Digital Audio

Principles of Dig Audio, Pohlmann

http://www.tc.umn.edu/~erick205/Papers/paper.html

http://www.earlevel.com/Digital%20Audio/index.html

Analog to Digital Conversion

- Analog = Continuous
- Digital = Discrete
- Analog becomes Digital via sampling
- Is information lost? No (band-limited)
  - A continuous signal can be reproduced from discrete samples
    - An analogy is a movie where our brain fills in missing information. With digital signals the missing information comes from interpolation

Draw graph of continuous signal…show sampling(samples in binary numbers)

Sampling Frequency

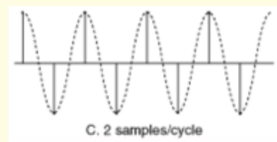
- How fast should we take samples?
  - Sampling theorem
    - A continuous **band-limited** signal can be replaced by a discrete sequence of samples without any loss of information.
    - the sampling frequency must be greater than or equal to twice the highest frequency in the original analog signal
    - Nyquist frequency, FS < 2F
      - Aliasing
        - » Example, wheels turning backwards in the movies

Caveat that the actual sample measurement must have infinite resolution

Frequency = cycles/time

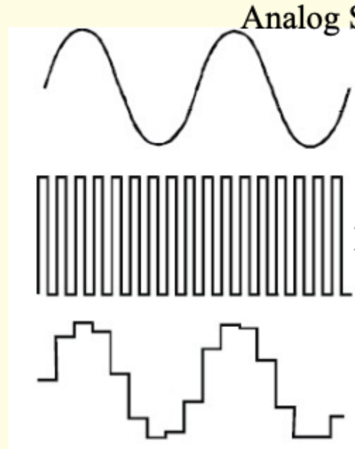
Nyquist Frequency

- **Intuitive understanding:** At least 2 uniformly spaced samples per cycle (of the max frequency) guarantees at least 1 detectable zero crossing between each sample pair
- E.g. for sampling a sine wave of max freq:
  - Sample 1 is above x axis, then next sample (2) is below x axis (and vice versa)
  - Sample 1 is on x axis, then next sample is also on x axis
- If two consecutive samples are both the same sign, then must be a lower than max frequency

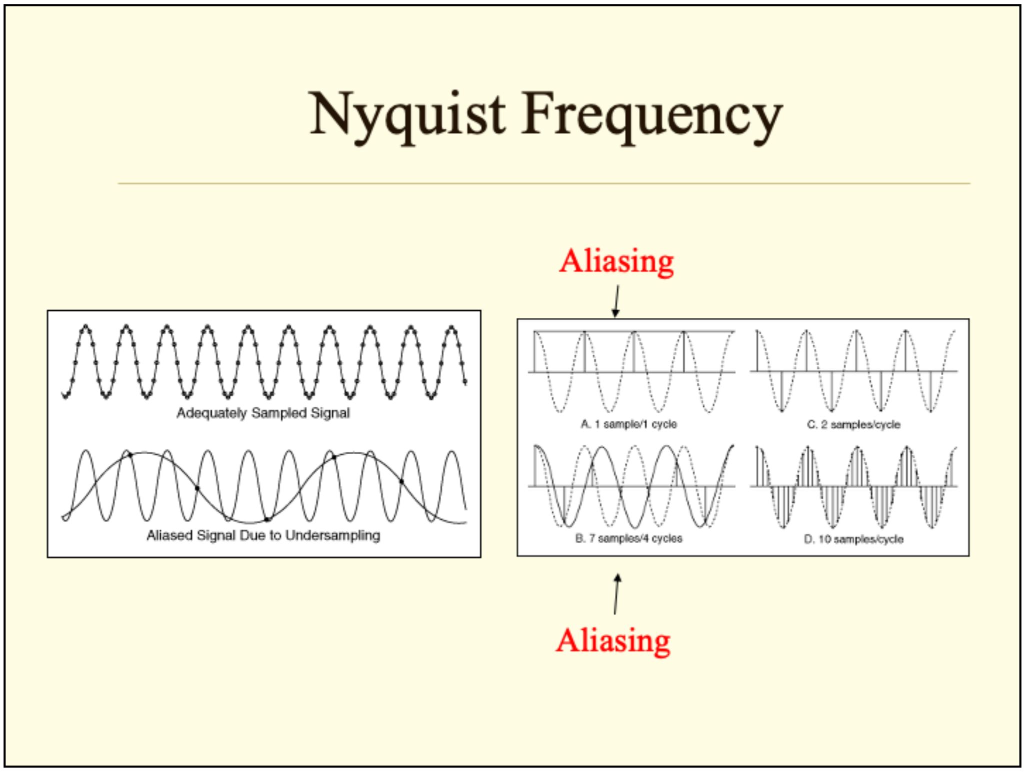


Sampling

- Digital audio does not look like a staircase when played!
- Playback is always analog! (No audio ”pixels”)
- Digital-to-Analog-Converters (DAC) rebuilds smooth curve



Digital Data (NOT the OUTPUT!)



http://zone.ni.com/reference/en-XX/help/371361H-01/lvanlsconcepts/aliasing/

the sampling frequency must be greater than or equal to twice the highest frequency in the original analog signal

Nyquist frequency, FS < 2F

In case A of the previous illustration, the sampling frequency f_s equals the frequency f of the sine wave. f_s is measured in samples/second. f is measured in cycles/second. Therefore, in case A, one sample per cycle is acquired. The reconstructed waveform appears as an alias at DC.

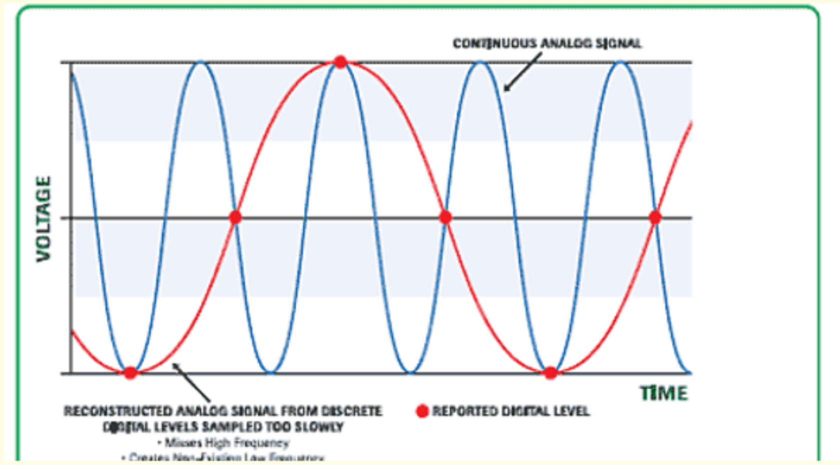
In case B of the previous illustration, f_s = 7/4f, or 7 samples/4 cycles. In case B, increasing the sampling rate increases the frequency of the waveform. However, the signal aliases to a frequency less than the original signal—three cycles instead of four.

In case C of the previous illustration, increasing the sampling rate to f_s = 2f results in the digitized waveform having the correct frequency or the same number of cycles as the original signal. In case C, the reconstructed waveform more accurately represents the original sinusoidal wave than case A or case B.

By increasing the sampling rate to well above f, for example, $f_s = 10f = 10$ samples/cycle, you can accurately reproduce the waveform.

Case D of the previous illustration shows the result of increasing the sampling rate to $f_s = 10f$.

Nyquist Frequency



More aliasing

Quantization

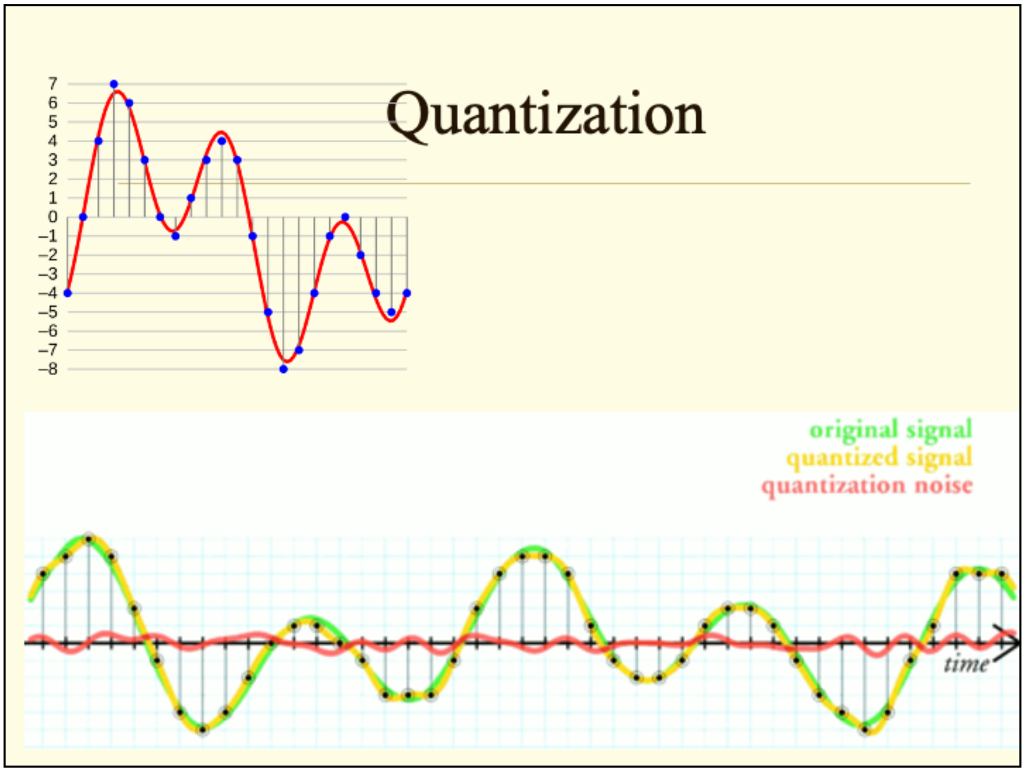
- A sample will represent the amplitude of the signal at that point in time
  - The analog value of the amplitude will be converted to a discrete value
    - Quantization
    - Sampling = Time, Quantization = Value, together they characterize an acoustic event
  - Therefore the bit resolution is very important
    - Finite word length means measuring error introduced
    - Sampling is lossless, but quantization is not
    - 16 bit = 65,536 different levels, 2^n is number of levels (n=bits)
    - Quantization Error will not exceed ½ of the interval

The measuring error can not be avoided

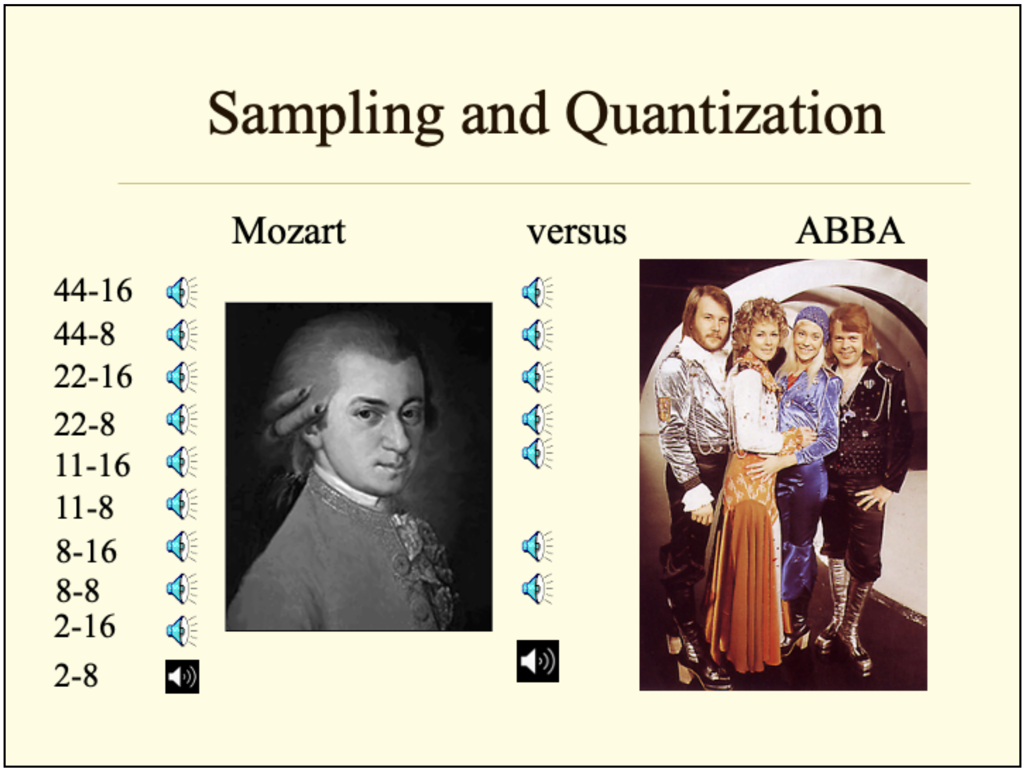
Speech can barely be intelligible with 1 bit

So adding a bit doubles the number of levels

24 bit accuracy of distance between NY and LA…accurate to 9 inches



There is no perfect quantization resolution. But the better the resolution, the lower the quantization noise floor



44-16

44-8

22-16

22-8

11-16

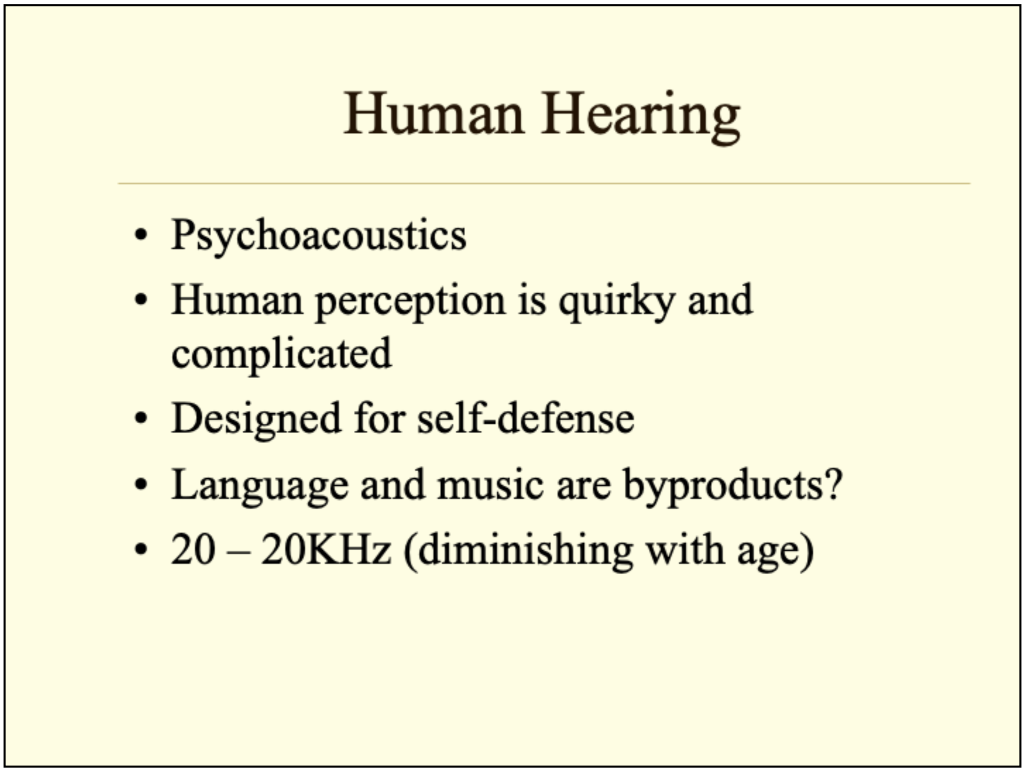
11-8

8-16

8-8

2-16

2-8

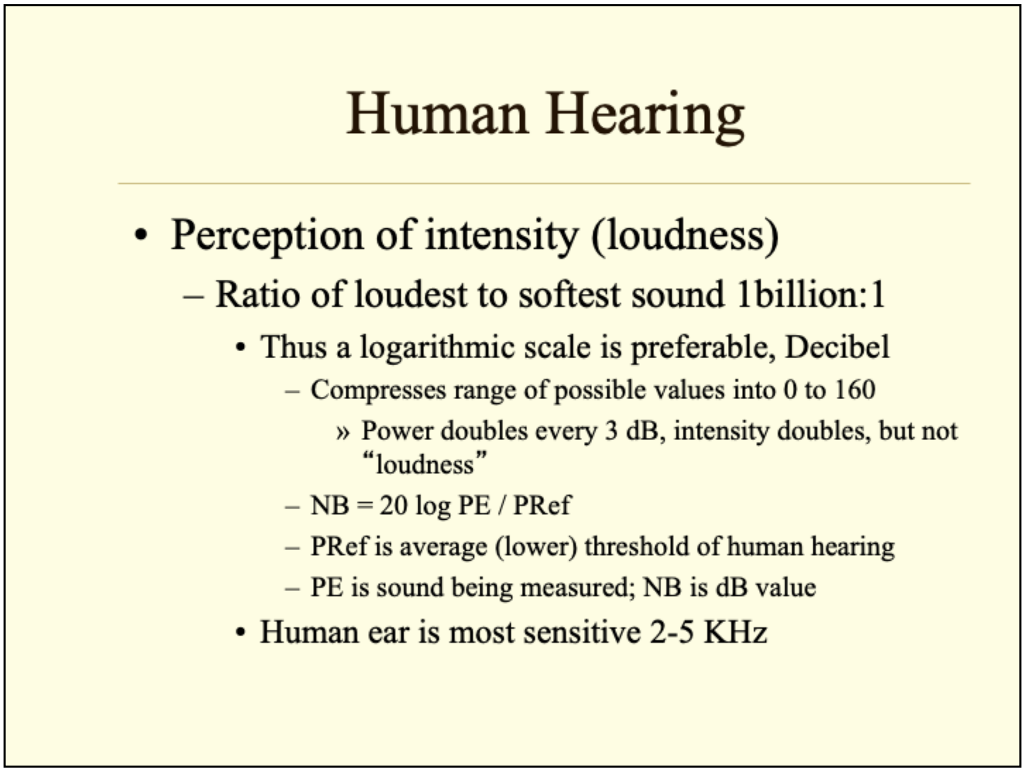


is the field of science which studies how we perceive sound and extract meaningful data from acoustical signals

Low level functions of human auditory system…not music or aesthetics

But important to sound compression, representation, production and processing, musicology, machine hearing, speech recognition and composition

Losing 1Hz a day after 15…so 70 year hold has range of 2-4KHz!



Name for Alexander Graham Bell

NB = number of decibels

PE = sound pressure to be expressed in decibels

PR = standard reference pressure, 0.0002 dynes/cm2, 10^ - 12

At PR eardrum moves distance smaller than diameter of hydrogen atom

Any lower and the sound is overwhelmed by the blood flow in the ears, and the the noise created by the molecules colliding with the eardrum

Talk about anechoic chamber.

20 whisper

60 normal conversation

120 rock concert

140 jet aircraft at takeoff…injury can occur

Dynamic range orchestra 20-30dB in quiet parts and up to 110 dB…so 80+

dynamic range required

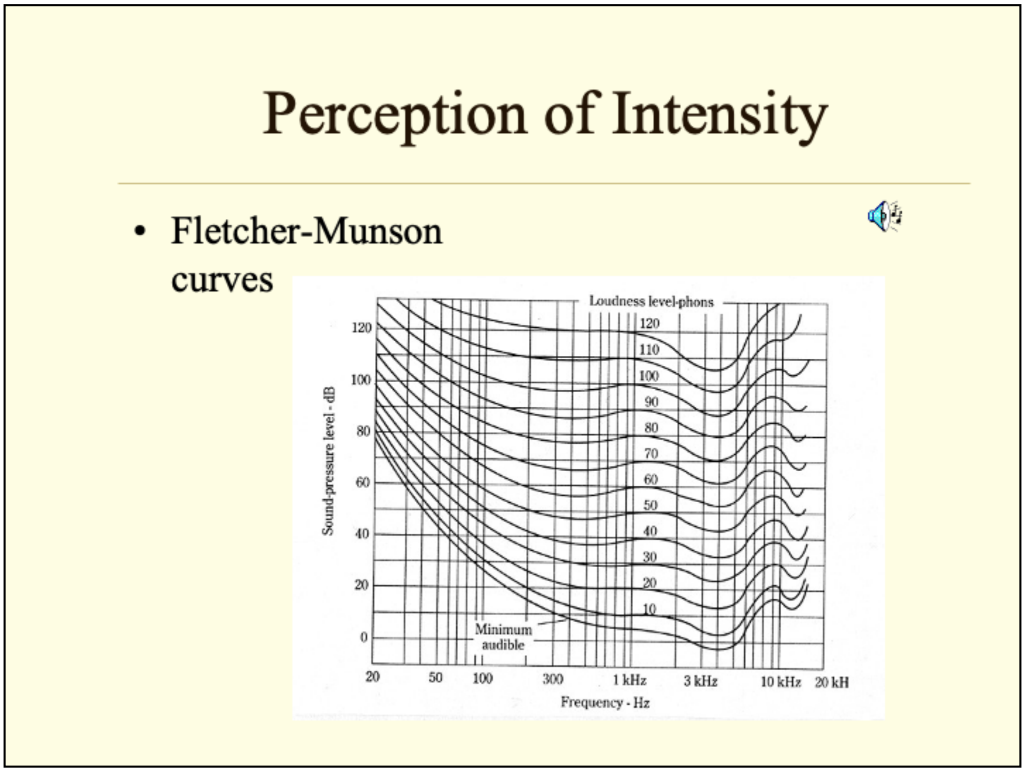
Vinyl = 50-70 dB

CD 96dB

DVD Audio 144 dB

Did hearing evolve for speech or did speech evolve to fit with our hearing? Baby cries are in this most sensitive band

2-5 is resonant frequency of the ear canal



Subject listen to 1000 HZ tone and are told to tune tones of other frequencies to that same loudness…

Phons a measure of loudness

Loudness setting on stereo

100,200,400,800,1600,3200,6400

http://www.audiotutorial.co.uk/psychoacoustics/

The Role of Two Ears

- Need two ears for perceiving distance and direction accurately
  - Monoaural (one ear) cues
  - Binaural (two ear) cues
- Ability to localize sounds is crucial to human and animal hearing

Dichotic: alphanumeric strings read into each ear through headphones…more digits that were read into right ear (and left auditory cortex) were reported correctly…right ear and left hemisphere are dominant for processing linguistic material.

Also true for “speech like sounds” like reversed speech

Reverse is true for music, memory of melodies, chords and notes from instruments, and pure tones.

And for nonverbal sounds (phone ringing, dog barking, clock ticking)

And for nonverbal vocalizations (cry, sigh, and laugh)

But the right and left dominance is not because brain specifically designed for verbal on the left and music on the right, instead the different hemispheres are specialized for a type of processing or analysis. Left being good for sequential, piecemeal, and analytical processing. Right is good and an integrative, holistic approach. Which happens to apply to speech and to music.

Music professionals recognize melodies better with right ear, probably because they use different listening strategies.

Perception of Space

- **Monaural cues** used for distance and crude localization
  - The louder the sound the closer the object
  - As the loudness changes so does its perceived location
  - Monaural direction sensed via head movements if the sound is repetitive and long lasting

an approaching sound seems to change more than a receding one. And this asymmetry increases with intensity.

The distance asymmetry makes sense biologically

Monaural cues for direction…the differing loudness can be used to figure out direction

Perception of Space

- Binaural cues important for direction perception
  - Time Differences
    - Interaural time difference
      - Precedence effect
        - » Filtering out of reflections and echoes, humans don't even hear echoes earlier than 35 msec
    - Phase differences
      - For frequencies <1000Hz
    - Interaural intensity difference
      - Shadowing, object must be large in relation to wavelength, <1000 Hz
      - Same tone in both ears, but of differing intensity, will be heard as coming from the direction of the more intense tone.

Two ears give two different "views" of a single sound

Draw diagram showing sounds from different sources hitting head, and "shadowing" reducing intensity

Difference can be as little as .0001 of a second

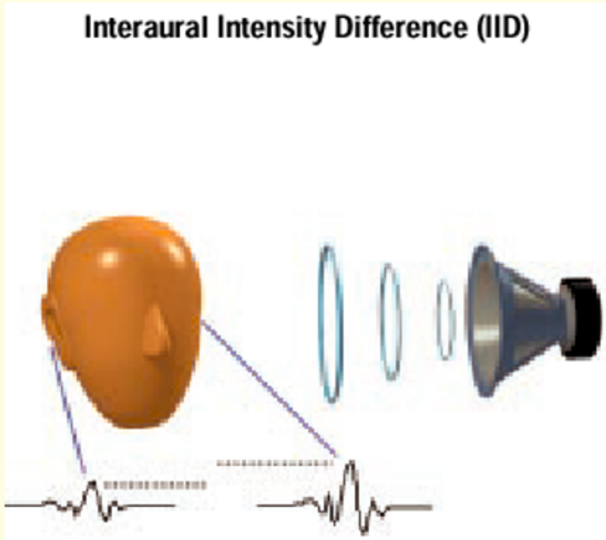
Phase differences… wavelength large in relation to head and head produces diffraction patterns

Trick to catch people feigning deafness in 1900s

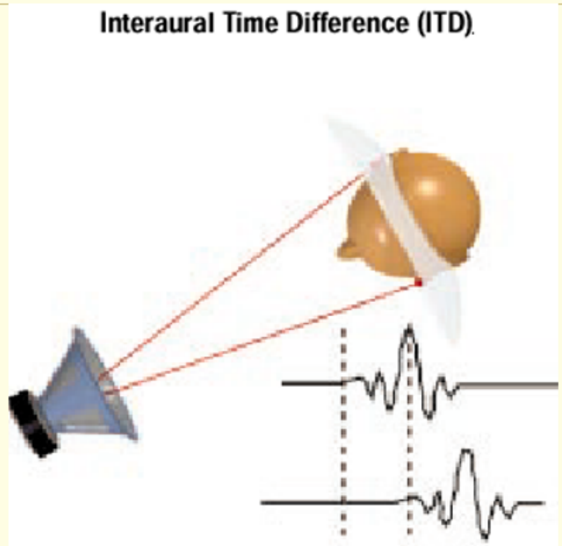
ITD for high freq is group delay rather than phase delay (roughly a distortion)

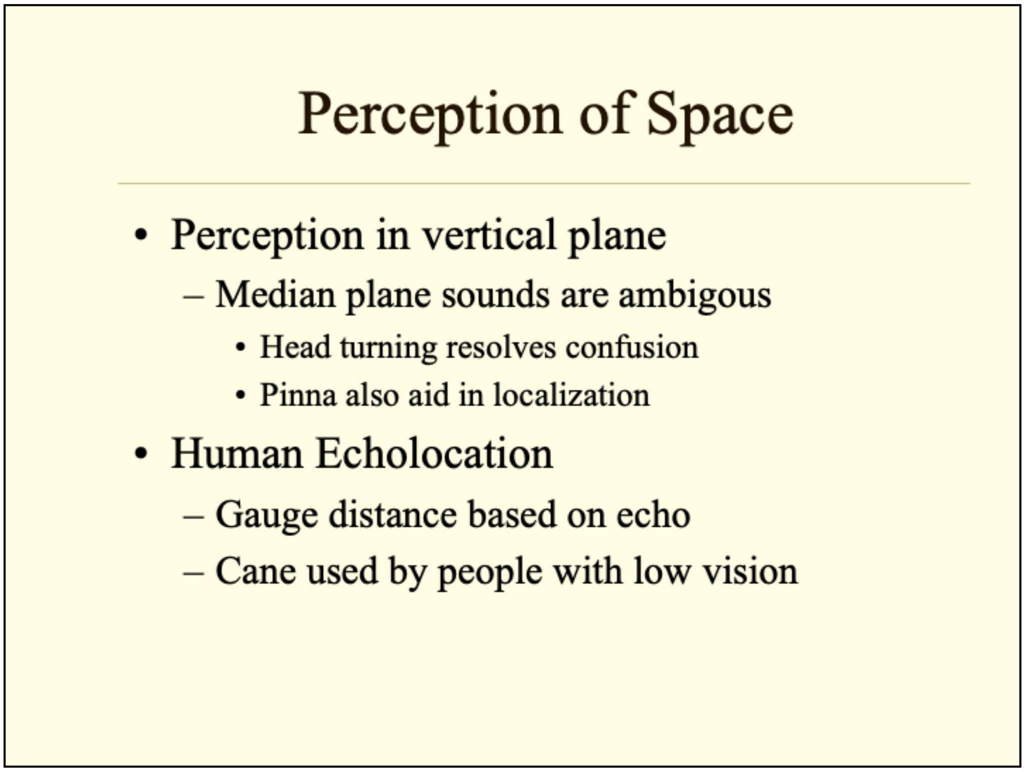
Time diff best for locatliing low freq….intensity best for high freq.

Interaural Intensity Difference



Interaural Time Difference





Big ears make for better localization ☺

Molds filling in the pinna caused for reduction in localization accuracy

But study also showed that this can be relearned somehow and accuracy returns

Move around pinna and notice change in sound…deeper and fuller if pulled forward…back flat…catch less sound waves and voice sounds thinner and higher in pitch….

Hard of hearing people cup hand around ear.

Echolocation: study with people listening to recorded audio of experimenter approaching an object. They were able to tell.

Tapping cane for the blind. Reflections also sound tells about surface.

Sighted people with blind folds are not very good at this, But blind people can figure out distance and size info based on self-produced sounds..also texture.

However, non-blind can also do stuff like this…estimating length of rod based on sound in makes when it hits the ground. And shape based on sound when object struck with hammer (circular, triangular, rectangular)

We also use sounds like…the frequency change as a container becomes filled with water.

Talk about clicker project

Parts of a Game Sound Track

- Dialog
  - Acoustics must be correct
- Sound Effects
  - Direct Narrative, can use sound to convey reality.
  - Subliminal Narrative, can also use sound to convey more complex themes that the viewer may not even be aware of
- Score
  - Must not drown out dialog and sound effects
  - Transitions and sequencing issues
  - Has market separate from game

Sound Effects

- Where do we get sound effects from?
  - Sound Libraries
  - Real Life
  - Foley, recorded sounds, probably not from the "real" event
  - Synthesized sounds

Foley Sound Effects

Sound Effects

- You don’t always want the “real” sound
- Sound must match viewers mental model of what the sound should sound like
  - This can be overdone
  - Subliminal Narrative
- And often the “real” sound doesn’t exist or is not available
- What is the sound of silence?
  - Total silence may sound strange or make people think there is a technical problem
  - “Even a small gap in an otherwise continuous noise can destroy the illusion of aural reality.”

Walkers: Machinists punch press

Tie: Elephant below…reversed and pitch shifted

R2 – mouth noises..electronics not human enough

Chewie – again, wanted emotion..pieced together lots of animals, like elephant seal

Laser – hammer on guide wire

Saber – waving mic behind TV

Sound in Video Games

- Audio capabilities used to lag behind graphics
  - Current consoles and computers have high quality audio
- Real-time issues (as with physics)
- How “correct” is your sound?
  - And is “correct” even possible?
- Acoustics
  - Dynamic simulation vs. Pre-rendered
    - Example: Environmental Audio Extensions (EAX)

Sound in Video Games

- Spatialization (filter to place mono sound in virtual space)
  - DirectSound
  - OpenAL
- How repetitious is your sound?
- How are multiple sounds handled (i.e. “ducking”)?
- How are transitions and unexpected changes in sound handled?
- Programmatic generation of sound effects
  - Physical modeling
    - car /plane engines, event-modeling
      - Road texture, tires, speed, doppler shift.
  - Dynamic selection of sound effects, filters, sound levels etc.
    - Layering and injecting randomness
  - Interactive music
    - DirectMusic
- Will your sounds become annoying in the first hour of play?
- Role is not just narrative, but functional as well
- Good sound can make your graphics look better

Game Audio Mistakes/Complaints

- http://www.gamasutra.com/blogs/AdamBishop/20090707/2288/Game_Audio_Mistakes.php
- Voices mixed too low
  - Employ proper equalization
- Confusing audio menu sliders
- Explosions too loud
- Audio integration later in dev
- Testing of audio with range of audio gear needed
- Others?