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 <u>creation</u>
- ***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
is most closely related to
Production

Frequency Domain
is most closely related to
Perception

we will see that many hybrids abound

Views of Sound: Time Domain

Sound is produced/modeled by physics, described by quantities of

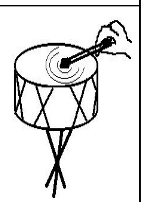
- Force — force = mass * acceleration
- Position — x(t) — actually < x(t), y(t), z(t) >
- Velocity — Rate of change of position dx/dt
- Acceleration — Rate of change of velocity dv/dt

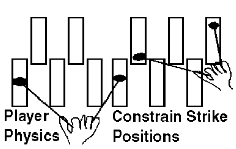
Examples: ***Mass+Spring+Damper***
Wave Equation

Views of Sound: Production

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

From our experience, certain <u>gestures</u> produce certain audible results





Examples:
Hit harder --> louder AND brighter
Can't move instantaneously
Can't do exactly the same thing twice

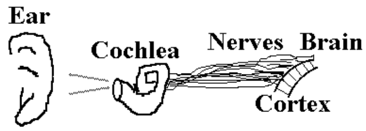
Sound Views: Frequency Domain

Frequency Domain:

- Many physical systems have <u>modes</u>
  (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

Views of Sound: Perception

Human sound perception:



Brain:
Higher level cognition, object formation, interpretation

Ear:
receive 1-D waves

Cochlea:
convert to frequency dependent nerve firings

Auditory cortex:
further refine time & frequency information

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 <u>PERCEPTION</u> related to <u>FREQUENCY</u>
Human Pitch Resolution is about 40 - 4000 Hz.

Limits of Human Hearing

Amplitude or Power???

- **"Loudness" is <u>PERCEPTION</u> related to <u>POWER</u>, not <u>AMPLITUDE</u>**
- **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

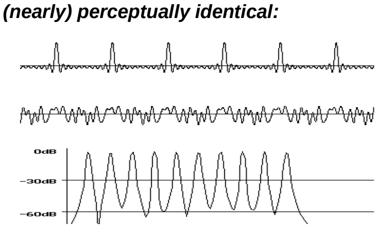
Waveshape or Frequency Content??

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**

Limits of Human Hearing

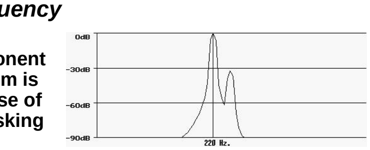
Masking in Time

- **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)**

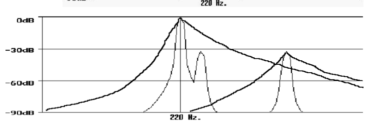
Limits of Human Hearing

Masking in Frequency

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



Each component exhibits a "masking curve"



Waveform Sampling and Playback

- *Sample and Hold*

  Sample Rate vs. Aliasing

- *Quantize*

Word Size vs.

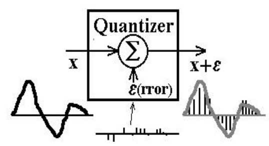
Quantization Noise

- *Reconstruct: (filter)*

  Hold and Smooth

  Filter Order vs.

  Error and Latency



References and Resources

General Psychoacoustics Books

Bregman, *Auditory Scene Analysis*, MIT Press, 1990.

Dowling and Harwood, *Music Cognition*, Academic Press, 1986.

Handel, *Listening: an Introduction to the Perception of Auditory Events*, MIT, Cambridge, MA, 1989.

McAdams and Bigand (eds.), *Thinking in Sound: the Cognitive Psychology of Human Audition*, Oxford Univ. Press, NY, 1993.

Pierce, *The Science of Musical Sound*, Freeman, New York, 1992.

Roederer, *Introduction to the Physics and Psychophysics of Music*, Springer-Verlag, New York, 1975.

Cook, ed. *Music, Cognition, and Computerized Sound*, MIT Press, 1999.

References and Resources

Critical Bands and Masking

Old Views

Zwicker, Flottorp, and Stevens, "Critical Bandwidth in Loudness Summation", J. Acoustical Soc. America 29, 1957.

Newer Views

Moore and Glasberg, "Suggested Formulae for Calculating Auditory-Filter Bandwidths and Excitation Patterns," JASA, 7, 4(3) 1983.