Digital Audio Processing

EECS 4462 - Digital Audio

November 24, 2018

Digital Audio Processing

- We have already developed a digital audio processing plugin: a digital delay
- A large variety of other plugins are possible
- **Volume-based**: Gain, Compression, Limiting
- **Time-based**: Delay, Time stretching, Reverse
- **Frequency-based**: Filters, Distortion, Pitch shifting
- **Other**: Reverb, Wahwah, Metering

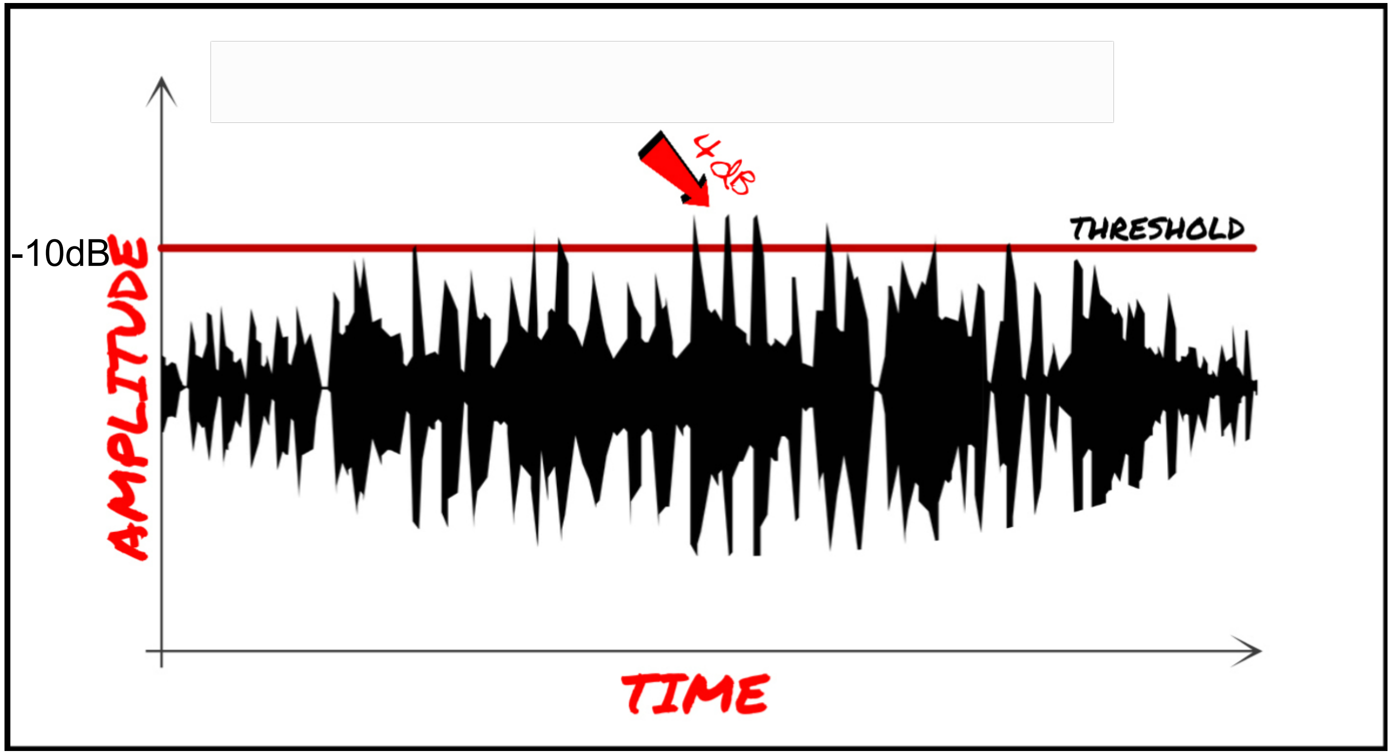
YORK UNIVERSITÉ UNIVERSITY

Volume-based processing

- Adjusting **gain** is rather simple
  - Multiply each factor by a gain factor
  - Care must be taken not to clip
- **Dynamic range compression** is a very common processing operation for both speech and music
- The dynamic range of a piece of audio is the difference in volume between the loudest part and the softest part of the signal
- A compressor makes that difference smaller by making the louder parts softer

Compressor Settings: Threshold

- The minimum level at which compression starts
  - E.g. everything above -10dB must get quieter

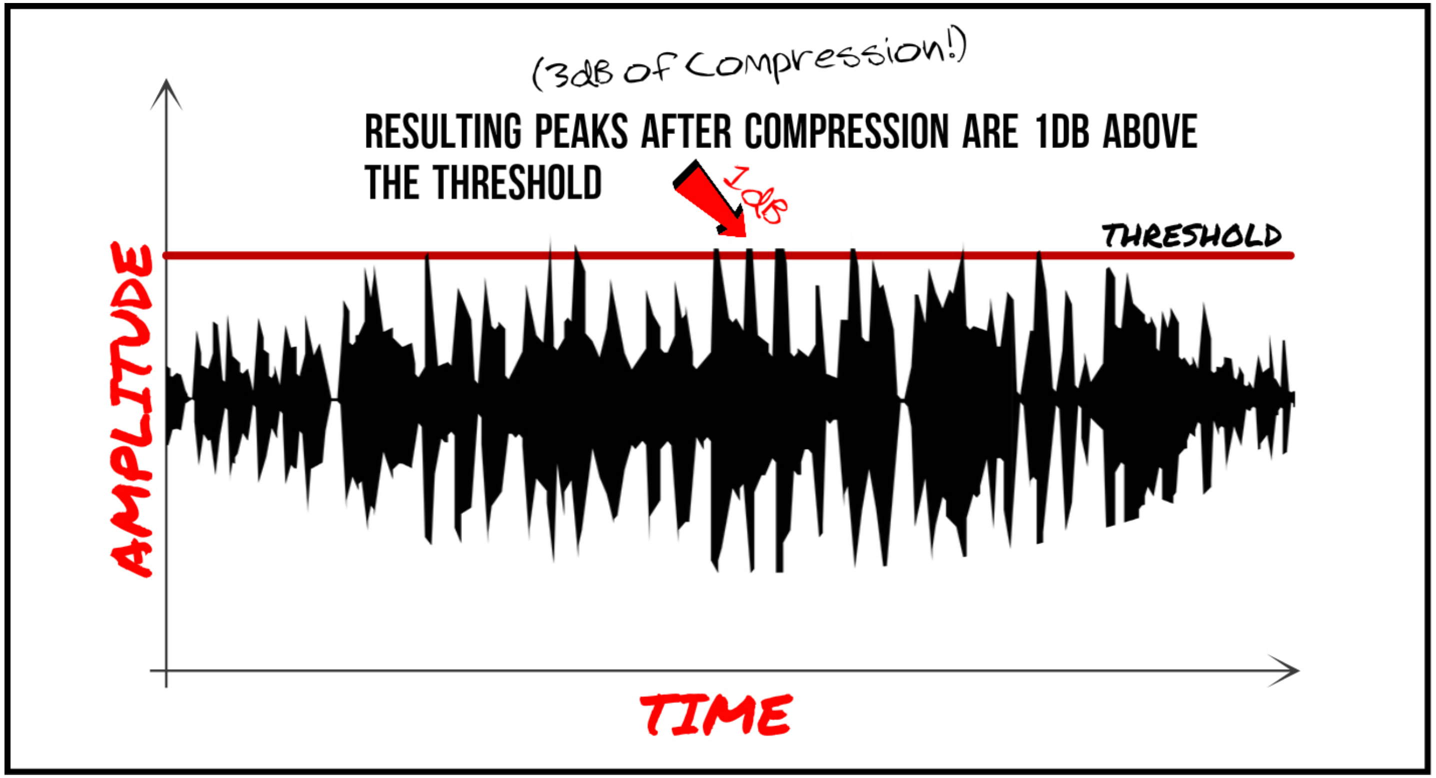


  - Bottom peaks will also be reduced (not shown)

https://urm.academy/visualising-compression/

Compressor Settings: Ratio

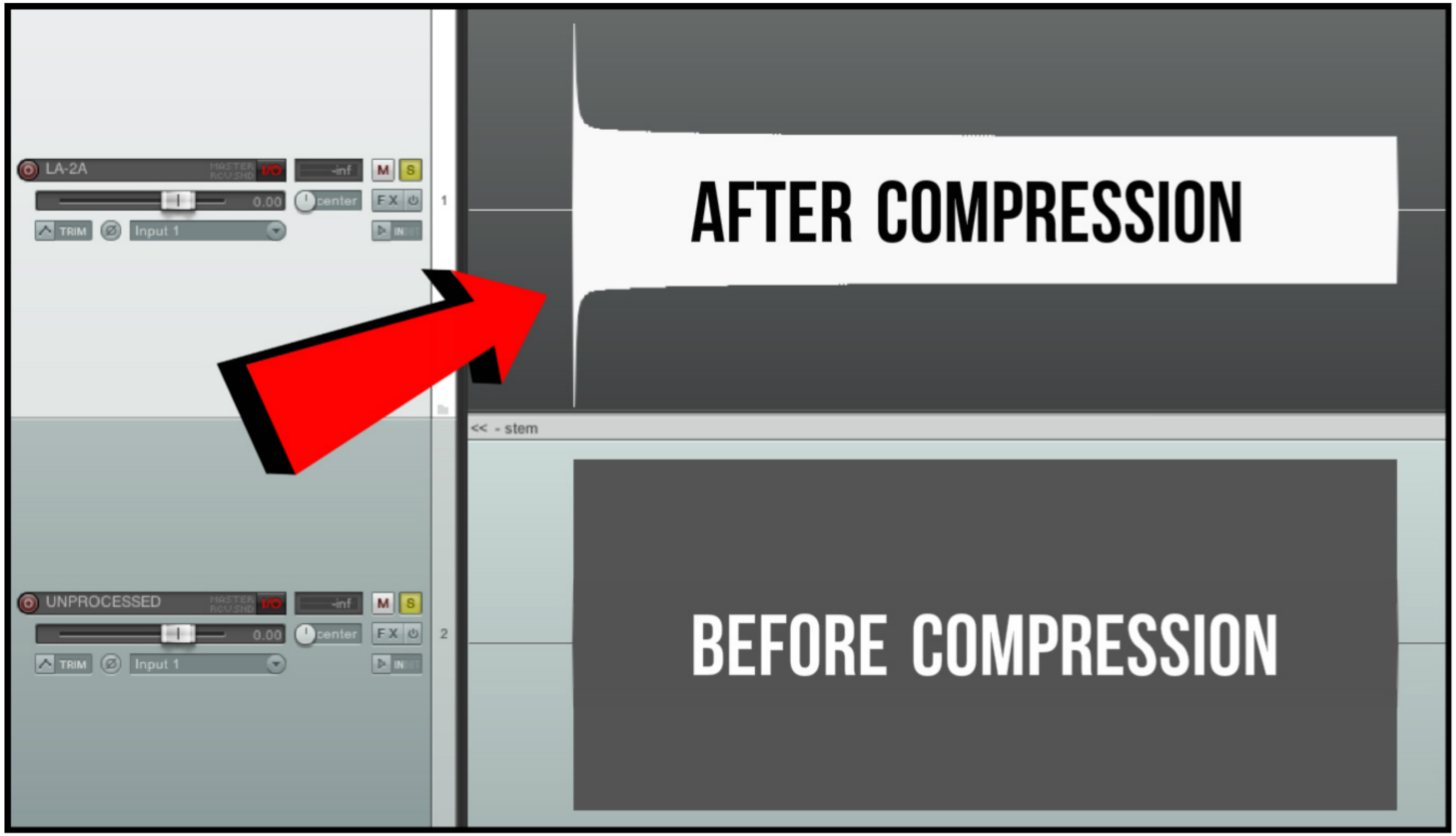
- The amount of compression applied
  - A ratio of 4:1 means that for every 4dB of level above the threshold, only 1dB will remain



https://urm.academy/visualising-compression/

Compressor Settings: Attack

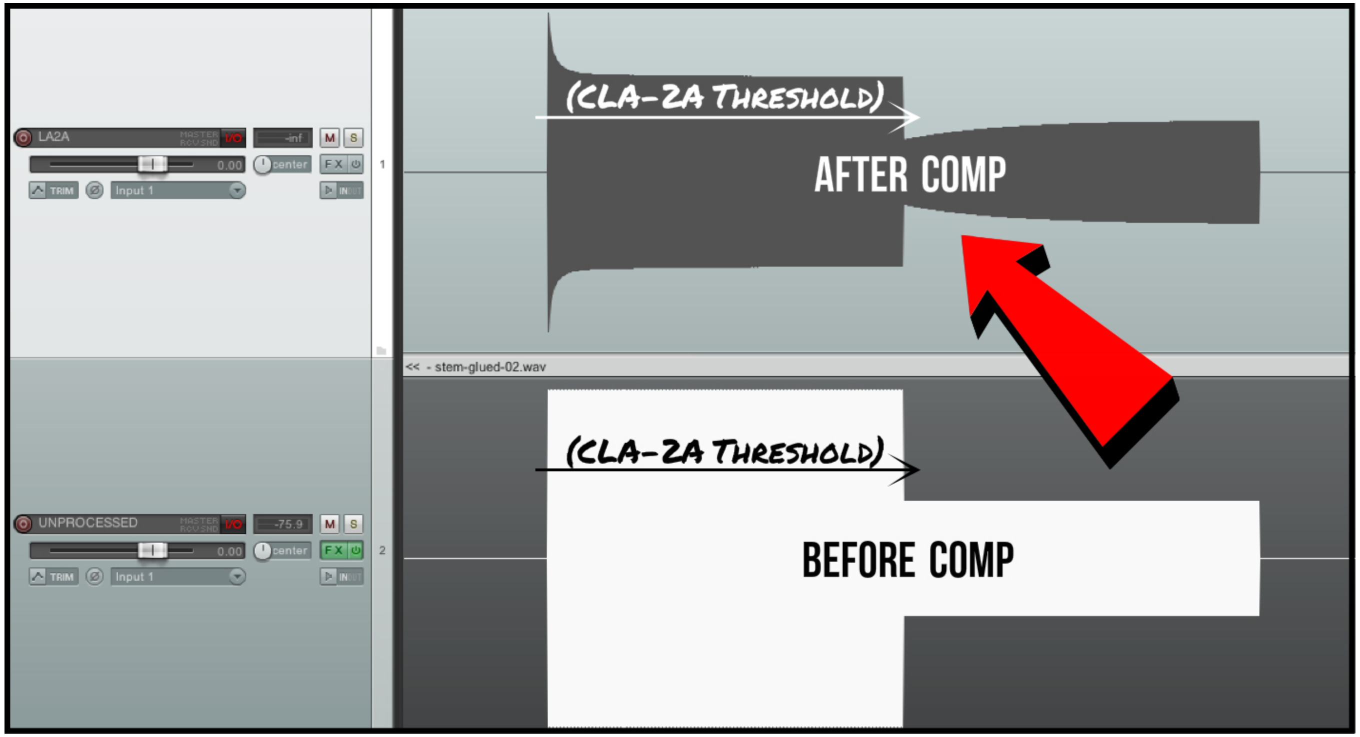
- How fast compression starts



https://urm.academy/visualising-compression/

Compressor Settings: Release

- How fast compression stops after the input signal goes below the threshold



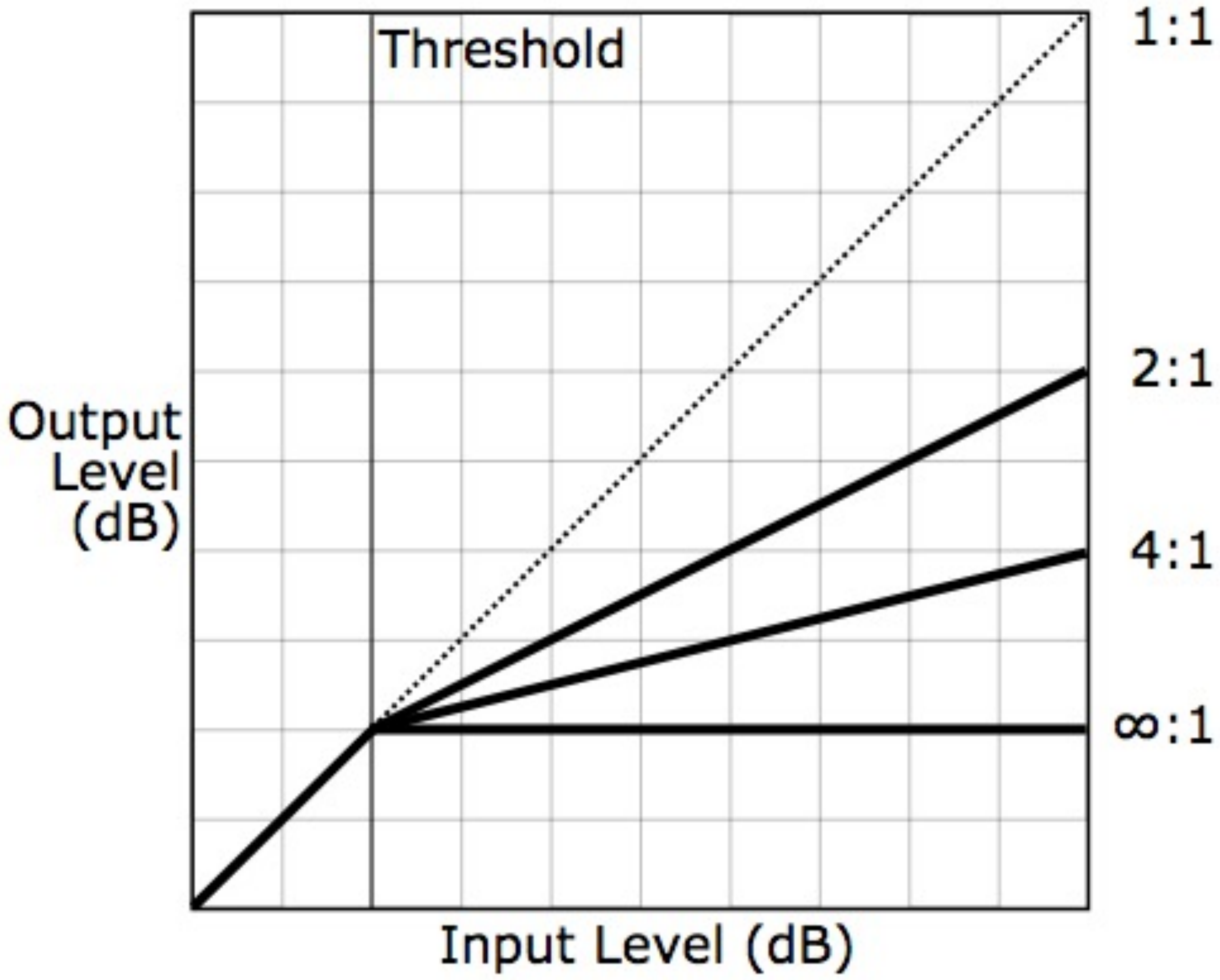
https://urm.academy/visualising-compression/

Other compressor settings

- **Makeup Gain**: Compression reduces the overall loudness of the input signal. If that is not desired, the output signal can be gained to compensate for it
  - Some compressor plugins have automatic makeup gain
- **Knee**: Whether the transition over the threshold is abrupt (hard knee) or gradual (soft knee)
- **Noise floor**: Parts of the signal below the noise floor will not be gained up
- **Peak vs RMS sensing**: Whether the peak level or the RMS level of the input signal is compared to the threshold

YORK UNIVERSITÉ UNIVERSITY

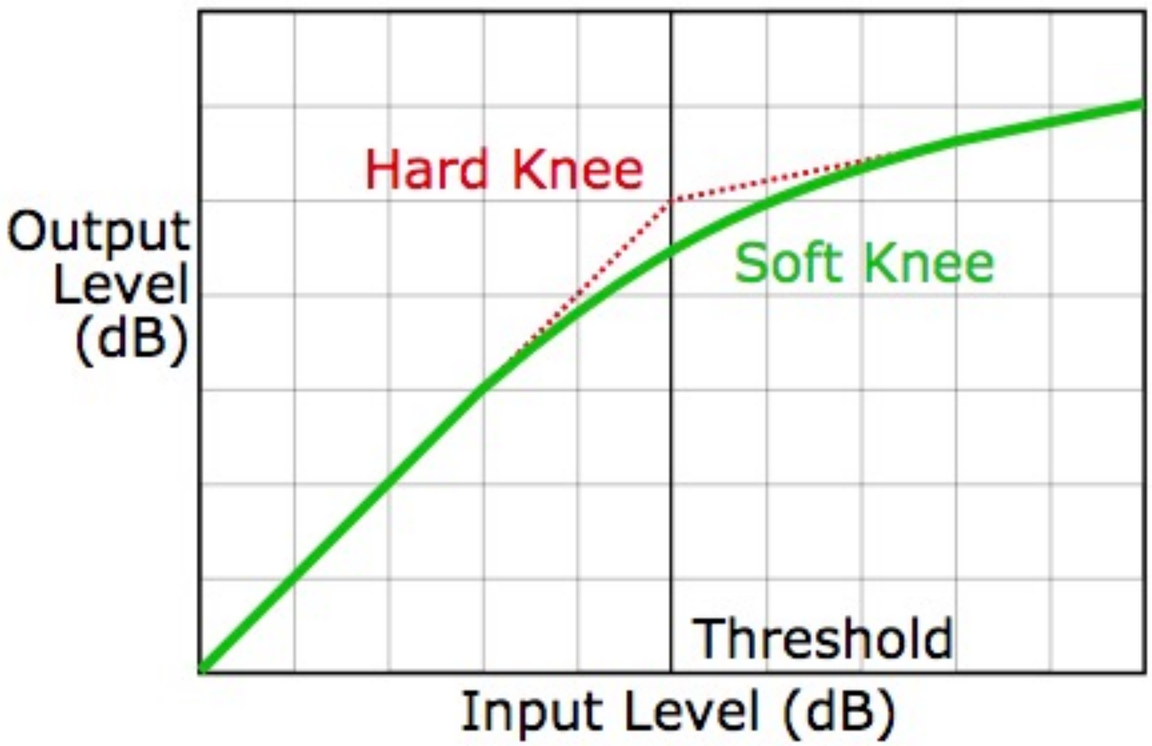
Compression transfer function



https://en.wikipedia.org/wiki/Dynamic_range_compression

Hard knee vs soft knee

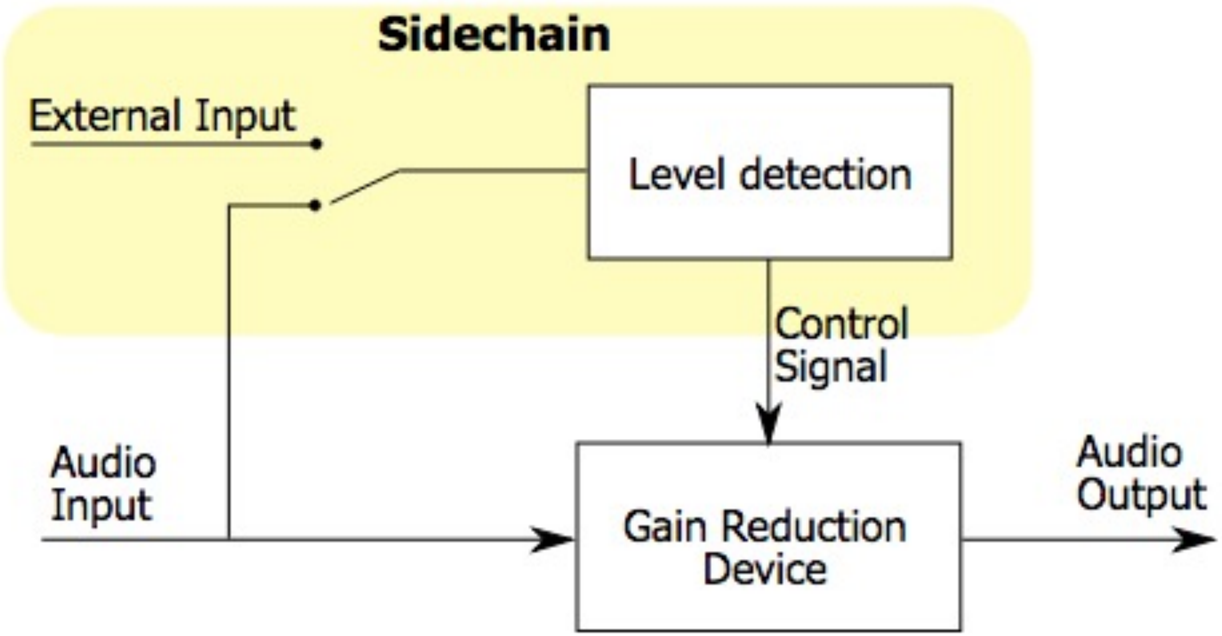
- Soft knee may be desired with high compression ratios to avoid having the compression effect being obvious



https://en.wikipedia.org/wiki/Dynamic_range_compression

Sidechaining

- Instead of compressing when the input signal is above the threshold, a compressor can be set to compress when an external signal is above the threshold
- Used to lower the music, when the DJ speaks



https://en.wikipedia.org/wiki/Dynamic_range_compression

Compression uses

- Even out the volume of a vocal track if the singer has been moving with regard to the microphone
- Increase the sustain of a musical instrument
  - By softening the onset of a note, the decay lasts longer
- Reduce the dynamic range of instruments with wide variation in loudness, such as drums
- Compressing each instrument ensures that they will all be heard in the final mix

Limiting

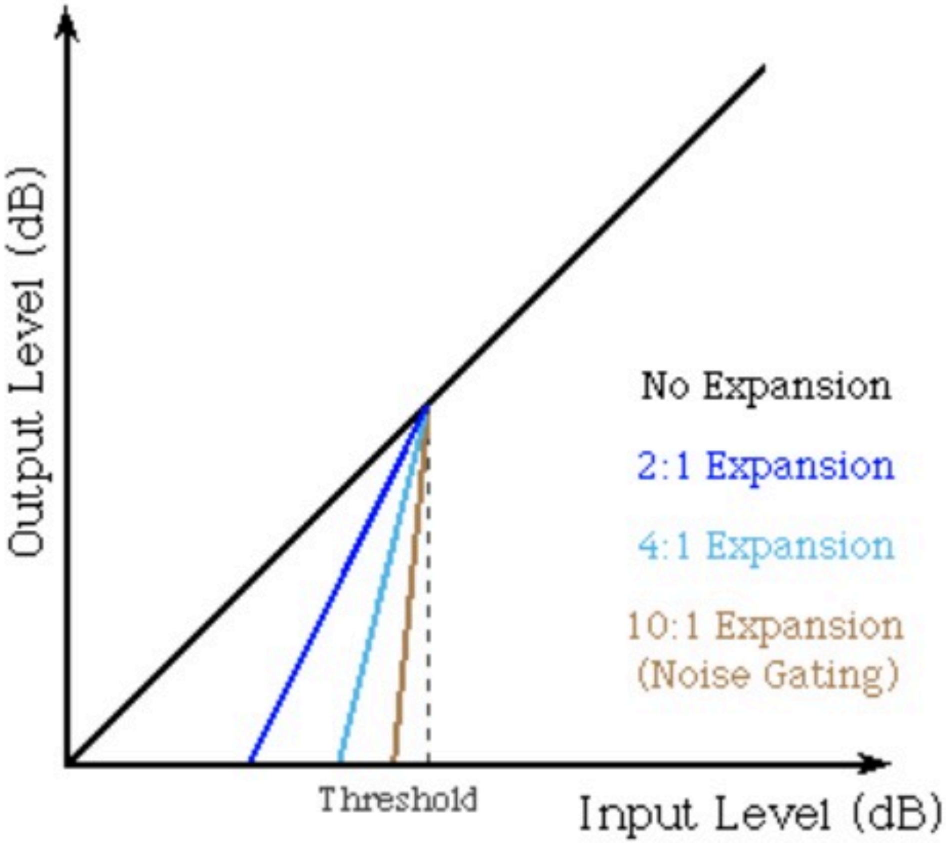
- A limiter is a compressor with a very high ratio (10:1 or more) and a fast attack time
- Brickwall limiters (ratios of 20:1 and higher) ensures audio is never above the threshold
  - Used as safety device in live sound and broadcast applications

Dynamic range expansion

- The dual of a compressor is an expander
- An expander makes the softer parts of a signal even softer
- Used to soften unwanted sounds, such as bleed from other instruments, or background noise
- Has similar settings to a compressor
  - Threshold, expansion ration, attack, release

YORK U UNIVERSITÉ UNIVERSITY

Expansion transfer function



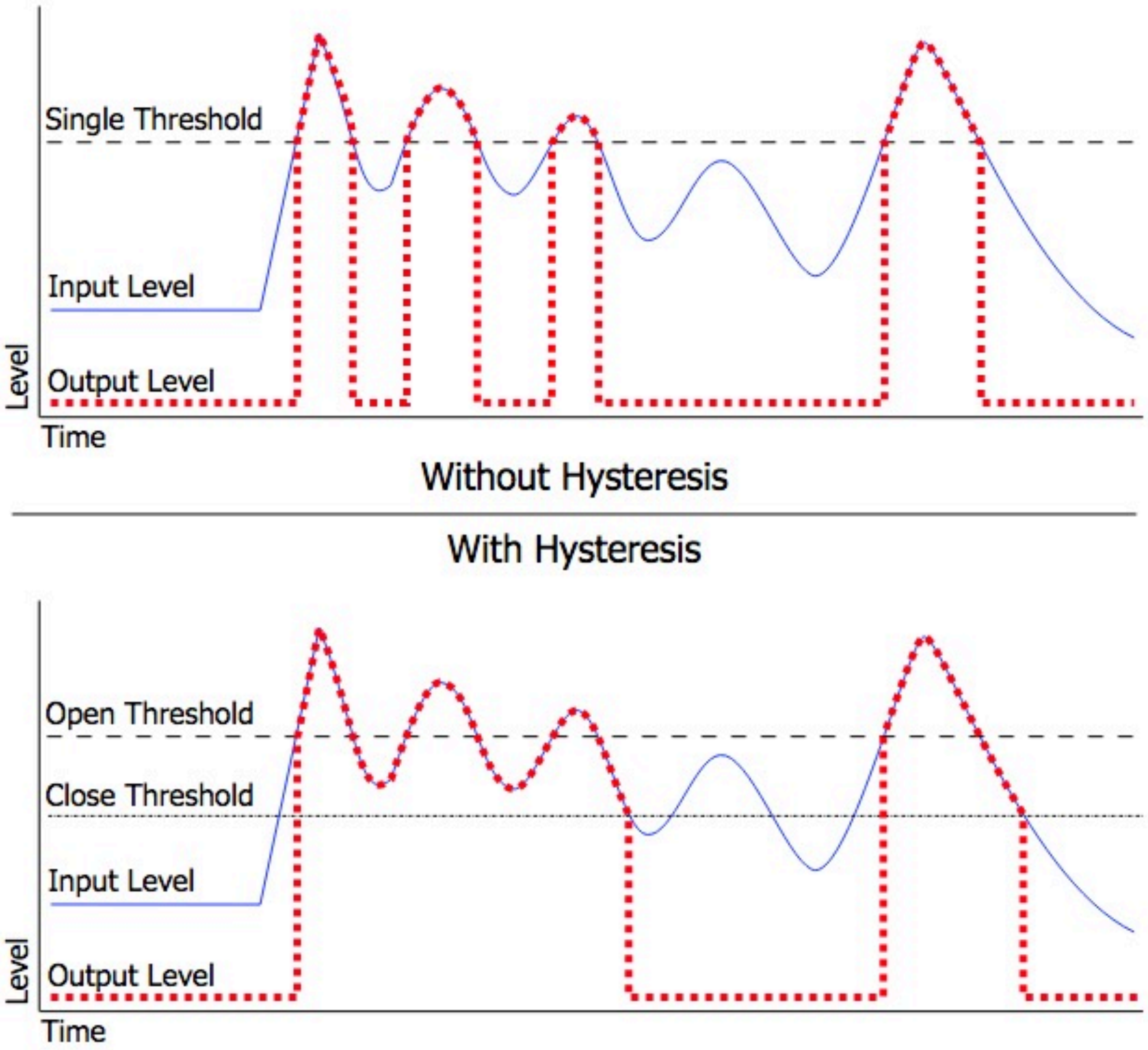
https://wamnation.wordpress.com/2009/02/13/channel-dynamics-part-2-limiters-expanders-and-gates/

Noise gates

- Expanders with a large expansion ratio are called noise gates, as no output is produced when the input signal is below the threshold
- Noise gates have an additional parameter: **Hold**
  - Specifies an amount of time that the gate will stay open even if the input signal is below the threshold
  - Used to avoid opening the gate on and off too often
- Some noise gates will have different open and close thresholds to achieve the same goal of not opening and closing the gate too often
  - This feature is called **hysteresis**

YORK U UNIVERSITÉ UNIVERSITY

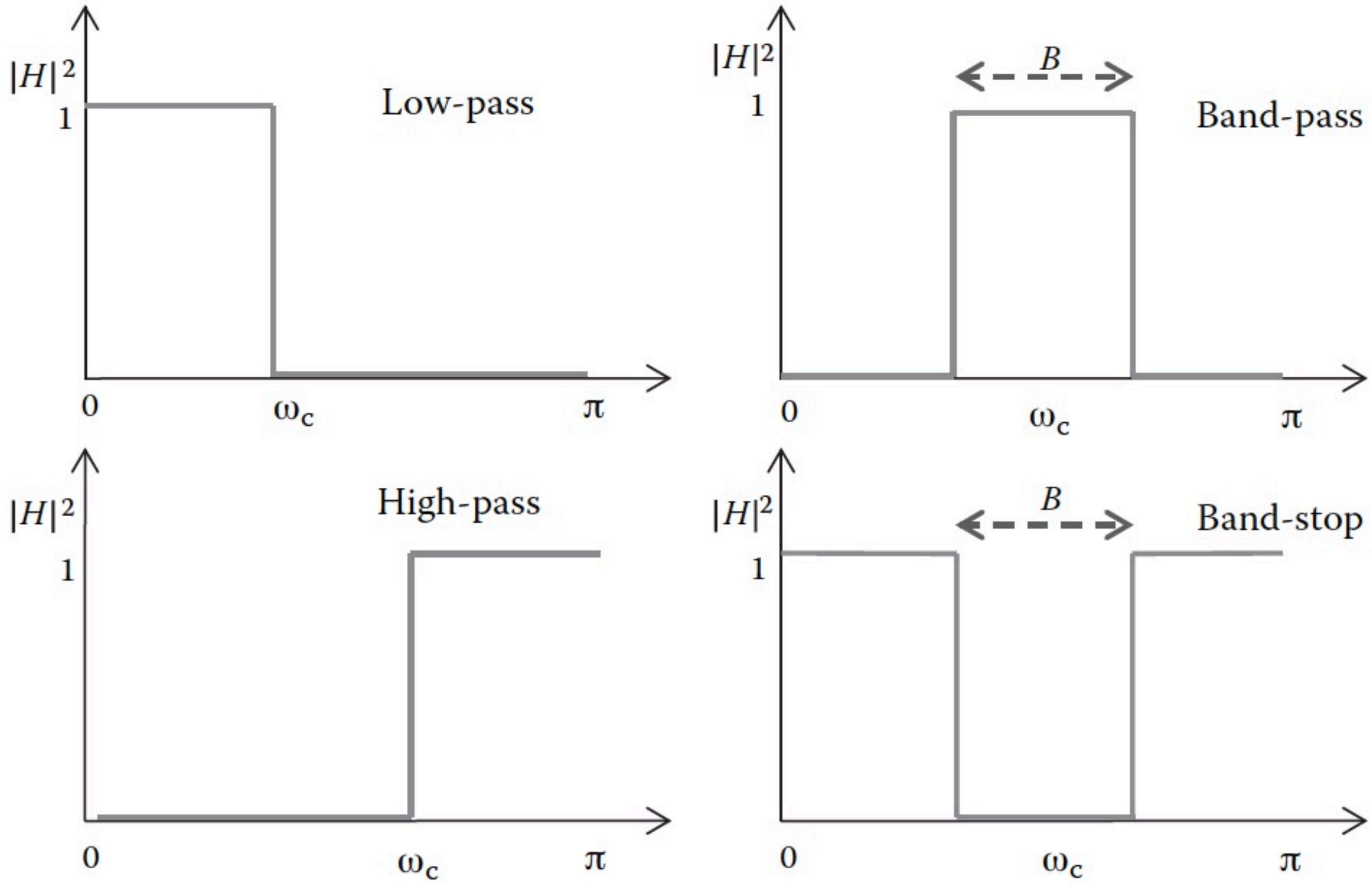
Hysteresis


Single Threshold
Input Level
Output Level
Level
Time
Without Hysteresis
With Hysteresis
Open Threshold
Close Threshold
Input Level
Output Level
Level
Time

Filters

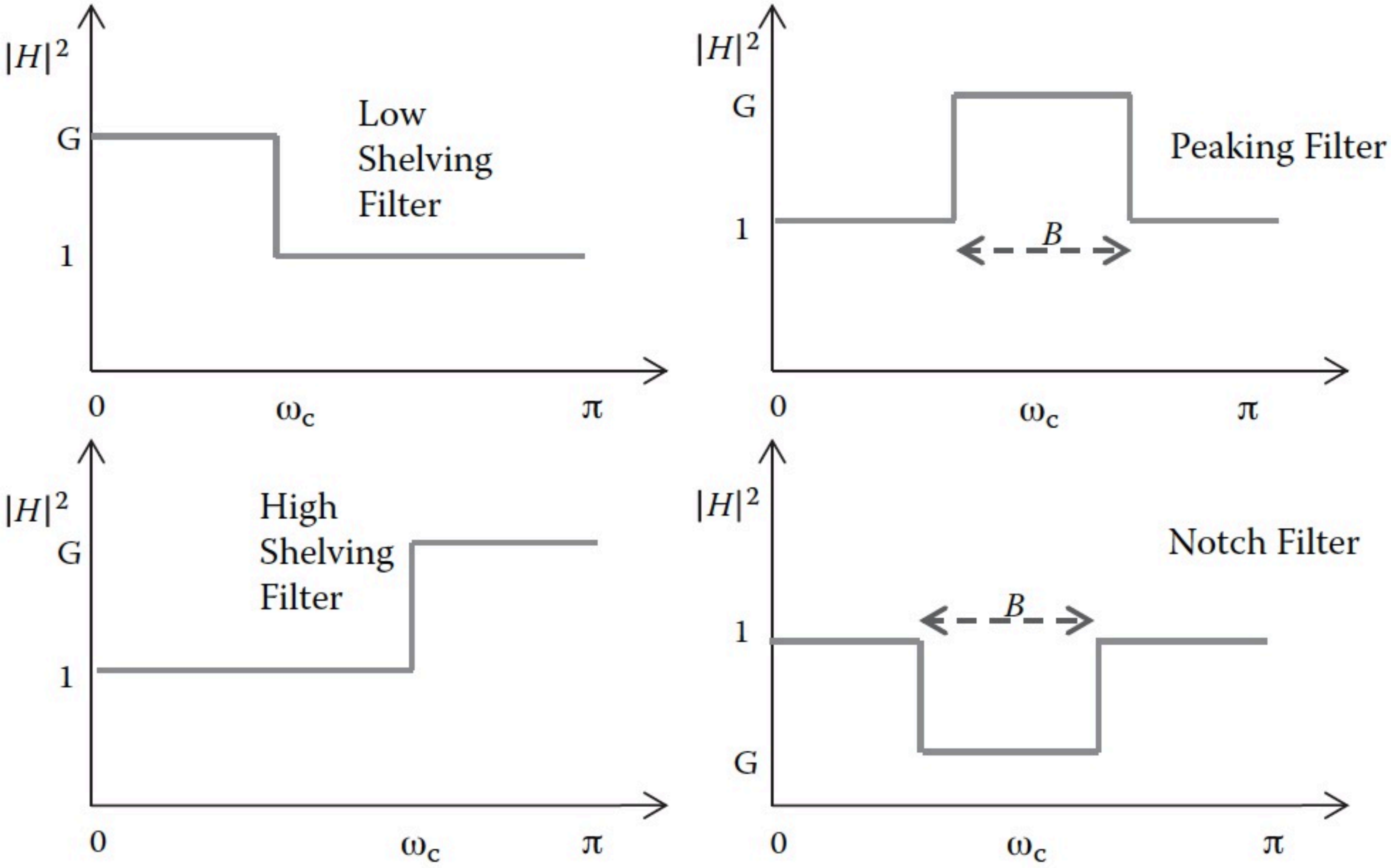
- Frequency-based effects that affect the distribution of magnitude to different frequencies
- Some filters aim to completely eliminate some frequencies
  - High-pass, Low-pass, Band-pass, Band-stop
- Others adjust the relative gain of different frequencies
  - Shelving, peaking, and notch filters

Ideal transfer functions 1



Audio Effects: Theory, Implementation, and Application, by Joshua Reiss and Andrew P. McPherson

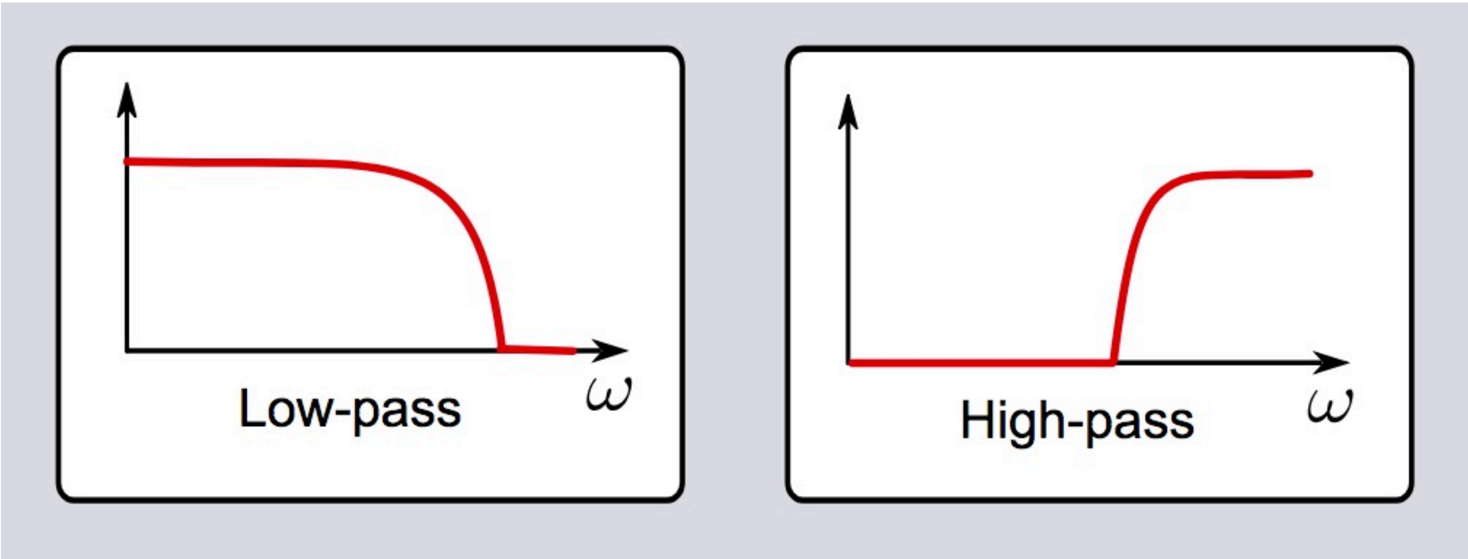
Ideal transfer functions 2



Audio Effects: Theory, Implementation, and Application, by Joshua Reiss and Andrew P. McPherson

Practical filters

- In practice, it is neither possible nor desirable to have such sharp cutoffs



- The cutoff frequency is where attenuation has reached 3dB
- The rolloff factor indicates how fast attenuation happens after the cutoff frequency
    - Often measured in dB per octave

https://en.wikipedia.org/wiki/Filter_(signal_processing)

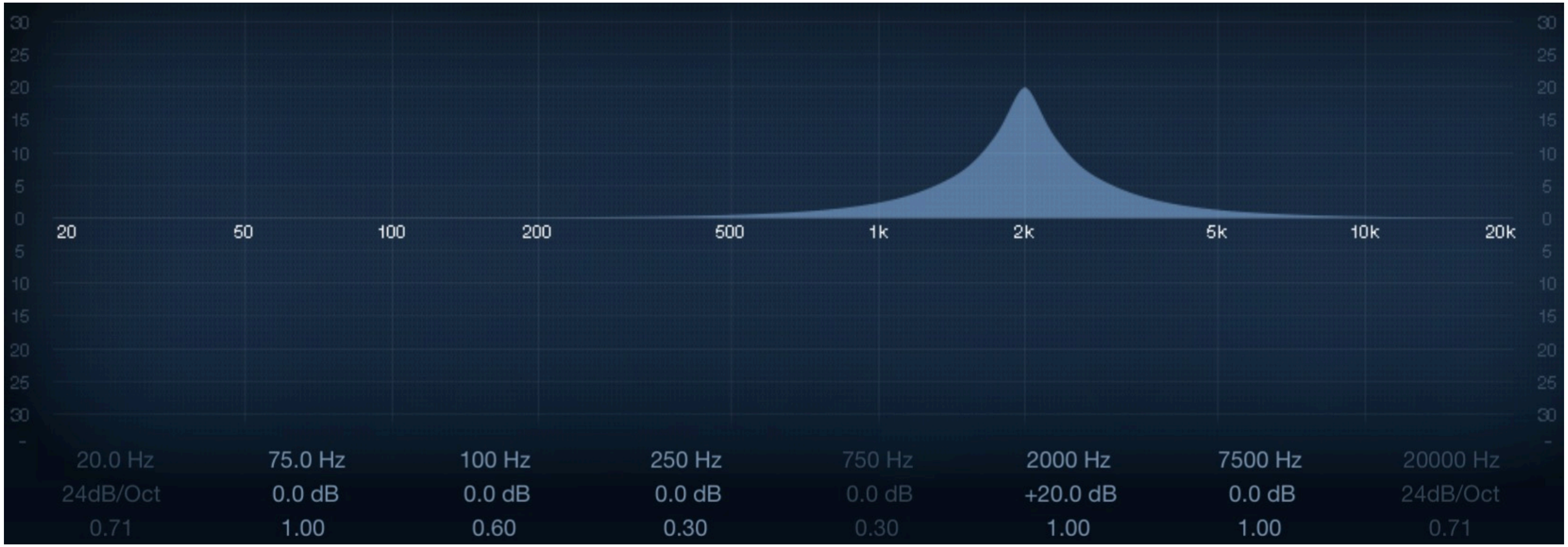
YORK U UNIVERSITÉ UNIVERSITY

Equalization

- Equalizers are used to balance the relative gain of different frequency bands
- An equalization operation has three parameters

1. **Center frequency**: The frequency of largest boost or attenuation
2. **Gain**: The amount of gain for the center frequency
3. **Bandwidth**: The total range of frequencies affected measured in octaves
    - In DAWs, bandwidth is often shown as Q
    - The smaller the Q, the wider the bandwidth

Equalization operation

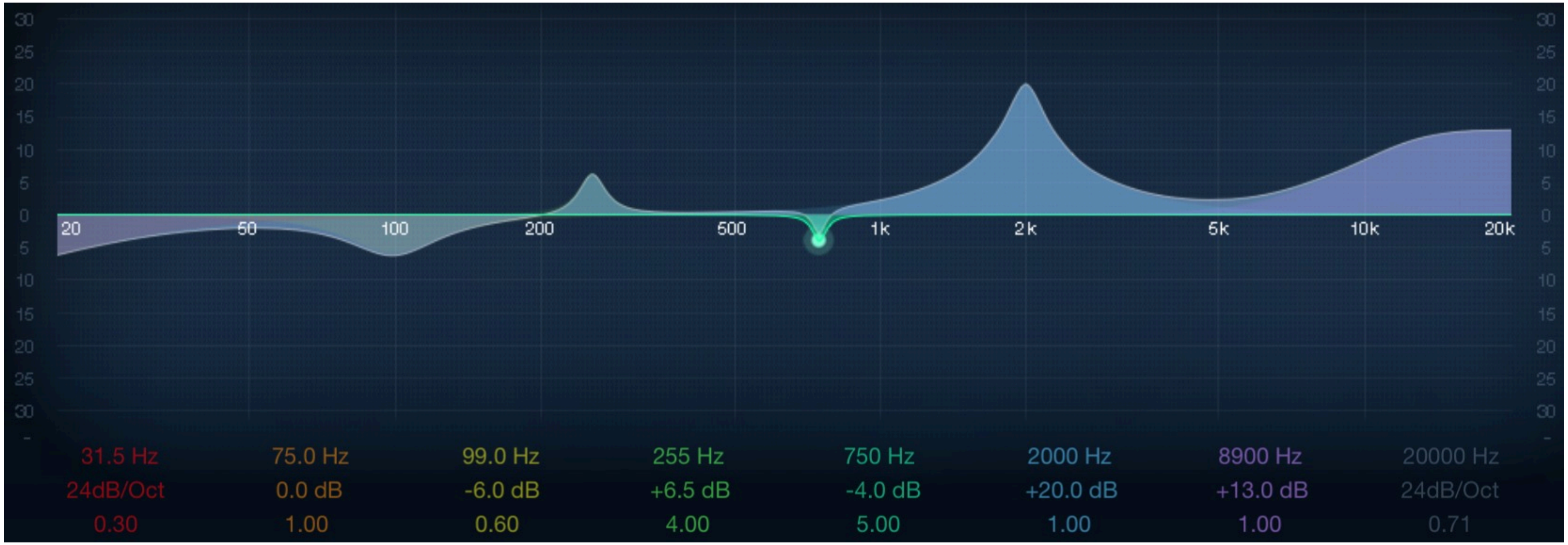


Center Frequency

Gain

Q (Bandwidth)

EQ Curve



Several filters may be combined to achieve the desired effect

Their combined effect provides the EQ curve of an equalizer effect

Wah-wah

- Essentially a band-pass or peaking filter whose center frequency is changed by a foot pedal
- In typical wah-wahs, the center frequency ranges from 300Hz to 1200Hz
- This simulates formants found in speech
  - The [u] sound has formants around 300Hz
  - The [a] sound has formants at 750Hz and 1200Hz
- **Formants** are stronger frequencies when we speak depending on how the vocal tract is shaped
  - They are the way we recognize different vowels

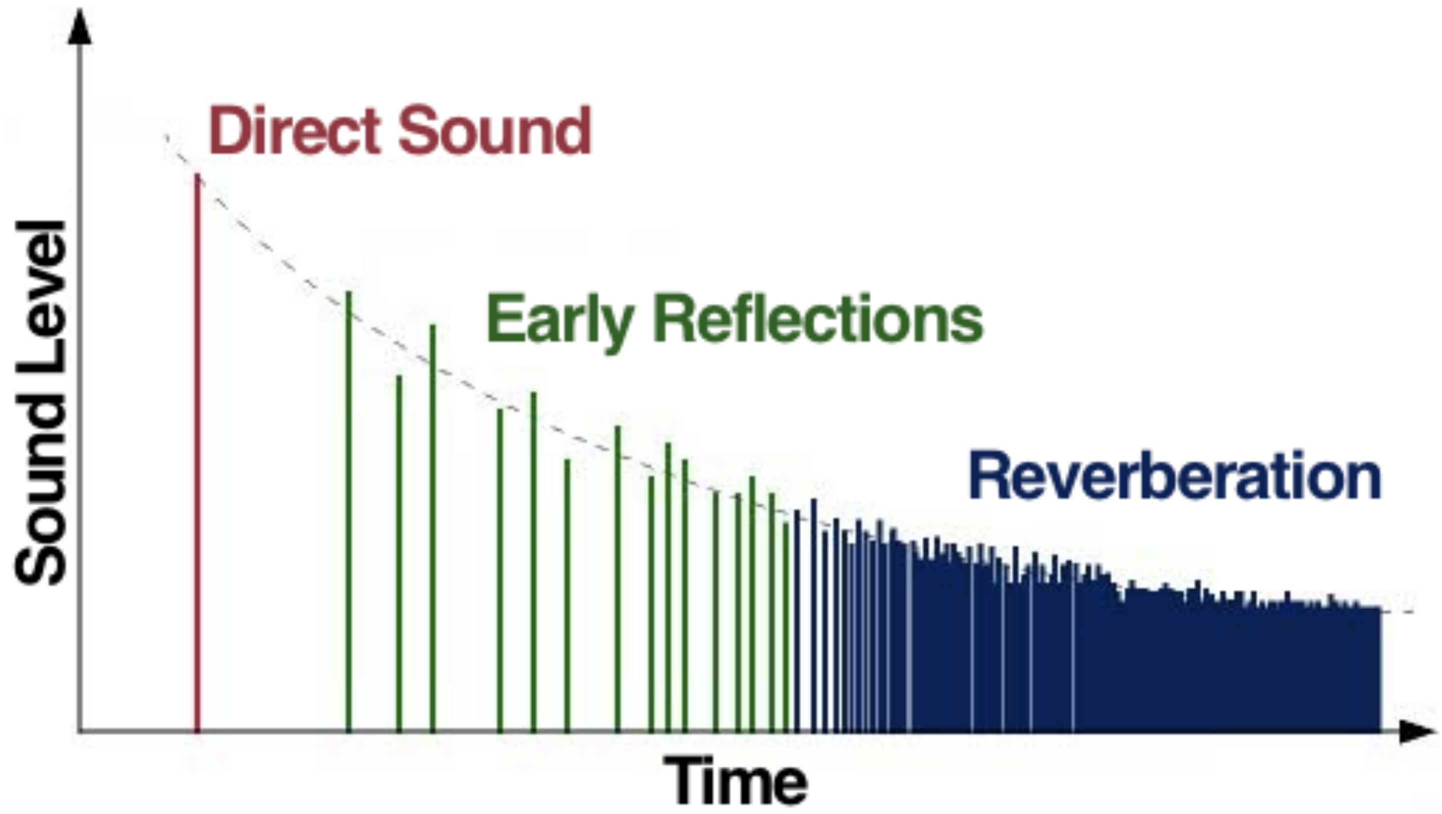
Wah variations

- **Auto wah**: Uses a low-frequency oscillator (**LFO**) typically at around 1-2Hz to adjust the center frequency
  - Some variations adjust the center frequency based on the amplitude of the input signal
- **Tremolo wah**: An auto wah with a second LFO to modify the amplitude of the output signal
  - The two LFO can be independent, synchronized, or even synchronized but out of phase

Reverberation

- Sound in a room reaches the listener in many ways:
  - Directly
  - After one reflection in any of the room surfaces
  - After many reflections in many different surfaces
- Reflected copies of the original sound are attenuated because they travel longer and they are partially absorbed by the surfaces
- As these reflections typically within 40ms of the direct sound, they are perceived as one sound
- The combination of all the copies gives us an impression of space

Room impulse response



Each vertical line is a copy of the original sound

The height of each of line is the amplitude of the copy

https://commons.wikimedia.org/wiki/File:Corrected_Room_Impulse_Response.png

Reverberation time

- To summarize the reverberation properties of a room, we use the notion of reverberation time RT_{60}
  - The time it takes for the sound to decay by 60dB
  - Small rooms have RT_{60} that is usually less than 1 sec
  - Larger rooms may have an RT_{60} of 2 sec or more
- Several factors affect the RT_{60} of a room
  - Room dimensions
  - Wall materials
  - Furniture or people in the room