ECS730

Revision Lecture

Queen Mary
University of London

Topics

- Week 1: Mixing Consoles, Software Mixers
- Week 2: Fundamentals, Delay, Vibrato
- Week 3: Flanging, Chorus
- Week 4: Filters, Tone Controls, Graphic EQ
- Week 5: Parametric EQ, Wah-Wah, Phasers
- Week 6: Distortion, Ring Modulation
- Week 8: Compression, Expansion
- Week 9: Reverb, Phase Vocoder
- Week 10: Spatial Audio

Queen Mary
University of London

Mixing consoles

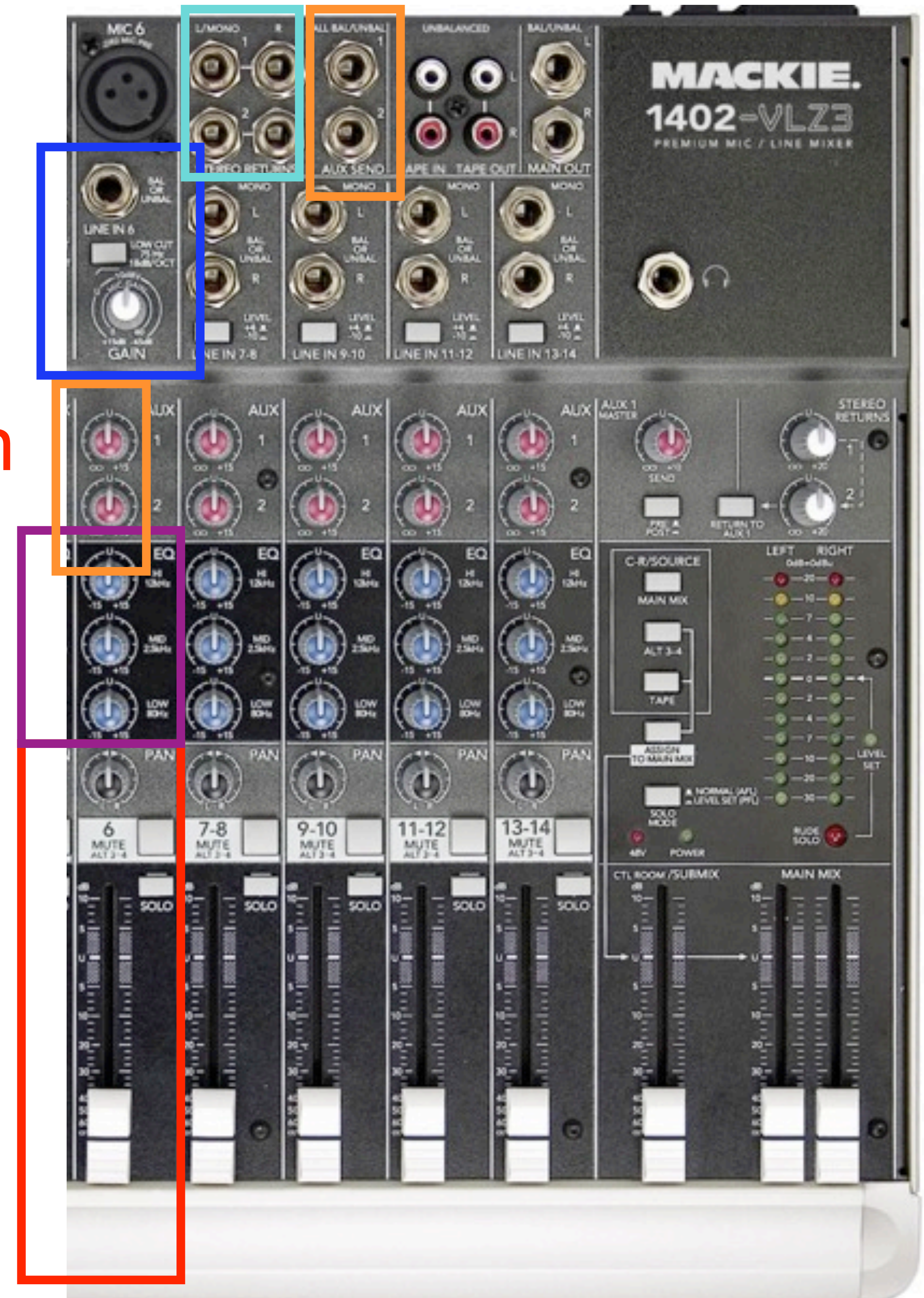
- What is a mixing console?
  - Electronic device for combining, routing, and modifying audio signals (including level, tone and dynamics)
  - Analogue or digital signals, depending on type
  - Modified signals are summed to produce output
- Three main functions:
  - Summing: combining audio signals
    - Many channels can be summed to stereo (or mono) via mix bus
  - Processing: modifying audio signal properties
    - Consoles often have on-board equalisers and sometimes dynamics processors
  - Routing: grouping, sending and receiving signals
    - Enable use of external processors and effects
    - Insert points, auxiliary sends and receives

Mixing consoles

- A mixing console has two main sections:
  - Channel section
    - Collection of channels organised in physical strips
    - Each channel can correspond to an individual track on a multitrack recorder
    - Most channels support mono input; some support stereo
    - Some channels will have different types of equalisation, etc.
  - Master section
    - Central control over console (global functionality)
    - Includes master aux sends, effect returns, control room level, ...
- Other terms to know
  - buss, group, effect vs. processor
  - tracks, mixer strips (software mixers)

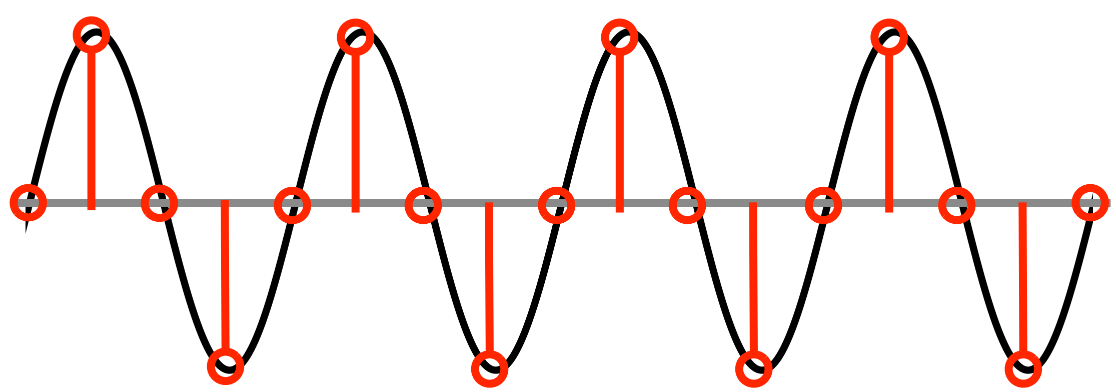
Signal flow in a mixing console

- Components
  - Faders, pan pots & cut switch
  - Line gains, phase invert & clip indicators
  - On-board processors
  - Insertion points *(not shown)*
  - Auxiliary sends
  - FX returns



Sampling

- Discrete sequences do not have to represent any continuous or physical quantity
  - But in the audio world, we often work with sampled audio
- Sampling = converting continuous time to discrete
- Reconstruction = discrete to continuous

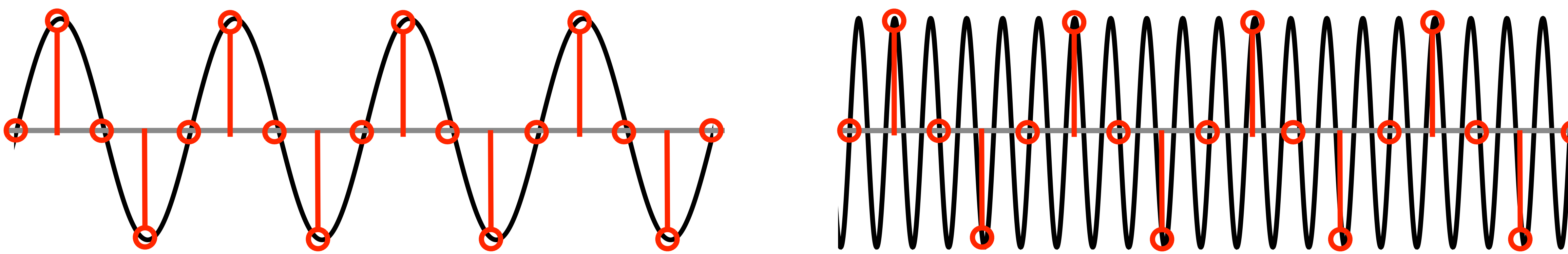


$$x[n] = x_a(nT), \quad -\infty < n < \infty$$

- T is called the sampling period
  - $1/T$ is the sampling frequency or sample rate (f_s)

Sampling 2

- Reconstruction is harder than sampling
  - Sampling inherently loses information
  - How do we distinguish between these two cases?



- Need a band-limited input signal
  - Limited frequency range
- Nyquist Sampling Theorem
  - Need to sample a signal at twice its highest frequency to be able to uniquely reconstruct it
  - This frequency is called the Nyquist frequency

Properties of digital systems

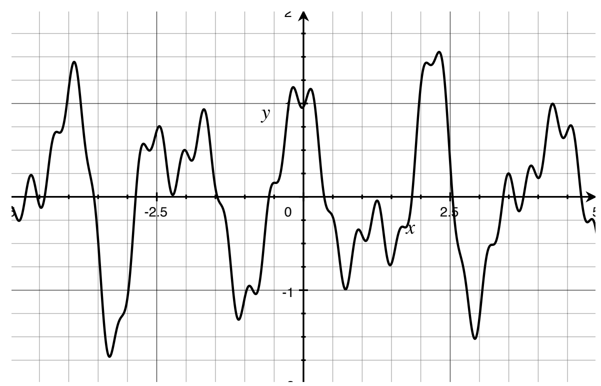
$$x[n] \longrightarrow \boxed{f} \longrightarrow y[n] = f(x[n])$$

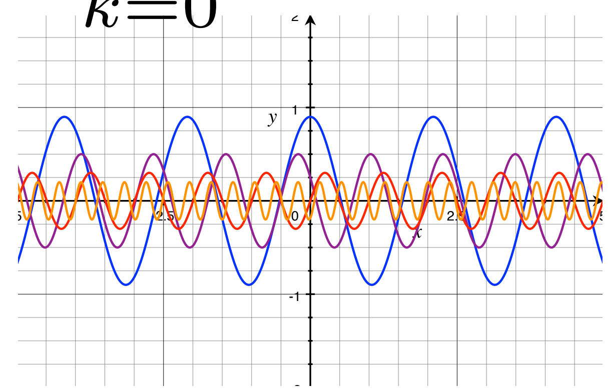
- Linearity: $f(x_1[n] + x_2[n]) = y_1[n] + y_2[n]$
- Time invariance: $f(x[n-N]) = y[n-N]$
  - Linear, Time-Invariant (LTI) systems are a common useful case
- Stability:
  - x[n] bounded $\longrightarrow$ y[n] bounded
- Causality:
  - x[n] = 0 for n < 0 $\longrightarrow$ y[n] = 0 for n < 0
  - f doesn't depend on any future sample (i.e. greater n)
- Memorylessness:
  - y[n] only depends on the current value of x[n]

Time and frequency domain

- Any x[n] can be represented as a sum of sinusoids...
  - So let's do that then!

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X_k e^{+j2\pi \frac{k}{N} n}$$



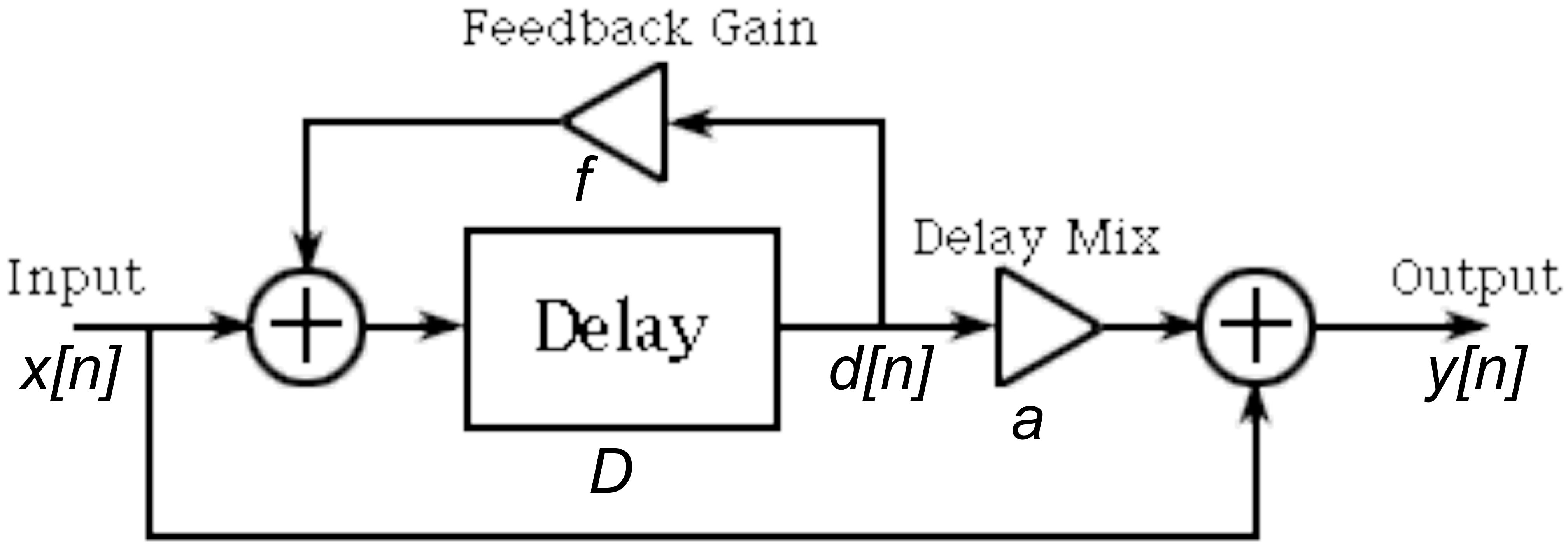


  - The second version is called the frequency domain representation of x
    - * or at least discrete samples of the frequency domain...
  - What frequencies are present, with what amplitude and what phase? From this we can exactly determine x[n].
- Things to know:
  - z-transform, relationship to frequency response
  - Converting between time & z-domains

Queen Mary
University of London

Delay (with feedback)

- Most delays have feedback control (regeneration)
  - Take the delay output, send it back to the input



$$d[n] = x[n-D] + f\,d[n-D] \qquad |f| < 1 \ \text{(strictly)}$$

$$y[n] = x[n] + a\,d[n] \qquad |a| \leq 1 \ \text{(usually)}$$

  - To derive *y[n]* in terms of *x[n]*:

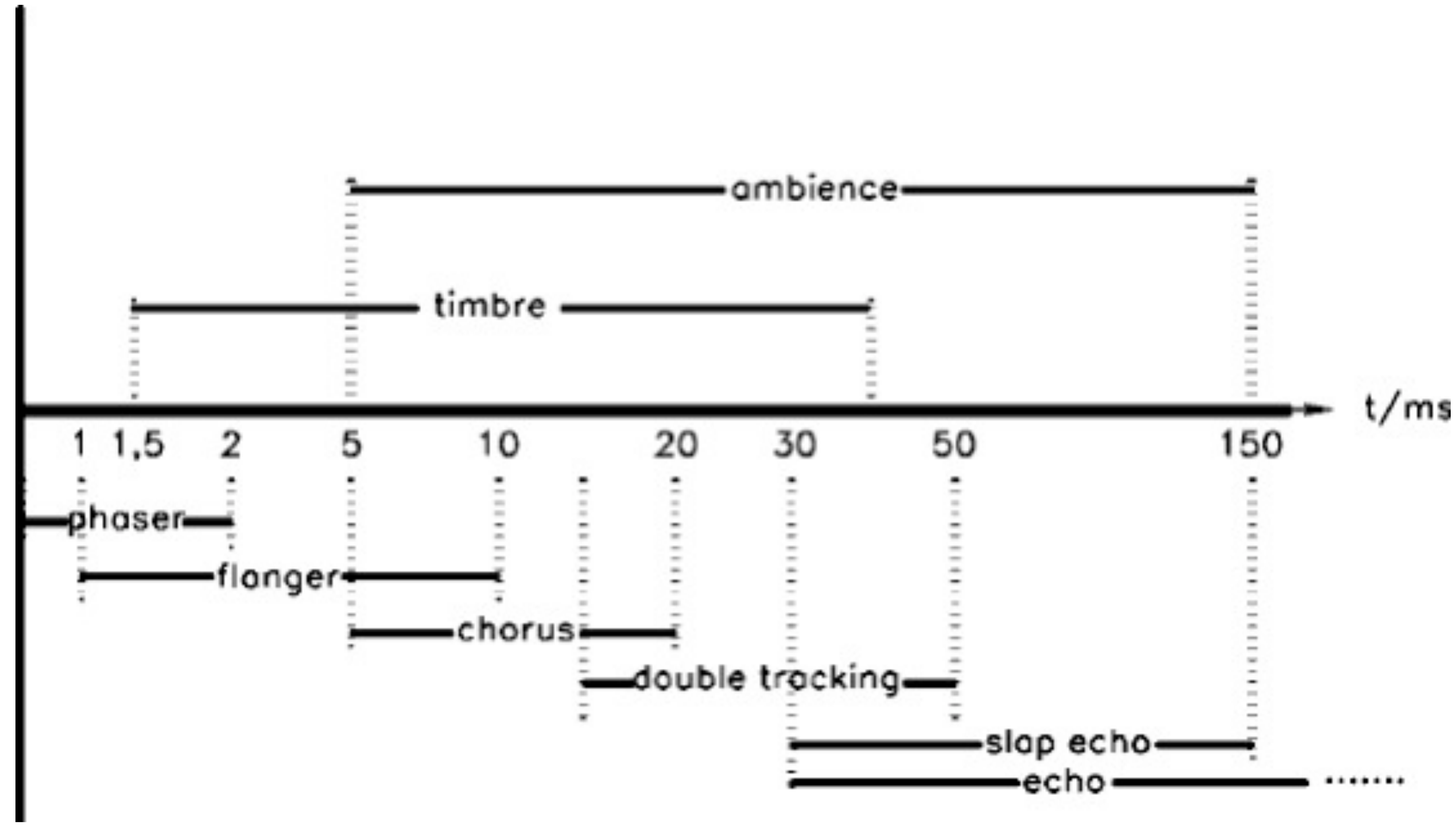
$$y[n-D] = x[n-D] + a\,d[n-D] \implies d[n] = \frac{f}{a}y[n-D] + (1-\frac{f}{a})x[n-D]$$

$$\implies y[n] = f\,y[n-D] + x[n] + (a-f)x[n-D]$$

Queen Mary
University of London

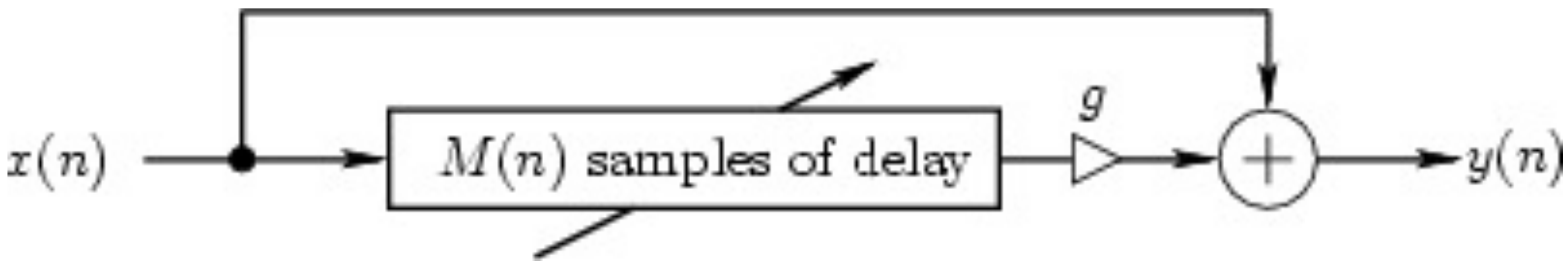
Delay 2

Delay range (ms) (Typ.)	Modulation (Typ.)	Effect name
0 ... 20	-	Resonator
0 ... 15	Sinusoidal	Flanging
10 ... 25	Random	Chorus
25 ... 50	-	Slapback
> 50	-	Echo



- Other things to know:
  - How to convert between block diagram & transfer function
  - Variations: multi-tap, ping-pong, etc.
  - Fractional delay
  - Implementation on a circular buffer (basics, not code)

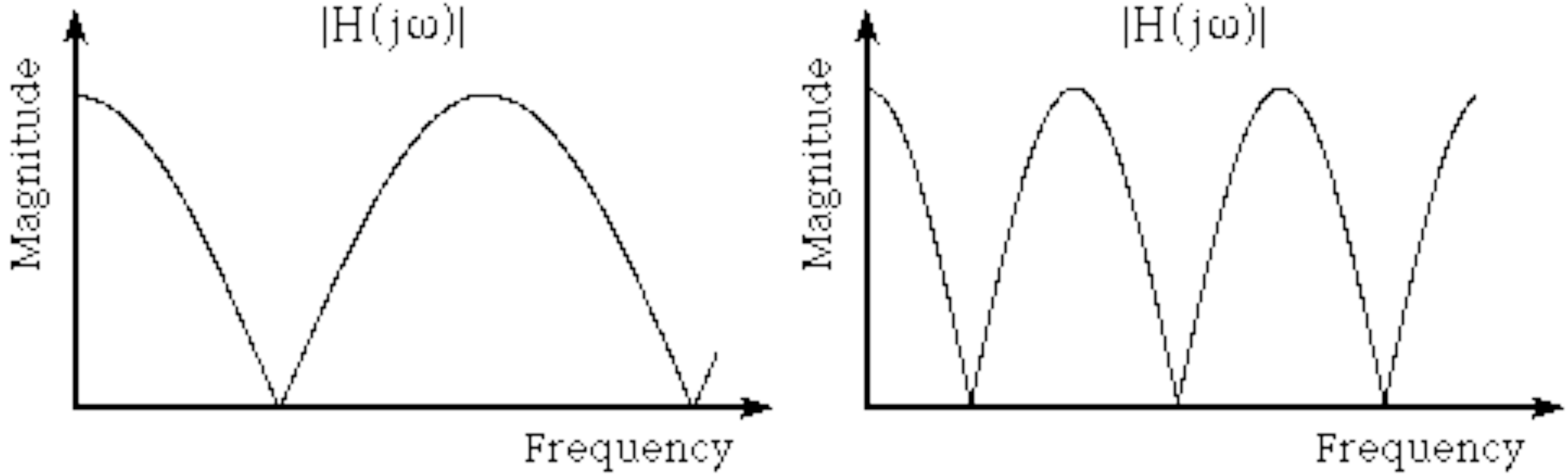
Flanging



- Feedforward comb filter
- Input-output relation: $y[n] = x[n] + gx[n - M(n)]$
  - x[n] - input signal
  - y[n] - output signal
  - g - depth of flanging effect
  - M(n) - length of delay line at sample n
- M(n) must vary smoothly over time
  - Interpolated (fractional) delay line allows noninteger M
- How are chorus and vibrato different?
- Also know:
  - Variations (feedback, stereo, etc.)
  - Parameters

Flanger frequency response

- Delay & add has filtering effect on signal
  - Creates series of notches in frequency response
  - Eliminates a periodically spaced set of frequencies
  - Other frequencies passed with amplitude change
- Comb filter (notches resemble teeth on comb)



Frequency response for two flangers with depth 1.
Which one has the smaller delay, and why?

IIR filters (also know FIR)

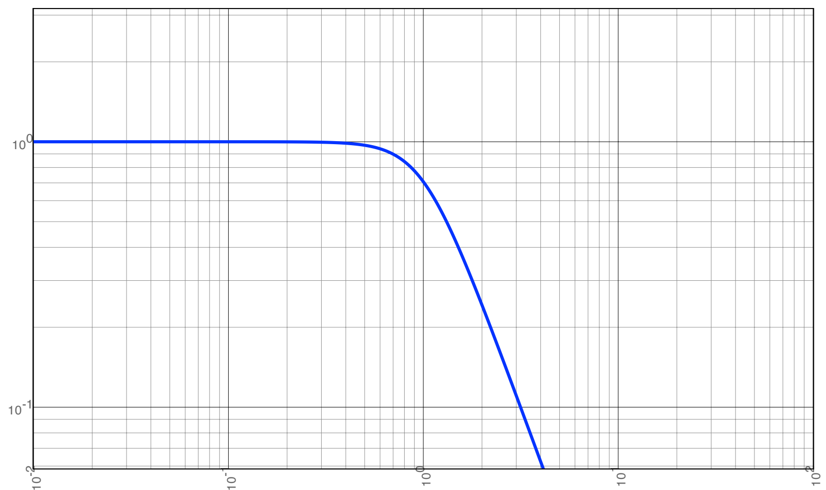
- More general form:

$$y[n] = \sum_{k=1}^{N} a_k y[n-k] + \sum_{k=0}^{M} b_k x[n-k]$$

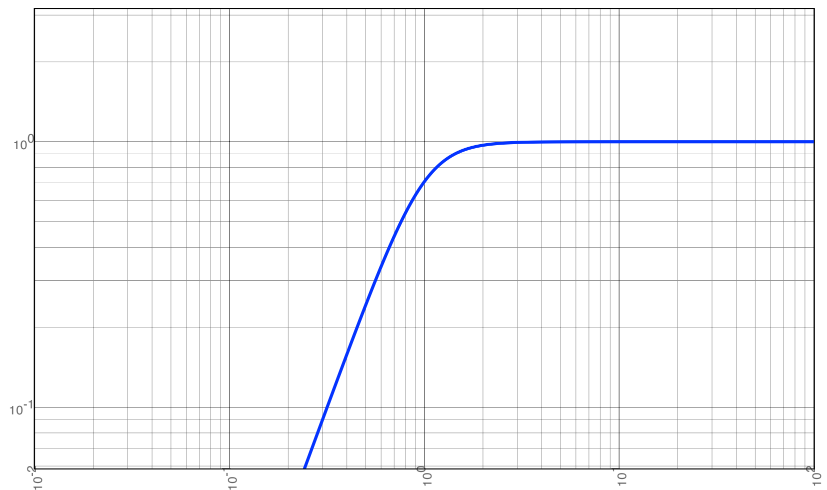
  - Output depends on last M inputs and N outputs
  - Infinite Impulse Response (IIR)
    - Unless all $a_k = 0$
  - Often N = M in practice, but this is not required
  - max(N,M) called order of filter
  - Properties of IIR filters
    - Linear
    - Time-invariant
    - Stability depends on coefficients a_k
- Most equalisers use IIR filters

Types of filters

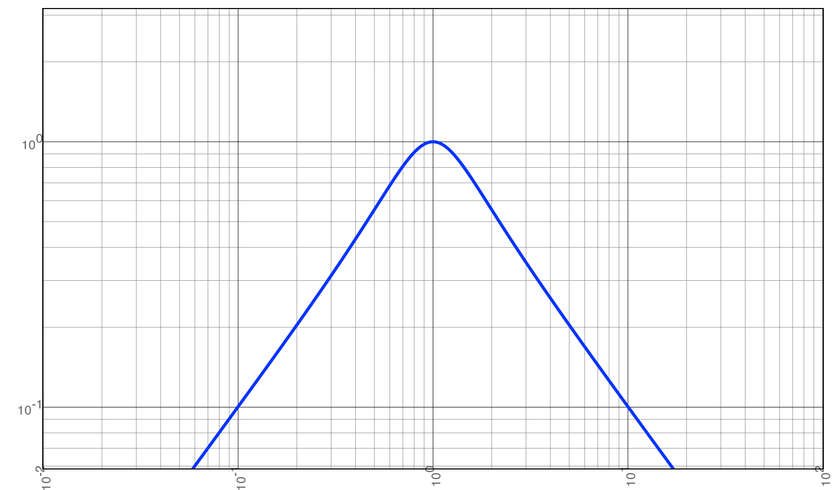
Types we frequently encounter in audio systems:



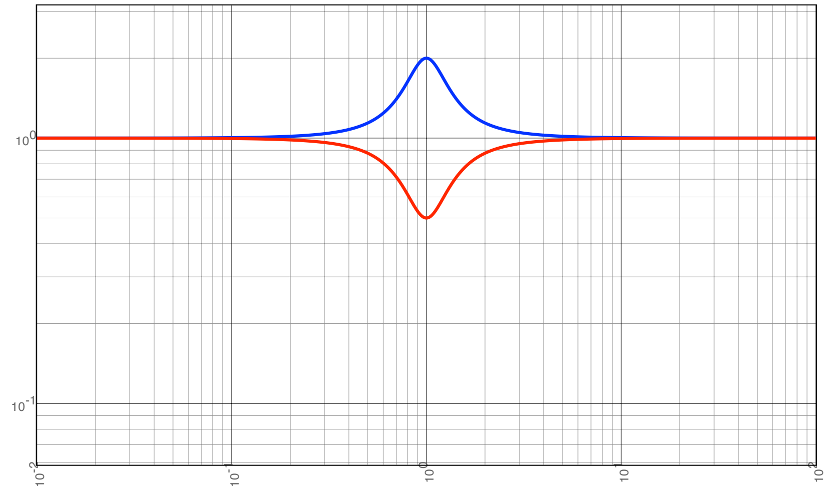
Lowpass



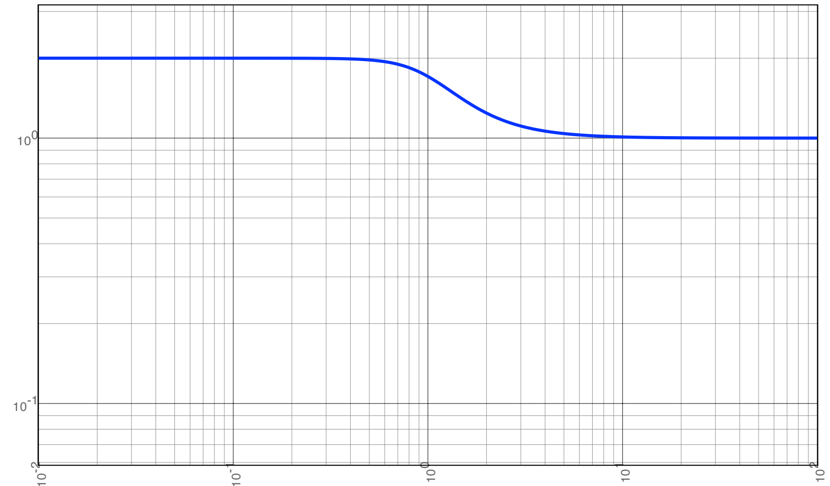
Highpass



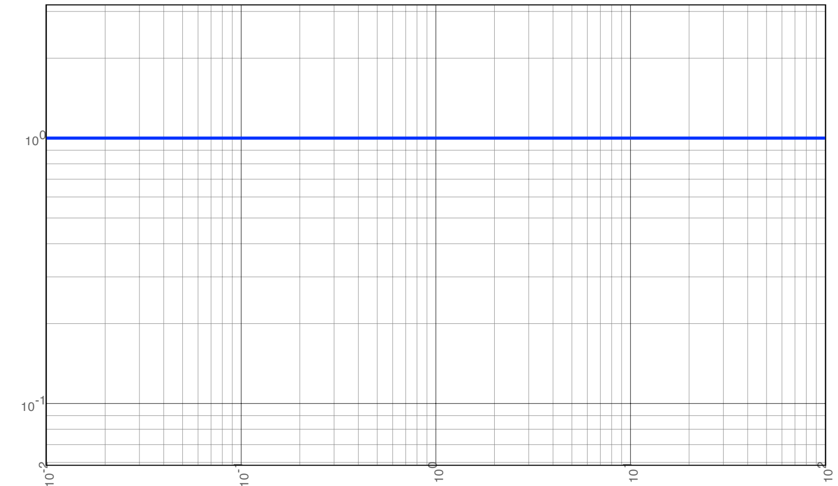
Bandpass



Peaking / Notch



Shelving

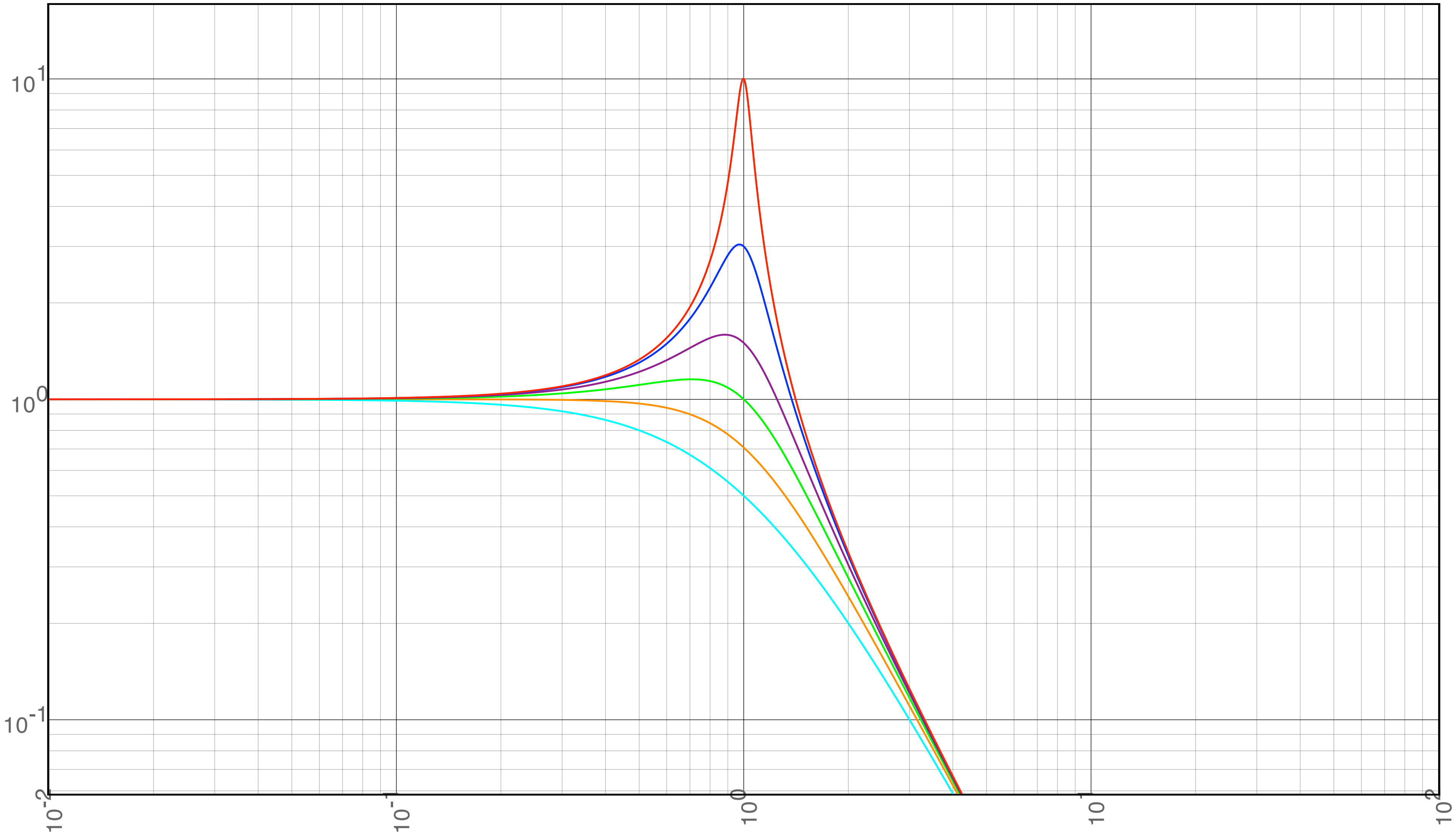


Allpass

Filter Q

- Filters can be described "quality factor" Q
  - Selectivity (bandwidth) relative to centre frequency

$$Q = \frac{f_0}{\Delta f}$$

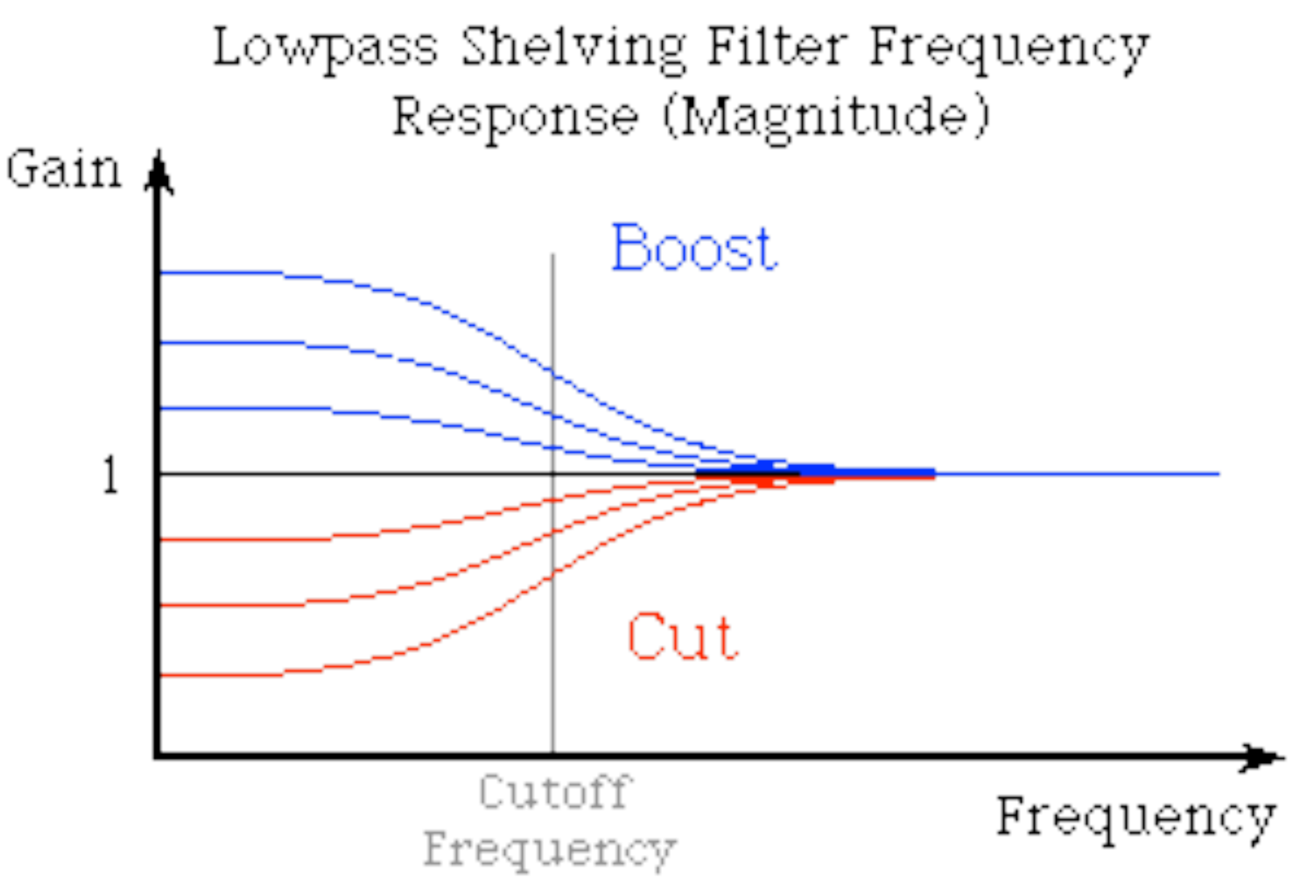


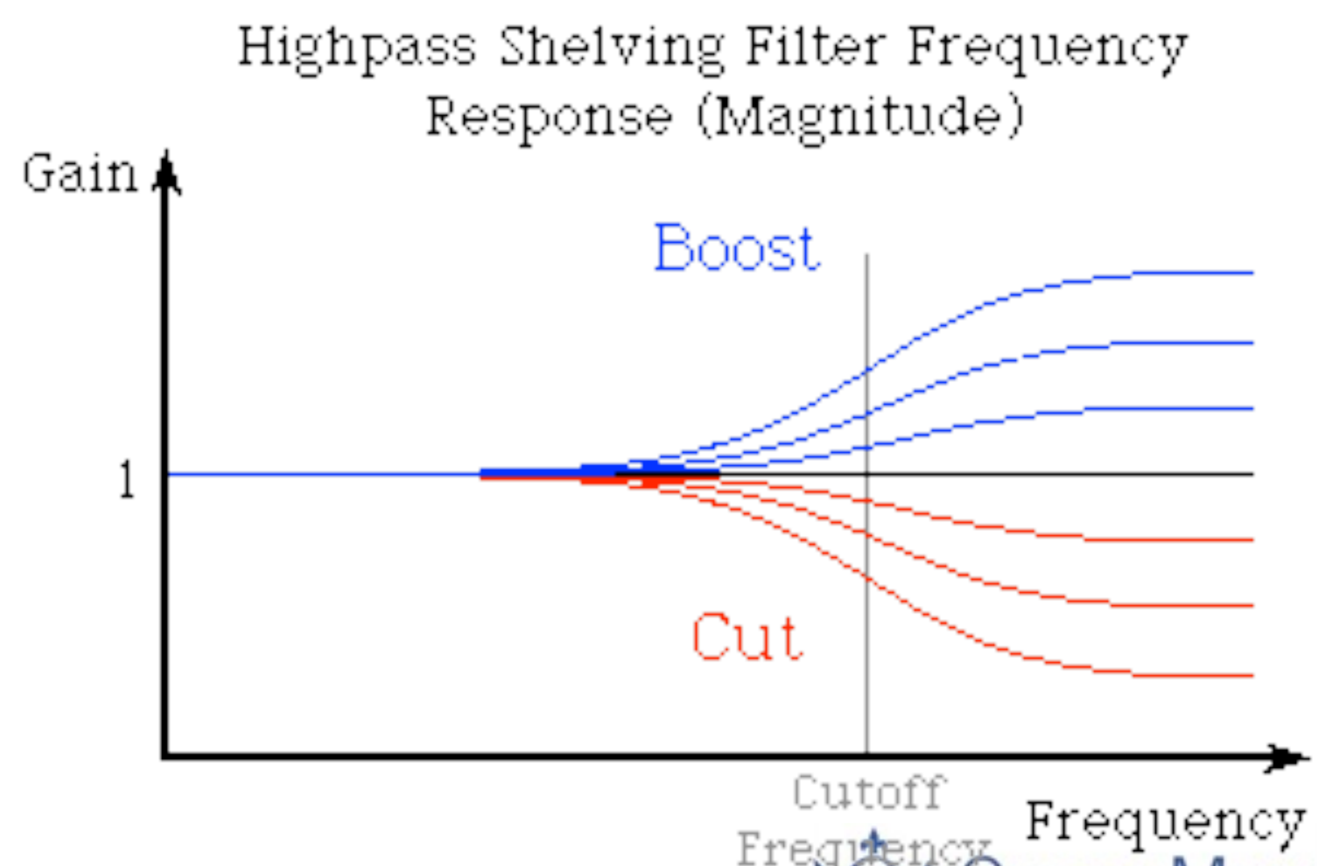
Second order lowpass filters (all)

Q = 0.5
Q = 0.71
Q = 1
Q = 1.5
Q = 3
Q = 10

Tone controls

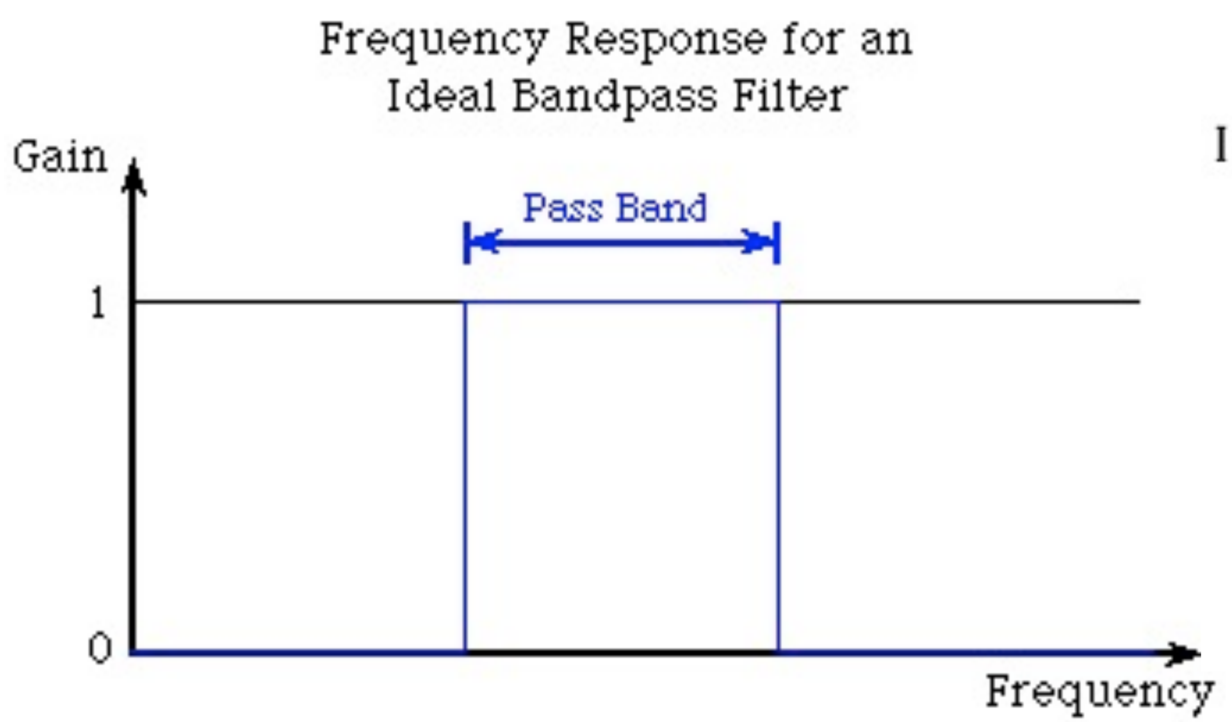
- Most common equalisation is the tone control
  - Controls: bass and treble
  - Adjust sound to suit your taste, or the room
  - Bass control = lowpass shelving filter
  - Treble control = highpass shelving filter
  - Control centred = unity gain = flat response

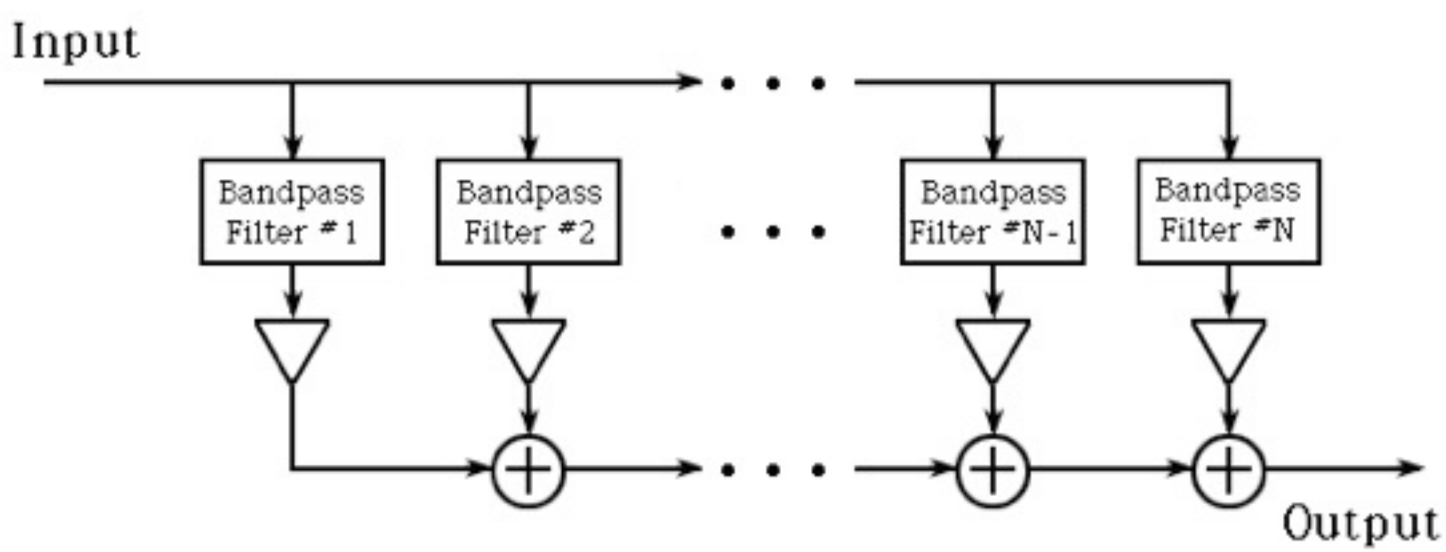




Graphic EQ (parallel implementation)

- Consider as bank of bandpass filters
  - Completely isolate individual frequency bands
  - Real bandpass filters have some ripple in passband and don't have a perfectly sharp cutoff



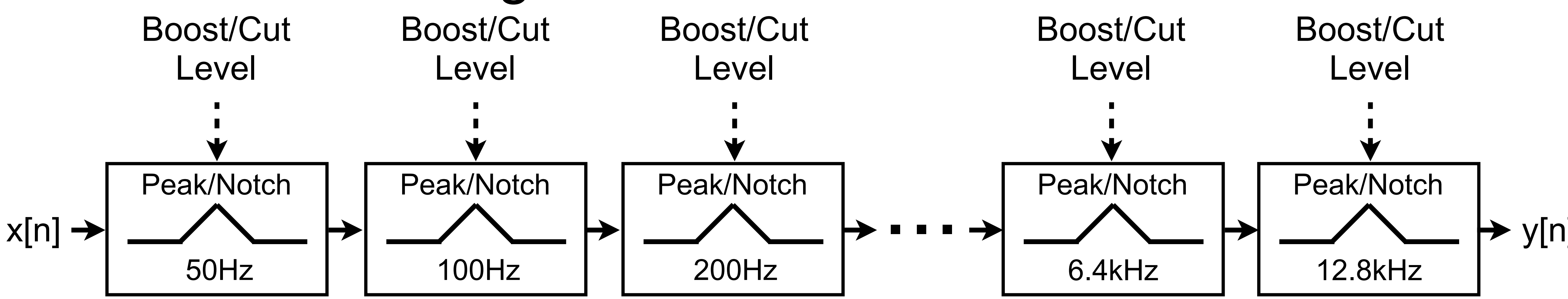


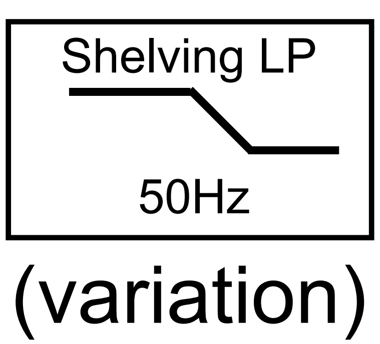
- Arrange filters in parallel
  - Each has the same input
  - Outputs are summed (mixed) after band gains adjusted
  - Also know: how band frequencies are chosen

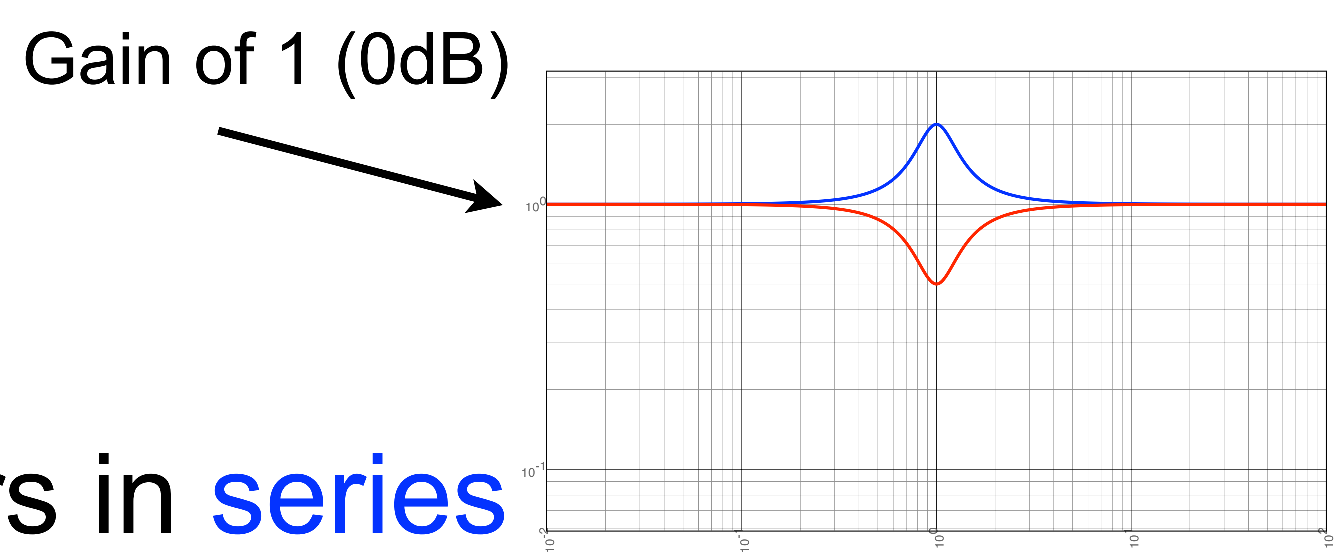
Queen Mary University of London

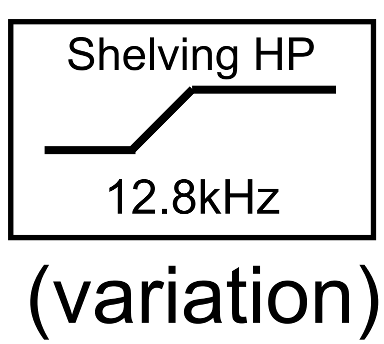
Graphic EQ (series implementation)

- Consider cascade of peaking or shelving filters
  - Add a boost or notch to a specific frequency band
    - Top and bottom bands may be shelving filters which extend to edge of spectrum
  - Each filter has a gain of 1 outside the band





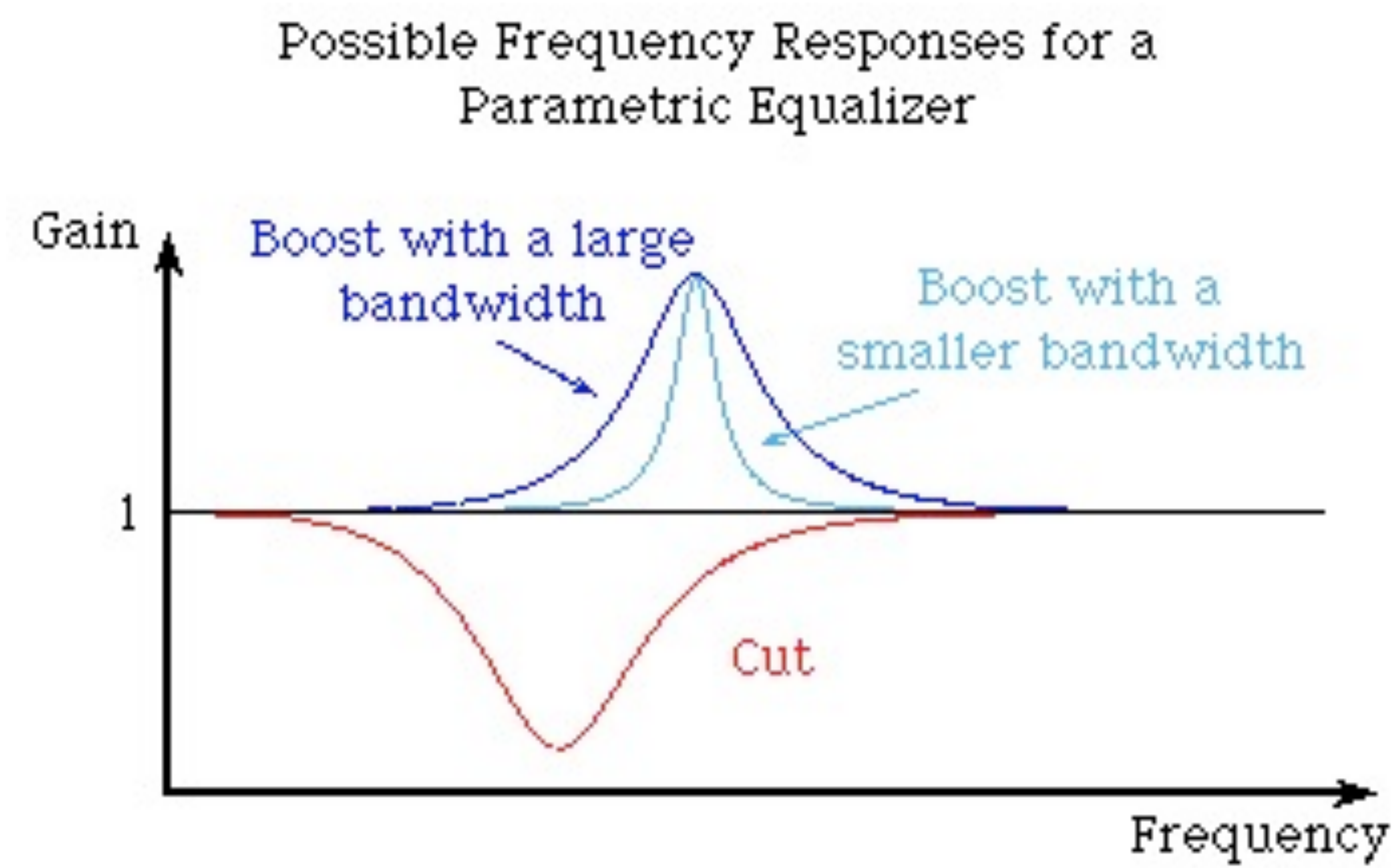




- Arrange filters in series
  - Each takes its input from the previous filter

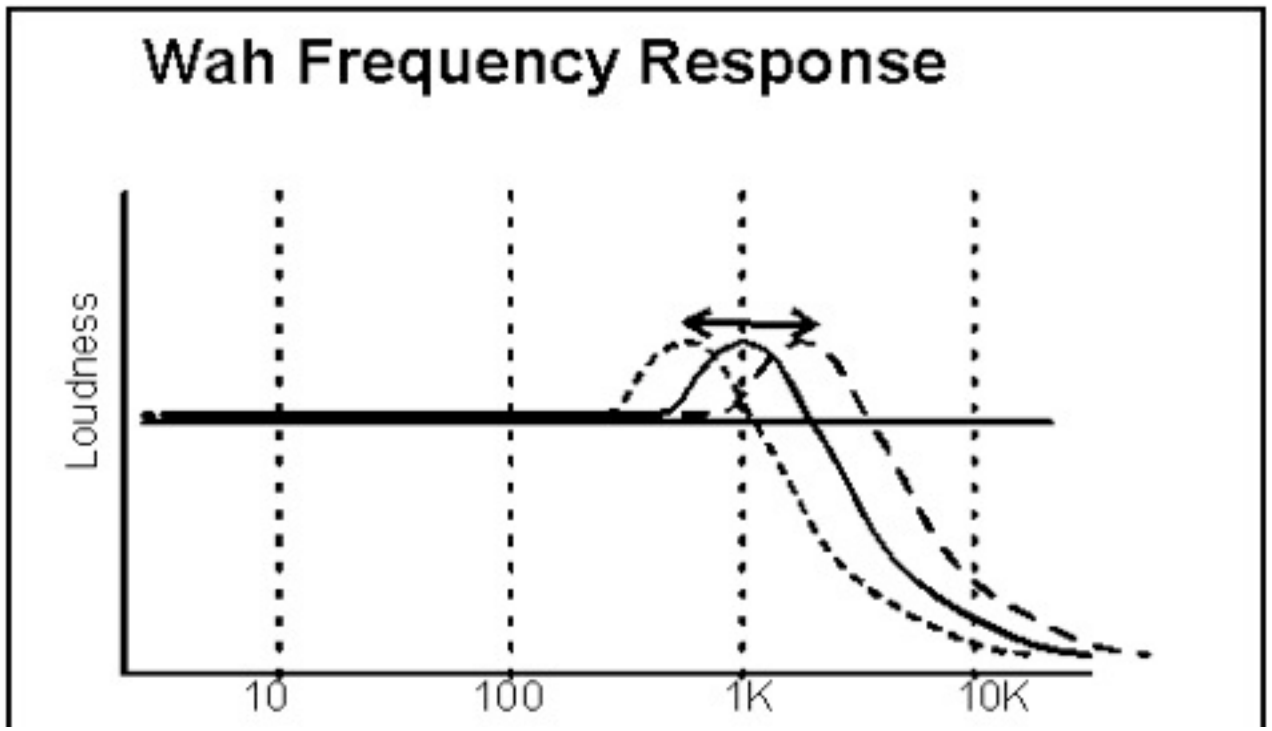
Parametric EQ plots

- Controls: centre frequency, gain, bandwidth
  - These plots for peaking filters (not shelving)



Wah-wah

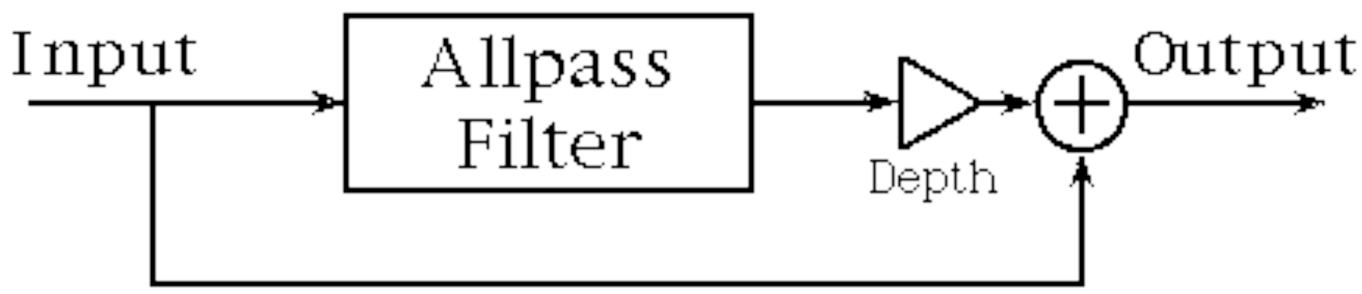
- Resonant (bandpass or peaking) filter
  - Variable centre frequency
  - Small bandwidth (i.e. high Q)
- Standard wah-wah
  - Centre frequency controlled by pedal (manual control)
- Auto-wah (2 effects)
  - Centre frequency controlled by LFO
  - Frequency of LFO usually around 1-2 Hz
  - OR, could depend on envelope of input signal



- Tremolo-wah: also amplitude variations
- Centre frequency can vary from near 0 to $f_s/2$
  - In practice, varies across the midrange frequencies only

Phasing

- Notches implemented indirectly by allpass filters
  - Allpass filter passes all frequencies equally
  - No amplification or attenuation: $|H(\omega)| = 1$
  - It's all about the phase
- Add filtered output to original
  - Amount of filtered output controlled by *depth* or *mix* control
- Where do the notches come from?



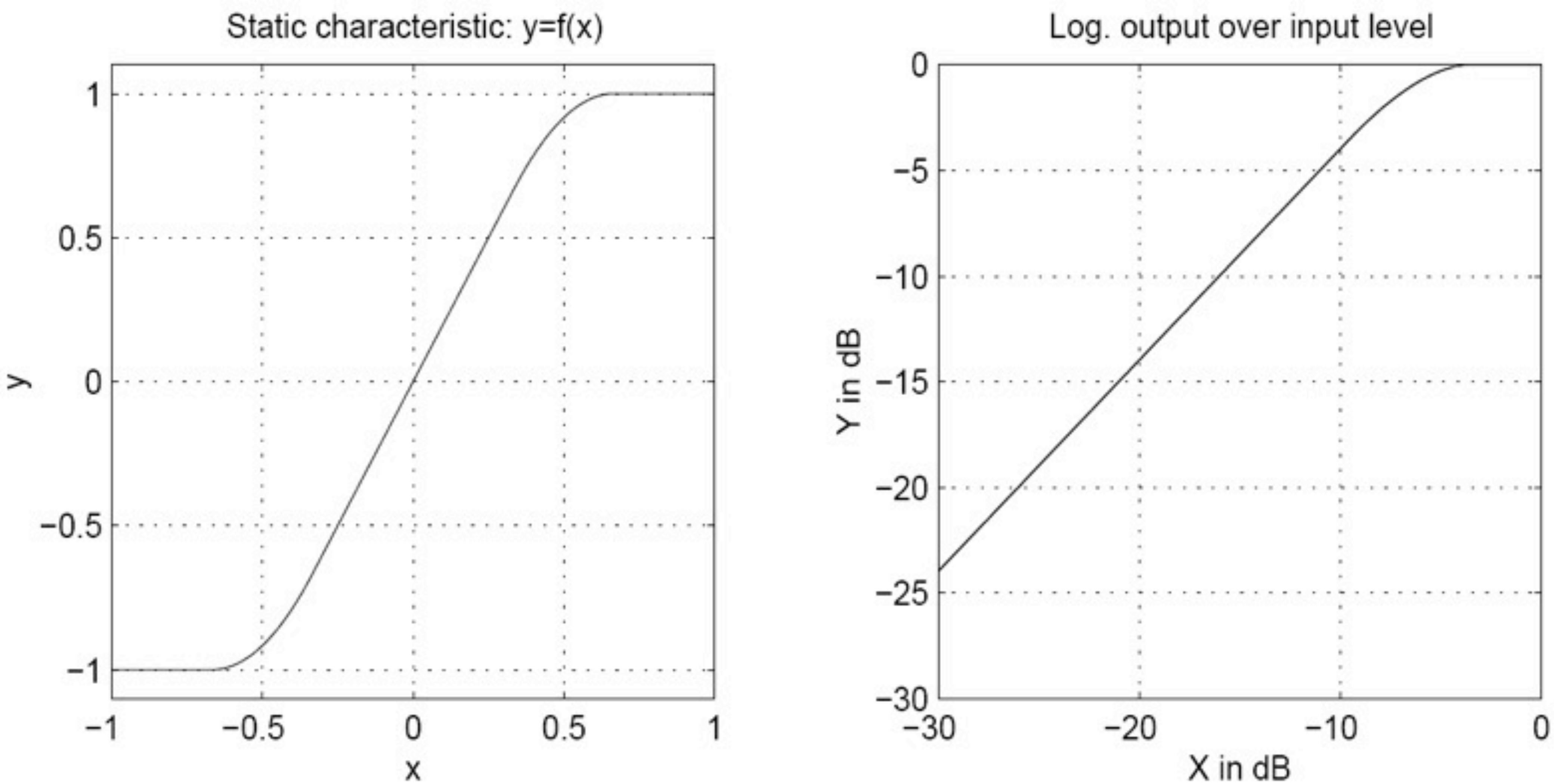
- Also know:
  - Parameters; variations (stereo, feedback, etc.)
  - Differences between phaser and flanger

Distortion (soft clipping variations)

- Most analogue systems have smooth transition from linear to non-linear
  - Rounded tops of waveforms known as soft clipping

$$f(x) = \begin{cases} 2x & \text{for} \quad 0 \le x \le 1/3 \\ \frac{3-(2-3x)^2}{3} & \text{for} \quad 1/3 \le x \le 2/3 \\ 1 & \text{for} \quad 2/3 \le x \le 1. \end{cases}$$

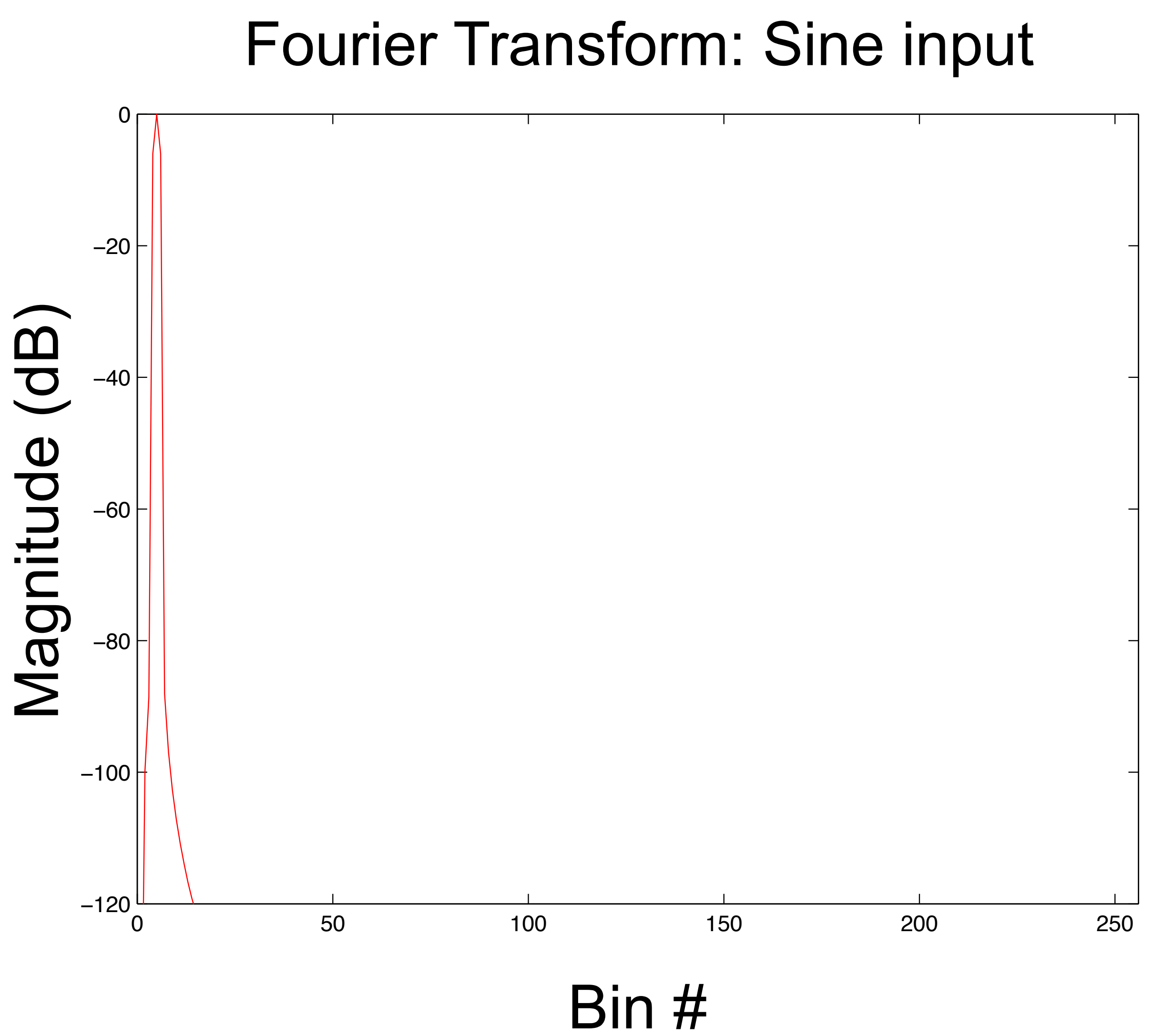
Notice: G = 2

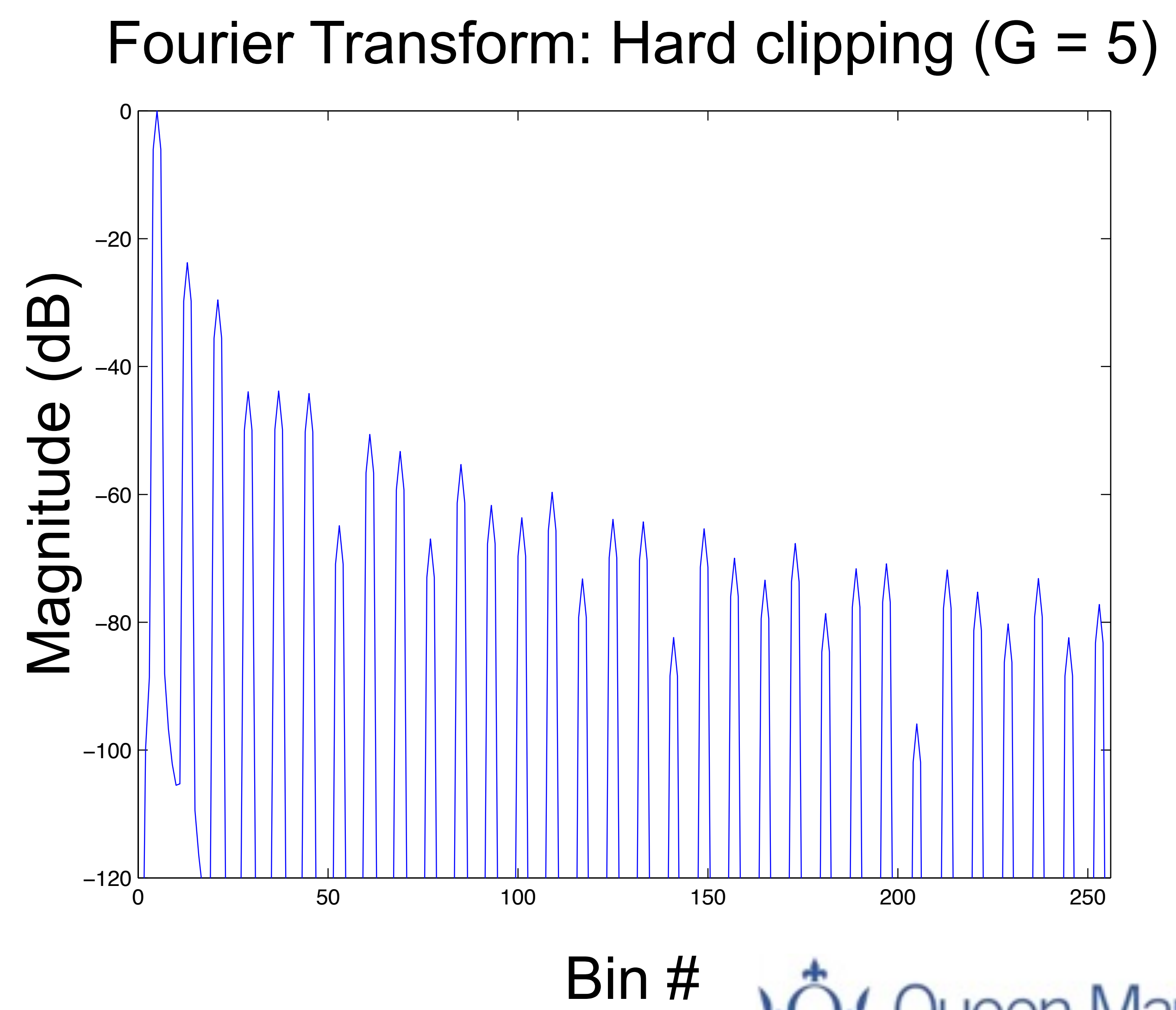


Queen Mary University of London

Distortion in the frequency domain

- Distortion is a non-linear effect (why?)
  - $f(x_1 + x_2) \neq f(x_1) + f(x_2)$
  - Non-linearity introduces new frequency components that weren't present in the original signal





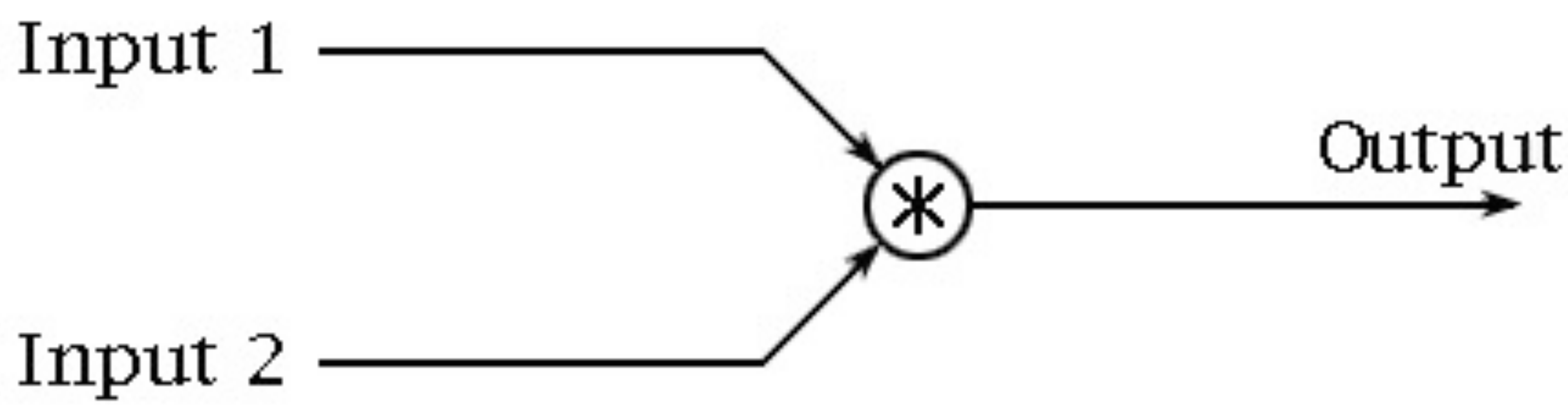
Intermodulation distortion

- The bane of fuzzbox technology
- Trig. identity: $\sin(x) + \sin(y) = 2\sin\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right)$
  - Sum and difference frequencies are products of intermodulation between inputs
  - When we apply non-linear distortion, intermodulation products will show up in the spectrum
  - If two tones are harmonically-related
    - Intermodulation products will be harmonically-related as well
  - If two tones are not harmonically-related
    - Intermodulation produces will be harsh and dissonant
    - Exact guitar tuning is *especially* important when using distortion!
- Also know:
  - Aliasing (and addressing it); hard clipping; (a)symmetry
  - Rectification, other variations

Queen Mary
University of London

Ring modulation

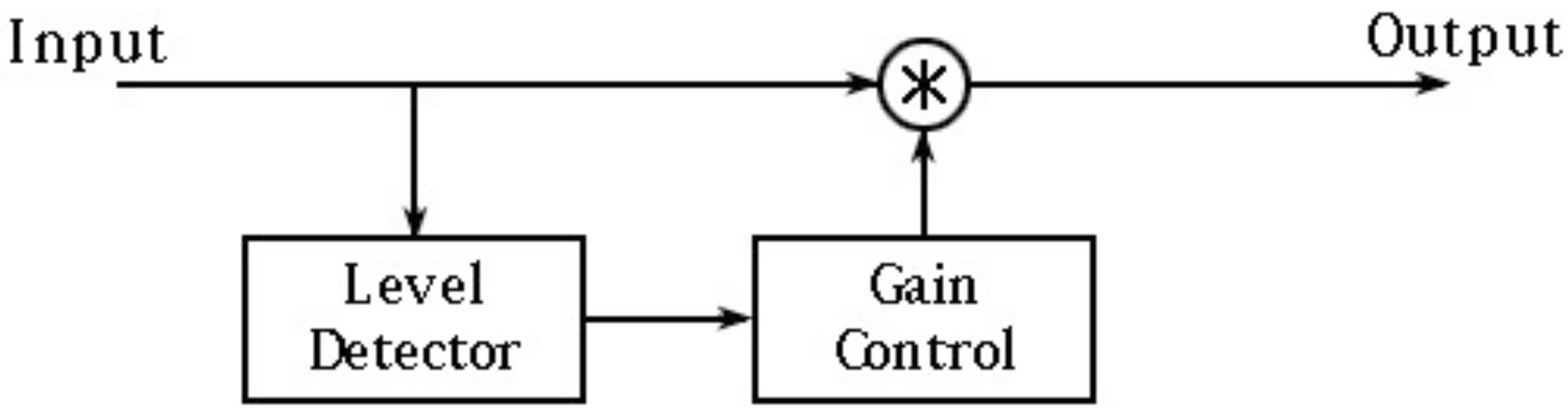
- Audio-frequency amplitude modulation



  - Multiply signal by a sinusoid
- Resulting frequencies will differ from input
  - Sum and difference frequencies
    - Known as sidebands
  - Output not harmonic
    - Sounds can be quite dissonant, even for simple inputs
    - Not widely used for this reason

Compression

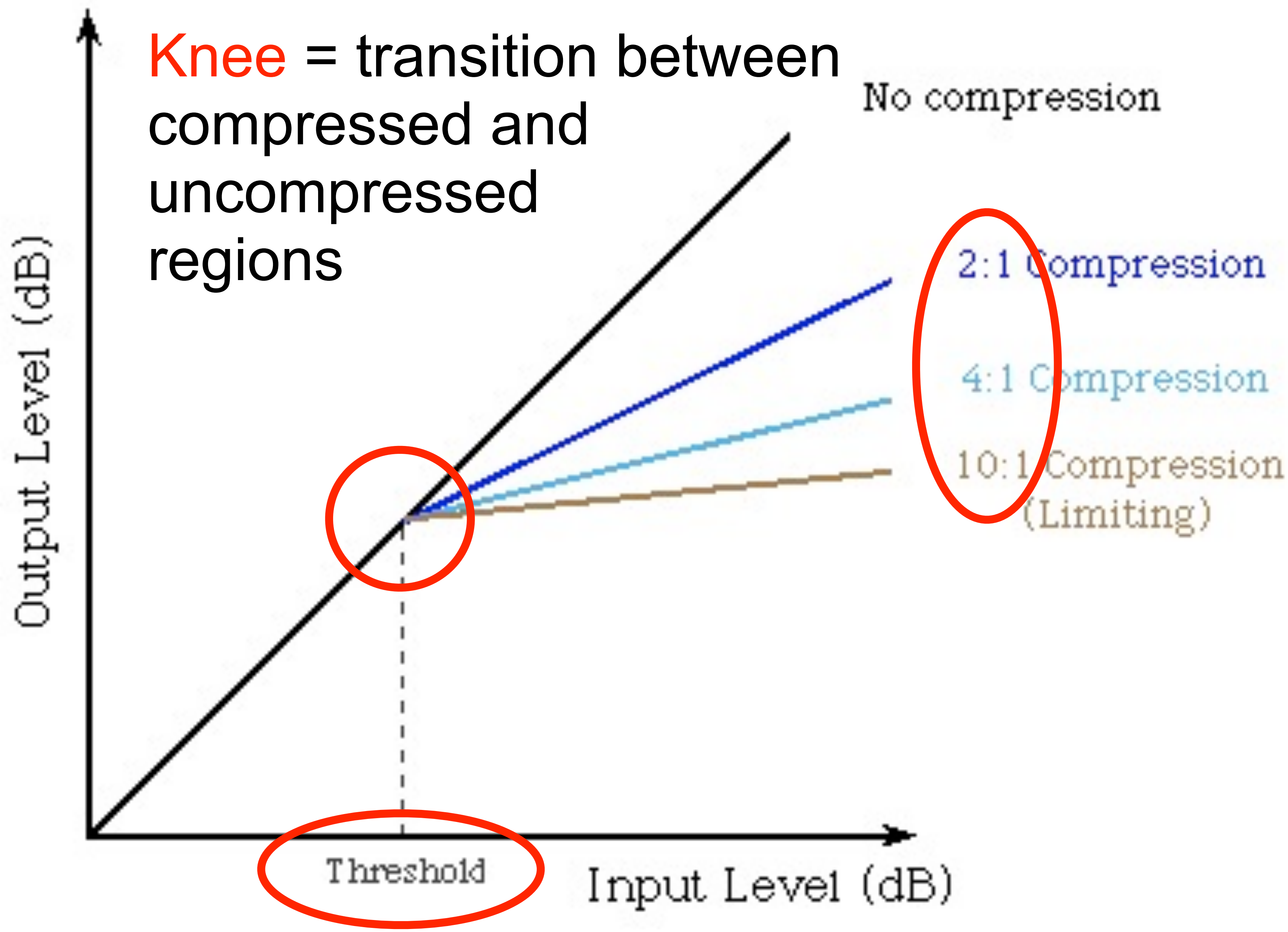
- Compressor = amplifier with variable gain control
  - Gain always ≤ 1
- Crucially: gain controlled by level of input
  - High signal level → gain ≪ 1
  - Low signal level → gain ≈ 1



- Two ways to calculate signal level
  - Before or after compression is applied
    - "Feedforward" and "feedback" strategies
    - We have mostly seen feedforward version

Compression terminology

Knee = transition between compressed and uncompressed regions



Threshold = input level at which compressor activates
(below this, gain = 1)

Ratio = amount that increasing input level affects output level
(not the same as gain -- why?)

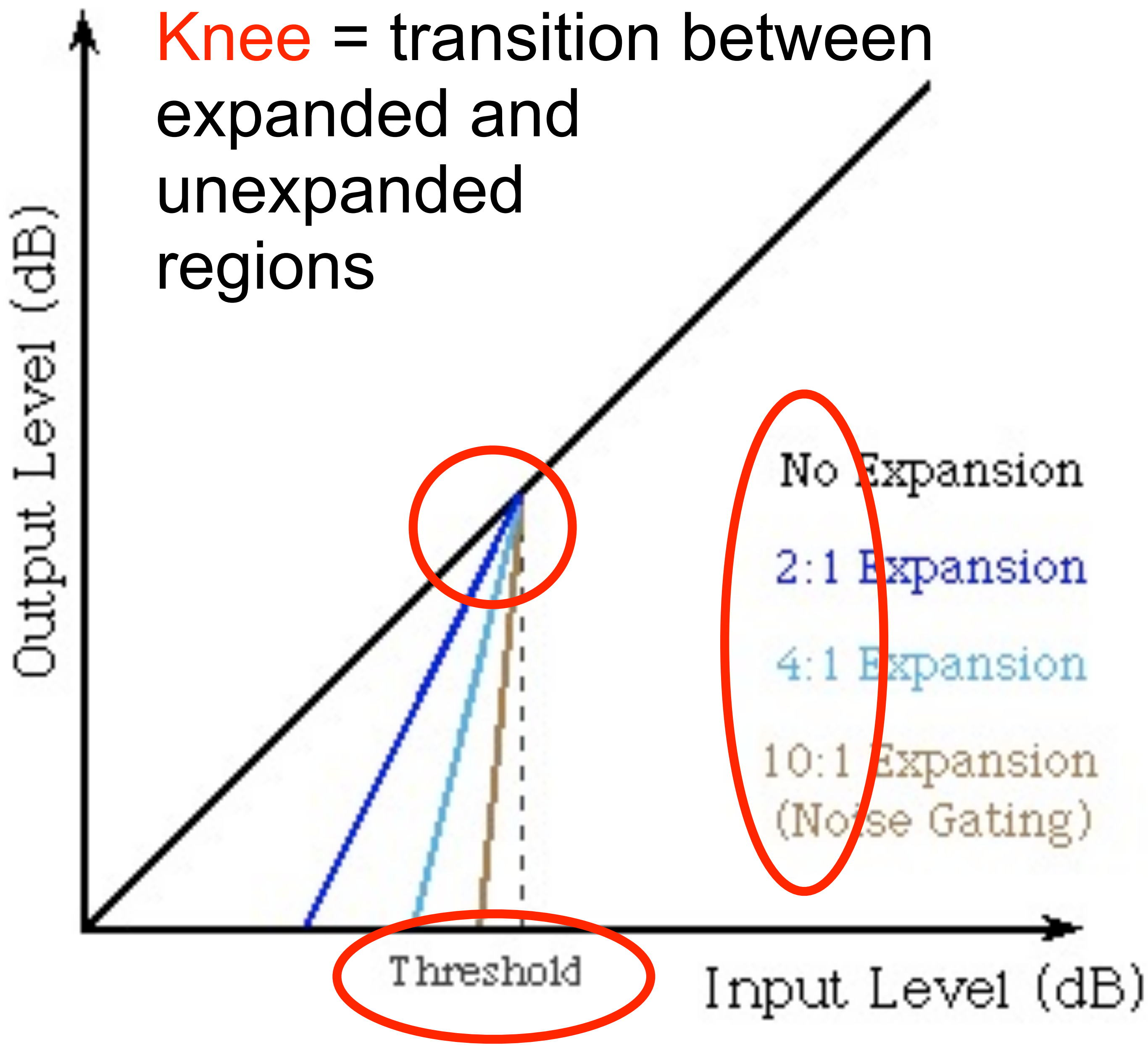
Queen Mary
University of London

Attack and release time

- Attack time and release time are properties of the level detector (not the gain control)
  - Assume level detector operates over a time window
  - Attack time = how long it takes to respond when input rises above the threshold
    - Usually short, less than 100ms
  - Release time = how long it takes to respond when input falls below the threshold
    - Usually larger than attack time, e.g. 1-2 seconds
- Also know:
  - Limiting, ducking, look-ahead, de-essing, etc.
  - Difference between compression and distortion
  - Pumping and breathing (and how to deal with them)

Expansion

Knee = transition between expanded and unexpanded regions



Threshold = input level at which expander activates
(above this, gain = 1)

Ratio = amount that decreasing input level affects output level
(not the same as gain)

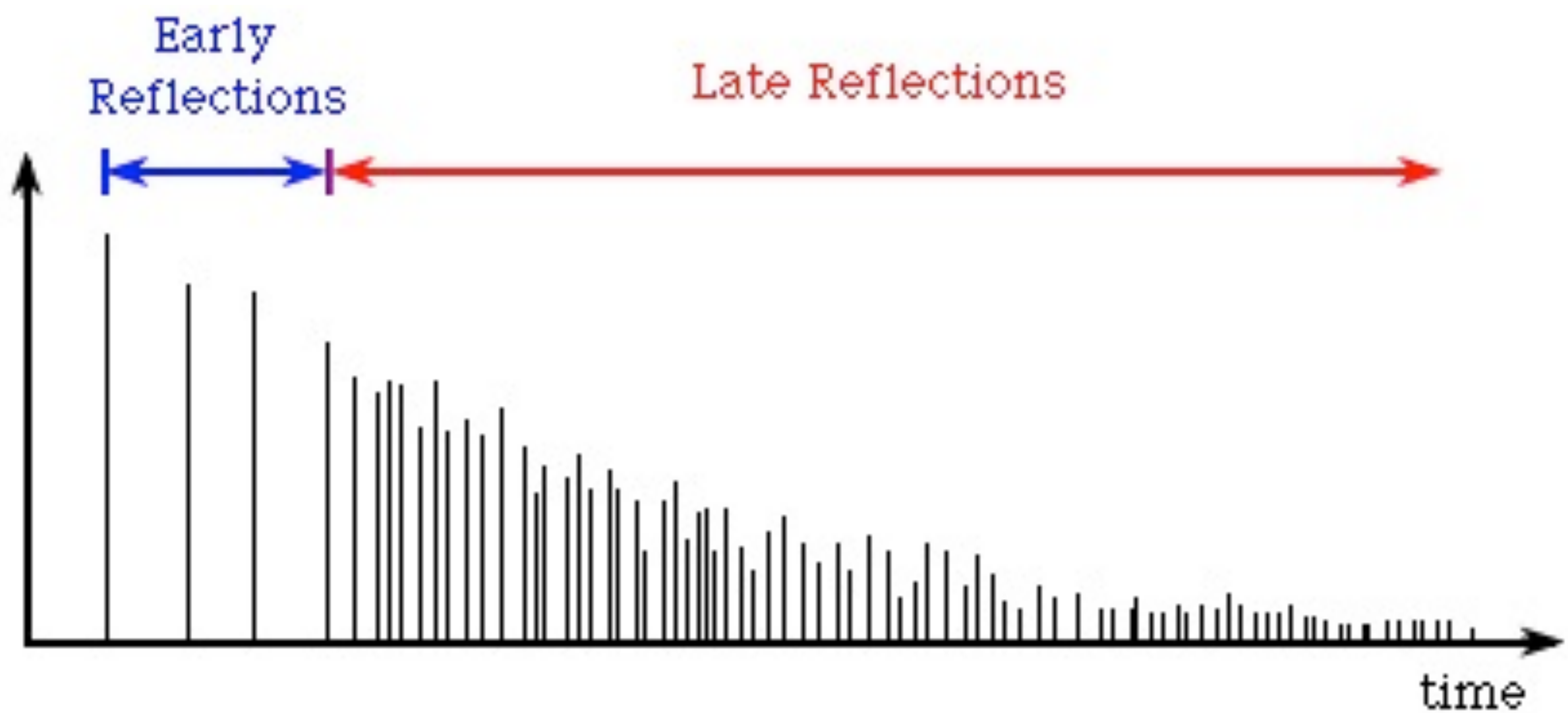
Also know: noise gating

Queen Mary
University of London

Reverb, echo and delay

- Reverb is not just a series of echoes
  - Echo: distinct, delayed version of sound
    - Delay more than 0.1-0.2 seconds
  - Reverb: many closely-spaced copies of a sound
    - Do not perceive reflections as echoes
    - However, the overall effect is highly audible
- Delay with feedback does not produce reverb
  - Delay: reflections with fixed time interval
  - Reverb: rate of arriving reflections changes over time
    - Greater spacing of initial reflections, closer spacing later

Reverb impulse response



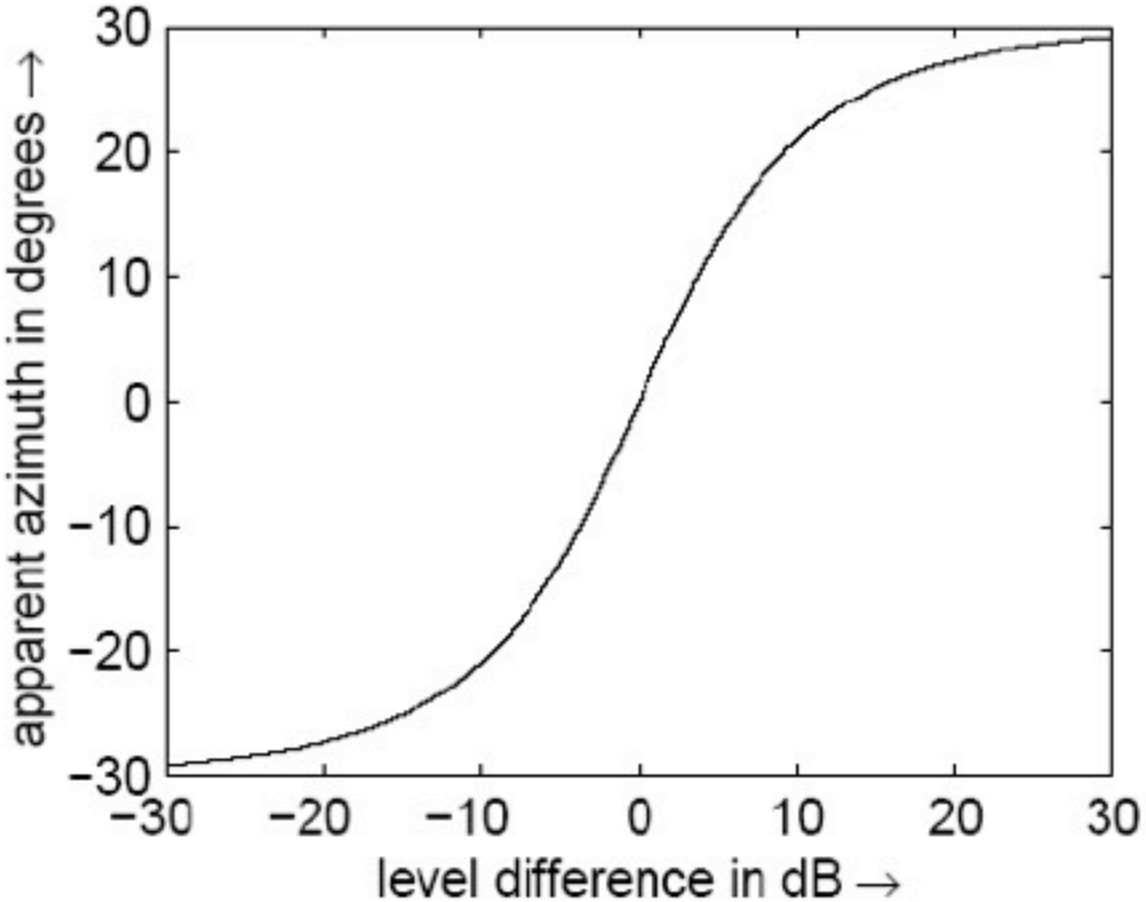
- Series of scaled delta functions
  - Each line represents a delayed, attenuated copy of sound
  - What would a simple feedback delay look like?

Phase vocoder

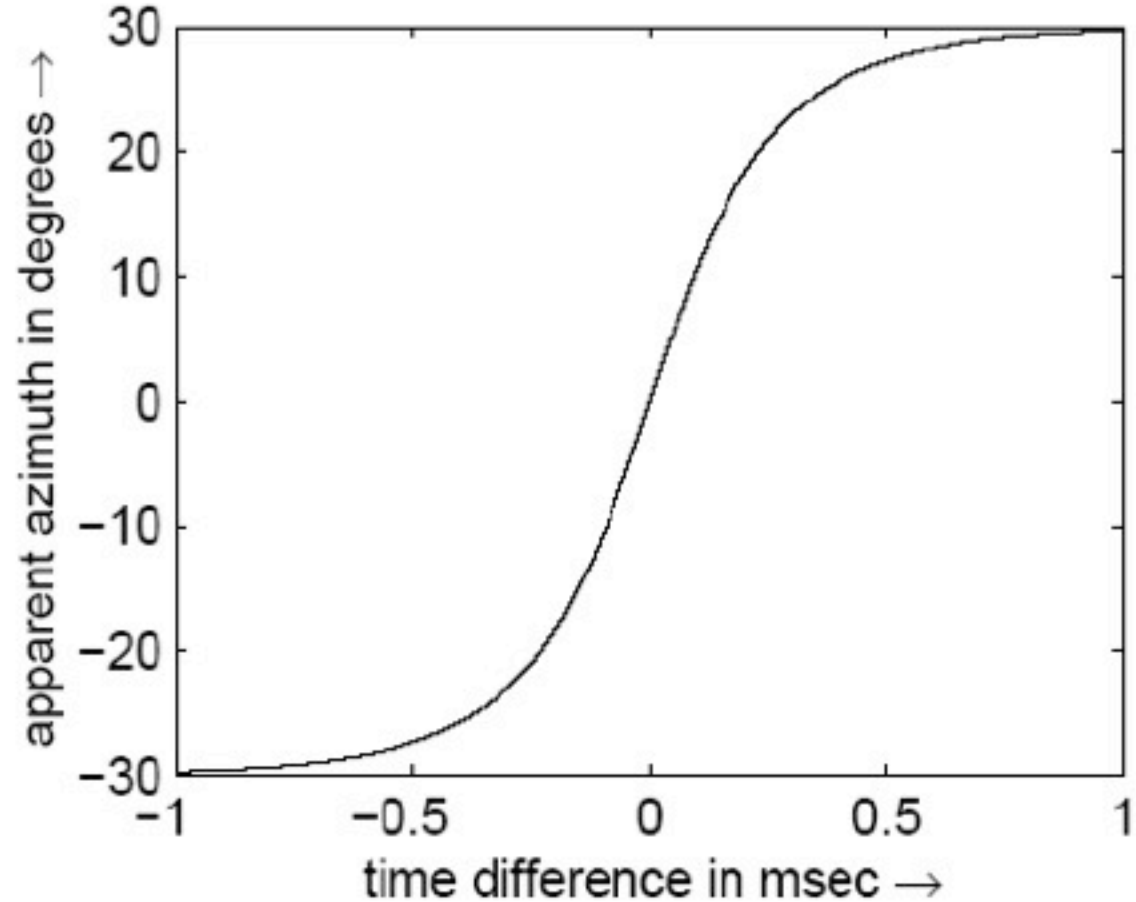
- Frequency-domain processing
  - Based on periodic windows processed by DFT/IDFT
- Overlap-Add processing method:
  1. Take *m*th segment (frame) of length M using windowing function
  2. Take DFT of length N ≥ M of segment
     - If N > M, zero-pad the segment (add zeros to end)
  3. Do something interesting to frequency data
  4. Take IDFT to get back to time domain segment
  5. Add the result to the output buffer containing the prior segments
  6. Advance by the hop size to *(m+1)*th frame and repeat
- Also know:
  - Robotisation, whisperisation, mutation, denoising, pitch-shift or time-stretch (basics only)
  - Latency considerations

Spatial audio: gain and delay

- Two primary tools for placing sounds in space



Two simultaneous sounds
with level difference



Two equal amplitude sounds
with time difference

- Interaural Level Difference and Interaural Time Difference
- > 1ms or > 30dB: sound localised to one source
- Also know: HRTF (basic concept, not equations)

Queen Mary
University of London