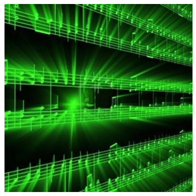
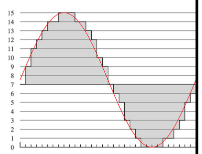
Penn Engineering ESE





Lecture #10 – Psychoacoustics: Masking

ESE 150 – DIGITAL AUDIO BASICS

1

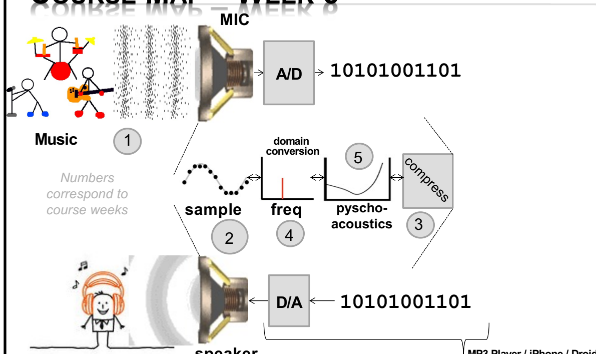
LECTURE TOPICS

- **Part 1**
  - Where are we on course map?
  - Review
  - Masking: Frequency or Simultaneous
- **Part 2**
  - Masking: Time-Domain or Temporal
  - Preview Use in Compression
- **References**

2

2

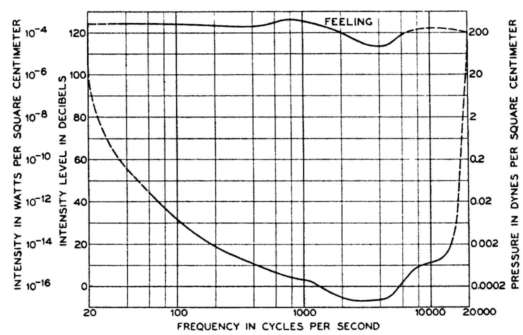
COURSE MAP – WEEK 6



3

3

LAB 5: MEASUR OWN FLETCHER CURVE



4

4

LAB 5: MEASURE MASKING

- See (hear) that loud tones could mask softer, nearby frequencies.

5

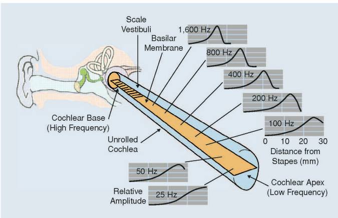
5

PSYCHOACOUSTICS

6

6

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.

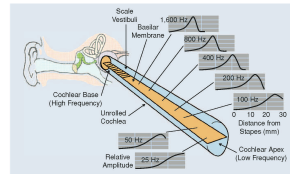
- **Cochlea**
  - directly senses frequencies
  - Captures frequency domain
  - …not time domain
- **Frequency sensitive locations**
  - activated by sound waves
- **Neurons sense activation**

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

7

7

Physical Ear to Engineering Model

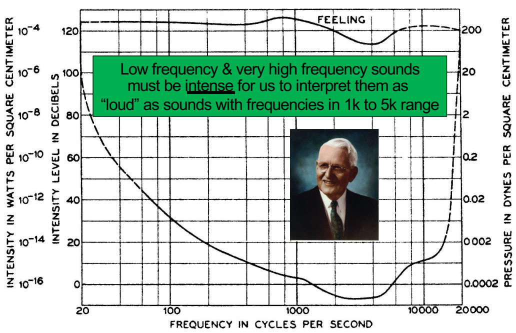


- **Limits of Human Hearing…easy to see from Cochlea**
  - Cochlea only so long…
    - lowest frequencies: 20 Hz
    - Highest frequencies: 20 kHz
- **Also helps us understand 'selectivity'**
  - Our brain can choose to 'listen' to output of various filters
  - *Example: At a party, but you can concentrate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

8

8

Auditory Thresholds – Measured by Fletcher



Low frequency & very high frequency sounds must be intense for us to interpret them as "loud" as sounds with frequencies in 1k to 5k range

9

9

Critical Frequency Bands

- **"Bark" scale –**
  - Maps frequency intervals into their respective critical band number

Number	Center Freq. (Hz)	Cut-off Freq. (Hz)	Bandwidth (Hz)
		20	
1	50	100	80
2	150	200	100
3	250	300	100
4	350	400	100
5	450	510	110
6	570	630	120
7	700	770	140
8	840	920	150
9	1000	1080	160
10	1170	1270	190
11	1370	1480	210
12	1600	1720	240

Number	Center Freq. (Hz)	Cut-off Freq. (Hz)	Bandwidth (Hz)
13	1850	2000	280
14	2150	2320	320
15	2500	2700	380
16	2900	3150	450
17	3400	3700	550
18	4000	4400	700
19	4800	5300	900
20	5800	6400	1100
21	7000	7700	1300
22	8500	9500	1800
23	10500	12000	2500
24	13500	15500	3500

10

10

Penn Engineering

ESE

Auditory Masking

11

11

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

12

12

Frequency Domain Masking

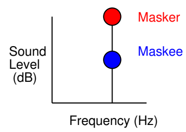
- **Masking illustrates the limits of ear selectivity**
  - In fact, we measure ear selectivity using masking!
- **Vocabulary:**
  - Masker – The noise 'masking' the maskee
  - Maskee – The signal being 'masked' by masker

13

13

On-Frequency Masking

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

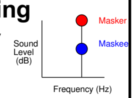


14

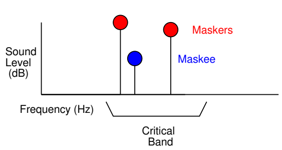
14

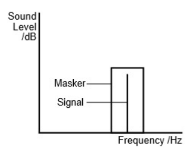
On-Frequency Masking

- **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)



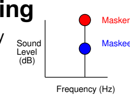


15

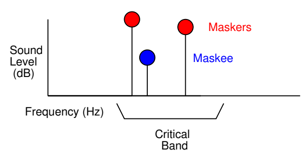
15

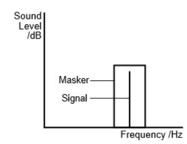
On-Frequency Masking

- **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

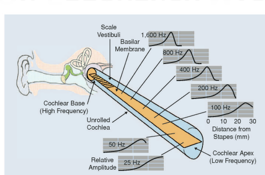




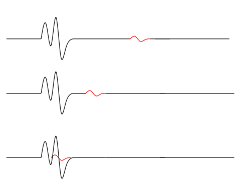
16

16

On-Frequency Masking



[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.



Think about Cochlea vibrating.

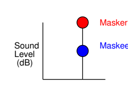
Nearby hair cells vibrating will cause some vibration.
Interpreted by brain as caused by large vibration, not a separate frequency.

17

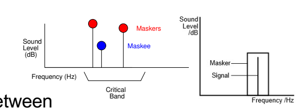
17

On-Frequency Masking

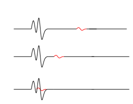
- **Greatest/worst form of frequency masking**
  - Occurs when maskee & masker are the same frequency



  - 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

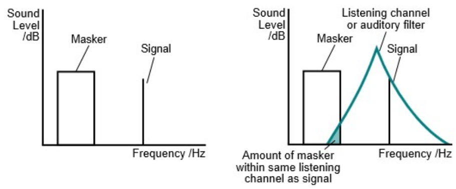


18

18

Off-Frequency Masking

- Off-frequency masking

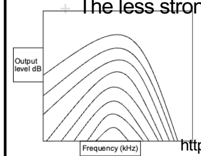


  - Masker has different frequency than signal
  - Masker still has effect…
    - …if it's in same auditory filter band as signal

19

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)

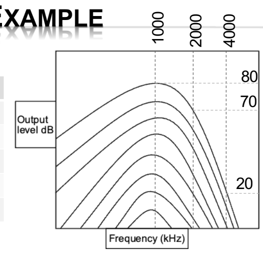


https://commons.wikimedia.org/wiki/File:OutputlevelMoore.svg

20

Frequency Masking Example

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



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

21

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

22

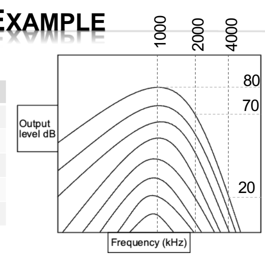
Demonstration

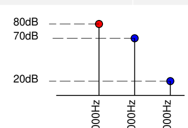
- **Generate 1000 Hz Tone (masker)** [band 9]
- **Sweep frequency 700Hz to 1600 Hz (masked)**
  - About 20% of level of masker
  - Bands 7--11
- **Both constant loudness**
- **Reference without Masker**
- **Play sound…**
  - When hear second signal?
- **See diminished masking effects as frequencies get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

23

Frequency Masking Example

Freq	Strength
1000	80
2000	70
4000	20



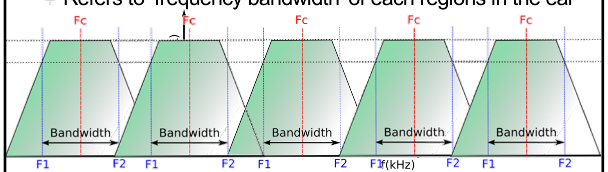


https://commons.wikimedia.org/wiki/File:OutputlevelMoore.svg

24

Physical Ear – Limits of Human Perception

- **Critical Frequency Bands**
  - Refers to 'frequency bandwidth' of each regions in the ear

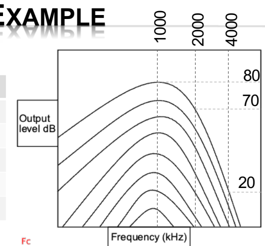


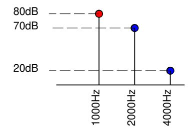
- A 'sharply tuned' filter has good frequency resolution
  - Allows frequencies in band pass well, but not others
  - Brain can then 'resolve' different frequencies

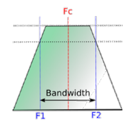
25

Frequency Masking Example

Freq	Strength
1000	80
2000	70
4000	20





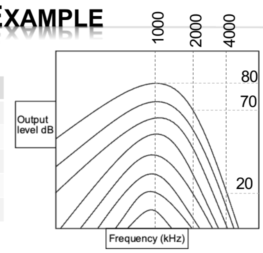


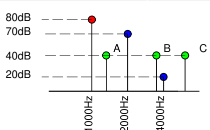
https://commons.wikimedia.org/wiki/File:OutputlevelMoore.svg

26

Frequency Masking Example

Freq	Strength
1000	80
2000	70
4000	20





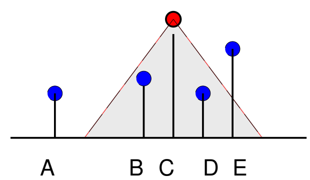
Which of A, B, C are masked by 1000Hz (red) signal?

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

27

Preclass 2

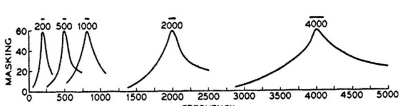
- **Which frequencies should we keep?**



28

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

29

Critical Frequency Bands – How many?

- **"Bark" scale –**
  - Maps frequency intervals into their respective critical band number
  - 24 frequency bins (or "barks"), get wider as frequency increases!

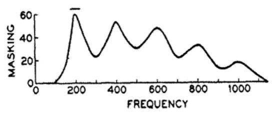
Number	Center Freq. (Hz)	Cut-off Freq. (Hz)	Bandwidth (Hz)
		20	
1	50	100	80
2	150	200	100
3	250	300	100
4	350	400	100
5	450	510	110
6	570	630	120
7	700	770	140
8	840	920	150
9	1000	1080	160
10	1170	1270	190
11	1370	1480	210
12	1600	1720	240

Number	Center Freq. (Hz)	Cut-off Freq. (Hz)	Bandwidth (Hz)
13	1850	2000	280
14	2150	2320	320
15	2500	2700	380
16	2900	3150	450
17	3400	3700	550
18	4000	4400	700
19	4800	5300	900
20	5800	6400	1100
21	7000	7700	1300
22	8500	9500	1800
23	10500	12000	2500
24	13500	15500	3500

30

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!

31

31

Penn Engineering

ESE

Part 2

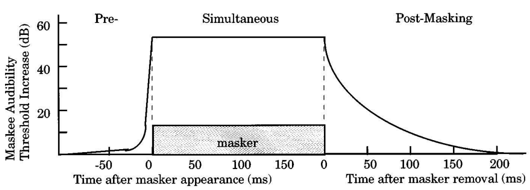
Temporal Masking

32

32

Time-Domain Masking (Temporal)

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

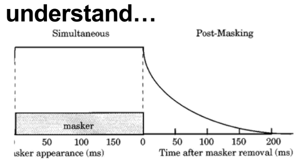


33

33

Temporal Masking - Forwards

- **Easier to understand…**

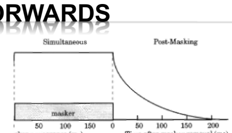


  - A sudden masker noise…
    - Makes inaudible other sounds following noise…for up to 200ms
      - Physical: hair cells in Cochlea don't stop vibrating instantly
        - Brain accounts for the fact their vibration will decay over time after incident sound goes away

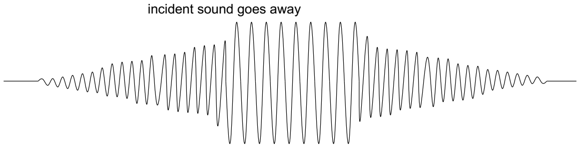
34

34

Temporal Masking - Forwards



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

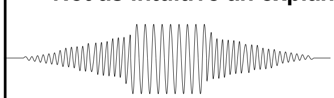


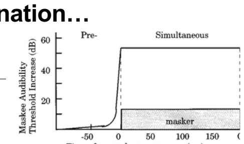
35

35

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

36

36

Penn Engineering ESE

Using Psychoacoustics in Digital Audio

37

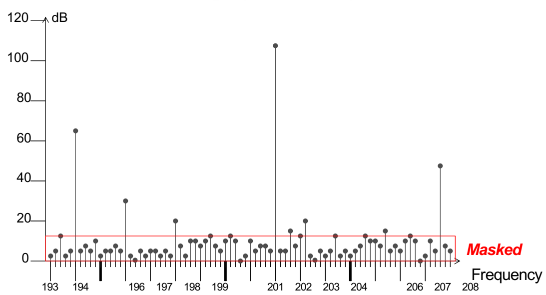
37

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

38

38

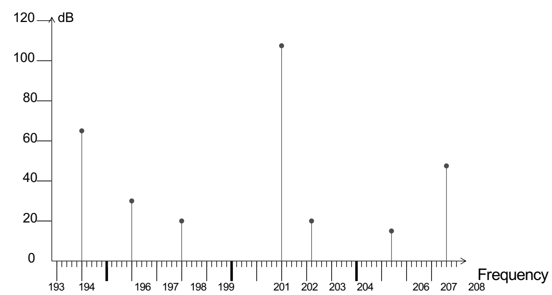
How do we use psychoacoustics in digital music compression? (Masking)



39

39

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

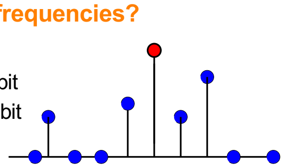


40

40

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

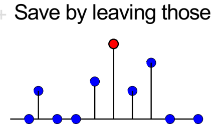


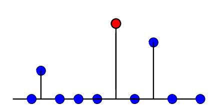
41

41

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





42

42

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

43

43

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

44

44

Admin / Coming up

- **Feedback**
- **Lab today: psychoacoustics**
- **Next Lecture**
  - Put this together to compress audio
  - Start deriving key features of MP3 (finish next Friday)

45

45

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

46

46