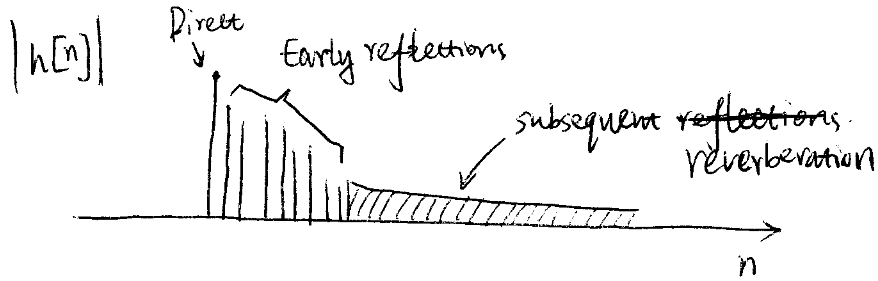
ECE 272/472 Lecture 9. Room Simulation and Reverberation

Room acoustics colors sound

View a room as an LTI filter. It has ~~an~~ impulse response



T_{60}: the time that takes $|h[n]|$ to decrease by 60 dB SPL.
(reverberation time)

$$T_{60} = 0.163 \frac{V}{\alpha S}$$

V ← Volume of room

S ← surface area.

α ← absorption coefficient

$$= 0.163 \frac{V}{\sum_n \alpha_n S_n} \quad \text{(for different materials)}$$

Eigenfrequencies (mode freqs) of a room [rectangular]

$$f_c = \frac{c}{2} \sqrt{\left(\frac{n_x}{l_x}\right)^2 + \left(\frac{n_y}{l_y}\right)^2 + \left(\frac{n_z}{l_z}\right)^2}$$

c ← sound speed

This is solved from 3D wave equation

l_x, l_y, l_z: dimensions of room.

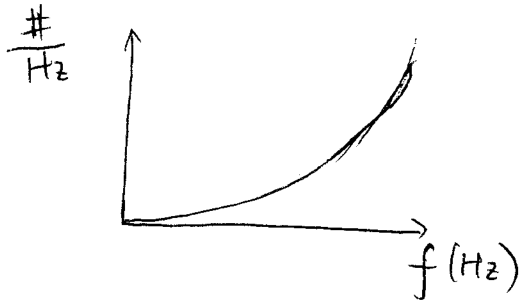
n_x, n_y, n_z: integers of half waves.

Basic idea:

Sound pressure always achieves largest variation at walls.

Question: How are the eigenfrequencies distributed?

Answer: eigenfrequency density $= \dfrac{\#\text{ eigenfreqs.}}{Hz} = \left(\dfrac{4\pi V}{c^3}\right) f^2$



Let's derive it.

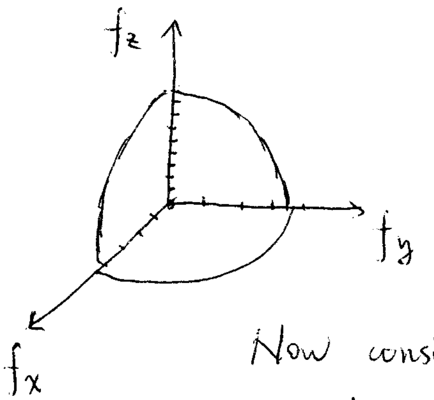
- When n_x, n_y, n_z are not too large, each combination of (n_x, n_y, n_z) gives us a unique eigenfreq. f_e.
- when n_x, n_y, n_z changes, we are considering different standing waves with different frequency.

$$f_e = \sqrt{\left(\frac{n_x c}{2l_x}\right)^2 + \left(\frac{n_y c}{2l_y}\right)^2 + \left(\frac{n_z c}{2l_z}\right)^2}$$

Let $f_x = \dfrac{n_x c}{2l_x}$, $f_y = \dfrac{n_y c}{2l_y}$, $f_z = \dfrac{n_z c}{2l_z}$.

then $f_e = \sqrt{f_x^2 + f_y^2 + f_z^2}$

Frequency space (3D)



each grid corresponds to a ~~org~~ combination of (n_x, n_y, n_z), hence corresponds to an f_e.

Consider the "frequency volume" of each grid.

$$\frac{c}{2l_x} \cdot \frac{c}{2l_y} \cdot \frac{c}{2l_z} = \frac{c^3}{8V}$$

Now consider how many grids are there ~~within~~ whose f_e is smaller than f, i.e., # of grids within $\frac{1}{8}$ sphere with radius of f

Volume of the $\frac{1}{8}$ sphere $\frac{1}{8}\left(\frac{4}{3}\pi f^3\right)$

∴ # grids within this $\frac{1}{8}$ sphere:

$$\frac{1}{8}\left(\frac{4}{3}\pi f^3\right) \Big/ \frac{c^3}{8V} = \frac{4\pi V f^3}{3c^3}$$

$$\therefore \frac{\#\text{ eigenfreqs}}{\text{Hz}} = \frac{d}{df}\cdot\frac{4\pi V f^3}{3c^3} = \frac{4\pi V}{c^3} f^2 .$$

So as $f \uparrow$, the density of f_e goes up too.

But there is a critical frequency

$f_c = 2000\sqrt{\frac{T_{60}}{V}}$, when $f > f_c$, eigenfreqs will overlap.

Early reflections: important for spatial perception

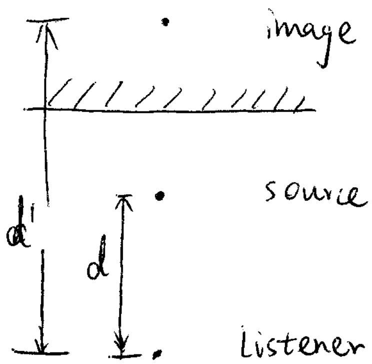
Craven hypothesis: human perception of room size is related to the strength and delay of early reflections from the direct sound.

gain: $g = \frac{d}{d'} < 1$

Delay: $T_D = \frac{d'-d}{c}$

$$\therefore T_D = \frac{d/g - d}{c} = \frac{d}{c}(g^{-1}-1)$$

$$\therefore d = \frac{cT_D}{g^{-1}-1}$$



A more realistic treatment (considering air absorption)

$g = \frac{d}{d'}\exp(-rT_D)$, r: absorption coefficient.

$T_D = \frac{d'-d}{c}$

$\Rightarrow d = \frac{cT_D}{g^{-1}e^{-rT_D}-1}$ (gain and time delay are used to infer distance

$\Rightarrow g = \frac{e^{-rT_D}}{1+cT_D/d}$. ~~f design~~

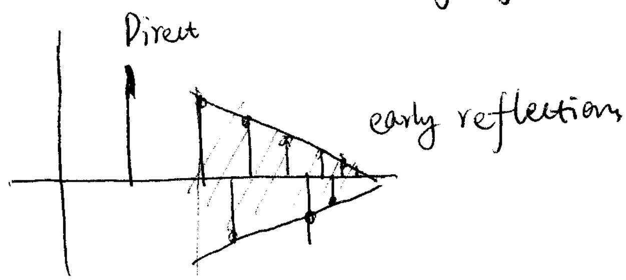
There are multiple early reflections.

How to simulate them?

Gerzon's Distance Algorithm.

Basic idea: ① Generate delays of the reflections.

② Generate a gain for each reflection using perious formular.
(weighting)



Notes:

1) Time density of early reflections should increase over time.

$$\frac{\#\text{ reflections}}{\text{second}} = \frac{4\pi c^3}{V} \cdot t^2$$

2) Stereo implementation should consider ~~a weighting factor~~ an angle (direction) factor for each reflection, to simulate reflections from different walls.

Subsequent Reverberation.

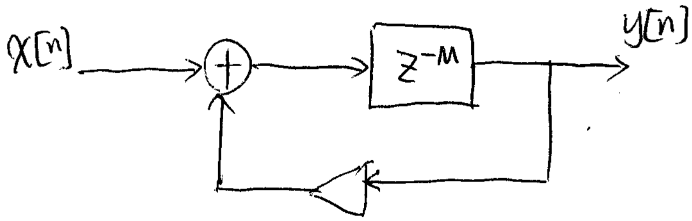
— For early reflections, we model each single reflection.

— For subsequent ~~reberation~~ reverberation, we can't do this in practice, as there are so many reflections and things get more complex. Instead, we will model the overall shape and frequency response of the impulse response.

Schroeder Algorithm: the first software implementation of reverb simulation, 1961.

Basic idea: Use parallels of comb filters and series of all-pass filters.

A comb filter: add feedback to a delay.



g: real, $0<g<1$.

$$y[n] = x[n-M] + g\,\cancel{y[n]}\,y[n-M]$$

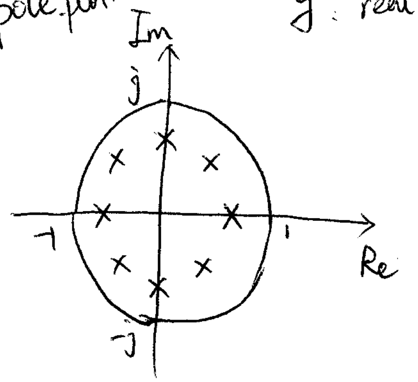
$$Y(z) = X(z)\cdot z^{-M} + g\,Y(z)\,z^{-M}$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{z^{-M}}{1-g z^{-M}} = \frac{1}{z^{M}-g}$$

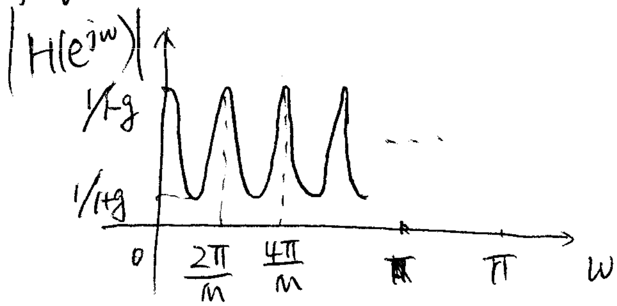
M poles: $g^{\frac{1}{M}}\cdot e^{j2\pi k/M}$, $k=0,1,\cdots,M-1$.

zero-pole plot.

M=8.



frequency response.



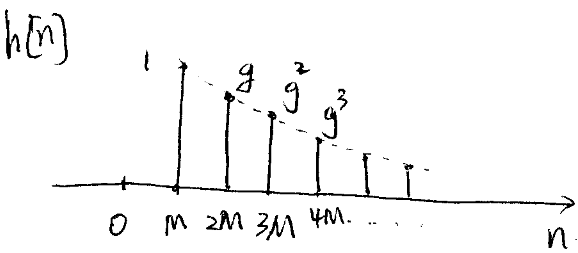
$$H(e^{jw}) = \begin{cases} \frac{1}{1-g} & \text{if } w = 0, \frac{2\pi k}{M},\ k=0,1,\dots,M \\ \frac{1}{1+g} & \text{if } w = \frac{\pi(2k+1)}{M} \end{cases}$$

eigenfreqs: $\frac{2\pi k}{M}$.

Frequency density: $\frac{M}{2}$ eigenfreqs between 0 and $\frac{f_s}{2}$, $\therefore D_f = M/f_s$. $\left(\frac{1}{Hz}\right)$

Impulse response

exponential decay.



Echo density: 1 impulse every M samples, i.e., M/f_s seconds.

$$\therefore D_t = \frac{1}{M/f_s} = f_s/M \quad \left(\frac{1}{s}\right)$$

$$\therefore D_f \cdot D_t = 1.$$

Clearly we can't make frequency density and echo density both high.

Parallels of comb filters with different delays.

$$H(z) = \sum_{p=1}^{P} \frac{z^{-M_p}}{1-g_p z^{-M_p}}$$

Let $M_1 : M_p = 1 : 1.5$

$M_1 < M_2 < \cdots < M_P$.

Eigenfreqs will be the union of ~~all~~ eigenfreqs of all the comb filters.

Frequency density: $D_f = \sum_{p=1}^{P} M_p/f_s = P\bar{M}/f_s$

where $\bar{M} = \frac{1}{P}\sum_{p=1}^{P} M_p$ mean delay length.

Echo density: $D_t = \sum_{p=1}^{P} \frac{f_s}{M_p} \approx \frac{P f_s}{\bar{M}}$

$\therefore\ P = \sqrt{D_f \cdot D_t}$.

$\bar{M}/f_s = \sqrt{D_f/D_t}$

e.g. $D_f = 0.15$, $D_t = 1000$.

then $P = 12$

$\bar{M}/f_s = 12\,ms$.

In practice, this is not enough. 10000 would be good.

Reverberation time of a single comb filter:

linear decay in terms of dB.

for every M samples, amplitue decays for $-20\log_{10} g$ dB

$$\therefore\ \frac{-20\log_{10} g}{M/f_s} = \frac{60}{T_{60}}.$$

$$\therefore\ T_{60} = \frac{3}{\log_{10}\left(\frac{1}{g}\right)} M_p/f_s.$$

To increase T_{60}, either ① increase M $\Rightarrow$ lower echo density.

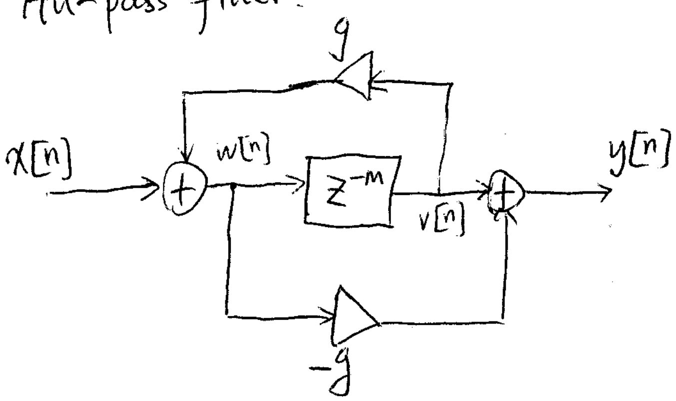
② increase g $\Rightarrow$ increase variation of frequency response

$\frac{1}{1-g} \uparrow \quad \frac{1}{1+g} \downarrow$.

"coloring" sound.

Question: How to increase echo density and T_{60} while keeping frequency response not varying too much?

All-pass filter:



$$y[n] = v[n] \cancel{*} - g\,w[n]$$
$$v[n] = w[n-M]$$
$$w[n] = x[n] + g\,v[n]$$

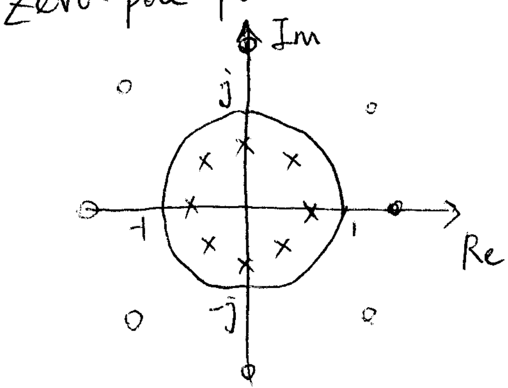
$$Y(z) = V(z) - g\,W(z)$$
$$V(z) = z^{-M} W(z)$$
$$W(z) = X(z) + g\,V(z)$$

$$\Rightarrow W(z) = X(z) + g z^{-M} W(z)$$
$$\Rightarrow W(z) = \frac{X(z)}{1-gz^{-M}}$$
$$\Rightarrow V(z) = \frac{z^{-M}}{1-gz^{-M}} \cancel{W}X(z)$$
$$\Rightarrow Y(z) = \frac{z^{-M}}{1-gz^{-M}} X(z) - \frac{g}{1-gz^{-M}} X(z)$$
$$\Rightarrow H(z) = \frac{Y(z)}{X(z)} = \frac{z^{-M} - g}{1-gz^{-M}} = \frac{g\left(\frac{1}{g} - z^{M}\right)}{z^{M} - g}$$

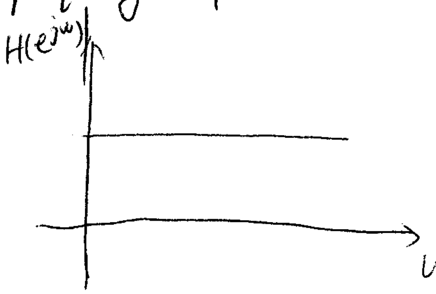
M zeros: $\left(\frac{1}{g}\right)^{\frac{1}{M}} e^{j2\pi k/M}$ $\quad k = 0, \dots, M-1$

M poles: $g^{\frac{1}{M}} e^{j2\pi k/M}$

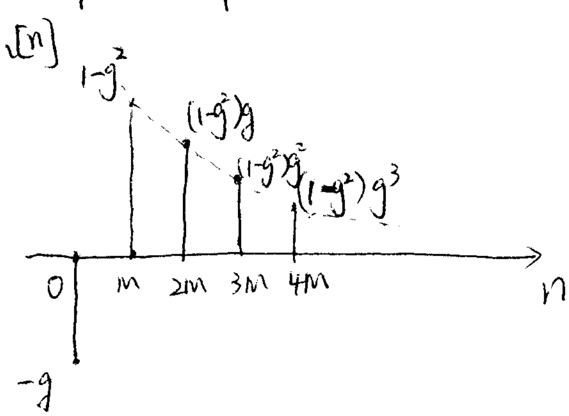
Zero-pole plot



frequency response



Impulse response



$$H(e^{jw}) = \frac{e^{-jwM} - g}{1 - g e^{-jwM}} = \frac{e^{-jwM}(1 - g e^{jwM})}{1 - g e^{-jwM}}$$

conjugate complex numbers.

$\therefore |H(e^{jw})| \equiv 1$ for all w.

Echo density: $D_t = M/f_s$.

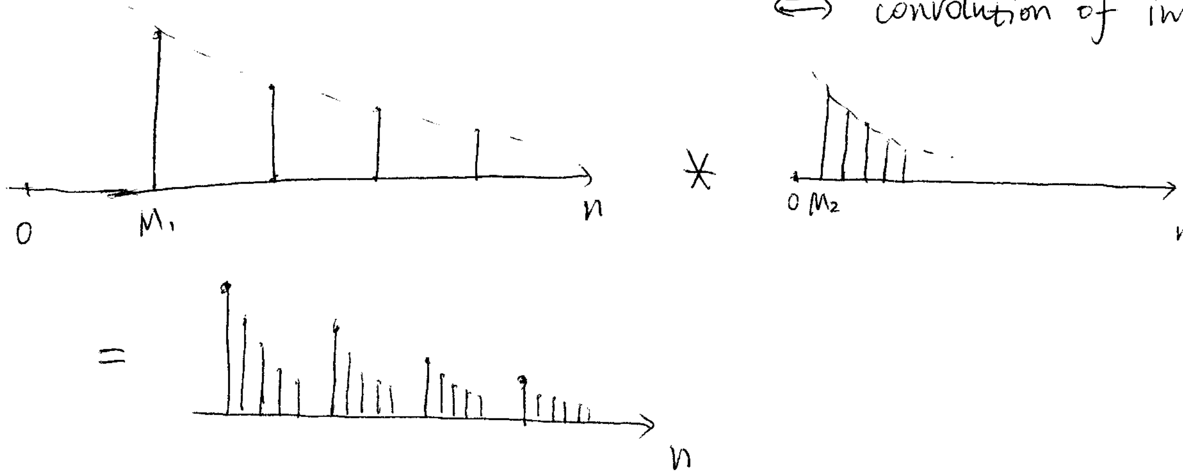
But varying M or g doesn't change frequency response!

How to increase echo density?

Series of all-pass filters with very different delays.

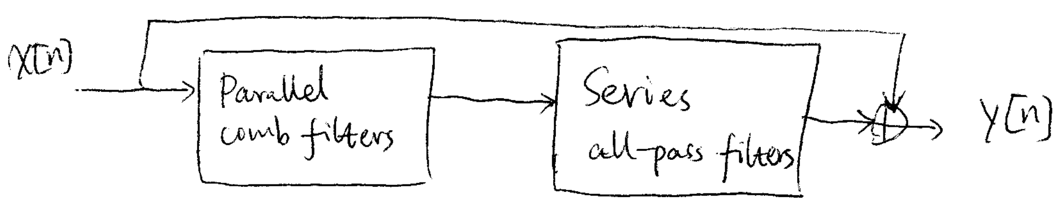
series combination of ~~flit~~ filters $\Leftrightarrow$ multiplication of freq. response

$\Leftrightarrow$ convolution of impulse response.



Schroeder's original paper suggested 5 all-pass filters with delay of 100, 68, 60, 19.7, 5.85 ms.

Combine comb filters and all-pass filters.



Schroeder' 1962. Figure 6.