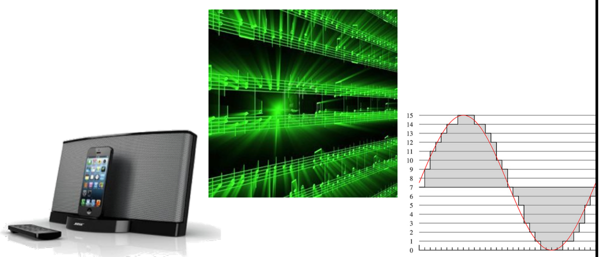


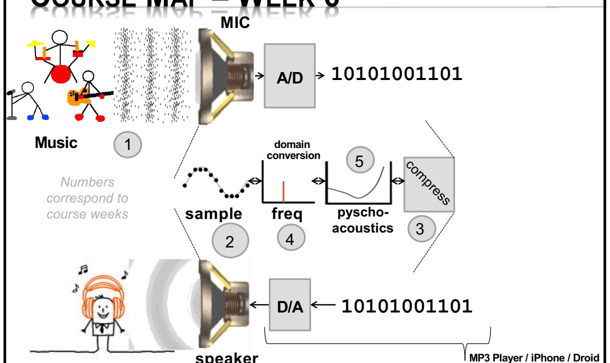
Lecture #10 – Psychoacoustics: Masking

ESE 150 – Digital Audio Basics

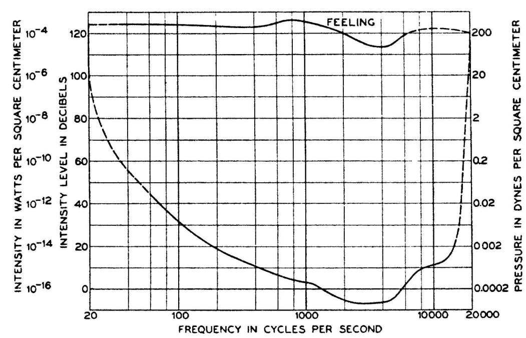
Lecture Topics

- Where are we on course map?
- Psychoacoustics
  - Review
  - Masking
- References

Course Map – Week 6



Lab 5: Measured own Fletcher Curve

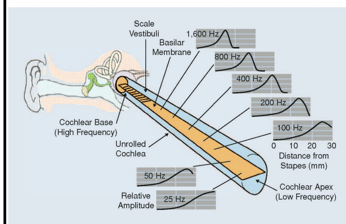


Lab 5: Measured Masking

- Saw (heard) that loud tones could mask softer, nearby frequencies.

Psychoacoustics

The Physical Ear – Take-Away

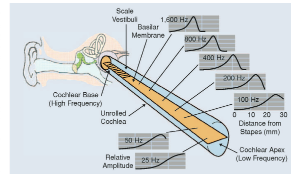


[FIG3] An illustration of an uncoiled cochlea. Due to the greater stiffness and smaller mass, the base of the basilar membrane is tuned to high frequencies while the apex resonates best with the low frequencies. The amplitude of the traveling waves across the membrane shows the frequency-to-place mapping.

- **Cochlea**
  - directly senses frequencies
  - Captures frequency domain
  - …not time domain
- **Frequency sensitive locations**
  - activated by sound waves
- **Neurons sense activation**

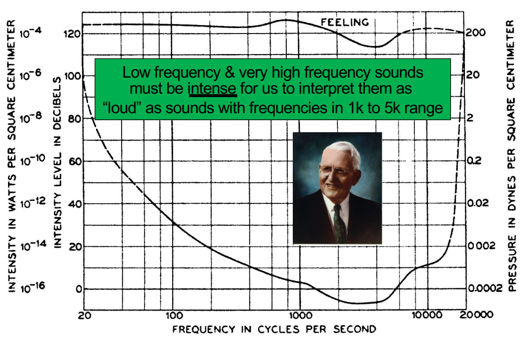
Picture above – uncoiled cochlea…
-- different stereovilli (Hairs) resonate at different frequencies
-- our ear performs Fourier Transform!

Physical Ear to Engineering Model



- **Limits of Human Hearing…easy to see from Cochlea**
  - Cochlea only so long…
    - lowest frequencies: 20 Hz
    - Highest frequencies: 20 kHz
- **Also helps us understand 'selectivity'**
  - Our brain can choose to 'listen' to output of various filters
  - *Example: At a party, but you can concentrate on conversation!*

Auditory Thresholds – Measured by Fletcher



Auditory Masking

Masking

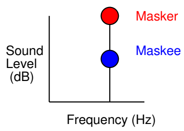
- **Auditory Masking**
  - When the perception of one sound is affected by the presence of another
    - Remember…perception
- **Two types:**
  - Frequency Domain Based:
    - Frequency Masking, simultaneous masking, spectral masking
  - Time Domain Based:
    - Temporal Masking / non-simultaneous masking

Frequency Domain Masking

- **Masking illustrates the limits of ear selectivity**
  - In fact, we measure ear selectivity using masking!
- **Vocabulary:**
  - Masker – The noise 'masking' the maskee
  - Maskee – The signal being 'masked' by masker

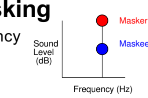
On-Frequency Masking

- **Greatest/worst form of frequency masking**
  - Occurs when maskee & masker are the same frequency

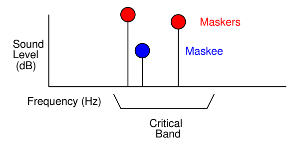


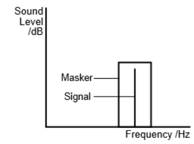
On-Frequency Masking

- **Greatest/worst form of frequency masking**
  - Occurs when maskee & masker are the same frequency



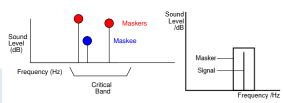
  - Also strong when masker and signal are within same auditory band (filter)

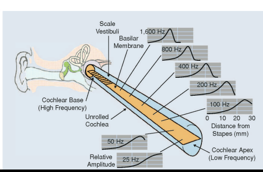




On-Frequency Masking

- **Greatest/worst form of frequency masking**
  - Occurs when maskee & masker are the same frequency
  - Also strong when masker and signal are within same auditory filter
  - Think about Cochlea vibrating:

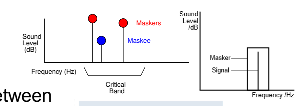


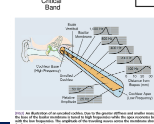


Nearby hair cells vibrating will cause some vibration.
Interpreted by brain as caused by large vibration, not a separate frequency.

On-Frequency Masking

- **Greatest/worst form of frequency masking**
  - Occurs when maskee & masker are the same frequency
  - Also strong when masker and signal are within same auditory filter
  - Think about Cochlea vibrating:

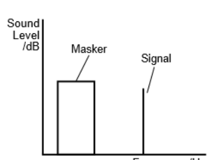


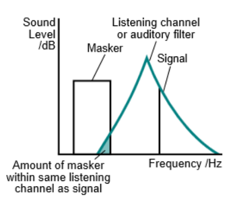


  - Listener cannot distinguish between them, perceived as one sound
  - Preclass 1: audience noise masks movie line

Off-Frequency Masking

- **Off-frequency masking**

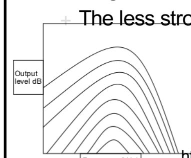




  - Masker has different frequency than signal
  - Masker still has effect…
    - …if it's in same auditory filter band as signal

Frequency Masking

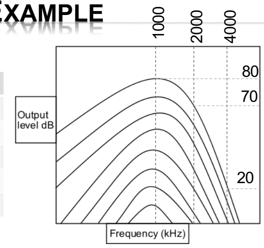
- **Given a signal at a frequency**
- **How strong must a signal (or noise) at a difference frequency be in order to be heard?**
- **General trend:**
  - Larger the frequency difference
  - The less strong it must be (the less masking)



https://commons.wikimedia.org/wiki/File:OutputlevelMoore.svg

Frequency Masking Example

Freq (Hz)	Strength (dB)
1000	80
2000	70
4000	20



https://commons.wikimedia.org/wiki/File:OutputlevelMoore.svg

Demonstration

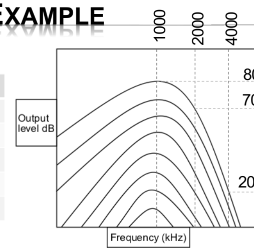
- **Generate 900 Hz Tone (left channel) (maskee)**
  - Turn gain all the way down (-36 dB)
- **Generate 1000 Hz Tone (right channel) (masker)**
  - Keep gain at 0 dB
- **Play sound…**
  - Bring intensity of 900 Hz tone up so we can hear both tones
  - Mute masker and play it again…
    - Maskee was always there, just couldn't hear it
    - Even though it was at different frequency of masker

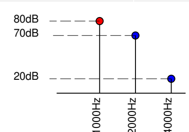
Demonstration

- **Generate 1000 Hz Tone (masker)** [band 9]
- **Sweep frequency 700Hz to 1600 Hz (masked)**
  - About 20% of level of masker
  - Bands 7--11
- **Both constant loudness**
- **Reference without Masker**
- **Play sound…**
  - When hear second signal?
- **See diminished masking effects as frequencies get further apart**

Frequency Masking Example

Freq	Strength
1000	80
2000	70
4000	20

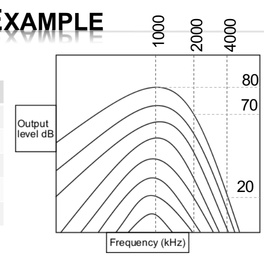


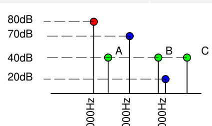


https://commons.wikimedia.org/wiki/File:OutputlevelMoore.svg

Frequency Masking Example

Freq	Strength
1000	80
2000	70
4000	20



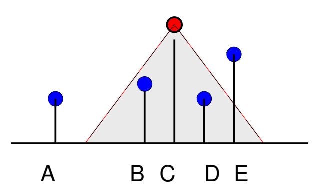


Which of A, B, C are masked?

https://commons.wikimedia.org/wiki/File:OutputlevelMoore.svg

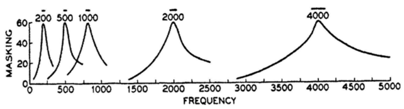
Preclass 2

- **Which frequencies should we keep?**



Frequency Masking @ Higher Frequencies

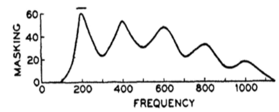
- **Plots of masking at several different frequencies:**



  - Effect of masking is 'worse' at higher frequencies
  - Masking band gets wider at higher frequencies

Frequency Masking and Harmonics

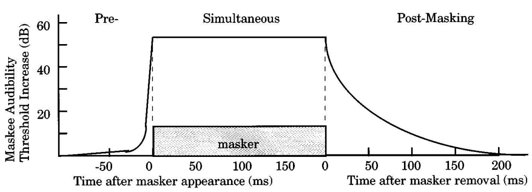
- **Masking can also occur at the harmonics of masker…**



  - Example has a masker at 200 Hz
  - While effect of masker is greatest at 200 Hz…
    - Also effects harmonics of masker signal!

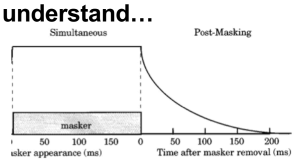
Time-Domain Masking (Temporal)

- **Two types:**
  - pre-masking (backwards)
  - post-masking (forwards)



Temporal Masking - Forwards

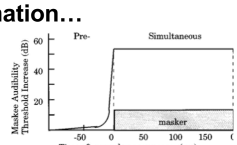
- **Easier to understand…**



  - A sudden masker noise…
    - Makes inaudible other sounds following noise…for up to 200ms
      - Physical: hair cells in Cochlea don't stop vibrating instantly
        - Brain accounts for the fact their vibration will decay over time after incident sound goes away
      - (imagine the compression possibilities!)

Temporal Masking - Backwards

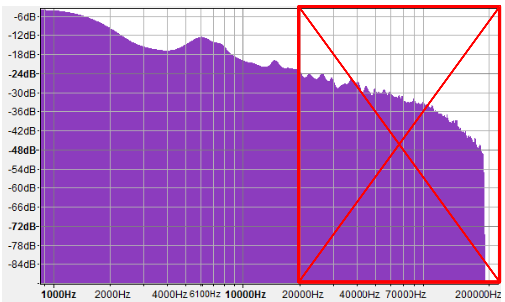
- **Not as intuitive an explanation…**



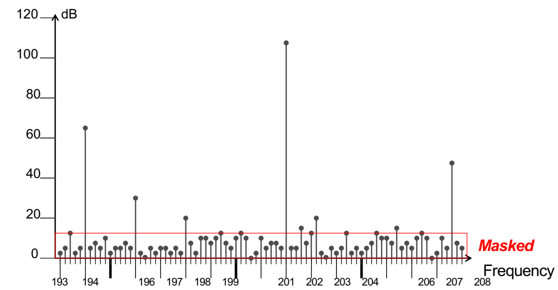
  - A sudden masker noise…
    - Makes inaudible other sounds preceding noise!
    - Why does this happen?
      - One thought: takes time for your brain to interpret sound
      - Think of it like a buffer…
      - Throws out contents of buffer when a loud sound comes in
        - to concentrate on only the loud sound (masker in this case)
      - Also, hair vibrations likely take time to come up to full amplitude

Using Psychoacoustics in Digital Audio

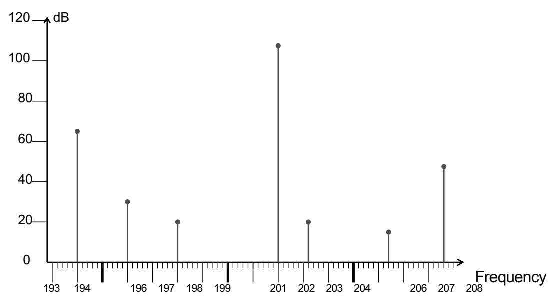
How do we use psychoacoustics in digital music compression? (Range)



How do we use psychoacoustics in digital music compression? (Masking)

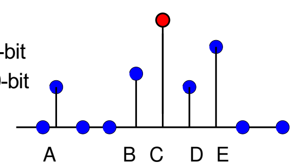


How do we use psychoacoustics in digital music compression? (Masking)



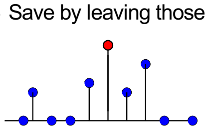
Preclass 3

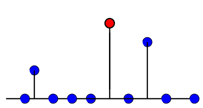
- **Same A, B, C, D, E from preclass 2**
  - With same masking effects
- **8b per frequency, 10 frequencies → bits?**
- **Non-zero, non-masked frequencies?**
- **Lossless encode Bits?**
  - Encode 0s/masked with 1-bit
  - Encode keep/non-0 with 9-bit



Basic Freq/Masking Compression Idea

- **Convert to frequency domain**
- **If few frequencies**
  - Cheaper to only represent those
- **Masking means can drop frequencies that are present, but not dominant**
  - Save by leaving those out





Big Ideas

- **Human hearing mechanism directly encodes frequency**
  - By position on Cochlea
- **Differential sensitivity by frequency**
  - Hear some frequencies louder than others
- **Frequency Masking**
  - Limit to what we can simultaneously perceive in critical bands – loud frequencies can hide others
- **Temporal Masking**
  - Loud signals can hide sounds that come after (or before) them

Learn More

- **BIBB417 – Visual Processing**
  - Same kind of look at physiology, but for vision
- **LING520 – Phonetics 1**
  - Focus on speech, includes both hearing and production

Midterm – Online Quiz

- **Open book, open notes**
- **Calculators allowed**
- **2 hours**
- **Anytime between 9am and 11pm EST Wed. 3/3**
- **5% of grade**
  - prepare for final
- **Last 3 year's midterm and answers**
  - on 2018, 2019, 2020 syllabus
  - Were all in-person, closed book
  - 2020 final was Online Quiz, also on syllabus
- **Post "rules" for exam before Friday**

Topics

- Data representation in bits
- Sounds waves
- Sampling
- Quantization
- Nyquist
- Lossy/lossless compression
- Common case
- Frequency domain
- Psychoacoustics
- Perceptual coding

Coming Up

- **Feedback – including Lab 5**
- **Next Lecture**
  - Put this together to compress audio
  - Start deriving key features of MP3 (finish next Friday)
- **Lab 5 report this Friday as usual**
- **Lab 6 on Monday – start of big, formal lab**
  - Lab now posted
- **Midterm next Wednesday**
  - No lecture
- **No lab report next Friday**

References

- **Physical Ear:**
  - R. Munkong and B.-H. Juang. IEEE Sig. Proc. Mag., 25(3):98–117, 2008
- **Filter Bank:**
  - http://www.ugr.es/~atv/web_ci_SIM/en/seccion_4_en.htm
- **Bark Scale:**
  - [E. Zwicker. J. Acoust. Soc.Am., 33(2):248, February 1961]
- **DB Chart:**
  - http://www.dspguide.com/ch22/1.htm
- **Masking Discussion:**
  - Wikipedia: PsychoAcoustics Article