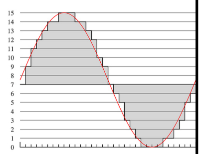


Lecture #5 – Psychoacoustics

ESE 150 – DIGITAL AUDIO BASICS

Based on slides © 2009—2020 DeHon, Koditschek
Additional Material © 2014 Farmer

1

TEASER

- Can hear whistle?
- Can you hear mouse eliminator?

2

OBSERVE

- There are sounds we cannot hear
  - Depends on frequency

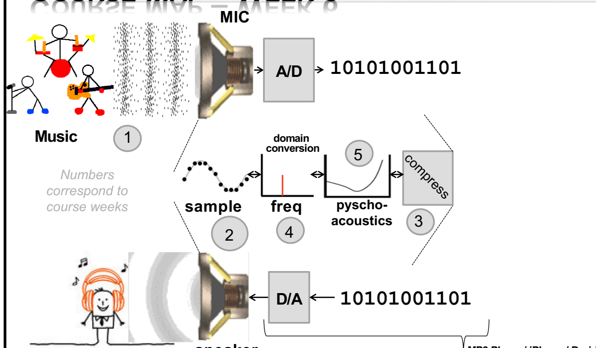
3

LECTURE TOPICS

- Where are we on course map?
- What we did in lab last week
- Psychoacoustics
  - Structure of Human Ear / encoding signals to brain
  - Human Hearing Limits
  - Critical Bands (Frequency bins)
  - Masking
- Next Lab
- References

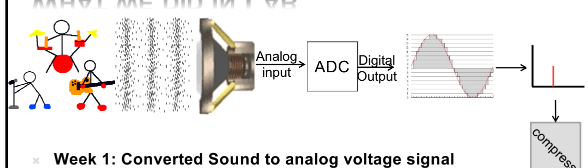
4

COURSE MAP – WEEK 6



5

WHAT WE DID IN LAB…



- **Week 1: Converted Sound to analog voltage signal**
  - a "pressure wave" that changes air molecules w/ respect to time
  - a "voltage wave" that changes amplitude w/ respect to time
- **Week 2: Sampled voltage, then quantized it to digital sig.**
  - Sample: Break up independent variable, take discrete 'samples'
  - Quantize: Break up dependent variable into n-levels (need 2^n bits to digitize)
- **Week 3: Compress digital signal**
  - Use even less bits without using sound quality!
- **Week 4: Frequency Domain Transform before we compress…**
  - Put our 'digital' data into another form…BEFORE we compress…less stuff to compress!

6

Psychoacoustics

What is Psychoacoustics?

- **Scientific study of sound perception**
  - Branch of science studying the psychological and physiological responses associated with sound
  - Also, considered a branch of: psychophysics
  - Human physical (and neurological) mechanism for sound perception
- **Why study sound & human's perception?**
  - Example: FREQUENCY vs. PITCH
    - Frequency of sound: "how often" air particles vibrate (Hz)
    - Pitch of sound: the sensation of frequency
      - How our brains "interpret" the frequency of a sound
- **Things may "sound" one way…**
  - …but be interpreted by our brains very differently!

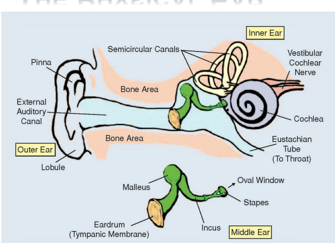
Psychoacoustics & Digital Music

- **How does psychoacoustics relate to MP3?**
- **The "consumer" of an MP3 is the human ear…**
  - Knowing more about brain's interpretation of sound…
  - …helps us remove things human's can't hear anyway
- **We've used some of this in our system already:**
  - Limit of human perception of sound: 20 Hz to 20,000 Hz
    - We put an anti-aliasing filter limiting incoming audio
  - Fixes our sampling rate, less data to store as a result!

Our Study of Psychoacoustics

- **Structure of Human Ear / encoding signals to brain**
- **Human Hearing Limits**
- **Critical Bands**
- **Frequency Bins**
- **Masking (Spatial vs. Temporal)**
- **Applied Psychoacoustics (mostly next lecture)**
  - Using all of the above to build...the "Psychoacoustical Model"
  - Perceptual Coding in MP3 (using the model to compress MP3s)

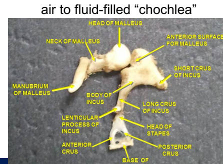
The Physical Ear



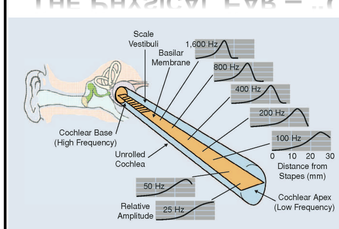
[FIG2] The cross section view of the ear displaying the components of the outer, the middle, and the inner ear.

Converts vibrations in air (sound) into mechanical motion in the middle-ear!

- **Outer Ear**
  - Guides sound waves into 'auditory canal'
- **Middle Ear**
  - Ear Drum – transmits sound from air to 3 bones in inner ear: ossicles (*hammer, anvil, stirrup*)
- **Inner Ear**
  - Ossicles – transmit sound from air to fluid-filled "chochlea"

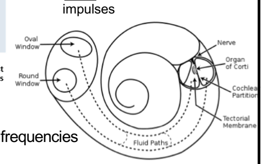


The Physical Ear – "Cochlea"



[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 – "snail shell"**
  - Fluid-filled "labyrinth"
  - Located in: "inner-ear"
  - Spiral Shaped (snail shell)
  - *Hair* inside cochlea 'resonates' according to incoming vibrations in the liquid
    - Stereovilli (name of hair)
  - *Hairs* convert vibration into nerve impulses



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

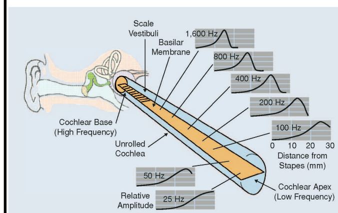
http://www.youtube.com/watch?v=zeg4qTnYOpw

Cochlea Animation

- **https://www.youtube.com/watch?v=dyenMluFaUw**

13

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 Fourier 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!

14

Take-away

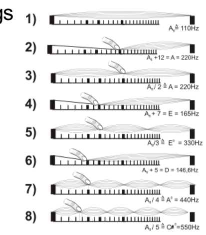
- **Our ear works in the frequency domain.**
- **We could consider devices that**
  - Directly recorded frequencies
    - Collection of resonators?
  - Directly produced frequencies
    - Collection of vibrators
      - Tuning forks
      - Strings
      - Pipes
      - …sound familiar?

15

Direct Frequency Generation

- **All traditional musical instruments work that way!**
  - Piano, guitar, violin – vibrating strings

Feliciano Guimarães from Guimarães, Portugal / CC BY (https://creativecommons.org/licenses/by/2.0)

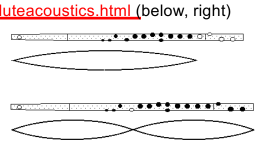


By user:Mjchael - made by me (Corel-Draw), CC BY-SA 2.5, https://commons.wikimedia.org/w/index.php?curid=1251597

More: Lhttps://www.zmescience.com/science/physics/guitar-strings-vibrate/ 16

Direct Frequency Generation

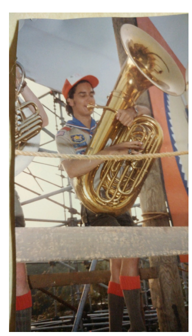
- **All traditional musical instruments work that way!**
  - Piano, guitar, violin – vibrating strings
  - Flute, trumpet, pipe organ – pipes
    - http://newt.phys.unsw.edu.au/jw/woodwind.html
    - http://newt.phys.unsw.edu.au/jw/fluteacoustics.html (below, right)

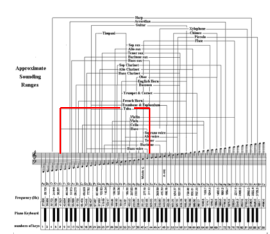


By Photo by Yasuhiko Sano, Nov 2005 - http://homepage2.nifty.com/iwatake/, Public Domain, https://commons.wikimedia.org/w/index.php?curid=435915

17

Time Domain Advantage?



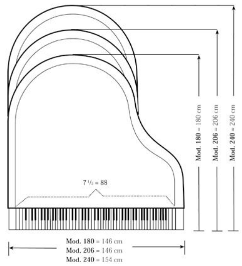


Larry Solomn:
http://solomonsmusic.net/insrange.htm

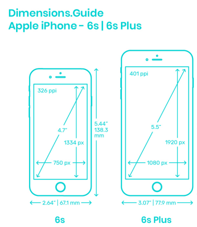
18

Time Domain Advantage?

1.5m x 2.4m



0.007m x 0.014m

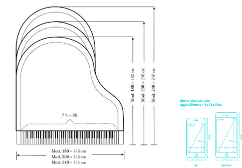


http://www.bluebookofpianos.com/types.html

https://www.dimensions.guide/element/apple-iphone-6s-6s-plus

Time Domain Advantage?

- **Can produce (receive) many frequencies**
  - Without large number of strings (vibrators)
  - Without large footprint for strings and resonant caveties
- **Smaller/cheaper**
  - Exploiting cheep processing from Moore's Law

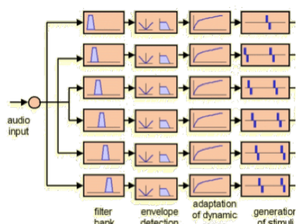


http://www.bluebookofpianos.com/types.html

https://www.dimensions.guide/element/apple-iphone-6s-6s-plus

Physical Ear to Engineering Model

- **With knowledge of structure/function of ear:**
  - We can model cochlea's behavior as bank of filters / bandpass filters
    - Cochlea breaks down auditory input into frequency ranges
    - Sends different frequencies down different nerve pathways!

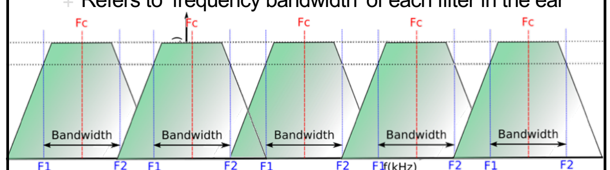


Each Frequency encoded independently on the auditory nerve

Brain ultimately "interprets" these Encoded signals as sound

Physical Ear – Limits of Human Perception

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

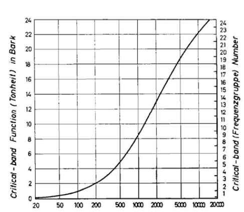


  - A 'sharply tuned' filter has good frequency resolution
    - Allows frequencies in band pass well, but not others
    - Brain can then 'resolve' different frequencies

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 frequencies Hz	Cut-off frequencies 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
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	10 500	12 000	2500
24	13 500	15 500	3500



OK, now…some tests…

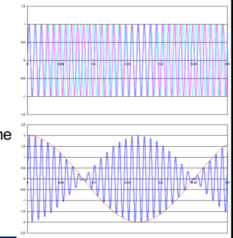
- **How well can you hear? (range)**
  - 20 Hz to 20kHz – frequency increasing over 20 seconds
  - Can you hear tone the entire time, or do is it appear to go silent at some point?
  - Tells you how high (and maybe how low) of frequencies you can hear
    - Probably need to switch to audacity on laptop to see when still playing…

OK, NOW…SOME TESTS…

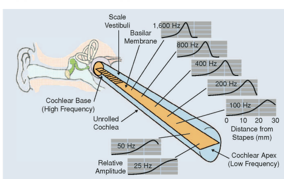
- **Can you hear two frequencies at once? (selectivity)**
  - Let's try: 400 Hz and 1000 Hz
  - First independent references
  - Then together

OK, NOW…SOME TESTS…

- **Frequency Resolution…(bands)**
  - In 1000 Hz to 2000 Hz octave…
    - Brain can't perceive changes in frequency
    - smaller than 3.6 Hz
  - Plays 1500Hz tone then 1502Hz
    - Aside from maybe a click in the middle, can you tell difference between tones?
    - Keep same tones, but add a 1500Hz tone on second track playing whole time.
    - Now hear interference demonstrating that first track did change tones.



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!*

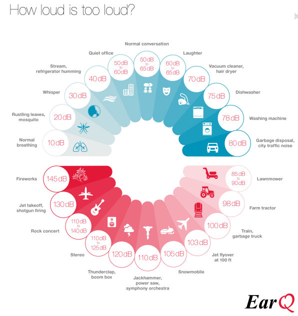
SOUND INTENSITY & LOUDNESS

SOUND INTENSITY – "LOUDNESS"

- **But first, we must discuss: deci-*Bels* (dB)**
  - Logarithmic unit in engineering: compare levels (fractions)
  - Compare two physical quantities: power, intensity, etc
  - Often compare quantity to a reference value
- **Sound and (dB)**
  - Sound is compression/expansion of air
  - We use (dB) to compare two air pressures in acoustics:
    - Lowest limit of human ear sensitivity: 20 μPascals (μPa)
      - We compare all sounds to this lower limit (reference sound pres.)

Loudness-Sound Pressure Level (L_{SPL}) = $20 * log_{10}\left(\frac{Sound\ pressure}{Reference\ Sound\ pressure}\right)$ in **dB**

SOUND INTENSITY IN (DB) – "LOUDNESS"



- **Preclass 4: Ratio of pressure between 20dB and 140dB?**

Sound Intensity in (dB) – "Loudness"

Decibels SPL	Example Sound
140 dB	Pain
130 dB	
120 dB	Discomfort
110 dB	Jack hammers and rock concerts
100 dB	
90 dB	OSHA limit for industrial noise
80 dB	
70 dB	
60 dB	Normal conversation
50 dB	
40 dB	Weakest audible at 100 Hz
30 dB	
20 dB	Weakest audible at 10 kHz
10 dB	
0 dB	Weakest audible at 3 kHz
-10 dB	
-20 dB	

- **If sound intensity level is:**
- **140 dB** **20 dB**

$20log\left(\frac{P2}{P0}\right) = 140dB$ $20log\left(\frac{P1}{P0}\right) = 20dB$

- **Divide both sides by 20:**

$log\left(\frac{P2}{P0}\right) = 7$ $log\left(\frac{P1}{P0}\right) = 1$

$\frac{P2}{P0} = 10^7$ $\frac{P1}{P0} = 10^1$

$\frac{P2}{P1} = 1{,}000{,}000$

- **Sound with intensity of 140dB:**
  - *has a sound pressure 10 million times greater than the quietist sound we can hear (which is 20 uPa) -- OUCH!*

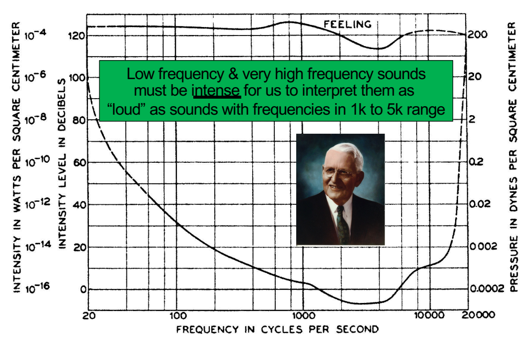
31

Sound Intensity in (dB) – "Loudness"

- **Loudness –**
  - subjective perception of intensity of sound
- **Intensity –**
  - Sound power per unit area
- **Does loudness change with frequency?**
  - Yes! Scientist: Harvey Fletcher (1940)
    - Measured loudness vs. frequency (Auditory Thresholds)
    - Same 'amplitude' sound can sound very quite or really loud
      - All depends on its frequency
  - Turns out…
    - We are very sensitive to frequencies from 1kHz to 5kHz
      - They don't have to be 'intense' for us to hear them…why??

32

Auditory Thresholds – Measured by Fletcher



33

Demonstration

- **Same demo as before: 1 Hz to 20kHz**
  - Instead of thinking about frequency cutoff (range)
  - Think instead about how "loud" the sounds at different frequencies are…
    - Which 'band' sounds 'loudest' to you?
    - Note: they are all at same amplitude, so equally intense
    - But we perceive sounds in 1 kHz to 5 kHz to be louder!

34

Auditory Masking

35

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

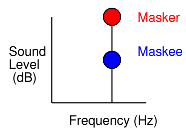
36

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

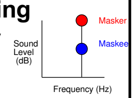
On-Frequency Masking

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

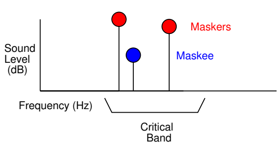


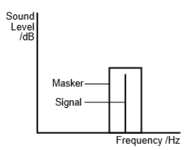
On-Frequency Masking

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



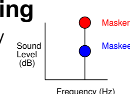
- Also strong when masker and signal are within same auditory filter



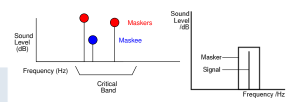


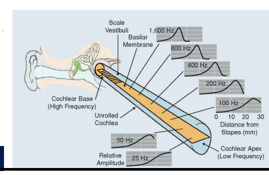
On-Frequency Masking

- **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:

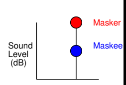




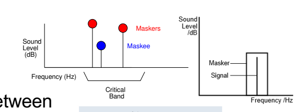
Nearby hair cells vibrating will cause some vibration.
Interpreted by brain as caused by large vibration not a separate frequency.

On-Frequency Masking

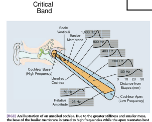
- **Greatest/worst form of frequency masking**
  - Occurs when maskee & masker are the same frequency



- Also strong when masker and signal are within same auditory filter
- Think about Cochlea vibrating:

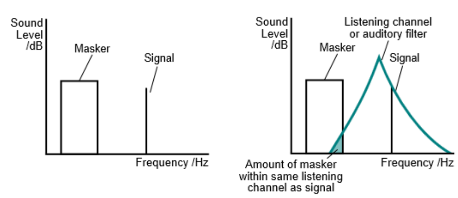


- Listener cannot distinguish between them, perceived as one sound
- Preclass 5: audience noise masks movie line



Off-Frequency Masking

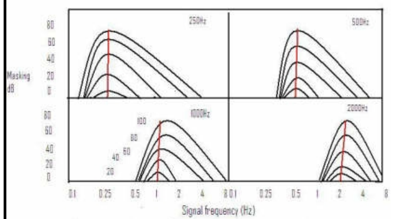
- **Off-frequency masking**



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

Frequency Masking and Intensity

- **Effect of masking as masker signal grows more intense**



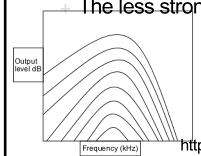
Plotting maskee amplitude vs bandwidth

With separate curve for each masker amplitude

  - More intense the masker, wider the band of masking
    - Note: These "masking patterns" are called: audiograms

Frequency Masking

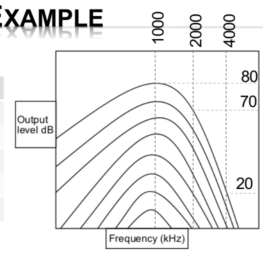
- **Given a signal at a frequency**
- **How strong must a signal (or noise) at a difference frequency be in order to be heard?**
- **General trend:**
  - Larger the frequency difference
  - The less strong it must be (the less masking)



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

Frequency Masking Example

Freq	Strength
1000	80
2000	70
4000	20



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

Demonstration

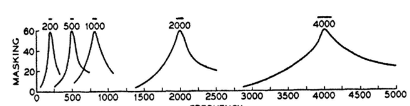
- **Generate 900 Hz Tone (left channel) (maskee)**
  - Turn gain all the way down (-36 dB)
- **Generate 1000 Hz Tone (right channel) (masker)**
  - Keep gain at 0 dB
- **Play sound…**
  - Bring intensity of 900 Hz tone up so we can hear both tones
  - Mute masker and play it again…
    - Maskee was always 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 @ Higher Frequencies

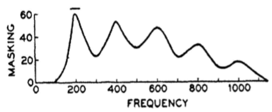
- **Plots of masking at several different frequencies:**



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

Frequency Masking and Harmonics

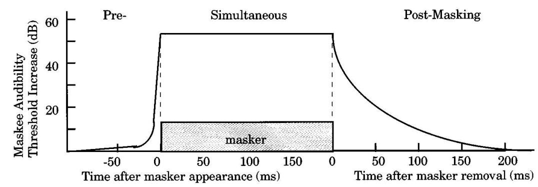
- **Masking can also occur at the harmonics of masker…**



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

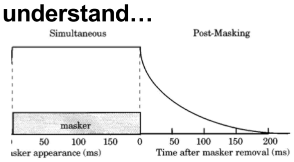
Time-Domain Masking (Temporal)

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



Temporal Masking - Forwards

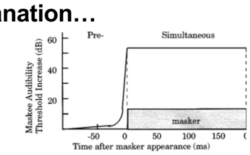
- **Easier to understand…**



- 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

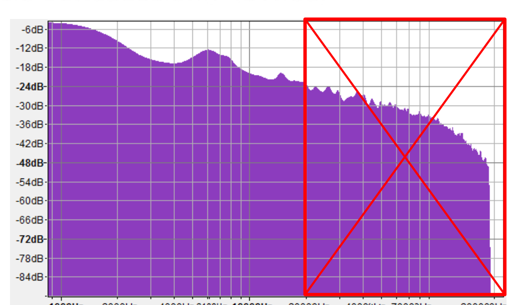
- **Not as intuitive an explanation…**



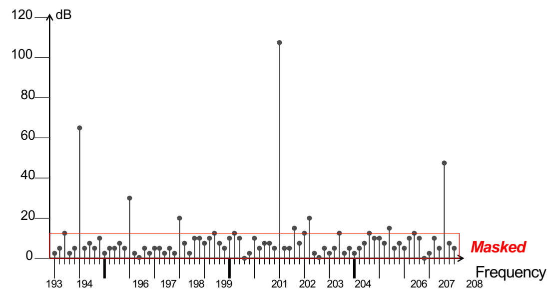
- A sudden masker noise…
  - Makes inaudible other sounds preceding noise!
  - Why does this happen?
    - One thought: takes time for your brain to interpret sound
    - Think of it like a buffer…
    - Throws out contents of buffer when a loud sound comes in
      - to concentrate on only the loud 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)

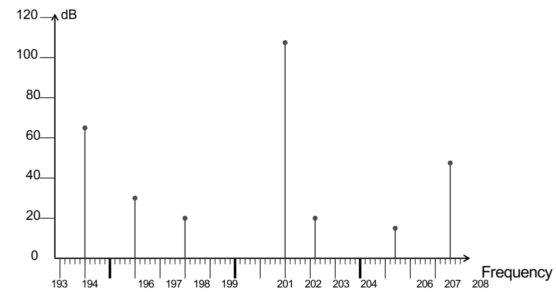


How do we use psychoacoustics in digital music compression? (Masking)



55

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



56

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

57

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

58

Coming Up

- **Feedback**
  - Note – one unique question about this lecture and last
- **In Lab**
  - Measure sensitivity and masking effects
  - Bring head phones
- **Next Lecture**
  - Put this together to compress audio
  - Derive key features of MP3

59

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

60