PsychoAcoustics Intro

Perry R. Cook
Princeton Computer Science
(also Music)

Views of Sound

- Sound is a recorded waveform
  PCM playback is all we need for
digital sound manipulation and creation
- Time Domain $x(t)$ (from physics)
- Frequency Domain $X(f)$ (from math)
- Production what caused it
- Perception our image of it

Views of Sound: Time Domain

Sound is produced/modelled by physics,
described by quantities of
- Force force = mass \times acceleration
- Position $x(t)$ actually $< x(t), y(t), z(t)>$
- Velocity Rate of change of position $dx/dt$
- Acceleration Rate of change of velocity $dv/dt$

Examples: Mass+Spring+Damper
Wave Equation

Views of Sound: Production

Throughout most of history, some physical mechanism was responsible
for sound production.

From our experience, certain gestures produce certain audible results

Examples:
- Hit harder $\rightarrow$ louder AND brighter
- Can’t move instantaneously

Sound Views: Frequency Domain

Frequency Domain:
- Many physical systems have modes
  (damped oscillations)
- Wave equation (2nd order) or
  Bar equation (4th order) need 2 or 4
  “boundary conditions” for solution
- Once boundary conditions are set
  solutions are sums of exponentially damped sines
  the sinusoids are Modes

Examples:
- Player
- Physics
- Constrain Strike
- Positions
Views of Sound: Perception

Human sound perception:

Ear: receive 1-D waves
Cochlea: convert to frequency dependent nerve firings
Nerves: further refine time & frequency information
Brain: higher level cognition, object formation, interpretation

Psychoacoustics

Limits of Human Hearing

- Time Domain Considerations
- Frequency Domain (Spectral) Considerations
- Amplitude vs. Power
- Masking in Time and Frequency Domains
- Sampling Rate and Signal Bandwidth

Limits of Human Hearing

Time and Frequency

Events longer than 0.03 seconds are resolvable in time
shorter events are perceived as features in frequency
20 Hz < Human Hearing < 20 KHz
(for those under 15 or so)

“Pitch” is PERCEPTION related to FREQUENCY
Human Pitch Resolution is about 40 - 4000 Hz.

Limits of Human Hearing

Amplitude or Power???

- “loudness” is PERCEPTION related to POWER, not AMPLITUDE
- Power is proportional to (integrated) square of signal
- Human Loudness perception range is about 120 dB,
  where +10 dB = 10 x power = 20 x amplitude
- Waveform shape is of little consequence.
  Energy at each frequency, and how that changes in time,
  is the most important feature of a sound.

Limits of Human Hearing

Waveshape or Frequency Content??

Here are two waveforms with identical power spectra, and which are (nearly) perceptually identical:

Wave 1

Wave 2

Magnitude Spectrum of Either

Limits of Human Hearing

Masking in Amplitude, Time, and Frequency

- Masking in Amplitude: Loud sounds ‘mask’ soft ones. Example: Quantization Noise
- Masking in Time: A soft sound just before a louder sound is more likely to be heard than if it is just after. Example (and reason): Reverb vs. “Preverb”
- Masking in Frequency: Loud ‘neighbor’ frequency masks soft spectral components. Low sounds mask higher ones more than high masking low.
Limits of Human Hearing

Masking in Amplitude
Intuitively, a soft sound will not be heard if there is a competing loud sound. Reasons:
• Gain controls in the ear
  *stapedes reflex and more*
• Interaction (inhibition) in the cochlea
• Other mechanisms at higher levels

Masking in Time
• In the time range of a few milliseconds:
  • A soft event following a louder event tends to be grouped perceptually as part of that louder event
  • If the soft event precedes the louder event, it might be heard as a separate event (become audible)

Masking in Frequency
Only one component in this spectrum is audible because of frequency masking

Each component exhibits a "masking curve"

References and Resources

General Psychoacoustics Books

Critical Bands and Masking
Old Views

Newer Views