

2018 Fall  
**CTP431: Music and Audio Computing**

# Digital Audio Effects

Graduate School of Culture Technology, KAIST  
Juhan Nam



# Introduction

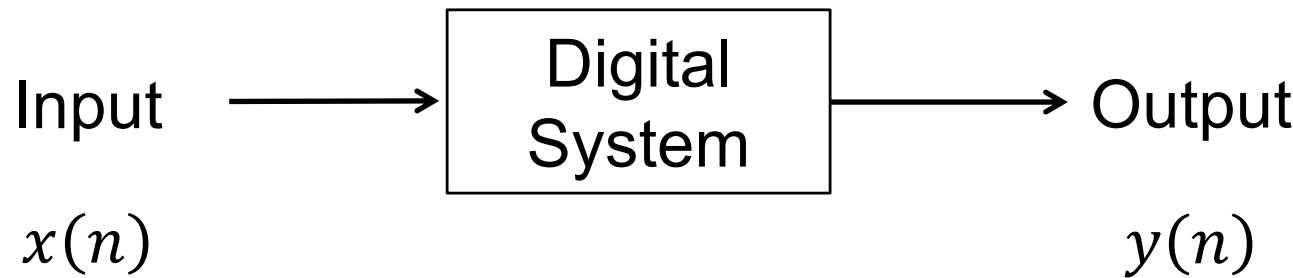
- Amplitude
  - Gain, fade in/out, automation curve, compressor
- Timbre
  - Filters, EQ, distortion, modulation, flanger, vocoder
- Pitch
  - Pitch shifting, transpose
- Time stretching
  - Timing change, tempo adjustment
- Spatial effect
  - Delay, Reverberation, panning, binaural (HRTF)



Source: <http://www.uaudio.com/uad/downloads>

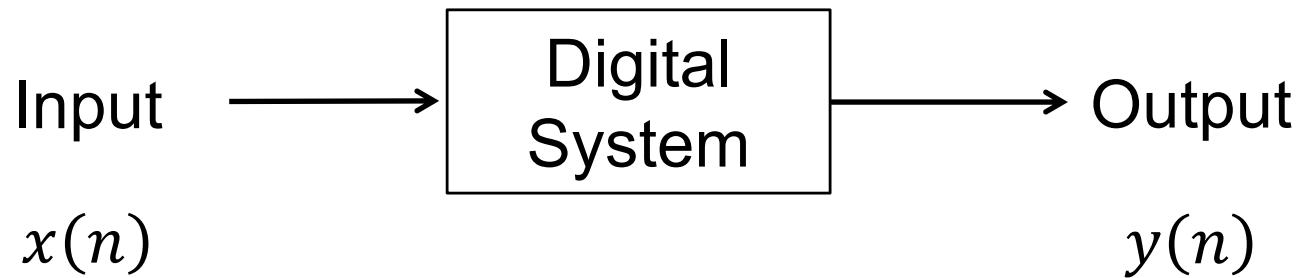
Source: <https://www.izotope.com/en/products/repair-and-edit/rx-post-production-suite.html>

# Digital System



- Take the input signal  $x(n)$  as a sequence of numbers and returns the output signal  $y(n)$  as another sequence of numbers
- We are particularly interested in **linear systems** that are composed of the following operations
  - Multiplication:  $y(n) = b_0 \cdot x(n)$
  - Delaying:  $y(n) = x(n - 1)$
  - Summation:  $y(n) = x(n) + x(n - 1)$

# Linear Time-Invariant (LTI) System



- Linearity
  - Homogeneity: if  $x(n) \rightarrow y(n)$ , then  $a \cdot x(n) \rightarrow a \cdot y(n)$
  - Superposition: if  $x_1(n) \rightarrow y_1(n)$  and  $x_2(n) \rightarrow y_2(n)$ , then  $x_1(n) + x_2(n) \rightarrow y_1(n) + y_2(n)$
- Time-Invariance
  - If  $x(n) \rightarrow y(n)$ , then  $x(n - N) \rightarrow y(n - N)$  for any  $N$
  - This means that the system does not change its behavior over time

# LTI System

- LTI systems in frequency domain
  - No new sinusoidal components are introduced
  - Only existing sinusoids components changes in amplitude and phase.
- Examples of non-LTI systems
  - Clipping
  - Distortion
  - Aliasing
  - Modulation

# LTI Digital Filters

- A LTI digital filters performs a combination of the three operations
  - $y(n) = b_0 \cdot x(n) + b_1 \cdot x(n - 1) + b_2 \cdot x(n - 2) + \dots + b_M \cdot x(n - M)$
- This is a general form of **Finite Impulse Response (FIR) filter**

# Two Ways of Defining LTI Systems

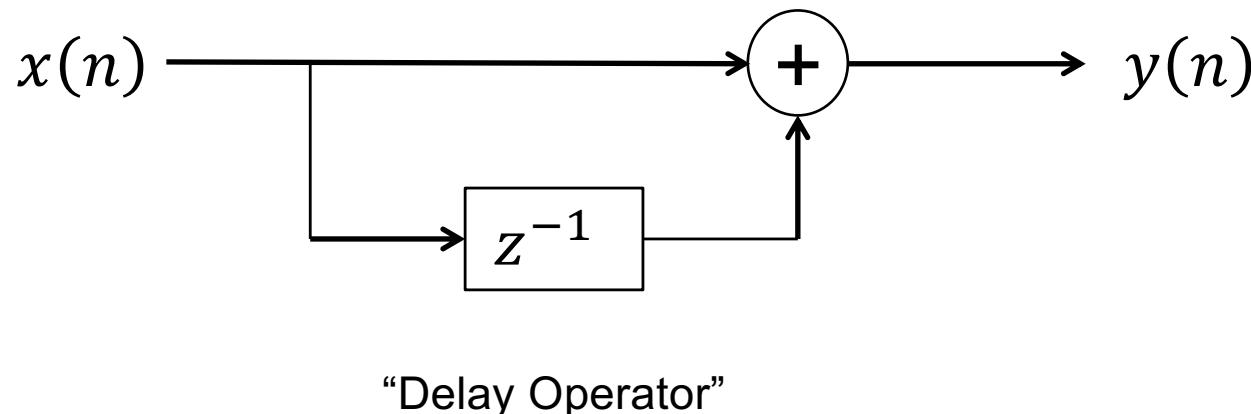
- By the relation between input  $x(n)$  and output  $y(n)$ 
  - Difference equation
  - Signal flow graph
- By the impulse response of the system
  - Measure it by using a unit impulse as input
  - Convolution operation

# The Simplest Lowpass Filter

- Difference equation

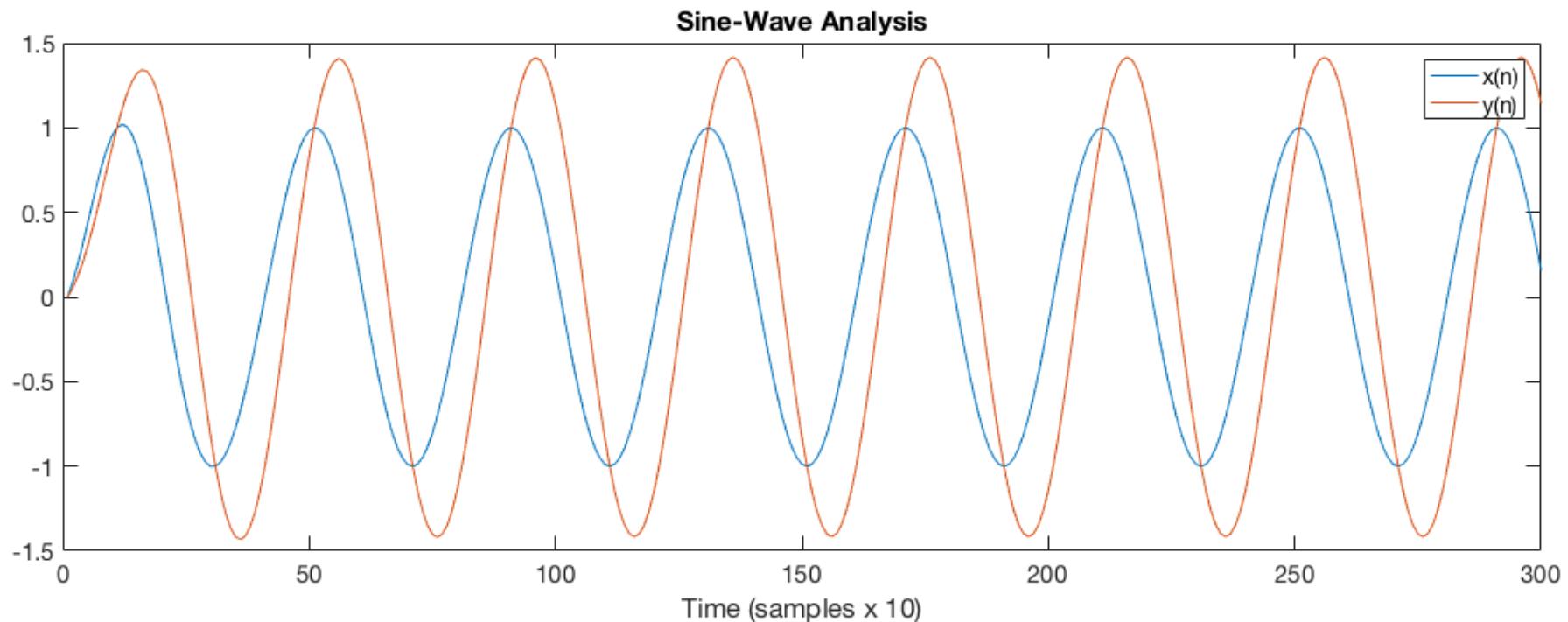
$$y(n) = x(n) + x(n - 1)$$

- Signal flow graph



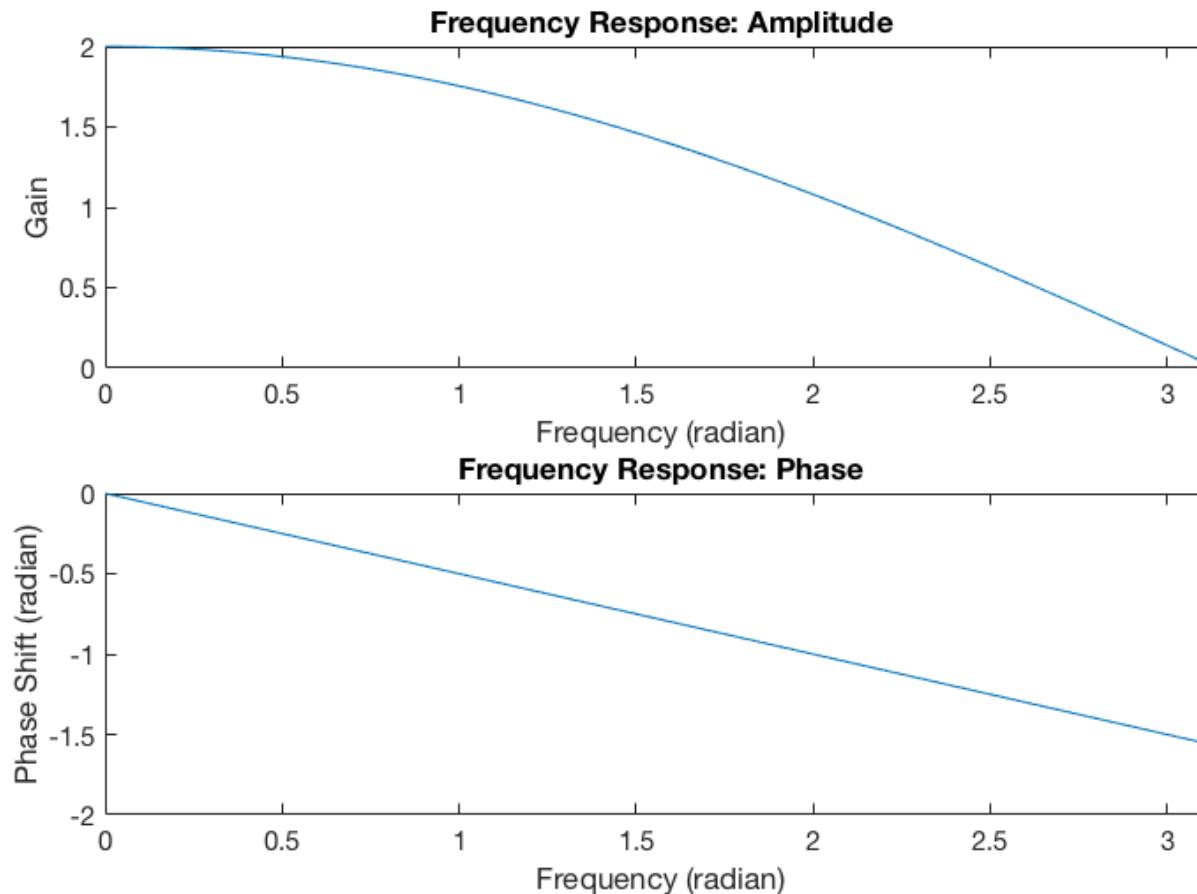
# The Simplest Lowpass Filter: Sine-Wave Analysis

- Measure the amplitude and phase changes given a sinusoidal signal input



# The Simplest Lowpass Filter: Frequency Response

- Plot the amplitude and phase change over different frequency
  - The frequency sweeps from 0 to the Nyquist rate



# The Simplest Lowpass Filter: Frequency Response

- Mathematical approach

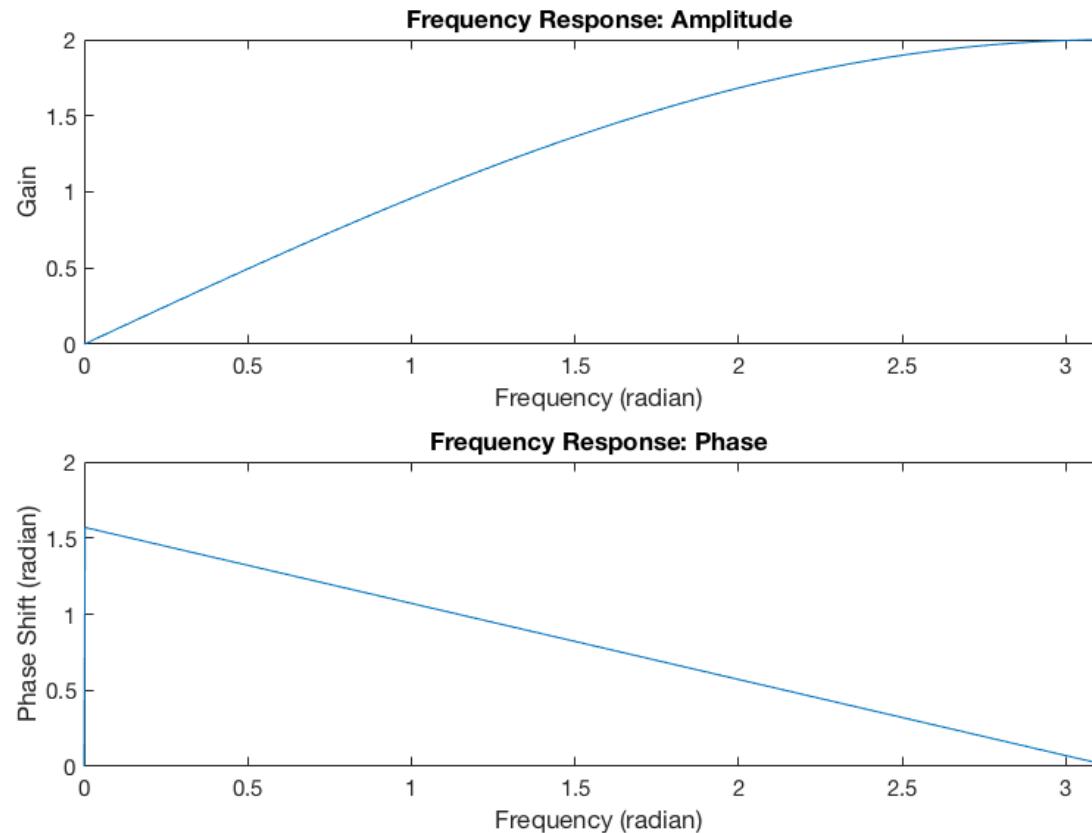
- Use complex sinusoid as input:  $x(n) = e^{j\omega n}$
  - Then, the output is:

$$y(n) = x(n) + x(n-1) = e^{j\omega n} + e^{j\omega(n-1)} = (1 + e^{-j\omega}) \cdot e^{j\omega n} = (1 + e^{-j\omega}) \cdot x(n)$$

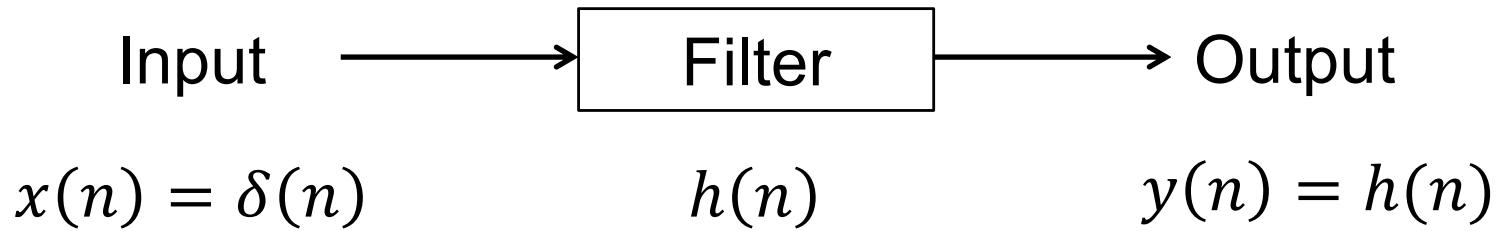
- Frequency response:  $H(\omega) = (1 + e^{-j\omega}) = \left(e^{j\frac{\omega}{2}} + e^{-j\frac{\omega}{2}}\right) e^{-j\frac{\omega}{2}} = 2\cos\left(\frac{\omega}{2}\right)e^{-j\frac{\omega}{2}}$
  - Amplitude response:  $|H(\omega)| = 2 \cos\left(\frac{\omega}{2}\right)$
  - Phase response:  $\angle H(\omega) = -\frac{\omega}{2}$

# The Simplest Highpass Filter

- Difference equation:  $y(n) = x(n) - x(n - 1)$
- Frequency response



# Impulse Response



- The filter output when the input is a unit impulse
  - $x(n) = \delta(n) = [1, 0, 0, 0, \dots] \rightarrow y(n) = h(n)$
- Characterizes **the digital system as a sequence of numbers**
  - A system is represented just like audio samples!

# Examples: Impulse Response

- The simplest lowpass filter
  - $h(n) = [1, 1]$
- The simplest highpass filter
  - $h(n) = [1, -1]$
- Moving-average filter (order=5)
  - $h(n) = \left[ \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5} \right]$
- General FIR Filter
  - $h(n) = [b_0, b_1, b_2, \dots, b_M] \rightarrow$  A finite length of impulse response

# Convolution

- The output of LTI digital filters is represented by **convolution operation** between  $x(n)$  and  $h(n)$

$$y(n) = x(n) * h(n) = \sum_{i=0}^M x(i) \cdot h(n - i)$$

- Deriving convolution
  - The input can be represented as a time-ordered set of weighted impulses
    - $x(n) = [x_0, x_1, x_2, \dots, x_M] = x_0 \cdot \delta(n) + x_1 \cdot \delta(n - 1) + x_2 \cdot \delta(n - 2) + \dots + x_M \cdot \delta(n - M)$
  - By the linearity and time-invariance
    - $y(n) = x_0 \cdot h(n) + x_1 \cdot h(n - 1) + x_2 \cdot h(n - 2) + \dots + x_M \cdot h(n - M) = \sum_{i=0}^M x(i) \cdot h(n - i)$

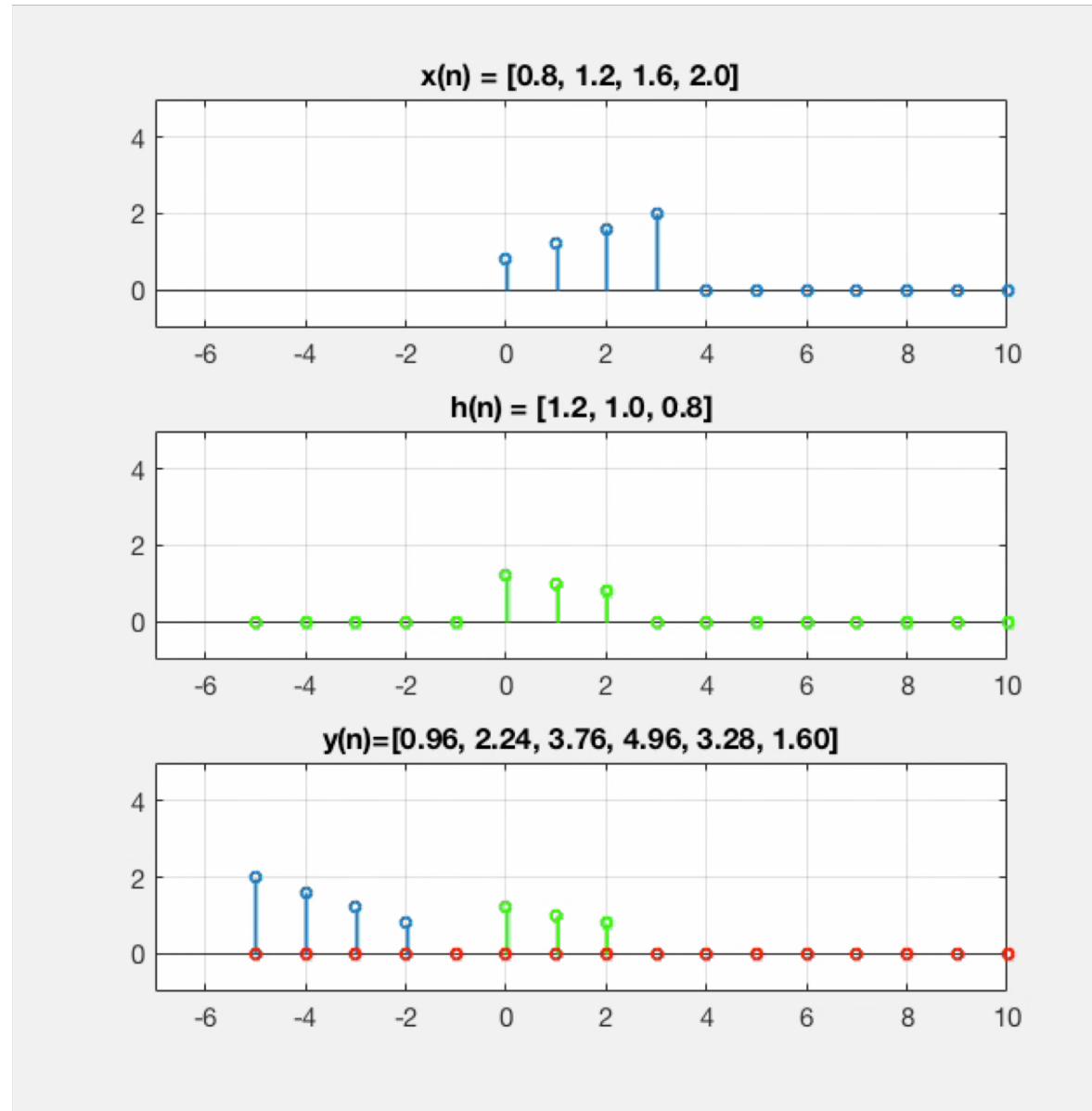
# Convolution In Practice

- The practical expression of convolution

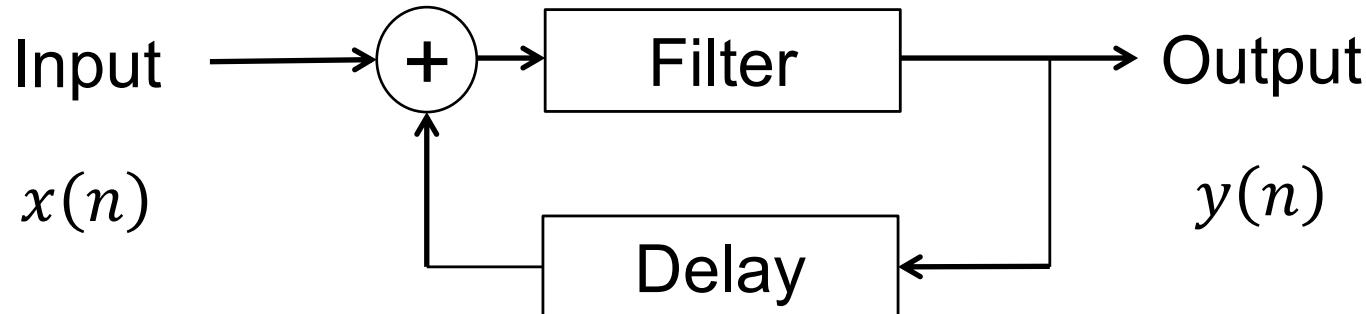
$$y(n) = x(n) * h(n) = \sum_{i=0}^M x(i) \cdot h(n-i) = \boxed{\sum_{i=0}^M h(i) \cdot x(n-i)}$$

- This represents input  $x(n)$  as a streaming data to the filter  $h(n)$
- The length of convolution output
  - If the length of  $x(n)$  is M and the length of  $h(n)$  is N, the length of  $y(n)$  is  $M+N-1$

# Demo: Convolution



# Feedback Filter



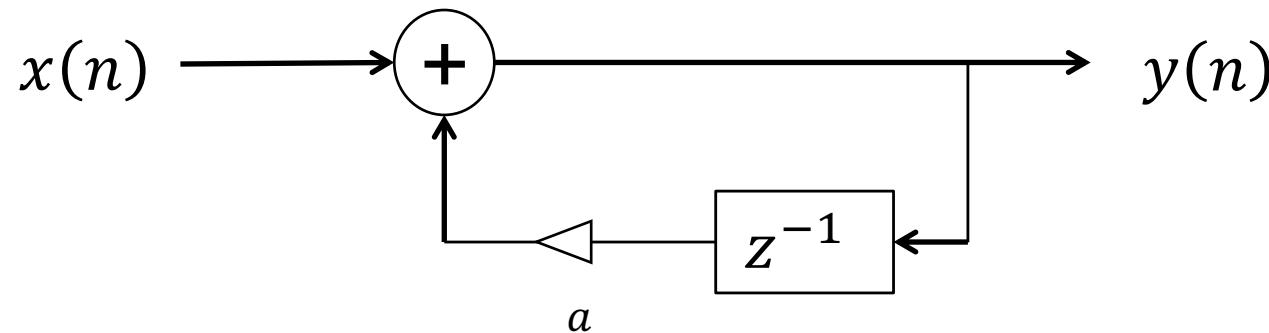
- LTI digital filters allow to use the past outputs as input
  - Past outputs:  $y(n - 1), y(n - 2), \dots, y(n - N)$
- The whole system can be represented as
  - $y(n) = b_0 \cdot x(n) + a_1 \cdot y(n - 1) + a_2 \cdot y(n - 2) + \dots + a_N \cdot y(n - N)$
  - This is a general form of **Infinite Impulse Response (IIR) filter**

# A Simple Feedback Lowpass Filter

- Difference equation

$$y(n) = x(n) + a \cdot y(n - 1)$$

- Signal flow graph



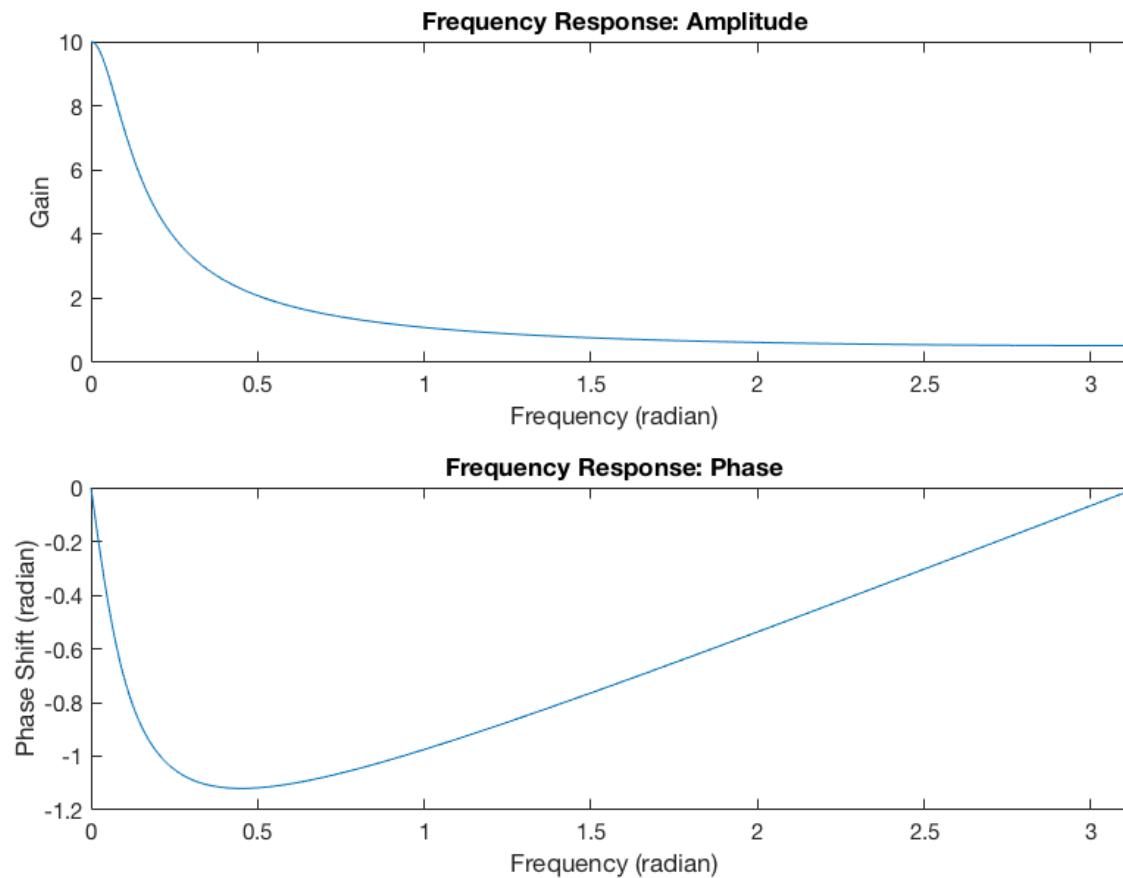
- When  $a$  is slightly less than 1, it is called “Leaky Integrator”

# A Simple Feedback Lowpass Filter: Impulse Response

- Impulse response
  - $y(0) = x(0) = 1$
  - $y(1) = x(1) + a \cdot y(0) = a$
  - $y(2) = x(2) + a \cdot y(1) = a^2$
  - ...
  - $y(n) = x(n) + a \cdot y(n - 1) = a^n$
- **Stability!**
  - If  $a < 1$ , the filter output converges (stable)
  - If  $a = 1$ , the filter output oscillates (critical)
  - If  $a > 1$ , the filter output diverges (unstable)

# A Simple Feedback Lowpass Filter: Frequency Response

- More dramatic change than the simplest lowpass filter (FIR)
  - Phase response is not linear



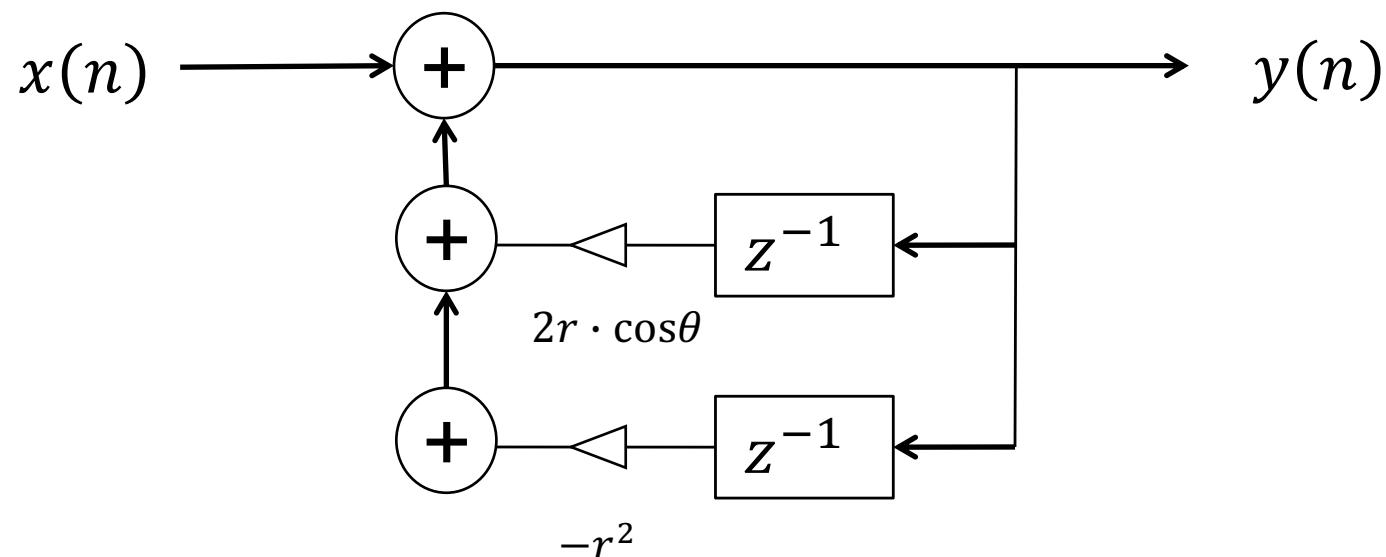
$$y(n) = x(n) + 0.9 \cdot y(n - 1)$$

# Reson Filter

- Difference equation

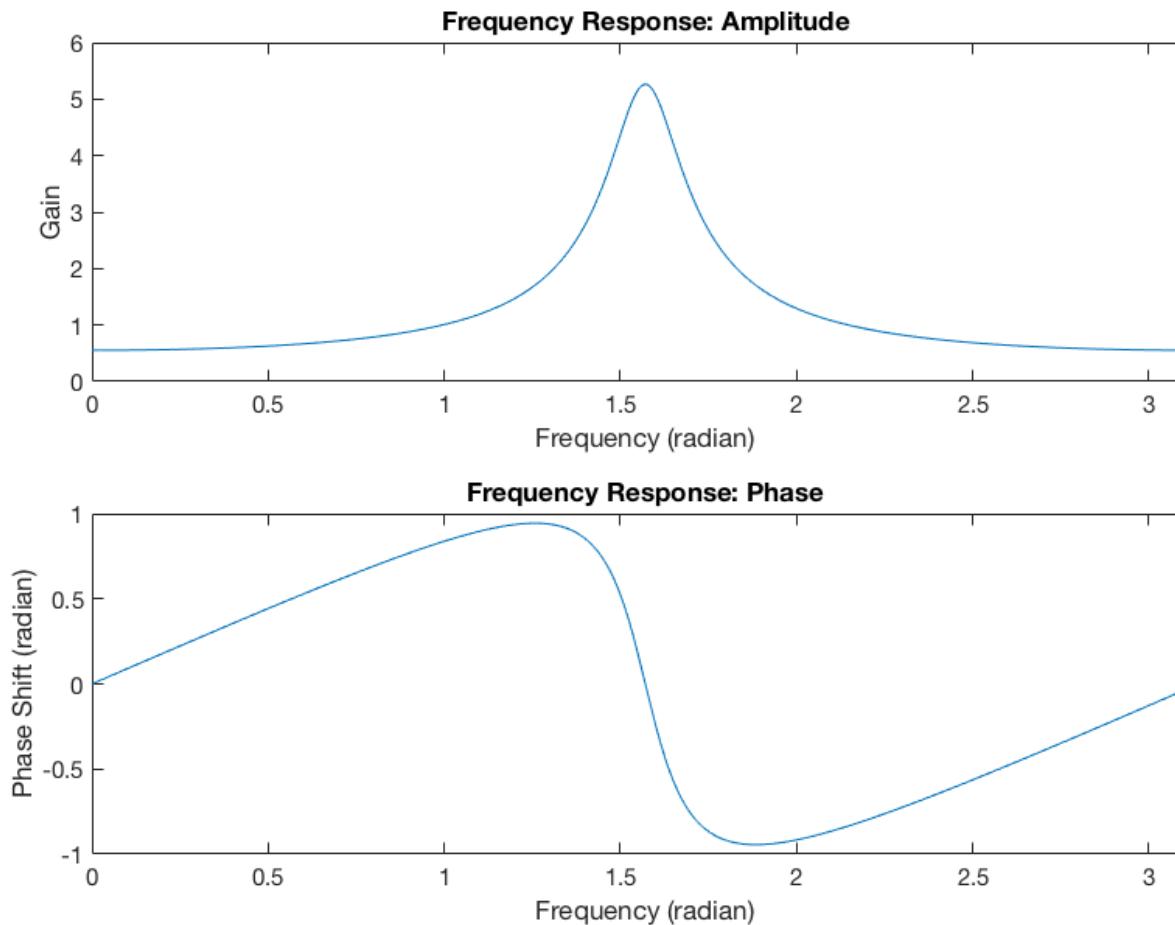
$$y(n) = x(n) + 2r \cdot \cos\theta \cdot y(n - 1) - r^2 \cdot y(n - 2)$$

- Signal flow graph



# Reson Filter: Frequency Response

- Generate resonance at a particular frequency
  - Control the peak height by  $r$  and the peak frequency by  $\theta$



For stability:  $r < 1$

# Filters

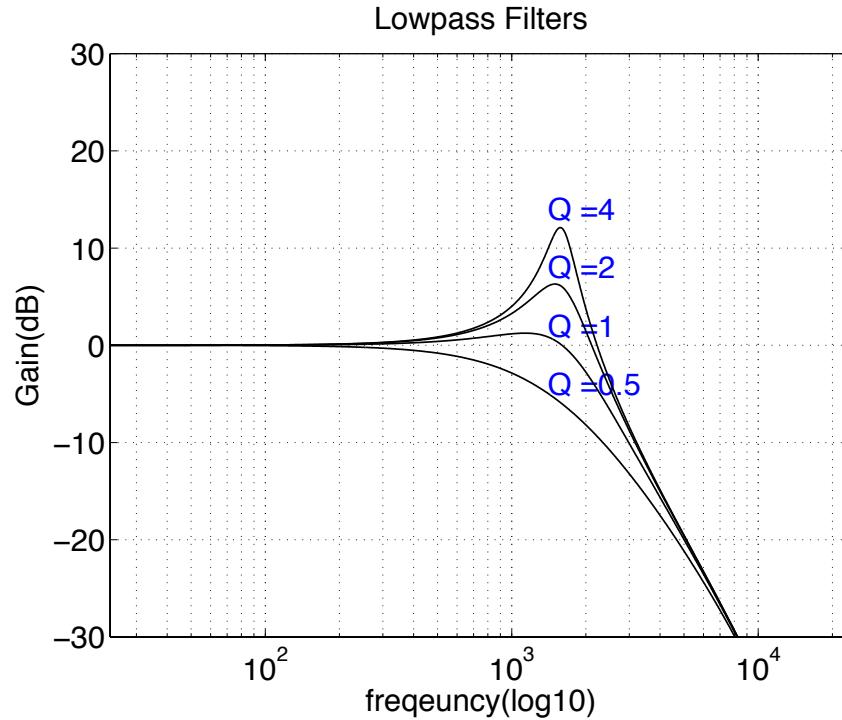
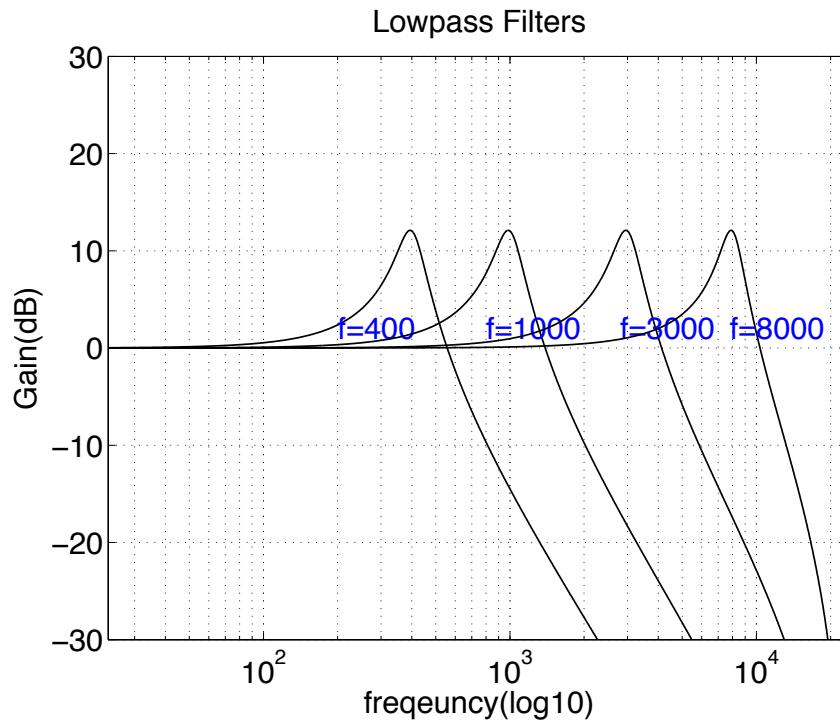
- Adjust the level of a certain frequency band
  - Lowpass
  - Highpass
  - Bandpass
  - Equalizer
- Parameters
  - Cut-off/Center Frequency
  - Q: sharpness/resonance

# Low-pass Filter

- Transfer Function

$$H(z) = \left(\frac{1-\cos\Theta}{2}\right) \frac{1+2z^{-1}+z^{-2}}{(1+\alpha)-2\cos\Theta z^{-1}+(1-\alpha)z^{-2}} \quad \alpha = \frac{\sin\Theta}{2Q} \quad \Theta = 2\pi f_c / f_s$$

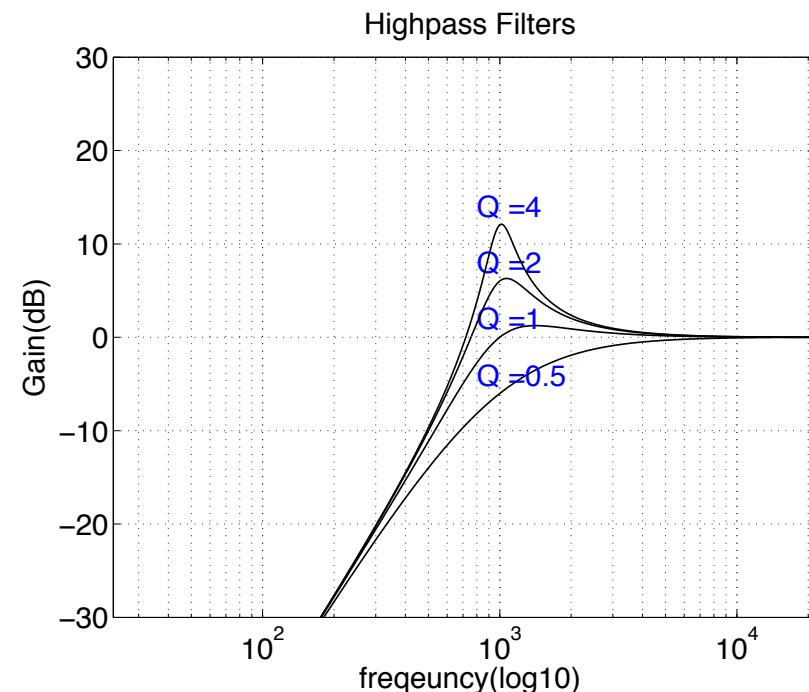
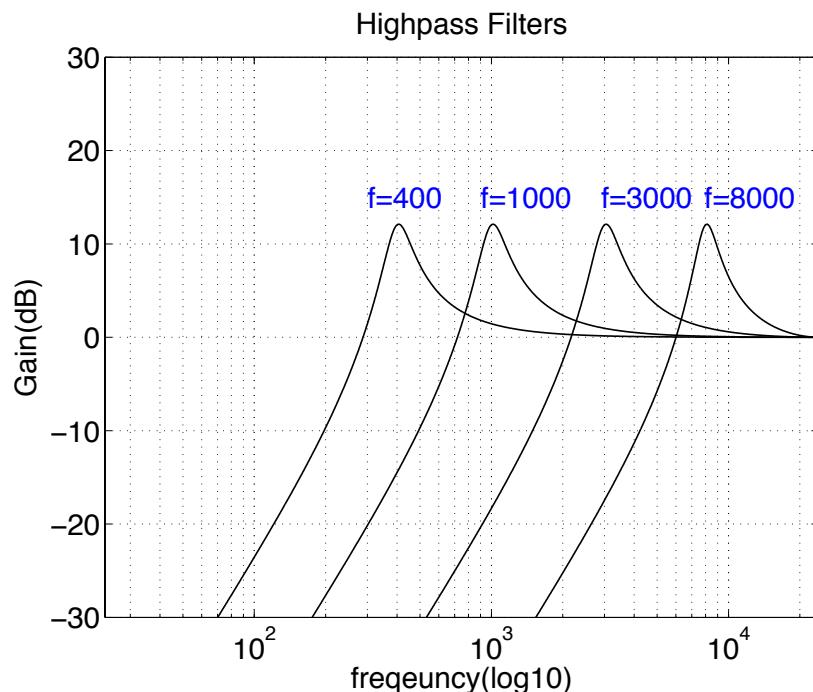
- $f_c$  : cut-off frequency,  $Q$ : resonance



# High-pass Filter

- Transfer Function

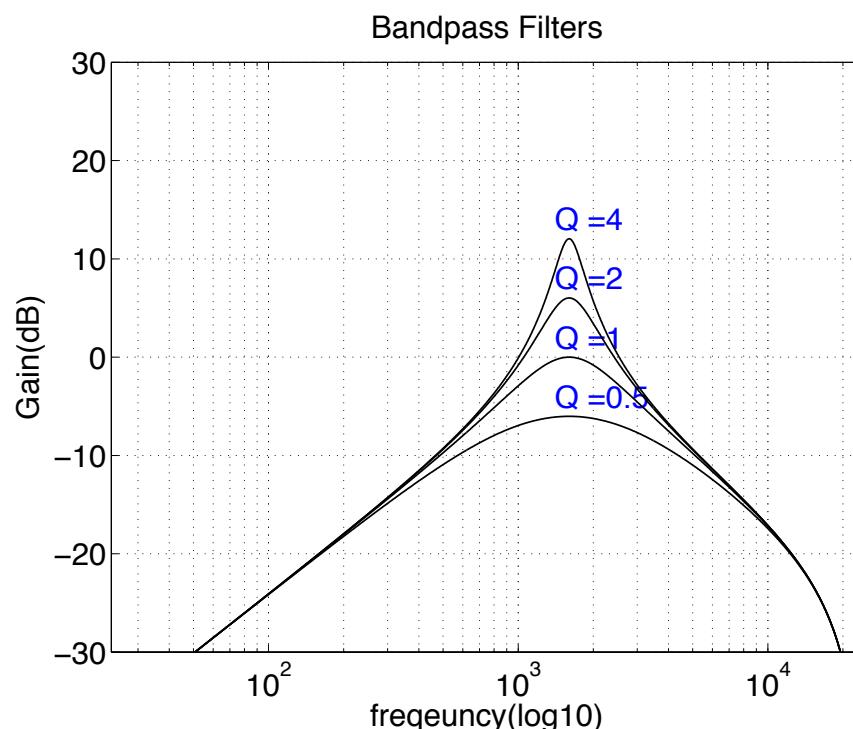
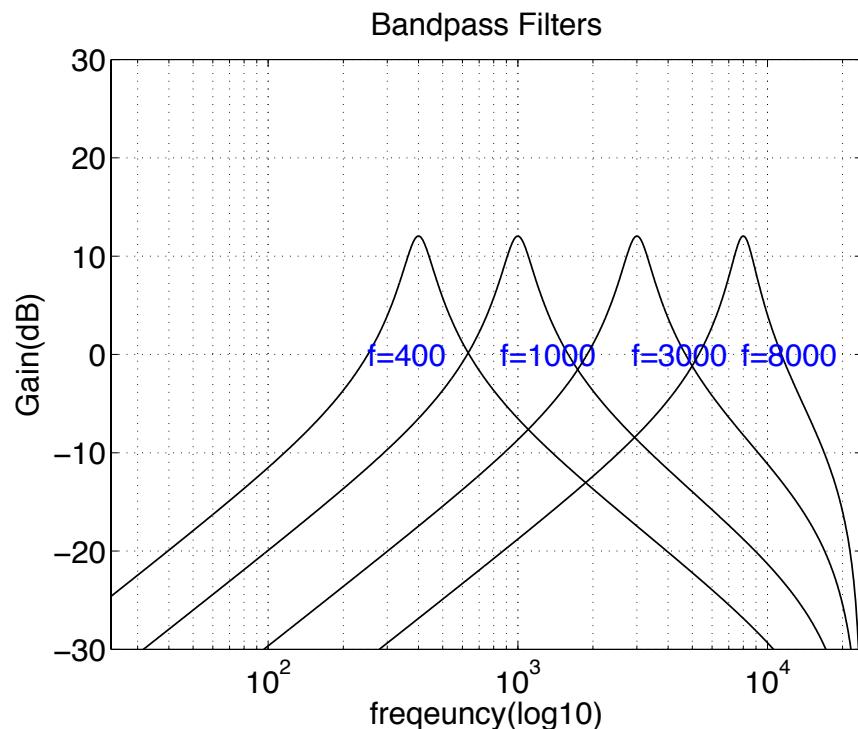
$$H(z) = \left(\frac{1+\cos\Theta}{2}\right) \frac{1-2z^{-1}+z^{-2}}{(1+\alpha)-2\cos\Theta z^{-1}+(1-\alpha)z^{-2}} \quad \alpha = \frac{\sin\Theta}{2Q} \quad \Theta = 2\pi f_c / f_s$$



# Band-pass filter

- Transfer Function

$$H(z) = \left(\frac{\sin \Theta}{2}\right) \frac{1 - z^{-2}}{(1 + \alpha) - 2 \cos \Theta z^{-1} + (1 - \alpha)z^{-2}}$$
$$\alpha = \frac{\sin \Theta}{2Q} \quad \Theta = 2\pi f_c / f_s$$

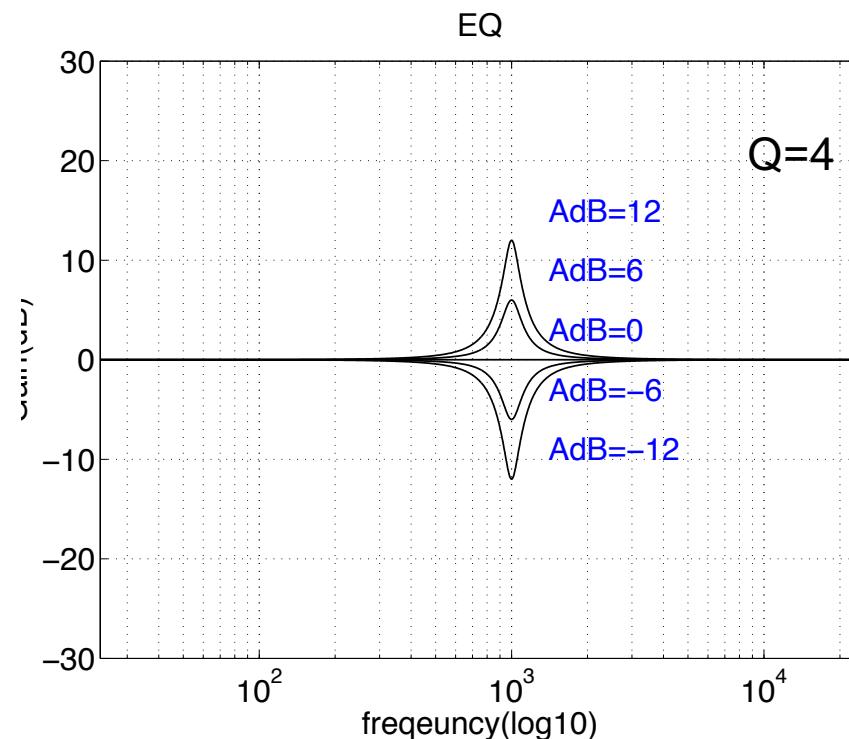
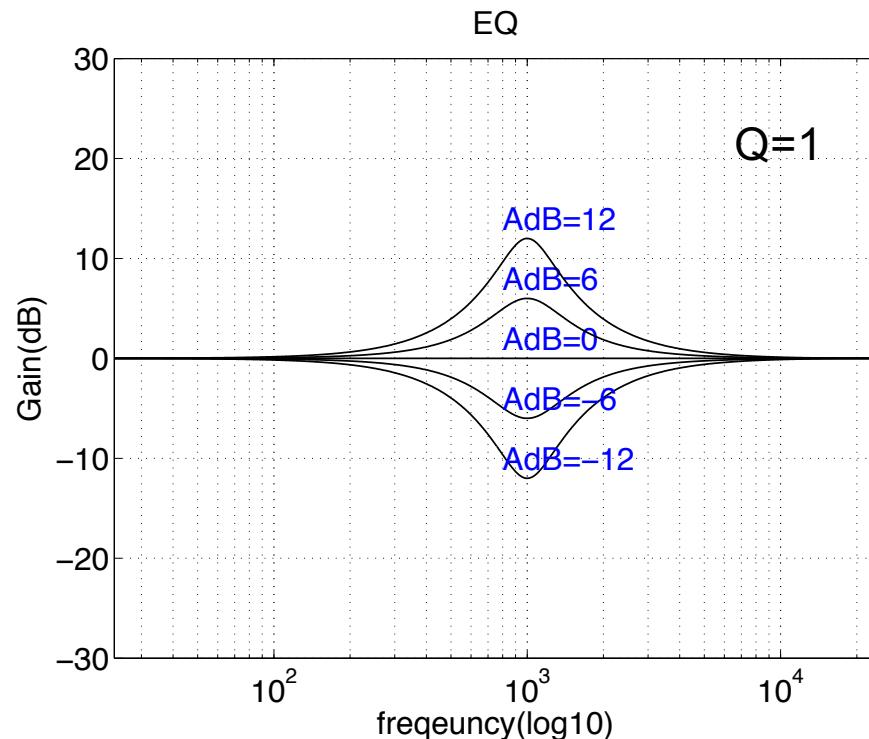


# Equalizer

- Transfer Function

$$H(z) = \frac{(1 + \alpha \cdot A) - 2 \cos \Theta z^{-1} + (1 + \alpha \cdot A)z^{-2}}{(1 + \alpha / A) - 2 \cos \Theta z^{-1} + (1 - \alpha / A)z^{-2}}$$

$$\alpha = \frac{\sin \Theta}{2Q} \quad \Theta = 2\pi f_c / f_s$$

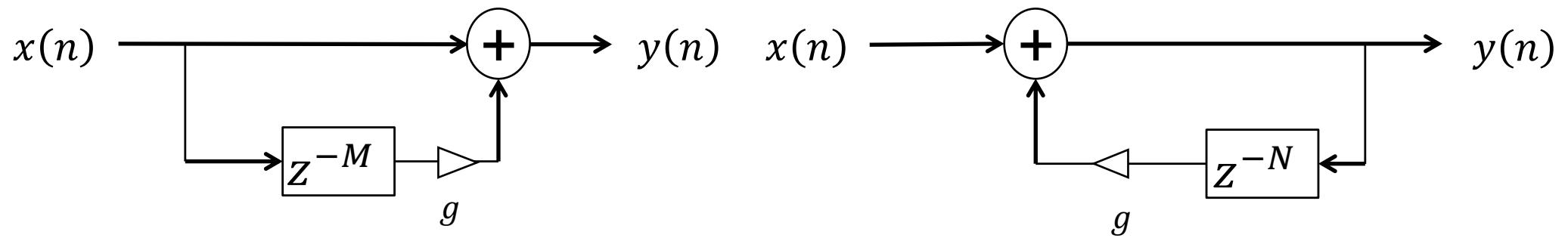


# Delay-based Audio Effects

- Types of delay-based audio effect
  - Delay
  - Chorus
  - Flanger
  - Reverberation



# Comb Filter



$$y(n) = x(n) + g \cdot x(n - M)$$

FIR Comb Filter

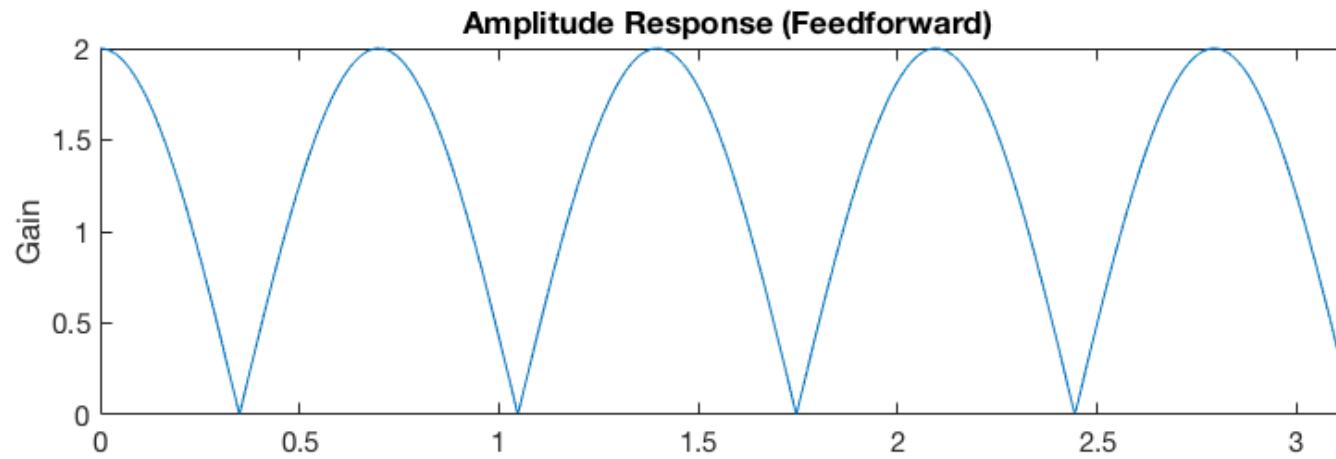
$$y(n) = x(n) + g \cdot y(n - N)$$

IIR Comb Filters

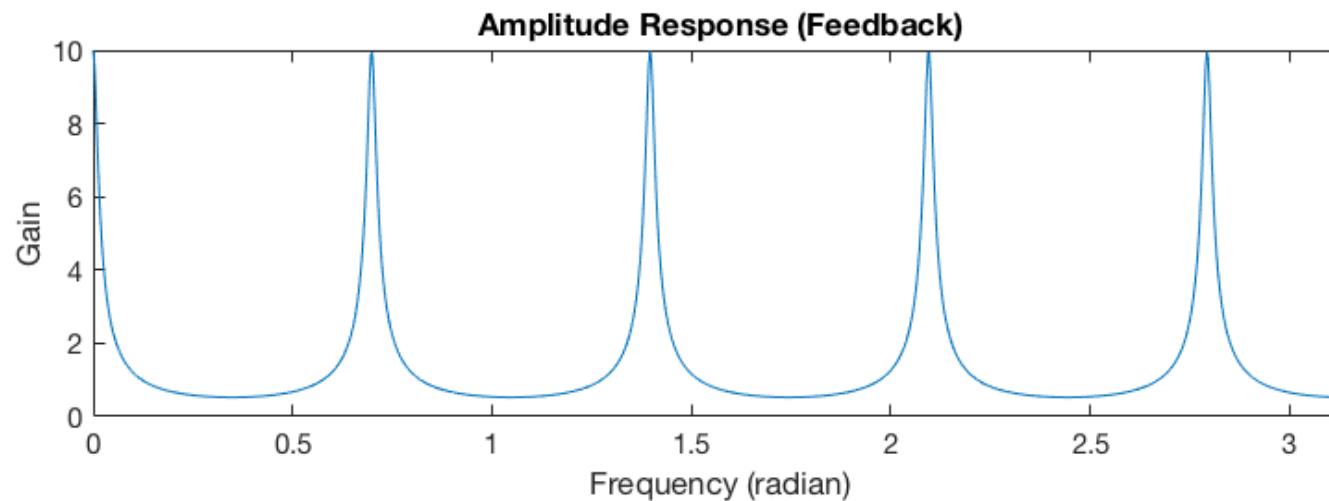
- Implemented by circular buffer: move read and write pointers instead of shift all samples in the delayline

# Comb Filter: Frequency Response

- "Combs" become shaper in the feedback type



$$y(n) = x(n) + x(n - 8)$$

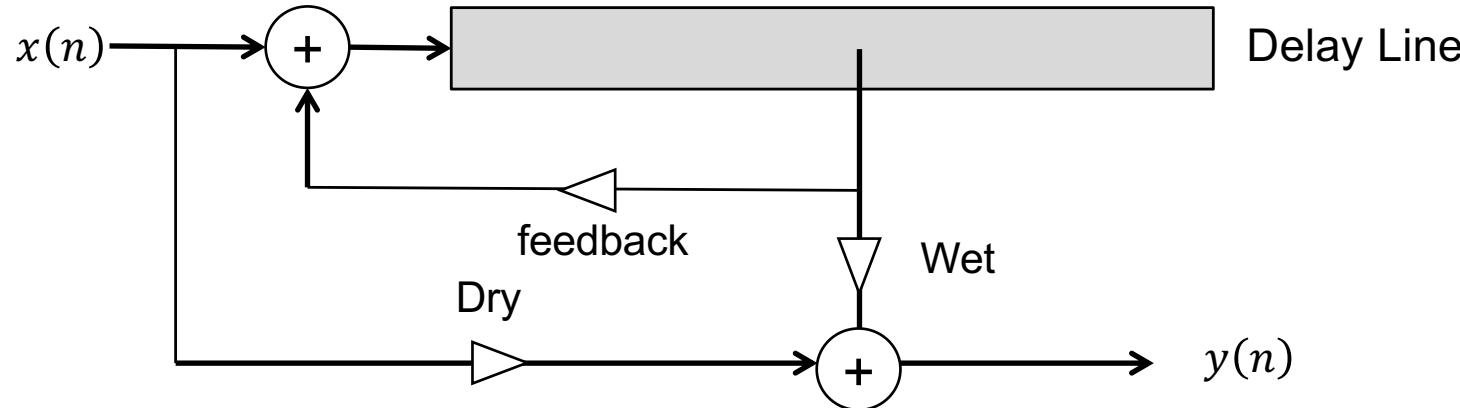


$$y(n) = x(n) + 0.9 \cdot y(n - 8)$$

# Perception of Time Delay

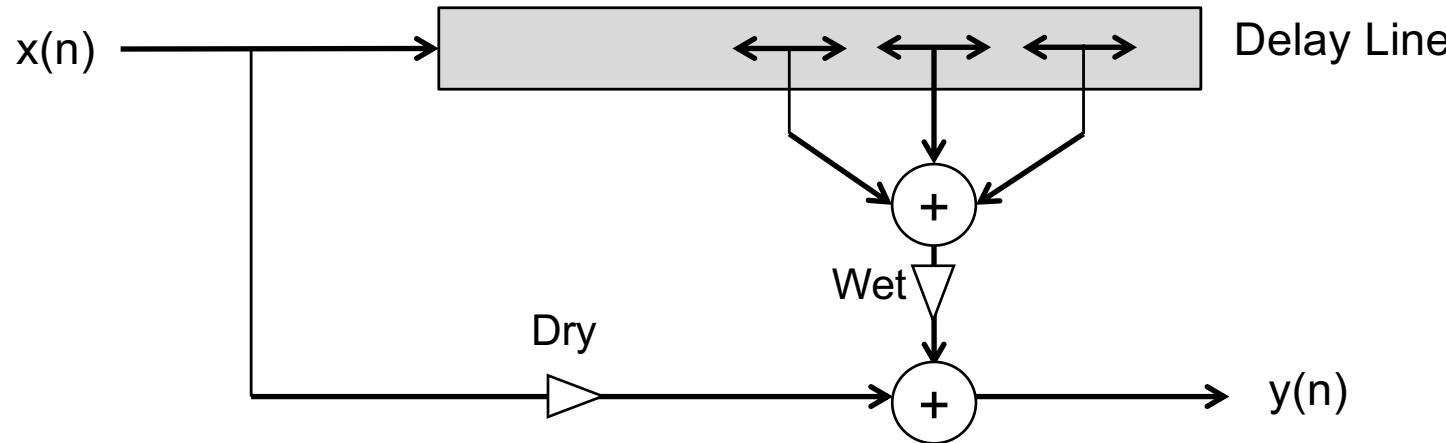
- The 30 Hz transition
  - Given a repeated click sound (e.g. impulse train):
    - If the rate is less than 30Hz, they are perceived as discrete events.
    - As the rate is above 30 Hz, they are perceive as a tone
  - Demo: [http://auditoryneuroscience.com/?q=pitch/click\\_train](http://auditoryneuroscience.com/?q=pitch/click_train)
- Feedback comb filter:  $y(n) = x(n) + a \cdot y(n - N)$ 
  - If  $N < \frac{F_s}{30}$  ( $F_s$ : sampling rate): models sound propagation and reflection with energy loss on a string (**Karplus-strong model**)
  - If  $N > \frac{F_s}{30}$  ( $F_s$ : sampling rate): generate a looped delay

# Delay Effect



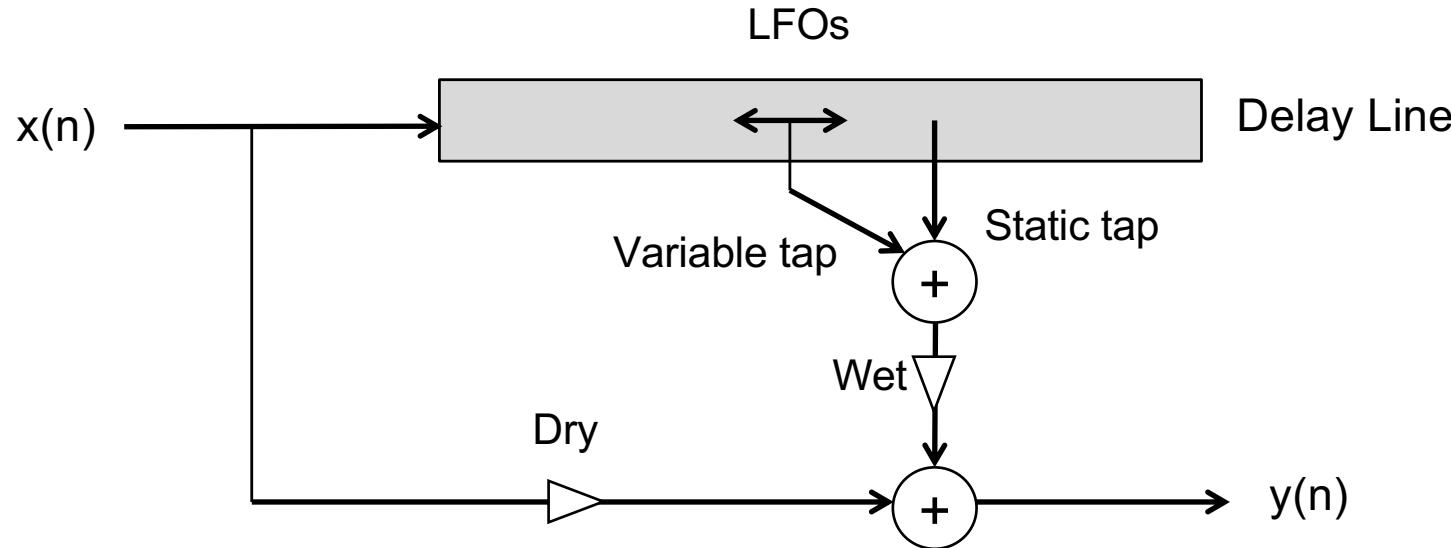
- Generate repetitive loop delay
  - Parameters
    - Feedback gain
    - Delay length
  - Ping-pong delay: cross feedback between left and right channels in stereo
  - The delay length is often synchronized with music tempo

# Chorus Effect



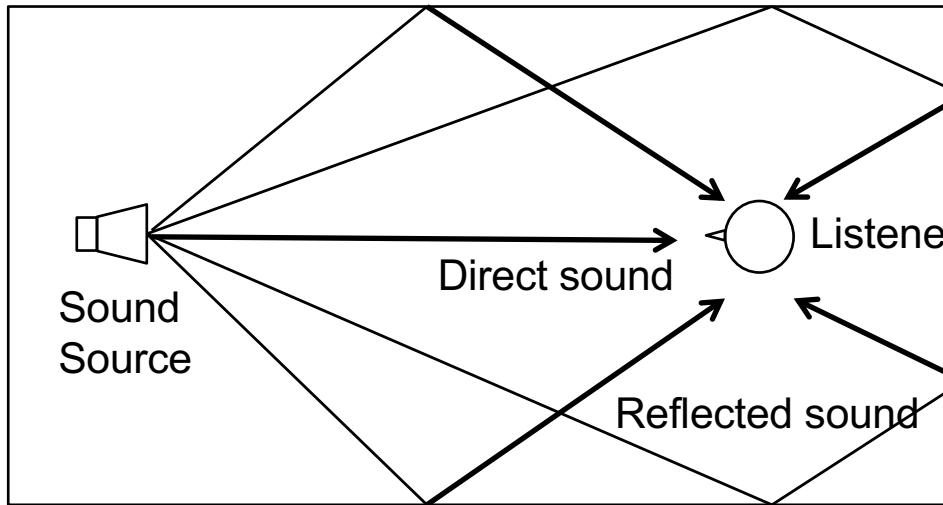
- Gives the illusion of multiple voices playing in unison
  - By summing detuned copies of the input
  - Low frequency oscillators (LFOs) are used to modulate the position of output tops
  - This causes pitch-shift

# Flanger Effect



- Emulated by summing one static tap and variable tap in the delay line
  - “Rocket sound”
  - Feed-forward comb filter where harmonic notches vary over frequency.
  - LFO is often synchronized with music tempo

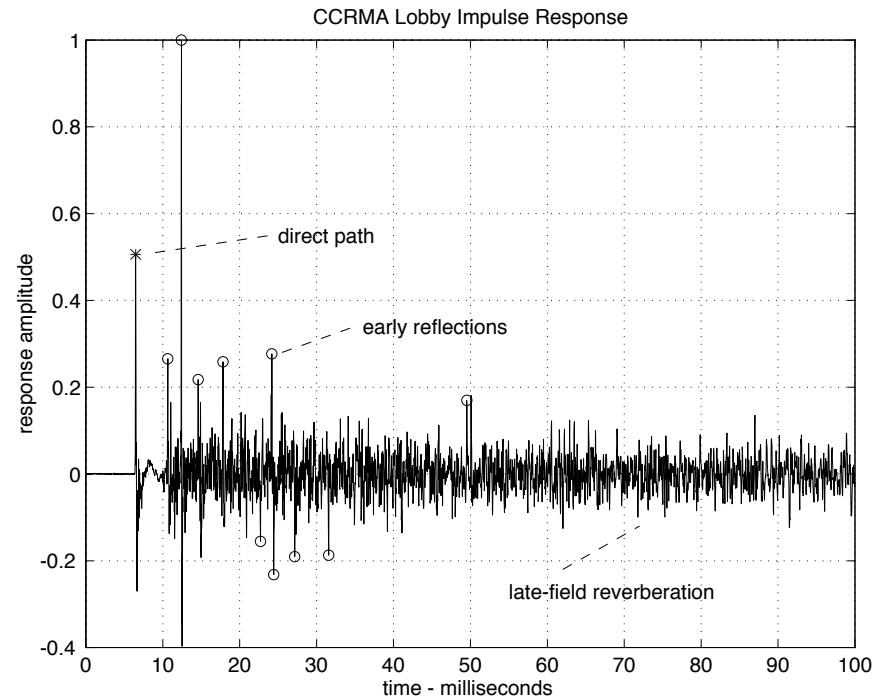
# Reverberation



- Natural acoustic phenomenon that occurs when sound sources are played in a room
  - Thousands of echoes are generated as sound sources are reflected against wall, ceiling and floors
  - Reflected sounds are delayed, attenuated and low-pass filtered: high-frequency component decay faster
  - The patterns of myriads of echoes are determined by the volume and geometry of room and materials on the surfaces

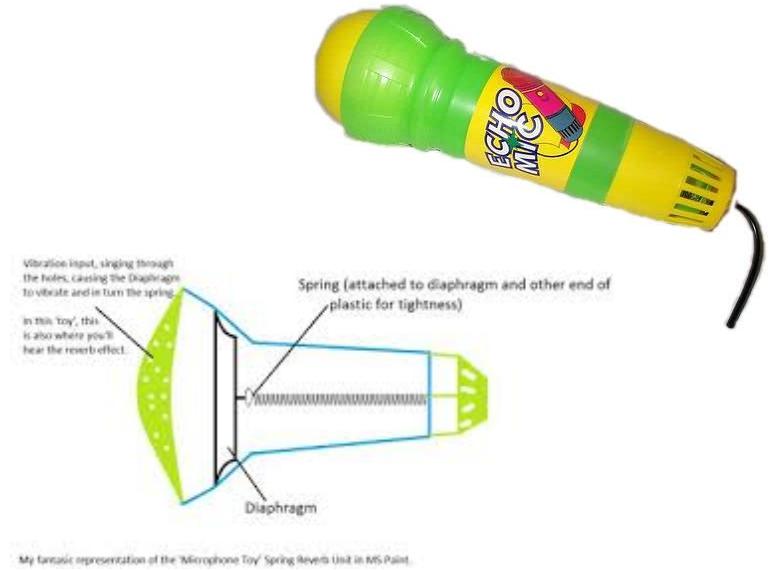
# Reverberation

- Room reverberation is characterized by its impulse response
  - e.g. when a balloon pop is used as a sound source
- The room IR is composed of three parts
  - Direct path
  - Early reflections
  - Late-field reverberation
- RT60
  - The time that it takes the reverberation to decay by 60 dB from its peak amplitude



# Artificial Reverberation

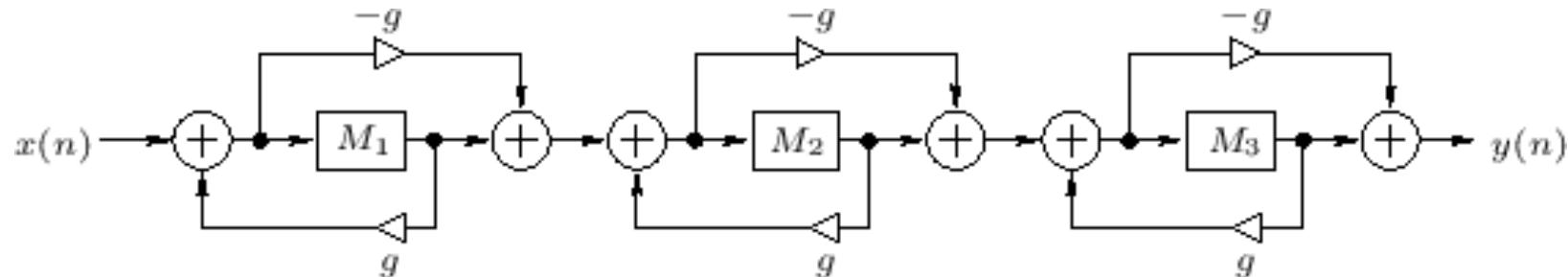
- Convolution reverb
  - Measure the impulse response of a room
  - Convolve input with the measured IR
- Mechanical reverb
  - Use metal plate and spring
  - EMT140 Plate Reverb: <https://www.youtube.com/watch?v=HEmJpxCvp9M>
- Delayline-based reverb
  - Early reflections: feed-forward delayline
  - Late-field reverb: allpass/comb filter, feedback delay networks (FDN)
  - “Programmable” reverberation



My basic representation of the 'Microphone Toy' Spring Reverb Unit in MS Paint.

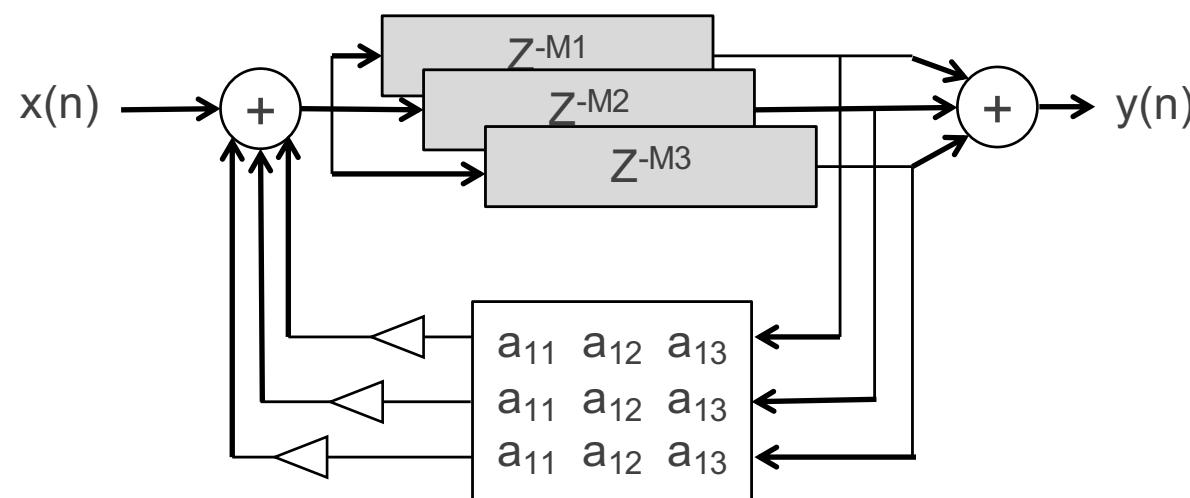
# Delayline-based Reverb

- Schroeder model



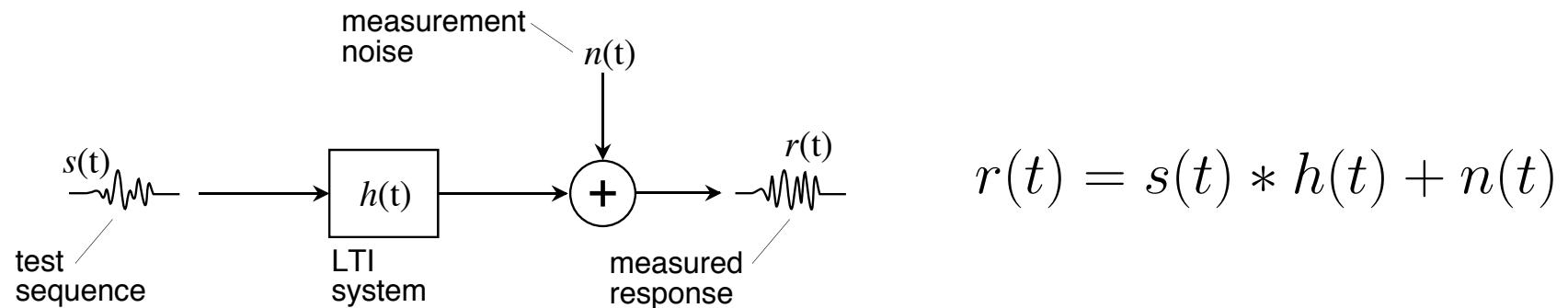
- Cascade of allpass-comb filters
- Mutually prime number for delay lengths

- Feedback Delay Networks



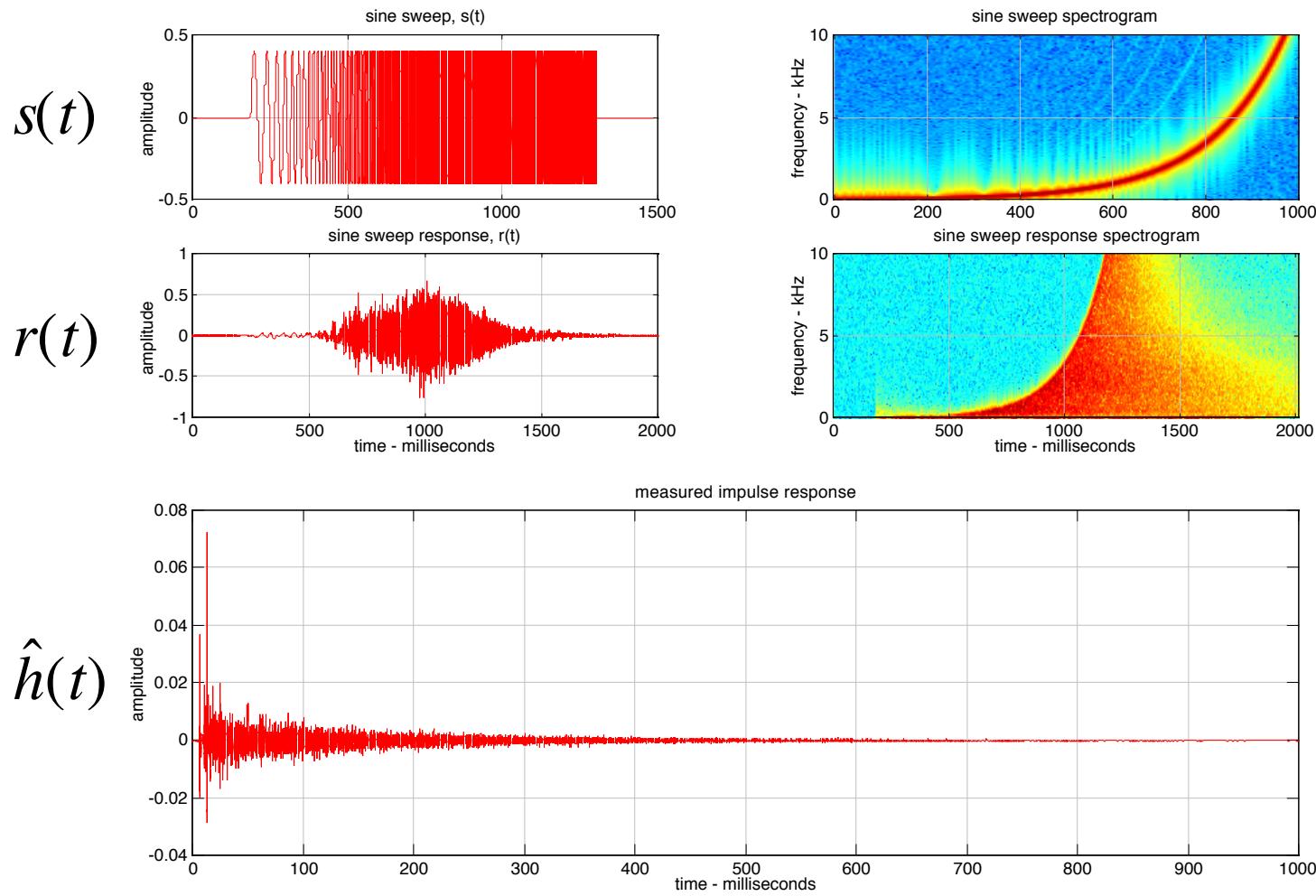
# Convolution Reverb

- Measuring impulse responses



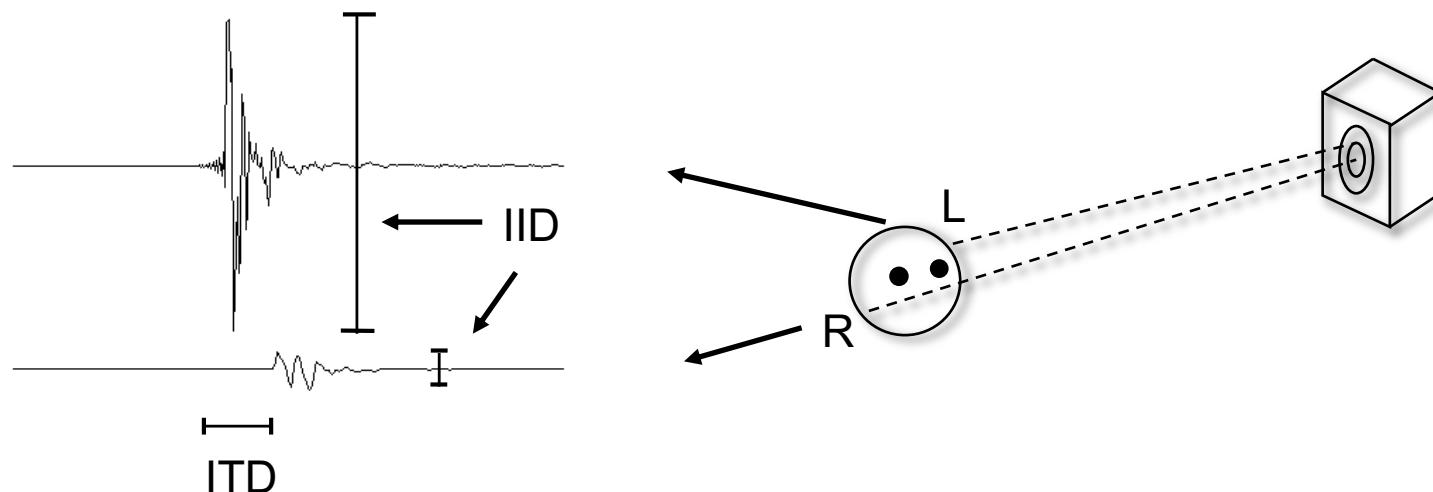
- If the input is a unit impulse, SNR is low
- Instead, we use specially designed input signals
  - Golay code, allpass chirp or sine sweep: their magnitude responses are all flat but the signals are spread over time
- The impulse response is obtained using its inverse signal or inverse discrete Fourier transform

# Convolution Reverb



# Spatial Hearing

- A sound source arrives in the ears of a listener with differences in time and level
  - The differences are the main cues to identify where the source is.
  - We call them **ITD** (Inter-aural Time Difference) and **IID** (Inter-aural Intensity Difference)
  - ITD and IID are a function of the arrival angle.



# Head-Related Transfer Function (HRTF)

- A filter measured as the frequency response that characterizes how a sound source arrives in the outer end of ear canal
  - Determined by the reflection on head, pinnae or other body parts
  - Function of azimuth (horizontal angle) and elevation (vertical angle)

