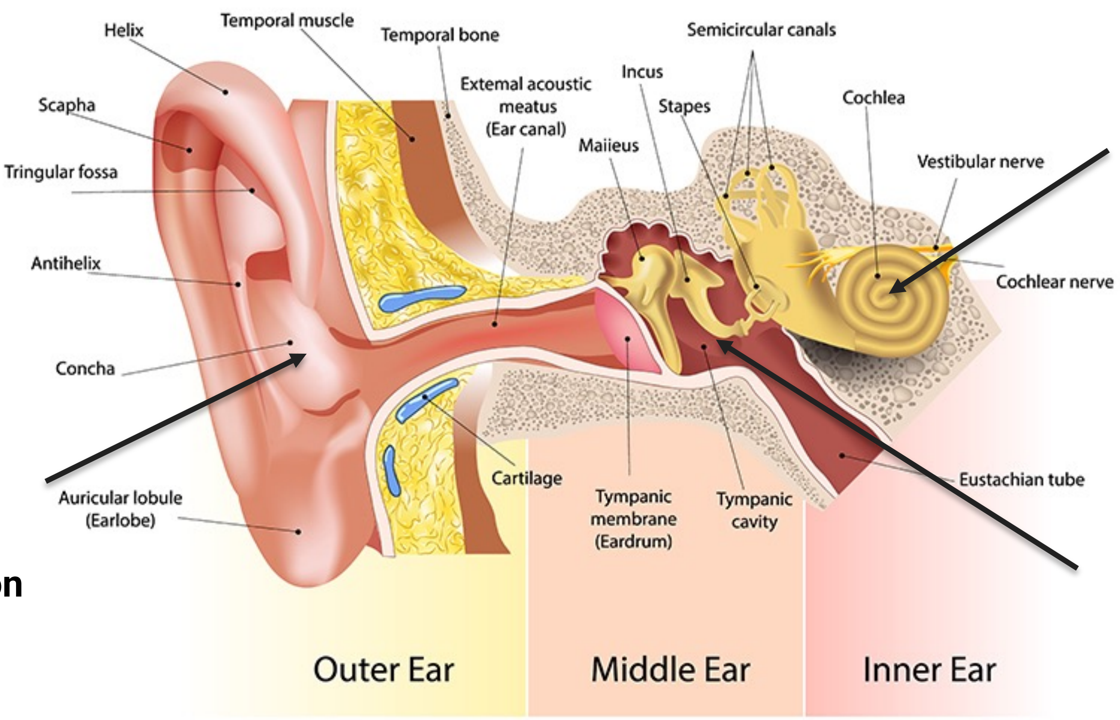
?

- Loudness
- Pitch
- Timbre





Human Ear

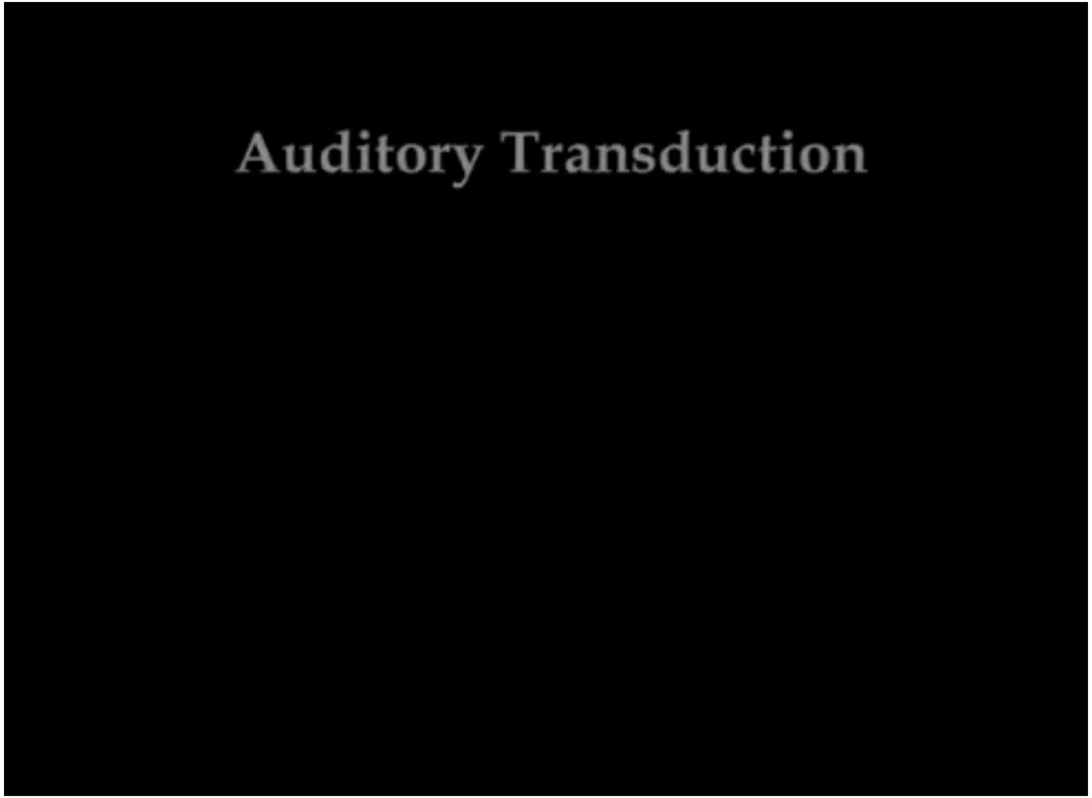


- ✓ **Protection**
- ✓ **Localization**
- ✓ **Amplification**

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

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

Auditory Transduction in Human Ears



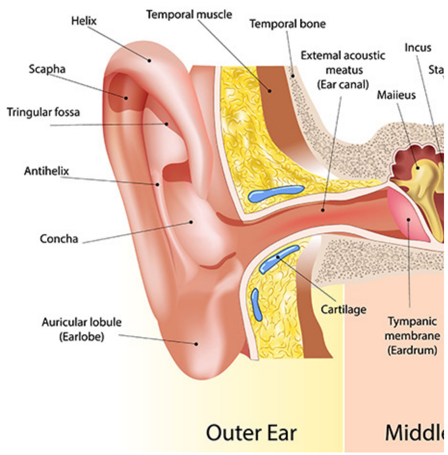
http://www.youtube.com/watch?v=PeTriGTENoc

Auditory Transduction in Human Ears

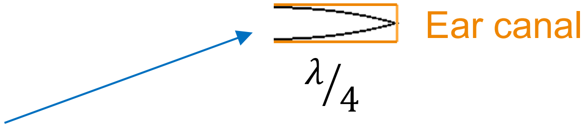
- Multiple steps of transduction
  - Air vibration (outer ear) → Mechanical vibration (middle ear) → Fluid vibration (inner) → Nerve firing (brain)
- High sensitivity by the amplification
  - Outer ear and middle ear
- Tonotopic organization
  - Inner ear can be regarded as a bandpass filterbank

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)



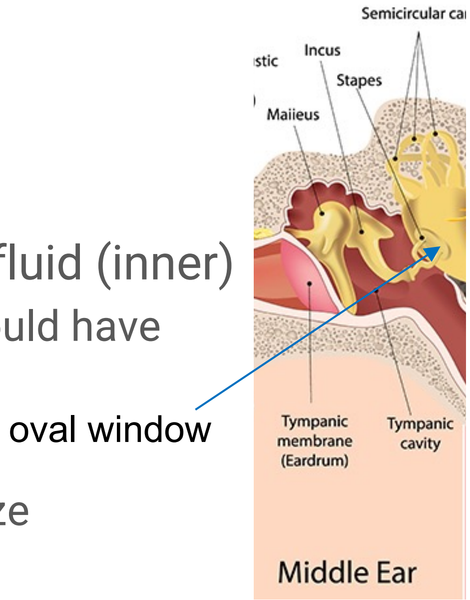
- Ear canal
  - Protect ear drums
  - Quarter-wave resonance: boost the vibration around 3kHz by 15-20 dB



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

Middle Ear

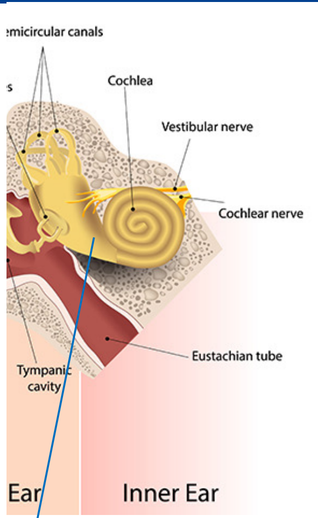
- Ossicles
  - malleus (hammer), incus (anvil) and stapes(stirrup)
  - The smallest bones in human body
  - **Impedance matching** between air pressure (outer) and fluid (inner)
    - Without ossicles, only about 1/30 of the sound energy would have been transferred to inner ear
  - **Amplification:** 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

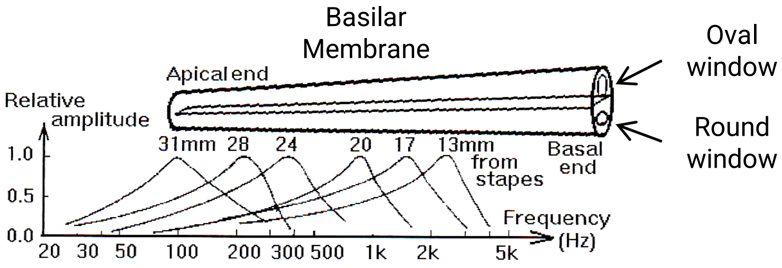


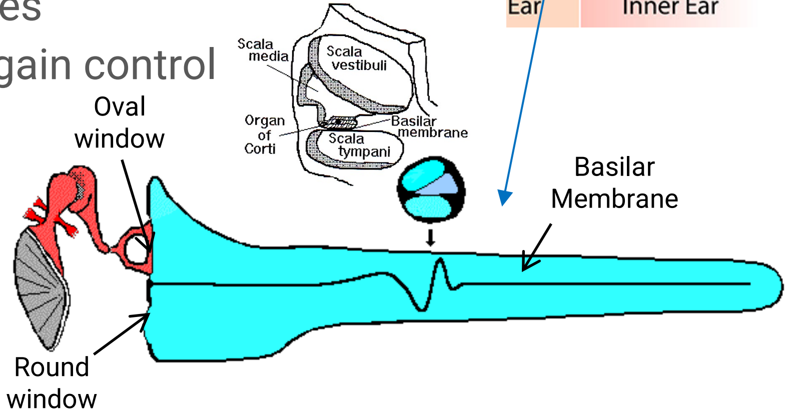
- Muscles
  - Reduce the sound transmission in response to loud sounds

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

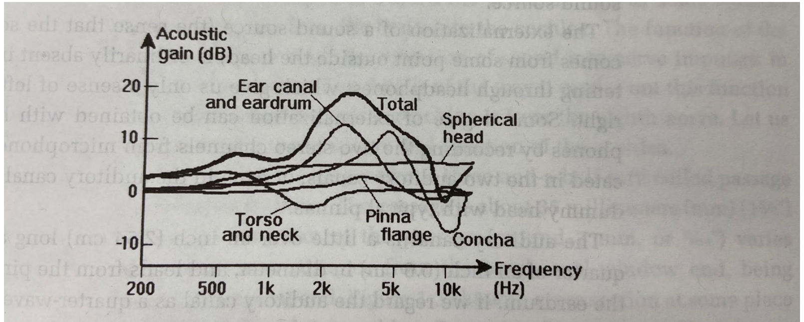
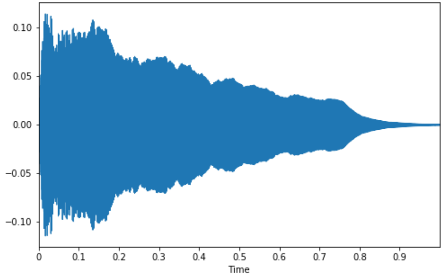
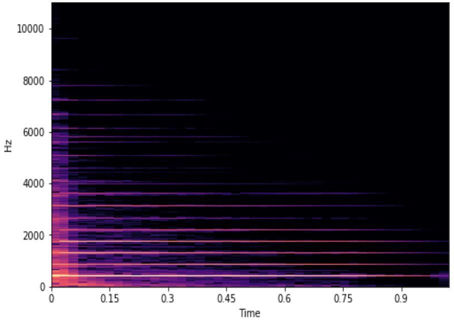


Figure 1.2 The acoustic effects of various body structures as a function of frequency. Sound is from 45 degrees off center in front of listener. (Modified from E. H. G. Shaw, 1974, *Handbook of Sensory Physiology*, vol. 5/1. Berlin: Springer, fig. 11.)

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

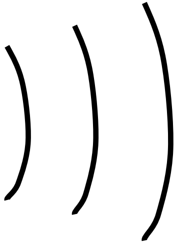
Perception of Sound

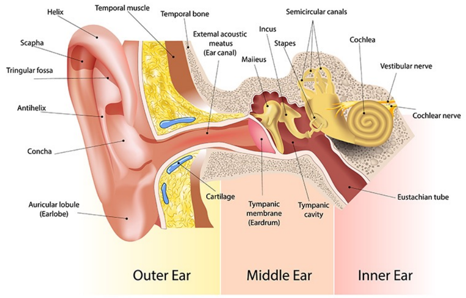




- Amplitude
- Frequency
- Envelope
- Duration
- Spectrum

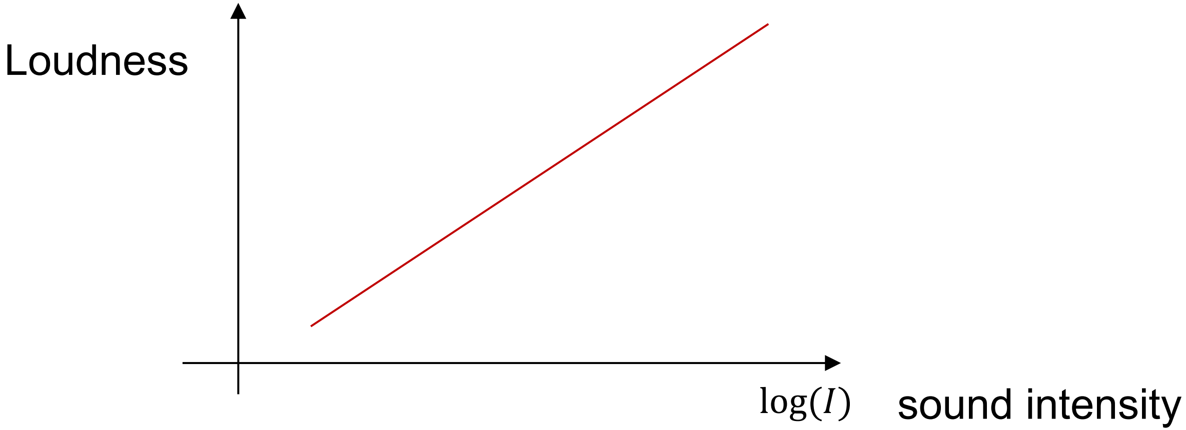
- **Loudness**
- **Pitch**
- **Timbre**





Loudness Perception

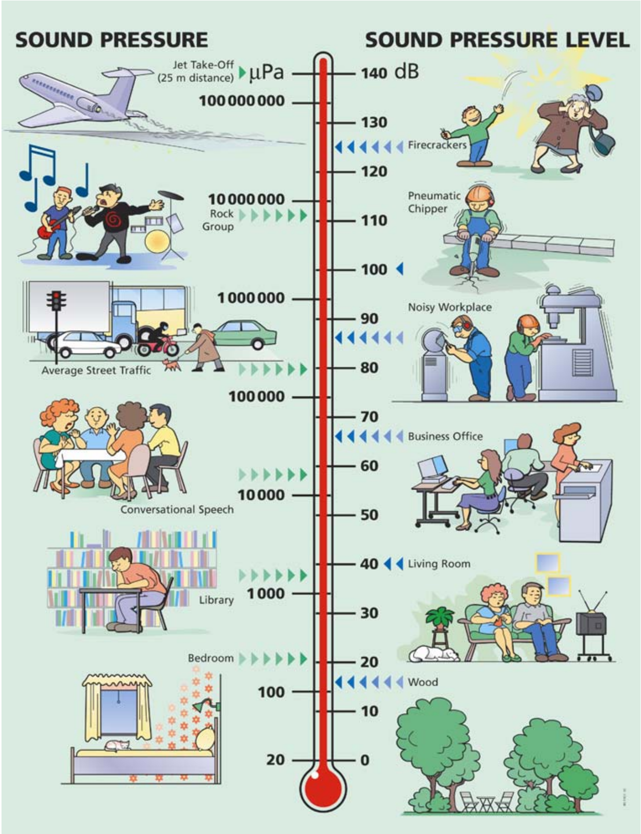
- Subjective perception of sound intensity (the squared amplitude)
- Weber's law: $\Delta I \propto I$
  - Just Noticeable Difference (JND) of intensity (ΔI) is proportional to the intensity (I) at the moment
  - A log function satisfies this rule



Loudness Perception

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

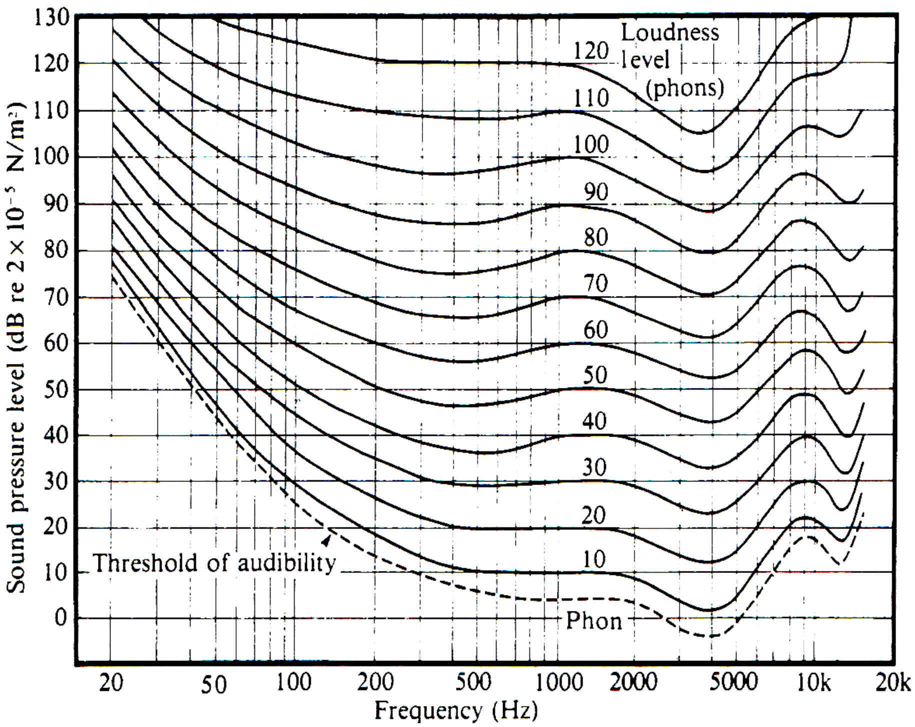
SPL meter



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

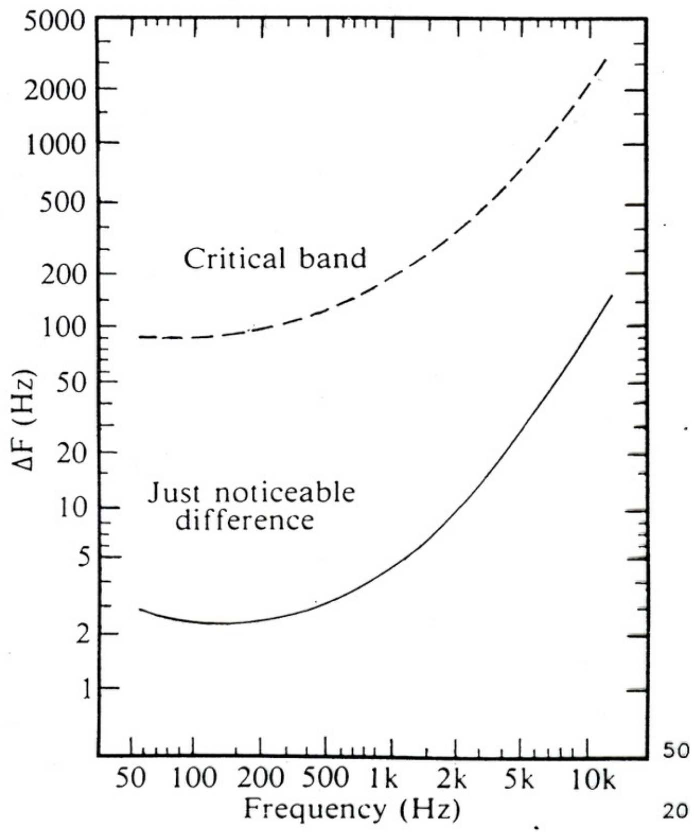
- Loudness depends on frequency (Fletcher and Munson, 1933)
  - 1kH is used as a reference
  - Most sensitive to 2-5 kHz tones
    - Resonance in the ear canal
  - See the threshold of audibility
    - Dynamic ranges are different



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

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

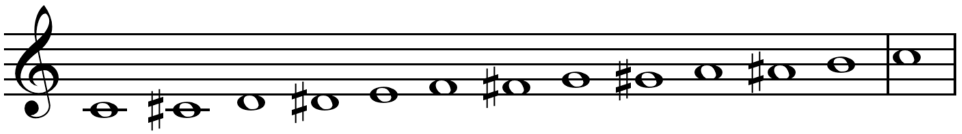


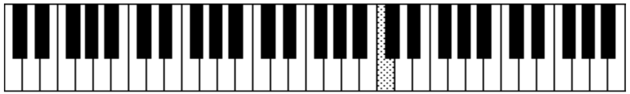
Pitch Scale

- 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 $Bark = 13\arctan(0.00075f) + 3.5\arctan((f/7500)^2)$
    - Critical band measurement by masking
    - Used in audio compression

Pitch Scale in Music

- Musical temperament
  - Determining the frequencies of all 12 notes within one octave
  - Tuning musical instruments

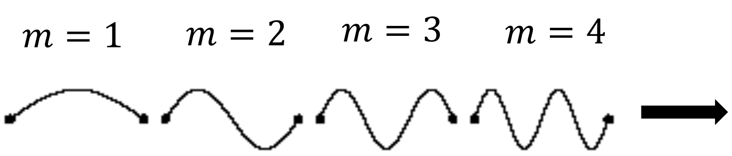


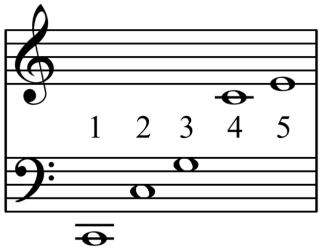


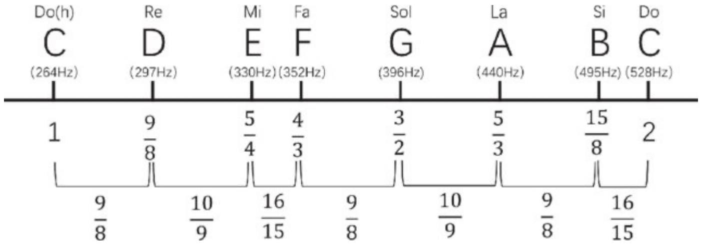
- Types
  - Pythagorean tuning
  - Just intonation
  - Equal temperament

Pitch Scale in Music

- Pythagorean tuning
  - Based on the 3:2 ratio for perfect fifth interval
  - Issue: “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 a tone
  - Issue: key transpose changes the temperament





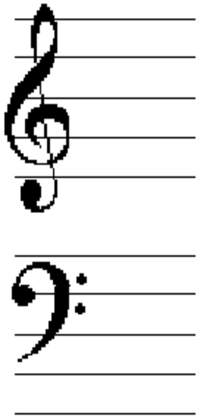


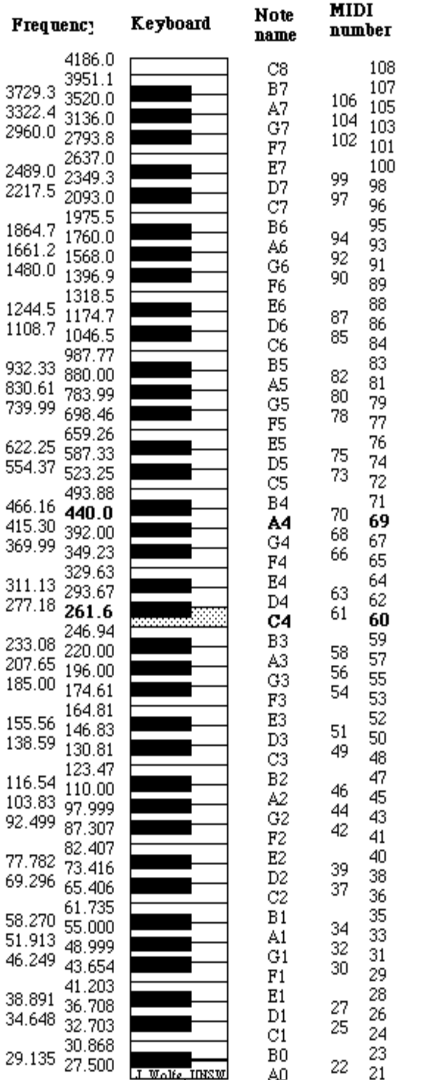
Pitch Scale in Music

- 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

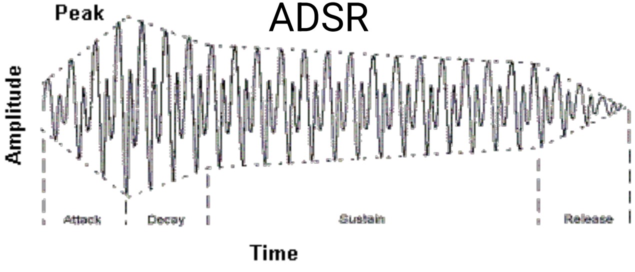
Source: https://newt.phys.unsw.edu.au/jw/notes.html

Timbre Perception

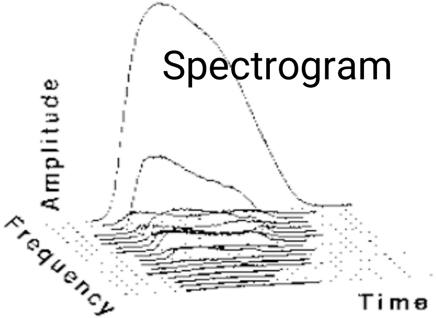
- Attribute of sensation by which a listener can judge two sounds having the same loudness and pitch are dissimilar (ANSI)
  - Tone color or quality that defines a particular sound
- Associated with classifying or identifying sound sources
  - Class: speech, piano, guitar
  - Identity: human voice
- 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



ADSR



Spectrogram

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

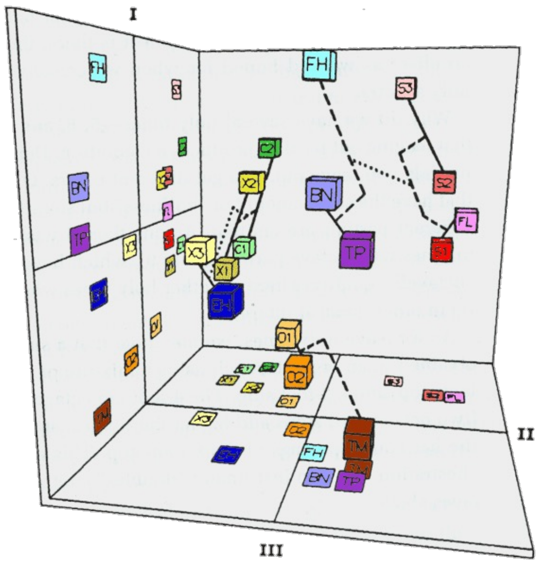


FIG. 1. Three-dimensional spatial solution for 35 similarity matrices generated by multidimensional scaling program INDSCAL (Carroll and Chang, 1970). Hierarchical clustering analysis (Johnson, 1967) is represented by connecting lines, in clustering strengths order: solid, dashed, dotted. Two-dimensional 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	*	*	*