Master MVA

Analyse des signaux Audiofréquences

Audio Signal Analysis, Indexing and Transformation

Lecture on
Sound effects and Reverberation

Gaël RICHARD
Image, Data, Signal department
January 2023

Content

- **Sound effects**
  - Compression/Expansion – Limitation
  - Flanging – Phasing – Wah/wah
  - Time-frequency modifications
    - Definition
    - Circular memory technic
    - PSOLA
  - Distortion

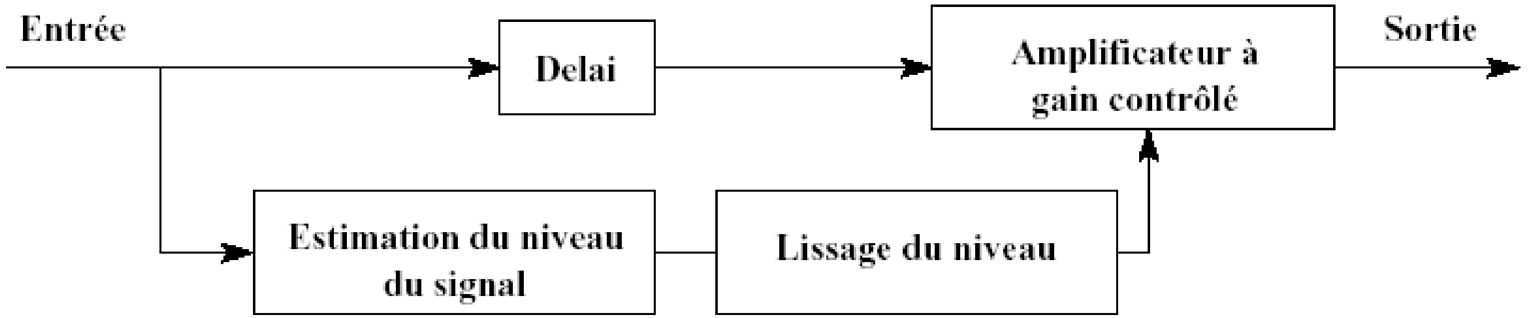
- **Artificial reverberation**
  - Introduction
  - Algorithms (perceptual, physics based)

Compression/Expansion – Limitation

- **Largely used in analog systems to control the dynamic range of signals**
  - **Dynamic range of signals:** ratio between the power of the louder signal to the power of the weakest signal
- **Order of magnitude:**
  - Dynamic of a room/hall: > 100 dB
  - Dynamic of a FM radio broadcast : < 70 dB
- **Principle:**
  - Compression before transmission
  - Expansion after transmission to restore the initial dynamic

Compression/Expansion – Limitation

- **Principle:**
  - Estimate the level of the input signal
  - Adjust the parameters of the dynamic control
  - Signal level modification according to the dynamic control gain (in logarithm scale).



Estimation of the input signal level

- **Root-mean square (RMS) value:**

$$P_{eff} = \sqrt{\frac{1}{T}\int_{t-T}^{t} x^2(u)du} \qquad P_{eff} = \sqrt{\frac{1}{N}\sum_{i=n-N+1}^{n} x_i^2}$$

- **Peak to Peak level :**

$$P_{crete} = \max_{n-M<i<n} |x_i|$$

- **For a sinusoidal signal:**

$P_{eff} = A/\sqrt{2}$

$P_{crete} = A$

Estimation of the input signal level

- **Crest factor: ratio of the peak to peak level and the root mean square value**
  - Always greater than 1
  - Measures the presence of peaks (e.g. impulses) in the signal
  - For a sinusoid: Crest factor= P_{crete} / P_{eff} = $\sqrt{2}$

- **Estimation of input signal level**
  - In RMS for **compression/expansion**,
  - Peak-to-peak value for **limitation**.

Real-time update of dynamic control parameters

- **Input level smoothing to avoid too abrupt variations of the amplification gain.**

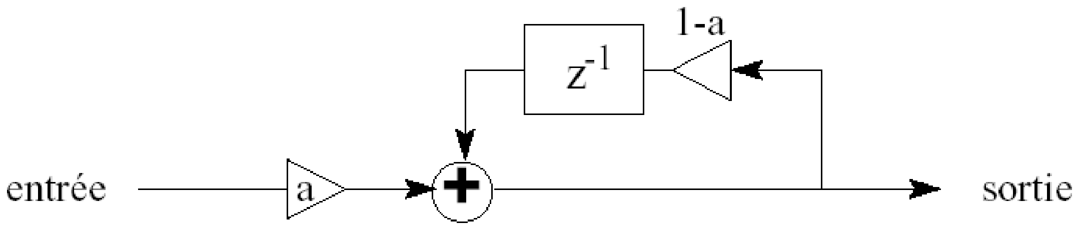
- **Dynamic control: 2 parameters:**

  - **Rising time (or Attack time)**: How soon the compressor starts to compress the dynamics after the threshold is exceeded.
    - *If volume changes are slow, a high value is possible. Short attack times will result in a fast response to sudden, loud sounds, but will make the changes in volume much more obvious to listeners.*

  - **Release time:** How soon the compressor starts to release the volume level back to normal after the level drops below the threshold.
    - *A long time value will tend to loose quiet sounds that come after loud ones, but will avoid the volume being raised too much during short quiet sections like pauses in speech.*

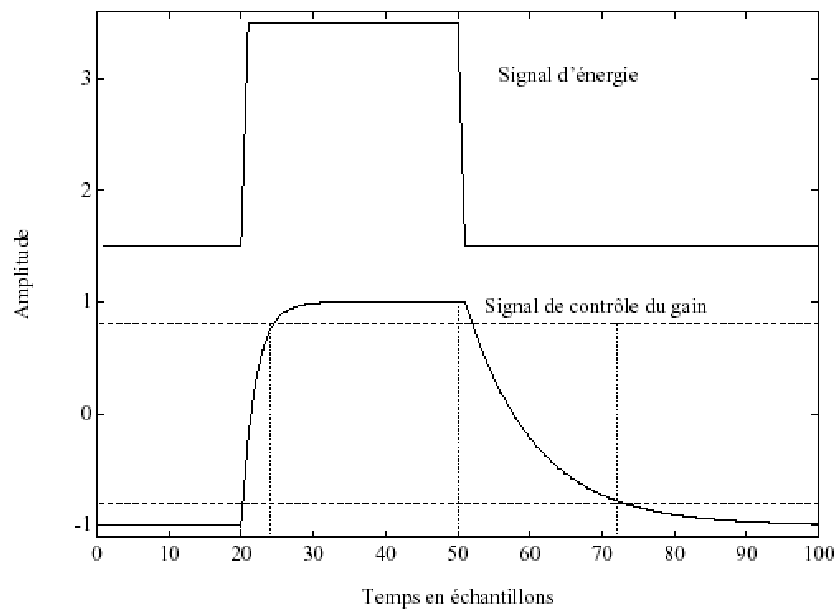
Real-time update of dynamic control parameters

- **Low-pass implementation**



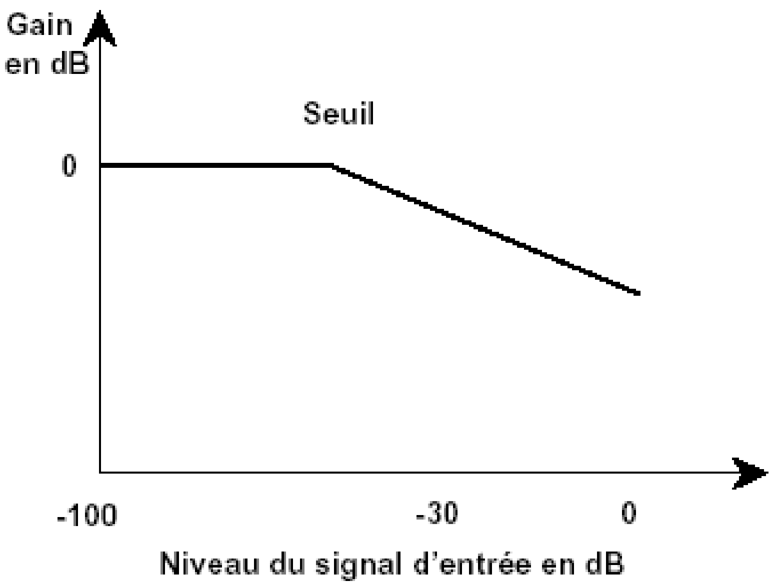
$$H(z) = \frac{a}{1-(1-a)z^{-1}}$$

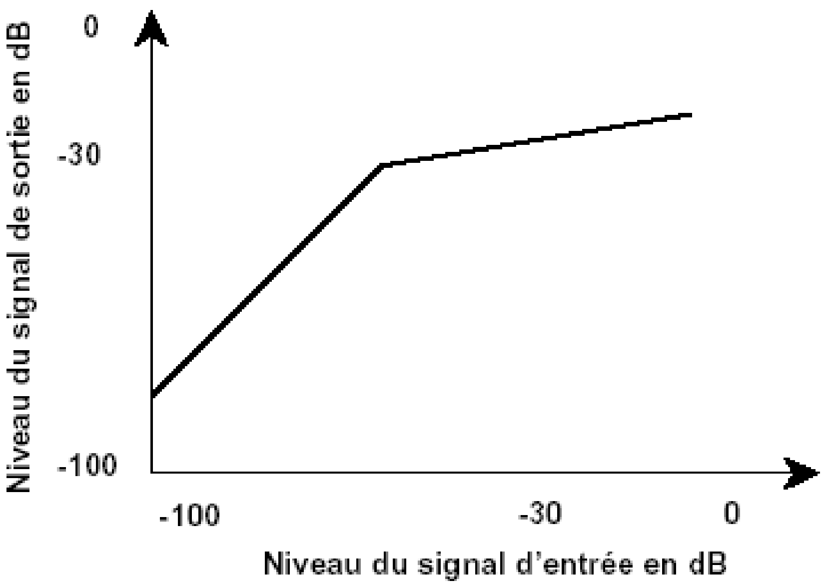
- **Examples:**



Dynamic compression (*or Compressor*)

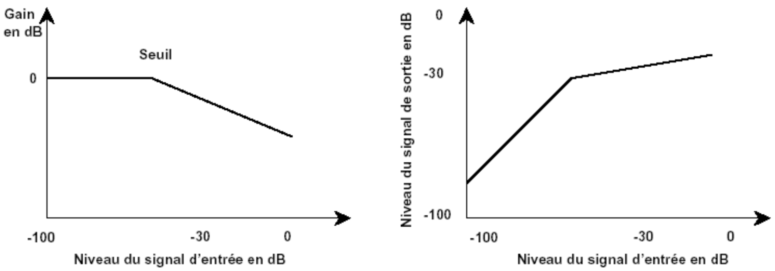
- **Compression is characterized by :**
  - A threshold
  - A compression ratio





Dynamic compression (*or Compressor*)

- **Amplification gain is given by:**



$$Y_{eff} = X_{eff} + g(X_{eff})$$

Where α is the gain slope

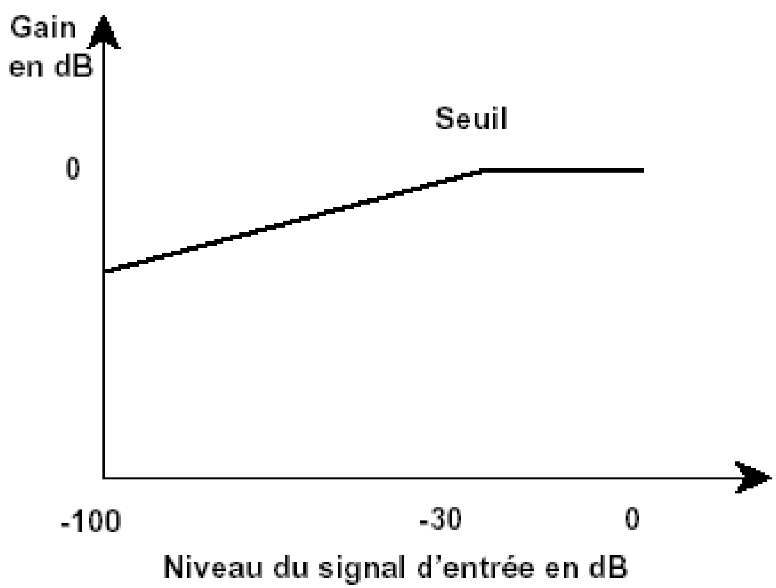
$$\frac{dY_{eff}}{dX_{eff}} = 1 + \frac{dg(X_{eff})}{dX_{eff}} = 1 + \alpha$$

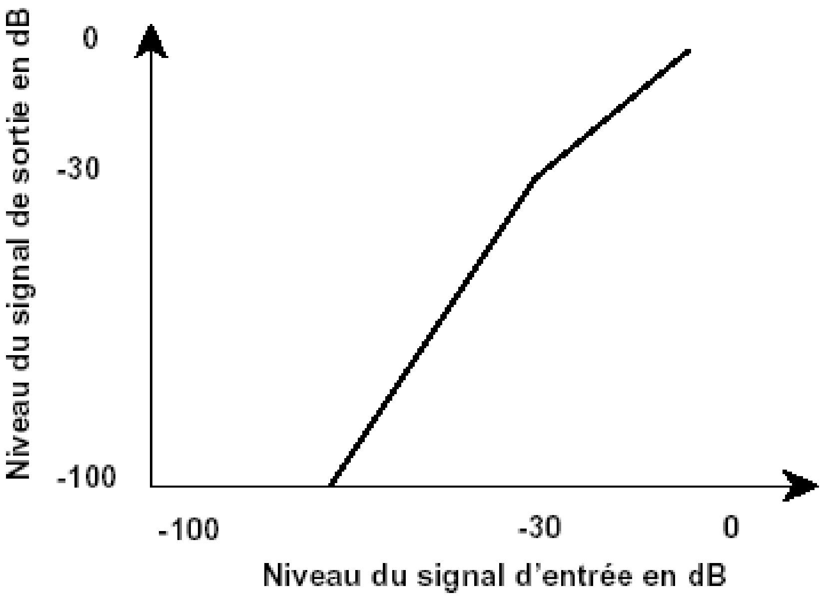
If the gain is linear in X_{eff} with slope α, the relation input/output is also linear with a slope equal to $1+ \alpha$

In practice : ratio of from 2:1 to 10:1

Dynamic expansion (Expander)

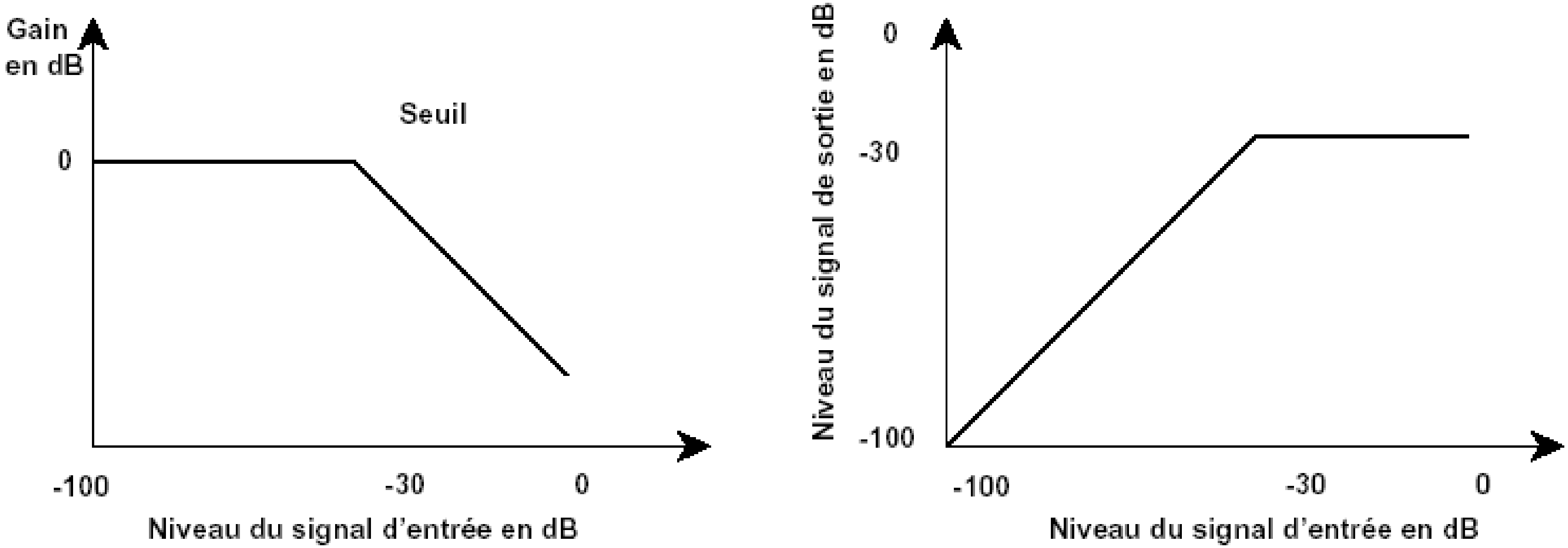
- **Expansion characteristics**





Limitation (*or Limiter*)

- **Limitation characteristics**



Use cases

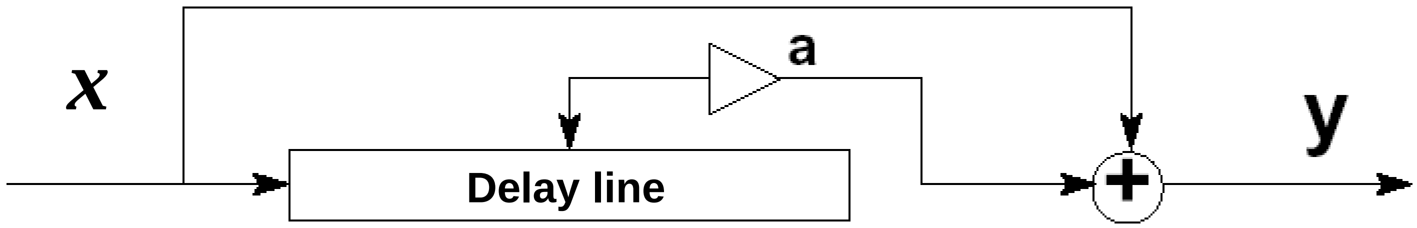
- Background noise ‘suppression’:
  – Dolby B, C…are based on reversible compression/expansion
- Recording of a signal with high dynamic (CD) on a traditional lower dynamic support (tape)
- High gain compression for producing a sustain effect (well appreciated by guitarists…
- High compression used by some radio .. which broadcast at high mean average level….
- Limitation used to avoid clipping, distortion (which are more detrimental to the signal quality than the effect of limitation).

Flanging / Phasing

- **Origin:**
  - Use of the thumb on two turntables playing the same tune.
  - **The thumb effect** : slows down one piece compared to the other one which are then sometimes synchronised and sometimes slightly out of phase.
- **Interpretation:**
  - The signals coming from the two turntables are delayed (with a variable delay)
  - When the two signals are added some sinusoidal components are cancelled because they are in opposition of phase (*e.g. phase inverted*)
  - **Effect:** Flanging then introduces "holes" in the spectrum (regularly spaced) and their positions change in time.

Flanging / Phasing

- **Implementation**



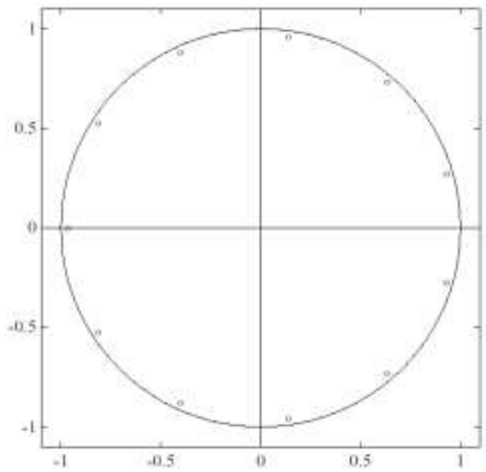
- Tranfer function

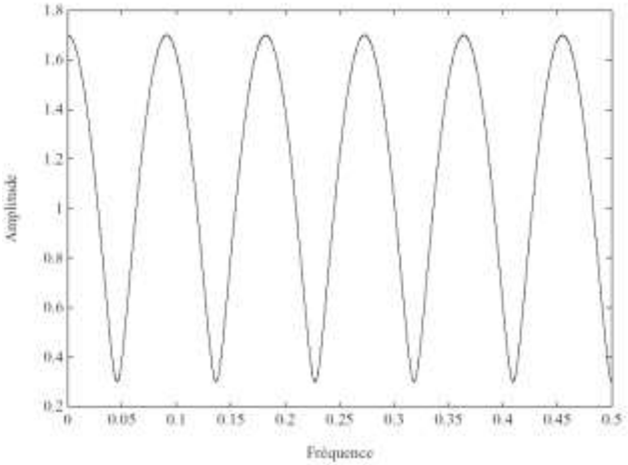
$$H(z) = 1 + a\ z^{-p} \qquad |H(e^{j\omega})| = \sqrt{1 + a^2 + 2a\cos(p\omega)}$$

- *p* minima at odd multiples of *Fe/2p*

Flanging / Phasing

- **Transfer function of the « fixed » (non variable) system**



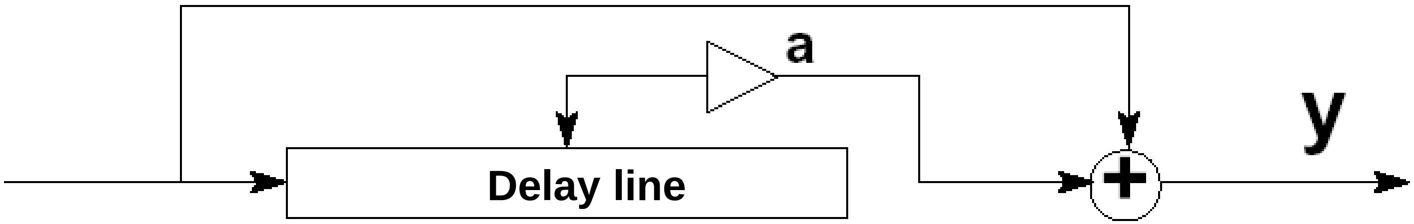


- This transfer function has a meaning only for slowly variations of the delay p ….
- Strength of the effect: parameter a

$$R_{dB} = 20\ (\log_{10}(1+a) - \log_{10}(1-a))$$

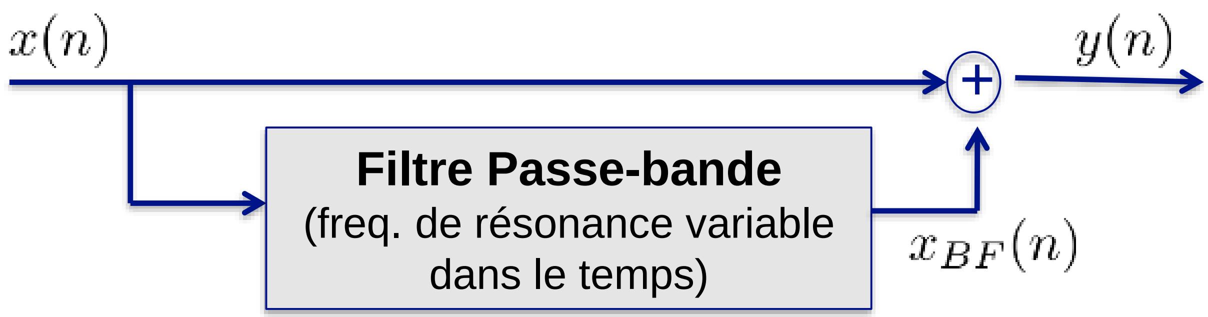
Phasing

- **Phasing is based on a variable comb filter but where here the gain a is variable (not the delay)**



“Wha-wha” effect

- **Effect “Wha-wha”: The delay (e.g. notch filter) is replaced by a time varying bandpass filter :**



- **Demonstration:**

Original

Wah-wah (modulation of resonance frequency with a triangle function between 500H and 3000Hz)

Time and pitch shifting

- **Time-shifting : to modify the duration of an audio recording without affecting the frequency content.**

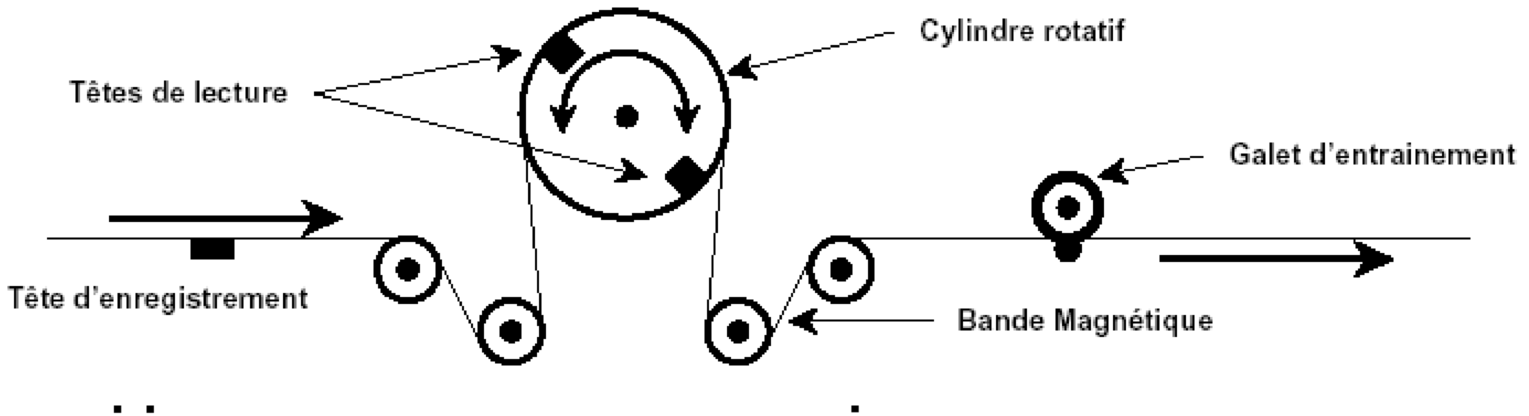
- **Pitch shifting: to modify the frequency content (e.g pitch) of an audio recording without affecting its duration.**

  - Pitch shifting without modifying the main resonances (e.g. formant for speech).
  - Straight pitch shifting (e.g. auto tune etc ..)

- **Various methods exist**
  - TD-Psola, Phase vocoders
  - Specific audio methods

Circular memory technique

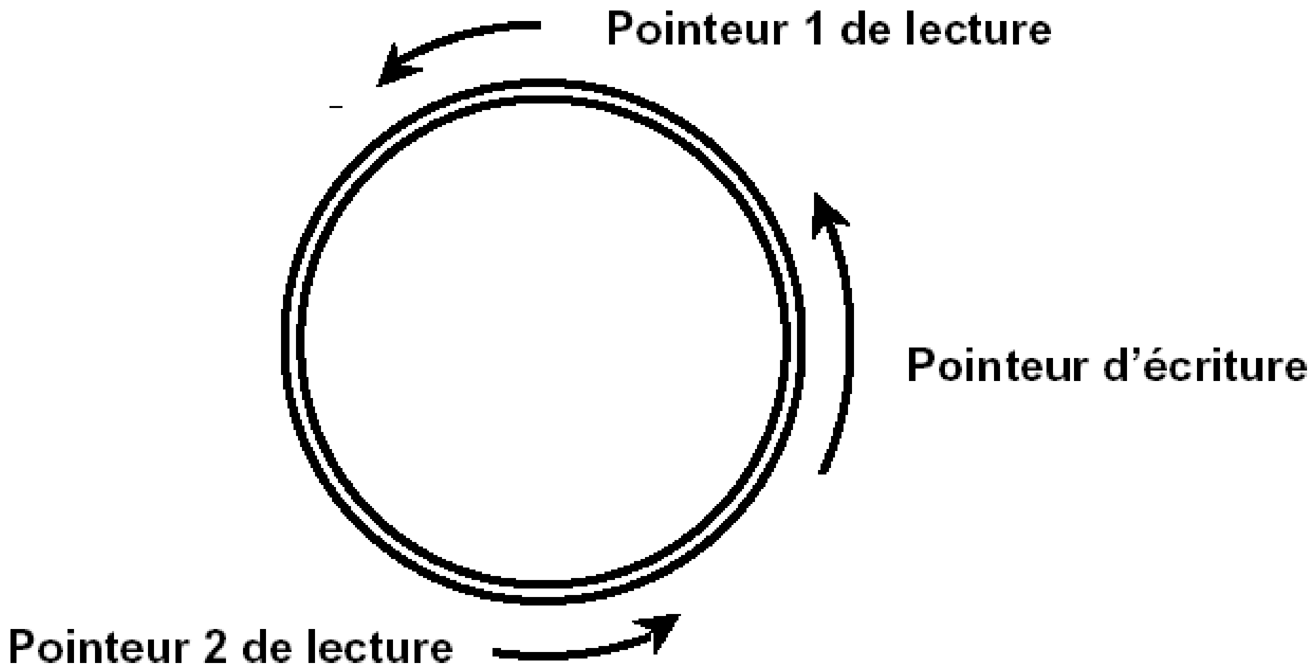
- **The old times …**



$$\alpha = \frac{V_r}{V_a} = \frac{V_a + R\ \Omega_{cylindre}}{V_a}$$

Circular memory technique

- **In digital …**



Circular memory technique

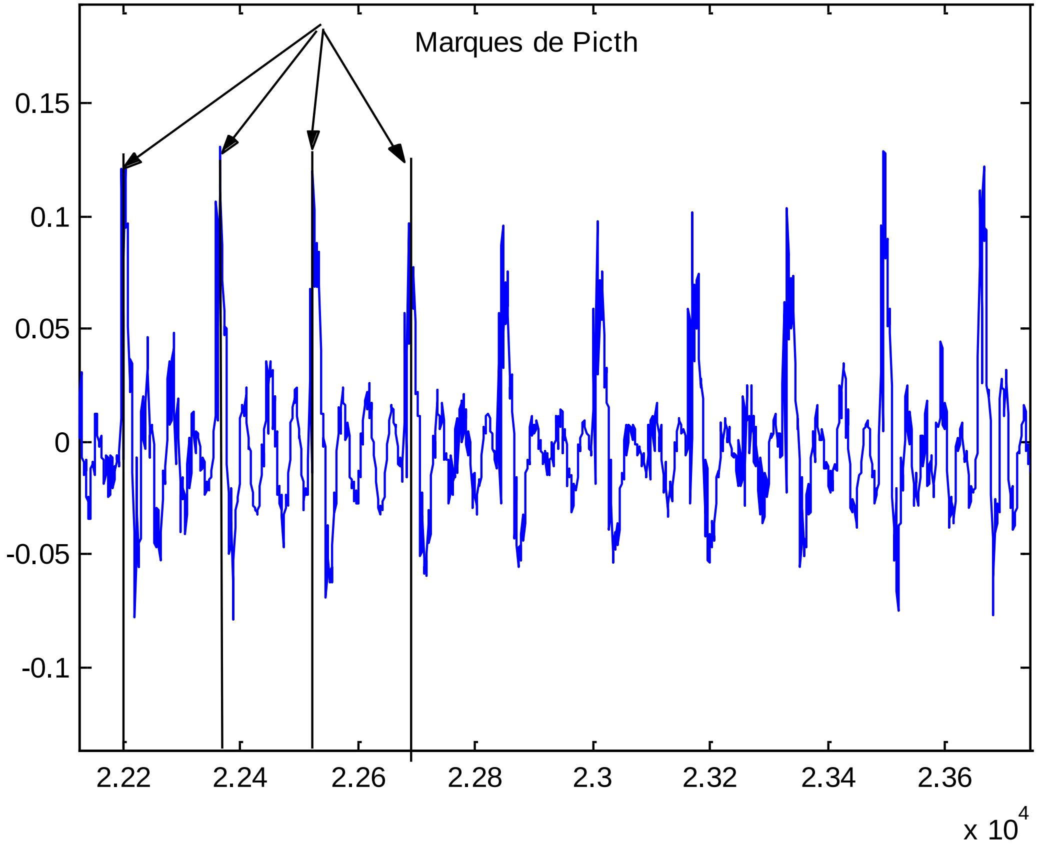
- **Time-shifting**

  - Change the sampling frequency by a factor α
  - Change the pitch using the circular memory technique

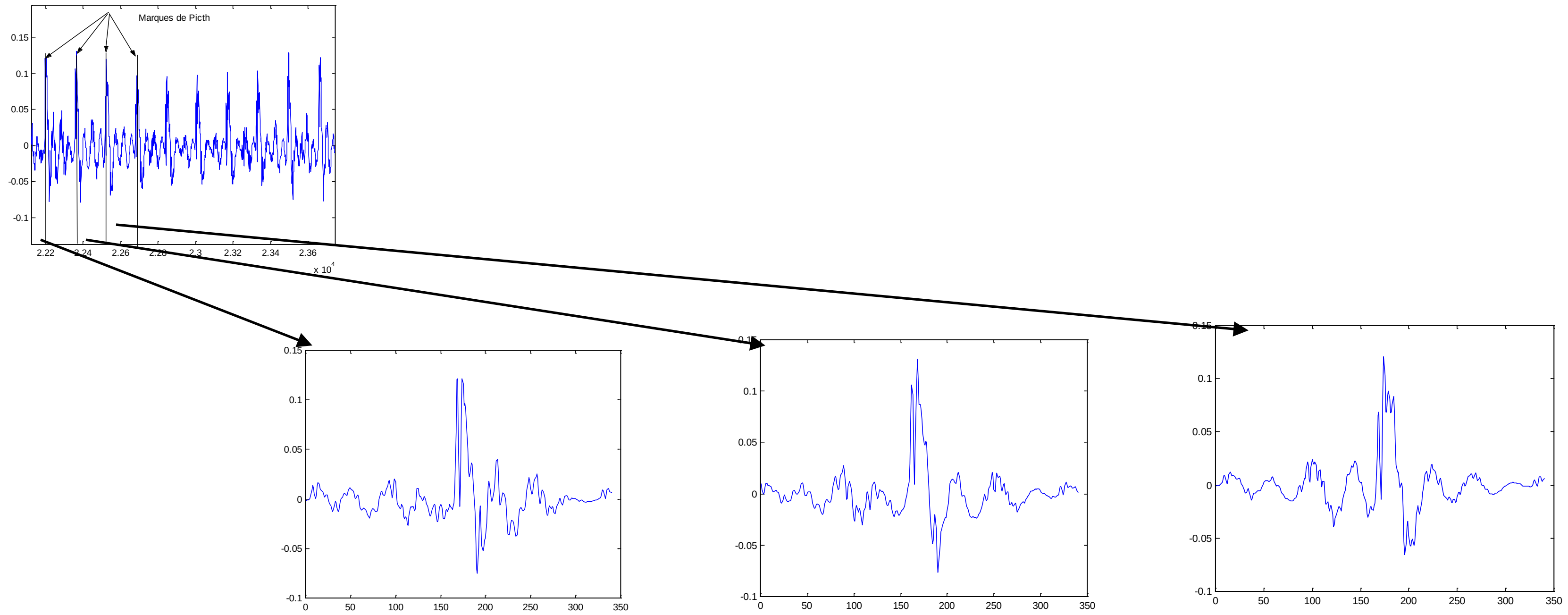
  - The two operations can be done simultaneously (we now get closer to the TD-PSOLA approach)

TD-PSOLA

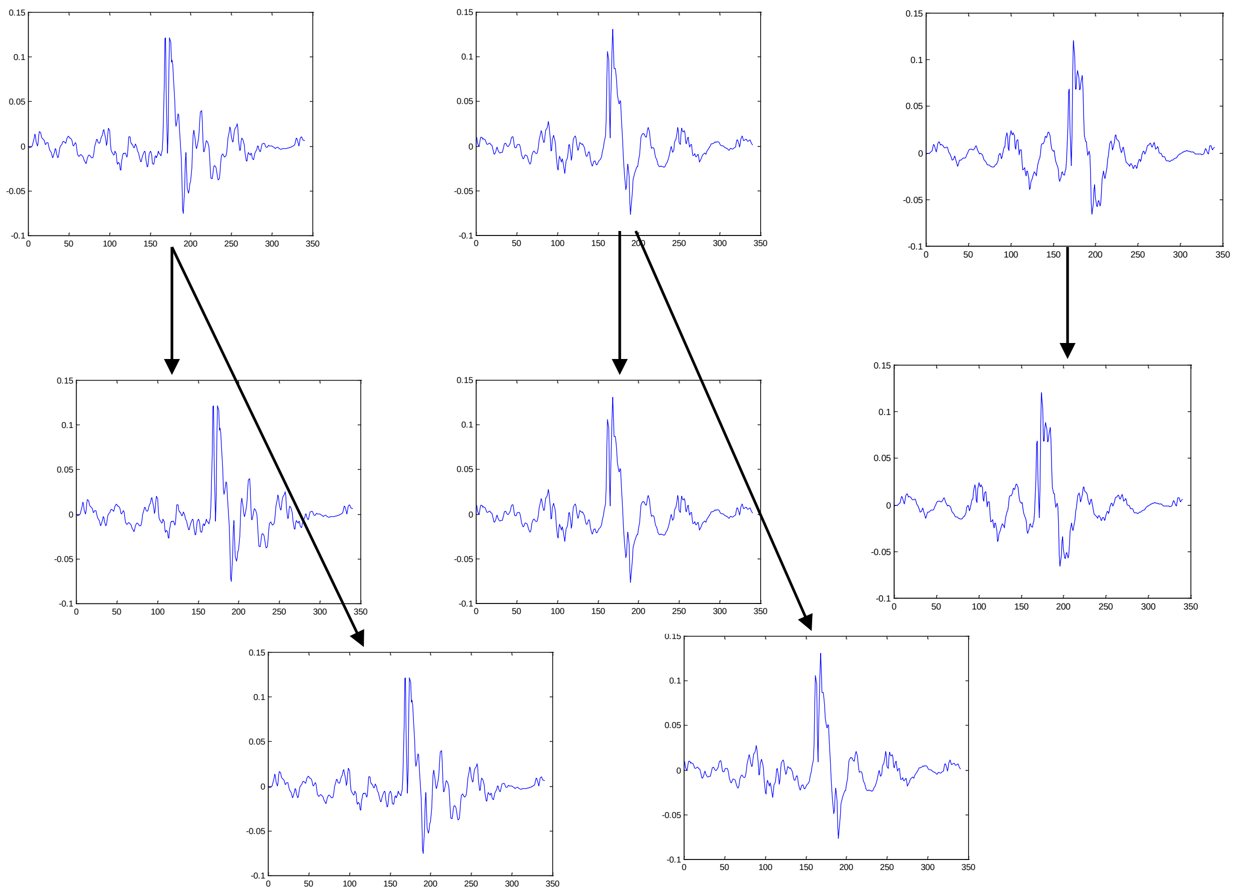
TD-PSOLA: Time Domain Pitch Synchronous OverLap and Add



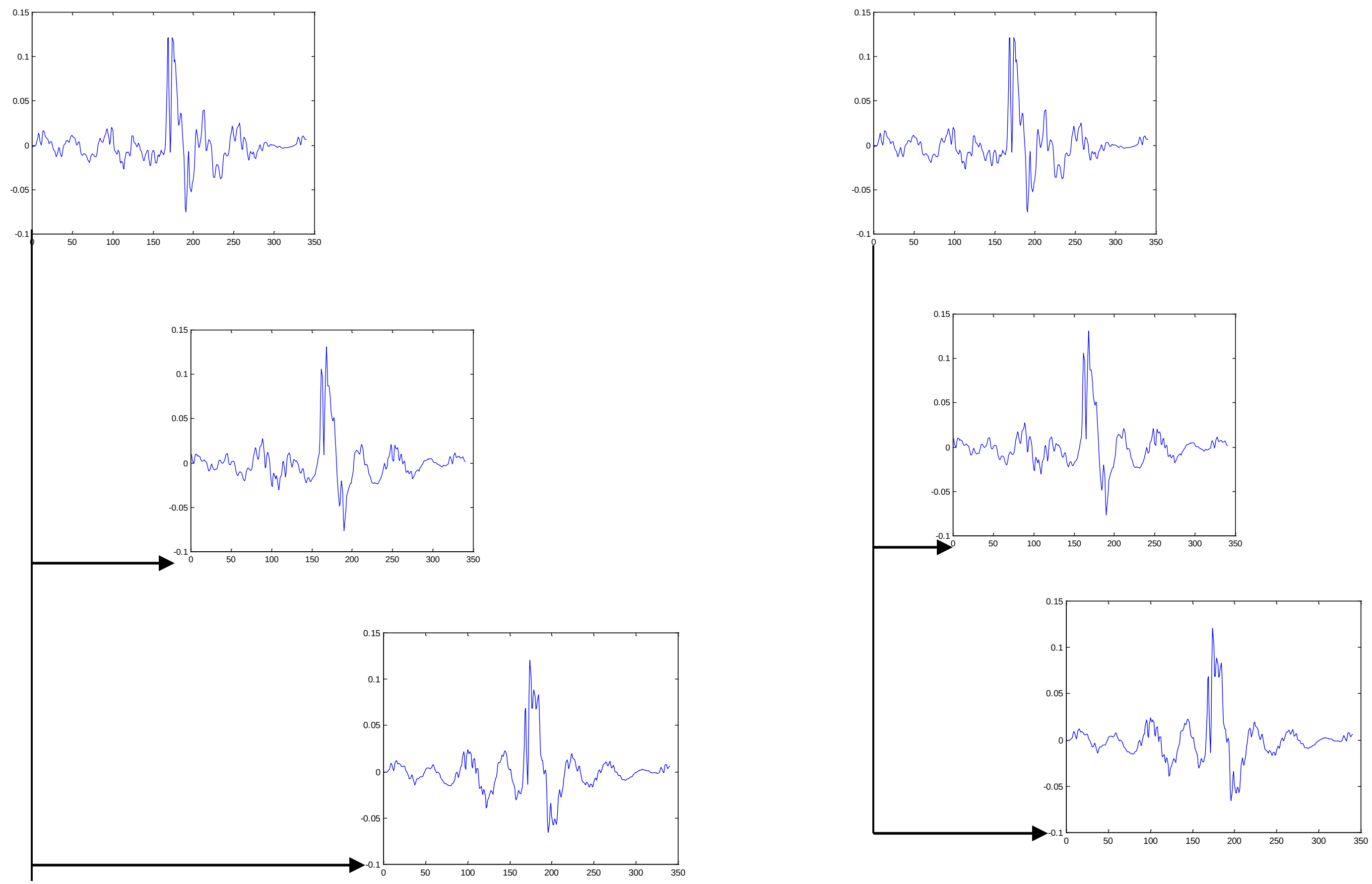
Short-term signals



Time-shifting



Pitch shifting



TD-PSOLA:Synthesis

- **Extraction of pitch synchronous short term waveforms**
  - Analysis window spanning 2 periods of the analysed signal (for quasi-periodic segments).
  - Fixed duration analysis window (around 10ms) for more complex segment (noise or high number of notes).

- **Synthesis by Overlap and Add**
  - Insertion / Suppression of short term waveforms to modify duration
  - Modification of short term waveforms spacing to modify pitch.

- **Sound example (singing voice)**

Transposed

Distorsion

- Usually, distorsion is not desired ….
- But it could also be an important artistic effect (for guitarist, for instance….)
- Distortion is traditionally obtained by applying a non-linear function to the audio signal.

La distorsion

- **It can be shown that distortion gives birth to new components in the signal :**
  - The development of the non-linear function is given by:

$$f(u) = f(0) + uf'(0) + \frac{u^2 f''(0)}{2!} + \frac{u^3 f'''(0)}{3!} + \ldots = \sum_{i=0}^{\infty} \frac{u^i}{i!} \frac{d^i f}{du^i}$$

  - Applying the fonction f to the audio signal $x_n = \sum_{i=-p}^{p} e^{j\omega_i n}$

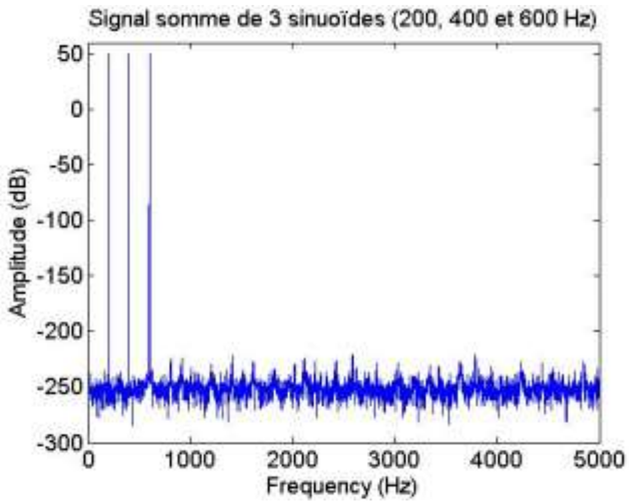
results in a weighted sum of its integer $y_n = \sum_{i=0}^{\infty} \alpha_i (x_n)^i$

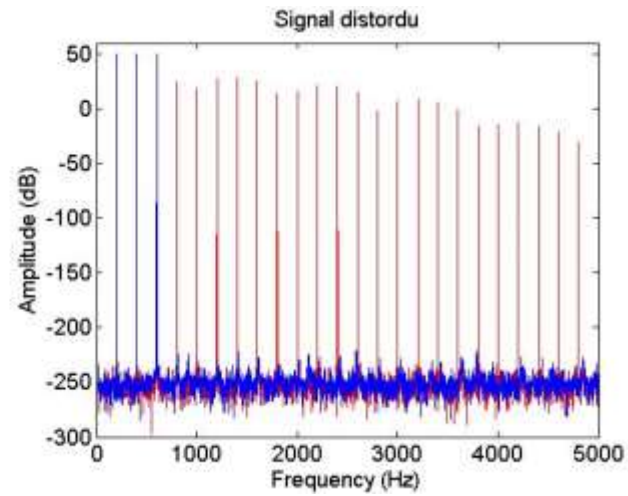
with $(x_n)^i = \sum_{-p \leq j(k) \leq p} e^{j(\omega_{j(0)} + \omega_{j(1)} + \ldots + \omega_{j(i-1)})}$

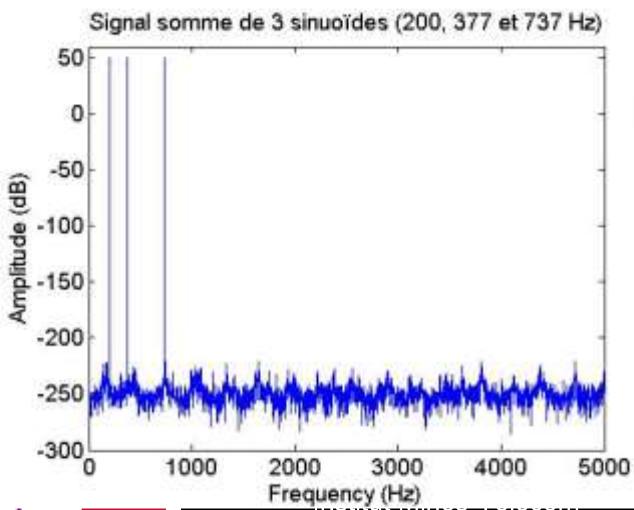
Droits d'usage autorisé

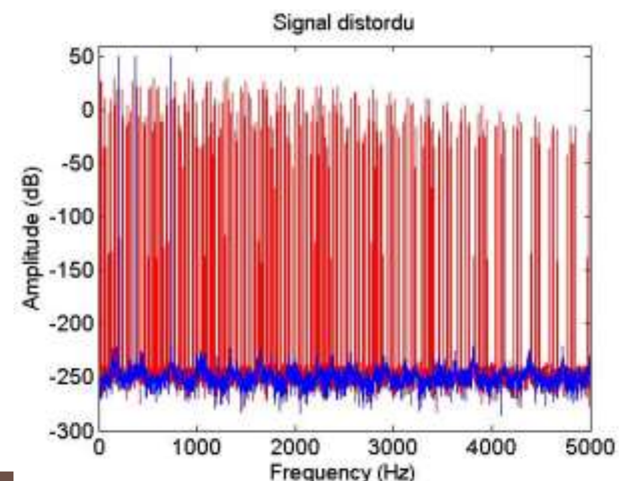
Examples

- Non linear function: $f(x) = x^9$ with $x = \sum_{i=1}^{3} A_i \sin(2\pi f_i t)$









Artificial reverberation

Reverberation

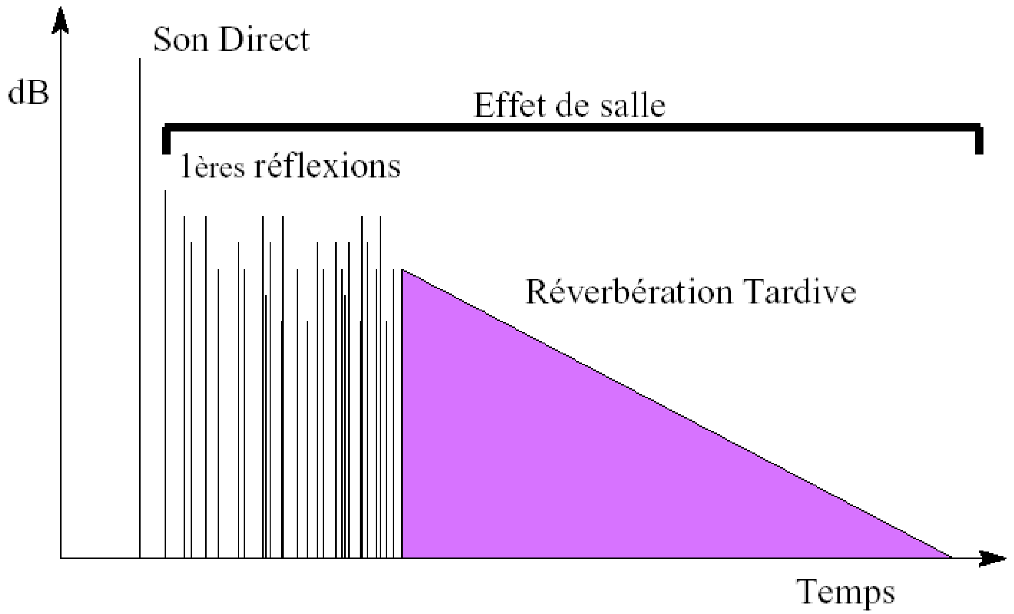
- **Introduction : the acoustic channel**

$$y(t) = \int_0^{\infty} x(t-u)h(u)du$$

$$y_n = \sum_{i=0}^{\infty} x_{n-i} h_i$$

Reverberation

- **Room effect can be decomposed in:**
  - A contribution due to **early echoes** or early reflexions (which depends on the room geometry and on the positions of the source and microphone)
  - A contribution due to **late reverberation** (which mainly depends on the volume and global absorption of the room)



Reverberation

- **Approaches for reverberation:**

  - **Based on Physics**
    - Aim at exactly reproduce the propagation of a sound source in a given room
    - Advantage: a direct link between the physical room characteristics and the corresponding reverberation effect.
    - Drawback: Rather complex approach (rather long room impulse responses).

  - **Perceptual approaches**
    - Aim at only reproducing the main perceptual characteristics of reverberation.

Reverberation

- **Advantages of the perceptual approaches:**
  - Low complexity algorithms (implementation with IIR filters)
  - Possible to define perceptual parameters that can be changed in real time.
  - Ideally, a unique algorithm could simulate the full space.
  - Some spaces can be very well simulated with perceptual approaches.
- **Drawback: not easy to link perceptual properties to specific physical properties of a room ….**

Late reverberation

- Characterised by a large density of echoes in time *(> 1000 per second in a large room)*
- Can be modelled as a Gaussian random process with exponential decay
- Similarly, the frequency response is characterised by a large density of modes (at least above a given frequency Fs called **Schroeder frequency**)
- Fs $= 2000\sqrt{\frac{T_{60}}{V}}$ *(T_{60} in seconds, V in m3, Fs in Hz)*

(Late) reverberation

- Polack Model (1998)

- The reverberation (e.g. the Room Impulse Response) is modelled as a non-stationary centered Gaussian process:

$$a(t) \sim \mathcal{N}(0, r^2(t))$$

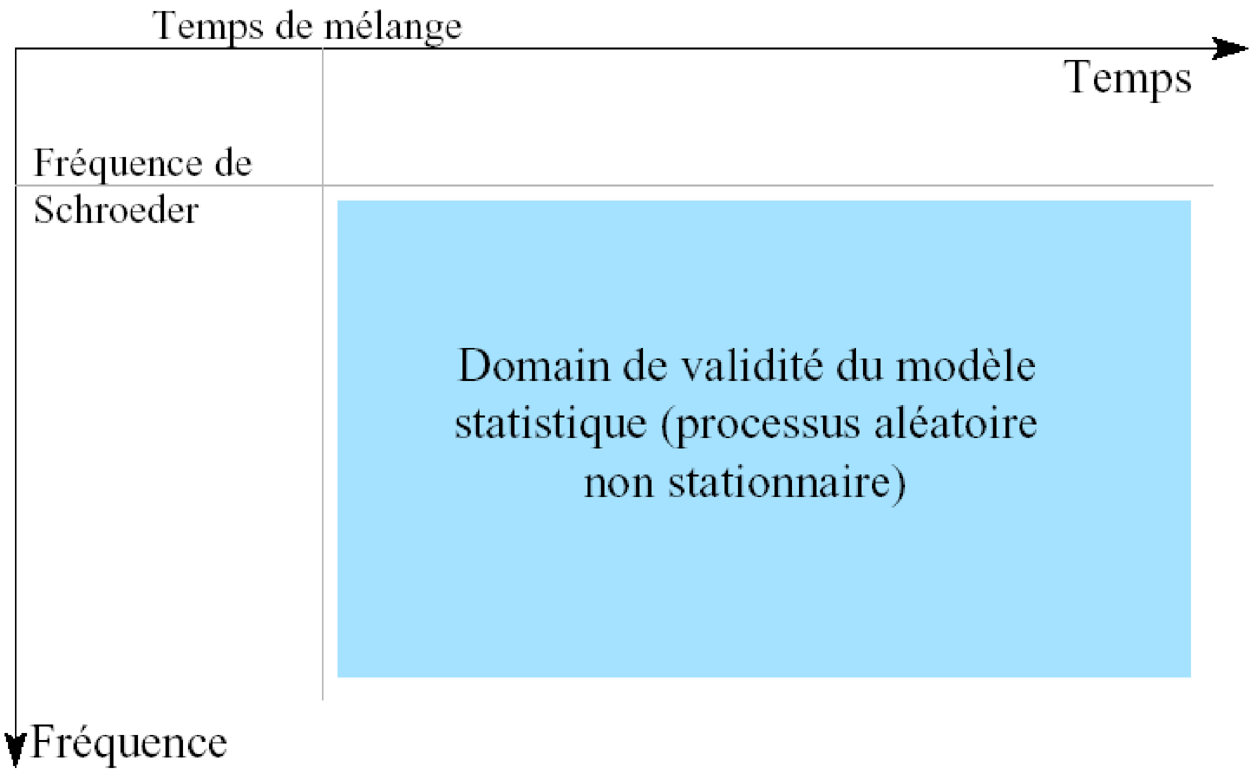
- with

$$r^2(t) = \sigma_r^2 \exp(-2t/\tau)) \qquad \tau = \frac{T_{60} F_s}{3 \ln(10)} \quad \text{samples}$$

- Often used to model the full Room Impulse Response, it is theoretically (and practically) only valid for late reverberation

Late reverberation

- Domain of validity of the statistical model: depends on the mixing time ($\sqrt{V}$) and Schroeder frequency.

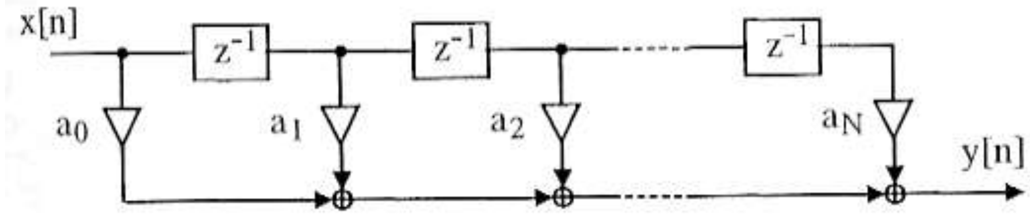


Artificial reverberation algorithms

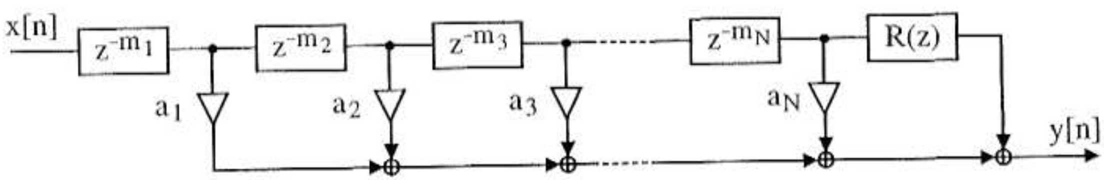
- **Used:**
  - Either to add a room effect to studio recordings
  - Or to modify acoustical properties of a listening room
- **Early systems used analog process :**
  - Network of springs, metallic plates, ..
- **First digital systems in 1960’s !!**

Early echoes models

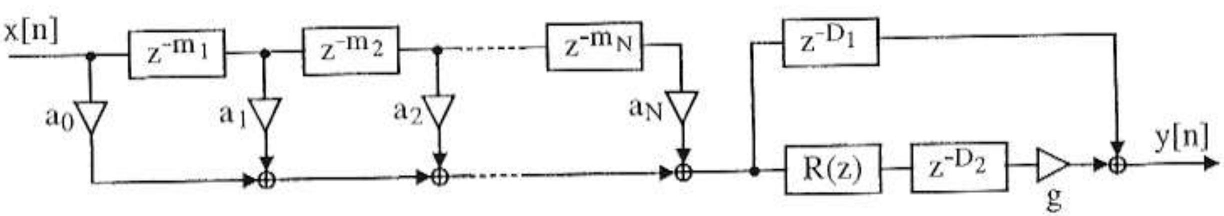
- **Simple model**



- **Schroeder (1970)**

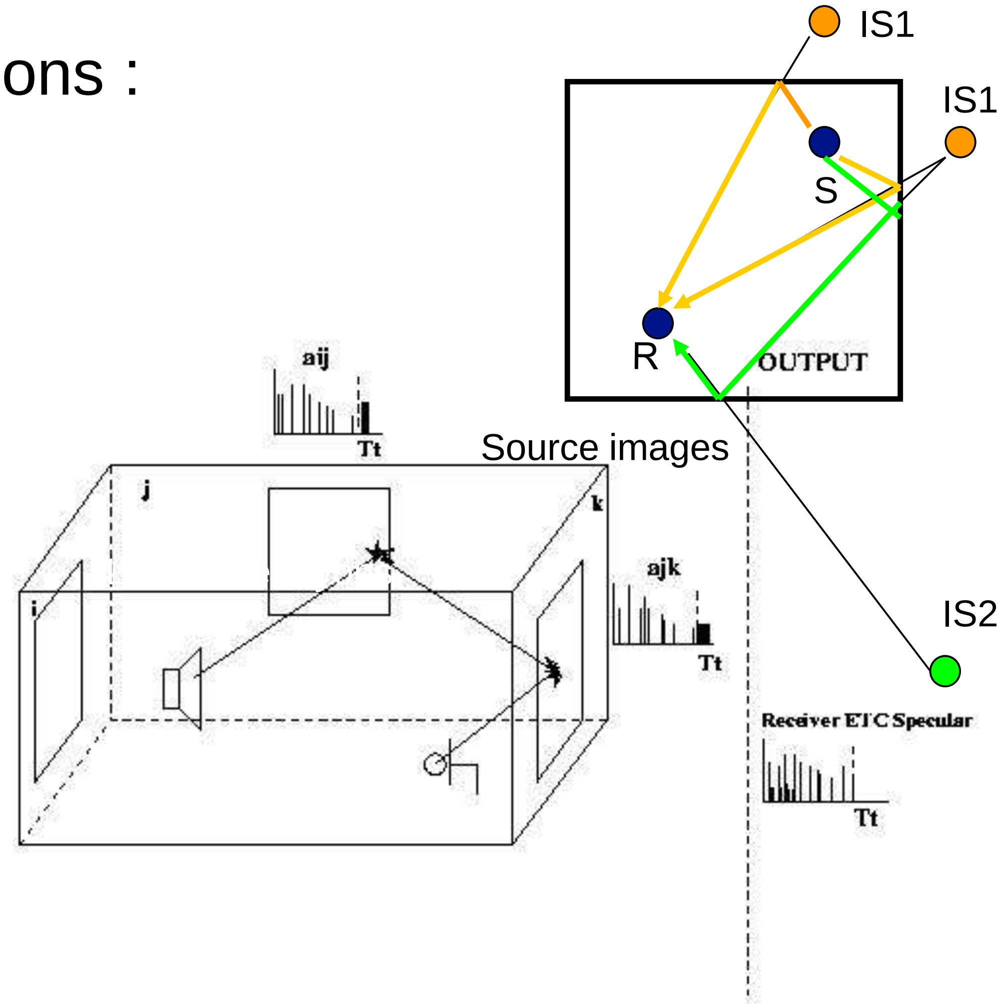


- **Moorer (1979)**

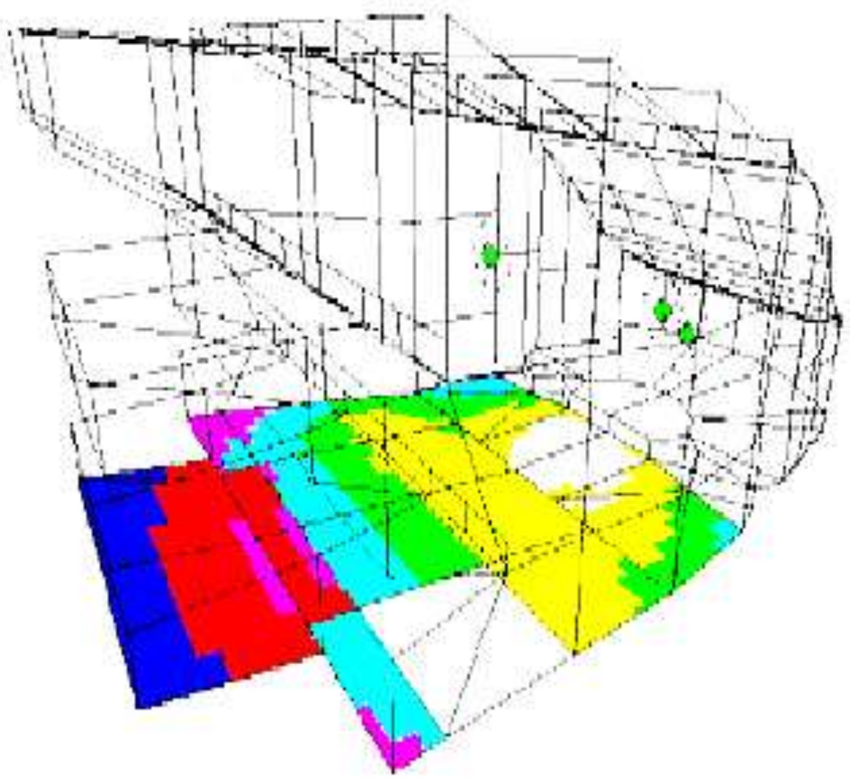


Early echoes models

- Simulate specular reflexions :
  - Ray beams
  - « cones » beams
  - Image sources



Application to any geometry



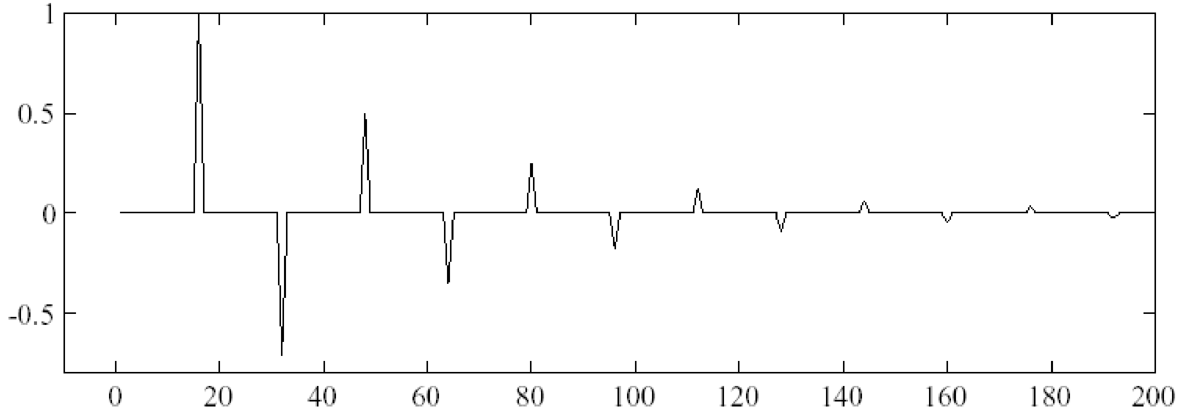
Late reverberation

- **Three main difficulties :**

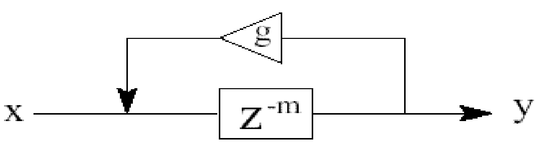
  - To obtain simultaneously a high density of modes and echoes
  - To avoid a potential artificially sounding output signal (e.g. avoid the « metallic » aspect)
  - To control, independently, the reverberation time and reverberated energy per frequency

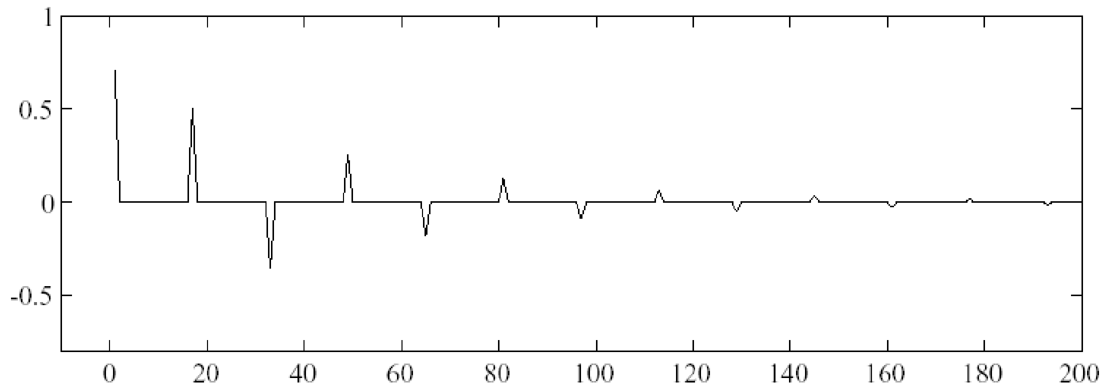
Reverberation

- **Comb filters *(m=16, g=0.707*)**

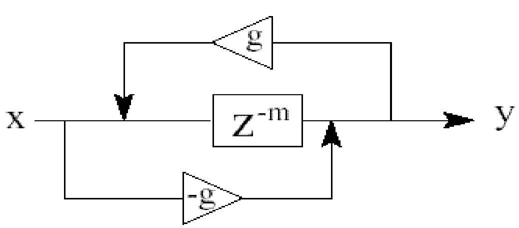


$$C(z) = \frac{z^{-m}}{1 - gz^{-m}}$$





$$A(z) = \frac{-g + z^{-m}}{1 - gz^{-m}}$$

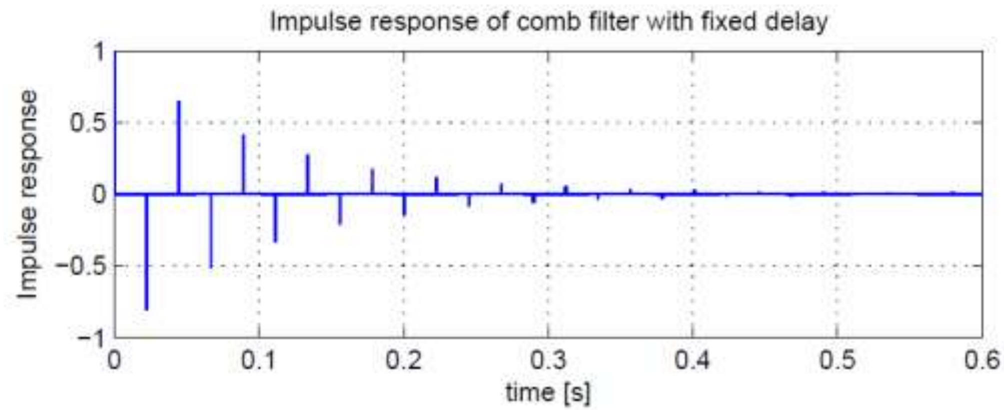


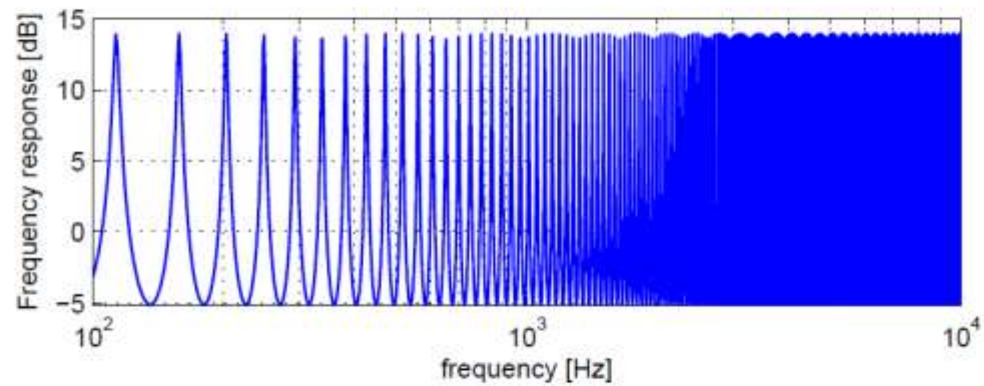
Reverberation

- Comb filters imply an harmonic coloration to the signal
  - *Resonances at frequencies:*

$$\omega_k = \frac{2\pi k}{m}$$

- All-pass filter also brings coloration when the signal is non-stationary (the usual case of audio signals)

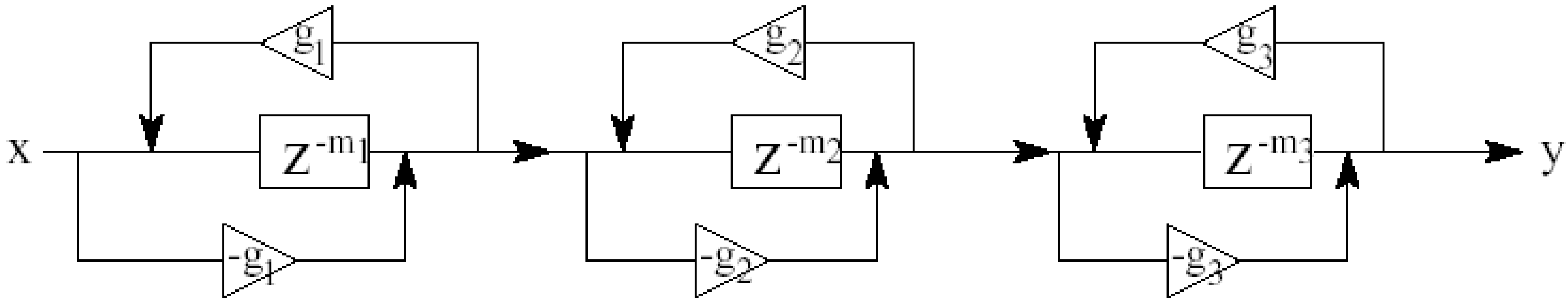




- **But main drawback is the low density of echoes**

Reverberation

- **A solution to augment echoes density: place several all-pass filters in serie**



- **In practice: it induces a « metallic » sonority for transients (not very natural)**

Comb filters

- Reverberation time of a comb filter

$$\frac{20\log_{10}(g_i)}{m_i T} = \frac{-60}{Tr}$$

- Transfer function of comb filter:

$$C(z) = \frac{z^{-m}}{1 - gz^{-m}}$$

- Modulus of the poles are given by :

$$\gamma_i = \sqrt[m_i]{g_i} = 10^{-3T/Tr}$$

Comb filters

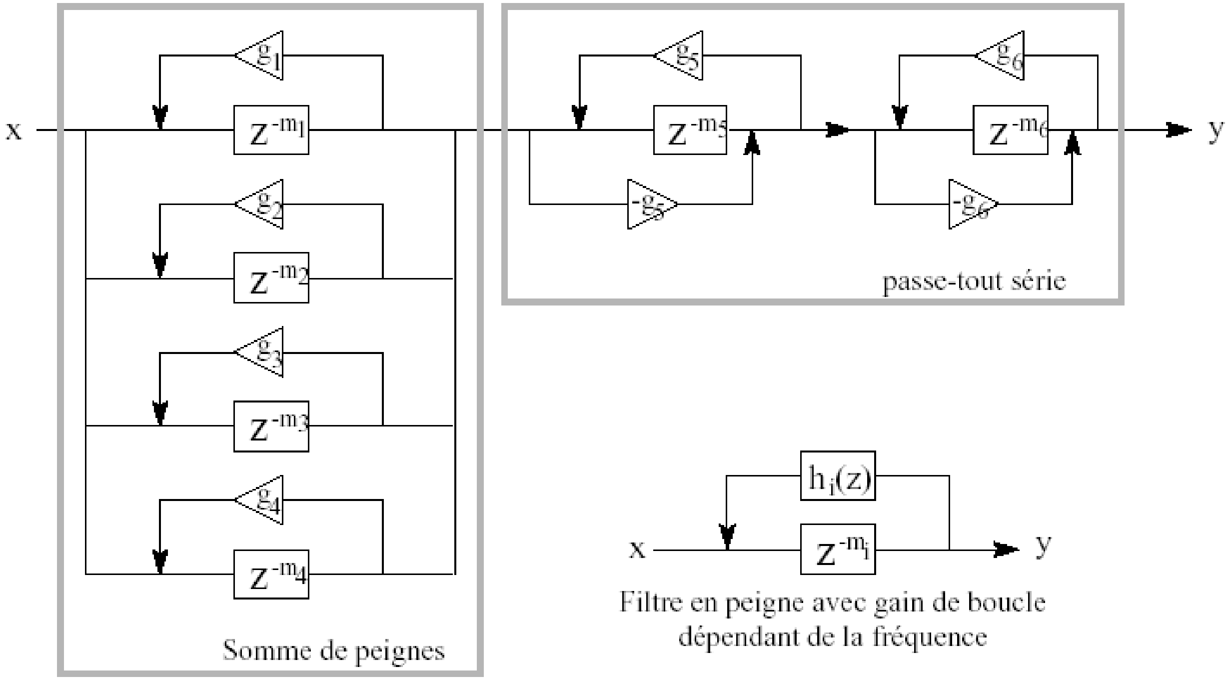
- **Importance of poles modulus:**

  - If all poles have the same magnitude, all resonances of the filters in parallel will decrease at the same speed.

  - If poles have difference amplitudes, a tonal coloration may be perceived.

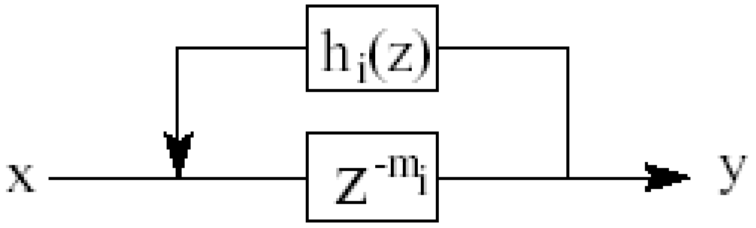
Schroeder reverberator

- **To avoid coloration, Schroeder proposed to combine all pass and comb filters.**



Schroeder reverberator

- **Improvements:**
  - Frequency dependant reverberation time (*taking into account air absorption*)



Comb filter with frequency dependant feedback gain

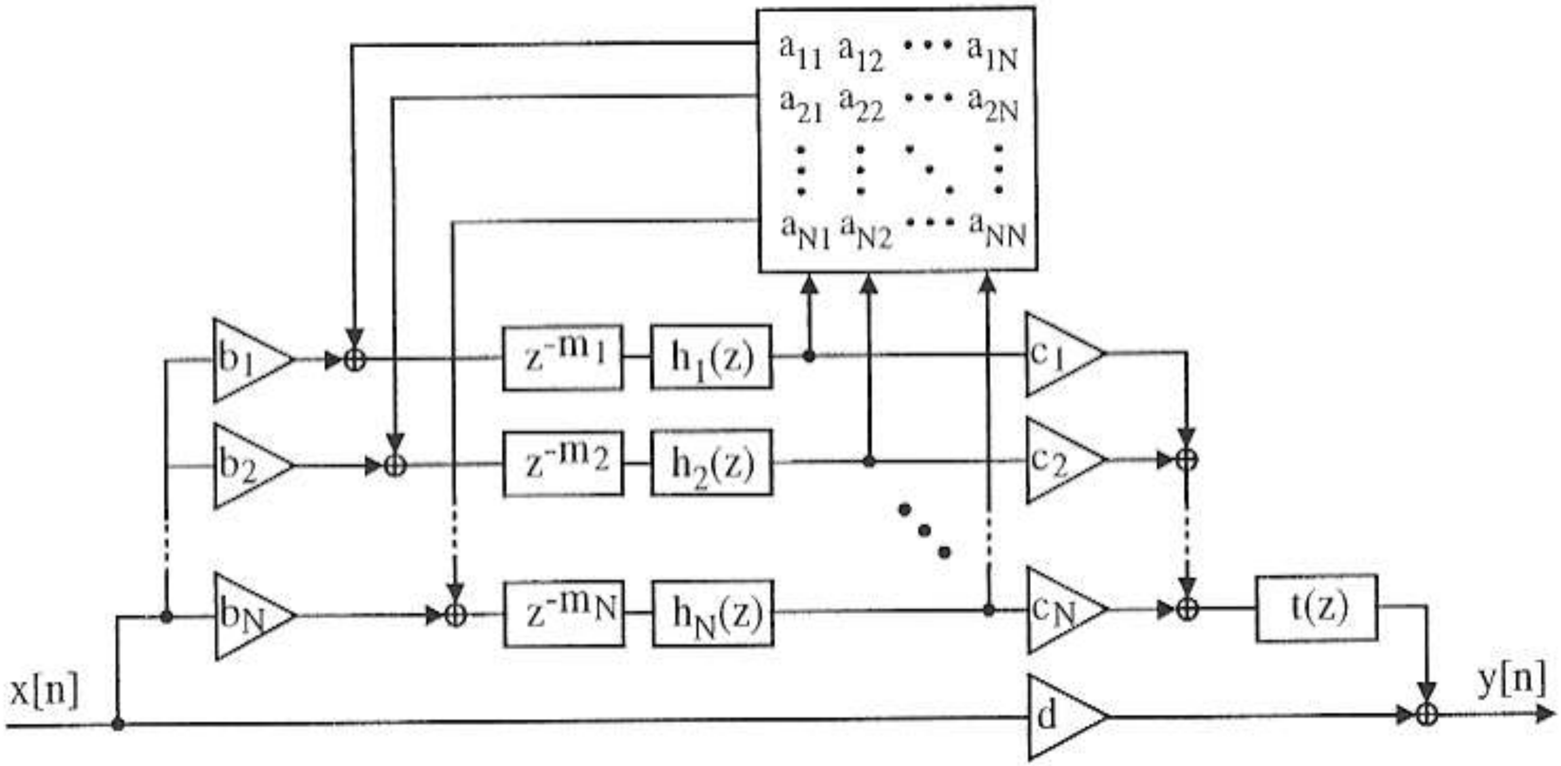
Schroeder reverberator

- **Example**

  - Tr=0.1 s (small room)
  - Tr=0.5 s (large room)
  - Tr=2 s (cathedrale)

Reverberant filter with unitary feedback

- **Generalisation (*Jot et Chaigne*)**



An alternative method (physical based):
Radiance Transfer Method (RTM) (*From H. Bai PhD thesis*)

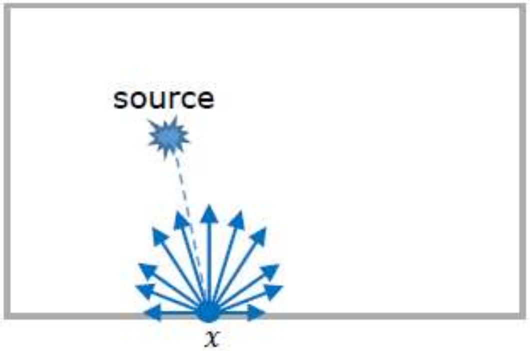
- **Analytical acoustic radiance transfer model**

$$I(x,t) = I_0(x,t) + \int_S R(x,x',t)I(x',t-\frac{|x-x'|}{c})dx'$$

Une approche alternative (physique):
Radiance Transfer Method (RTM)

- **Analytical acoustic radiance transfer model**

$$I(x,t) = \boxed{I_0(x,t)} + \int_S R(x,x',t)\,I\left(x',t-\frac{|x-x'|}{c}\right)dx'$$

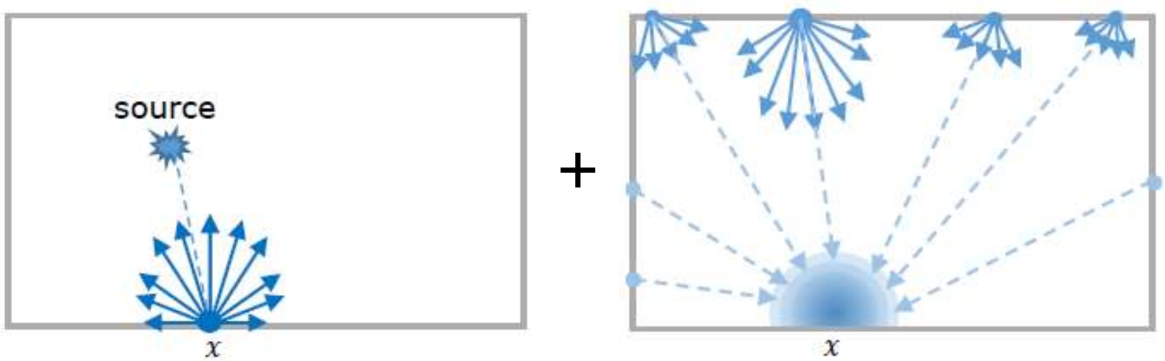


(a) Direct contribution

Une approche alternative (physique):
Radiance Transfer Method (RTM)

- **Analytical acoustic radiance transfer model**

$$I(x,t) = I_0(x,t) + \int_S \boxed{R(x,x',t)I(x',t-\frac{|x-x'|}{c})dx'}$$

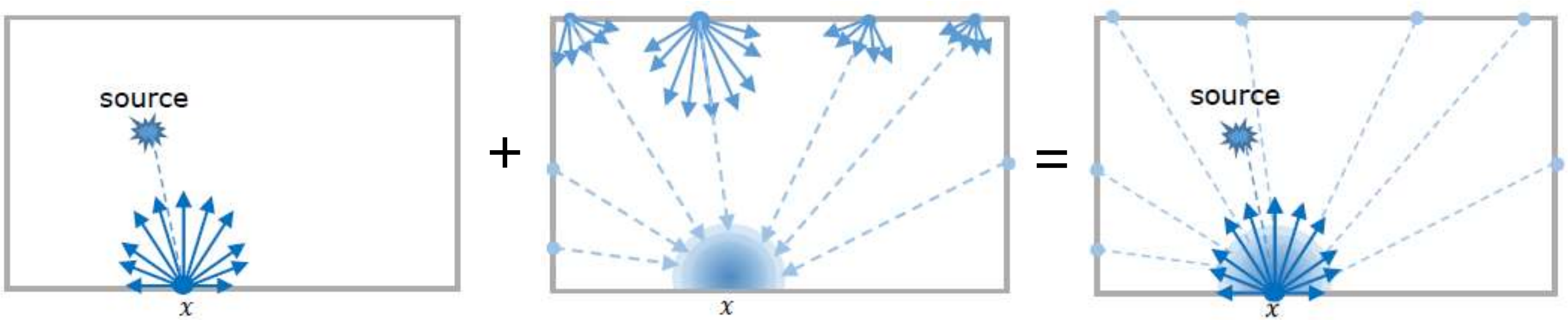


(a) Direct contribution (b) Indirect contribution

Une approche alternative (physique):
Radiance Transfer Method (RTM)

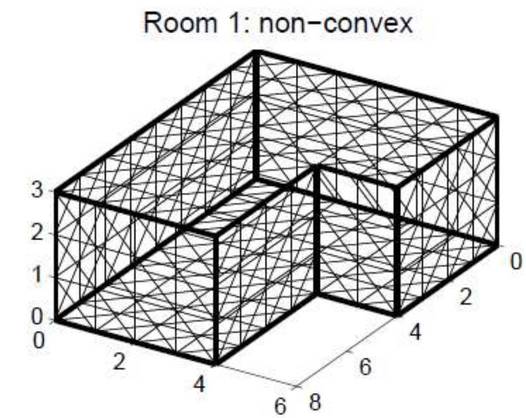
- **Analytical acoustic radiance transfer model**

$$I(x,t) = I_0(x,t) + \int_S R(x,x',t) I(x', t - \frac{|x-x'|}{c}) dx'$$



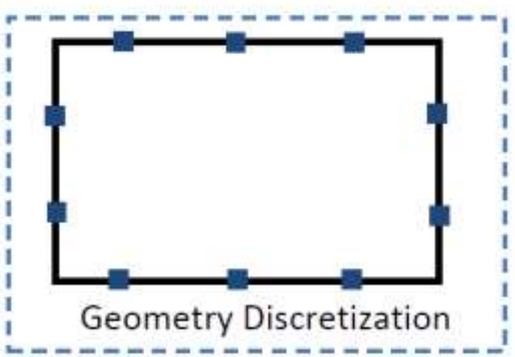
(a) Direct contribution (b) Indirect contribution (c) Total radiance

Radiance Transfer Method:
Digital simulation

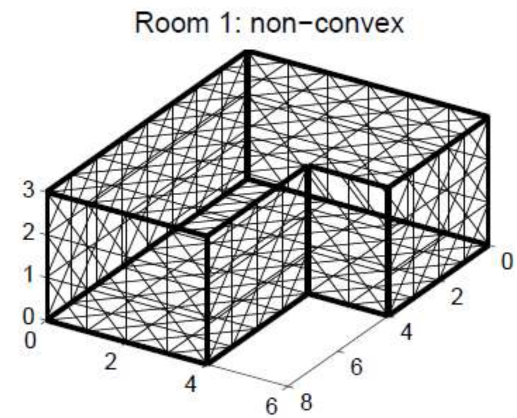


- **Room discretization**
  - Room is divised in patches
  - Iterative expression

$$I_i^{(n)}(t) = I_i^{(n-1)}(t) + \sum_{j=1, j\neq i}^{M} F_{i,j}^{(1)} I_j^{(n-1)}(t - \frac{r_{i,j}}{c}))$$

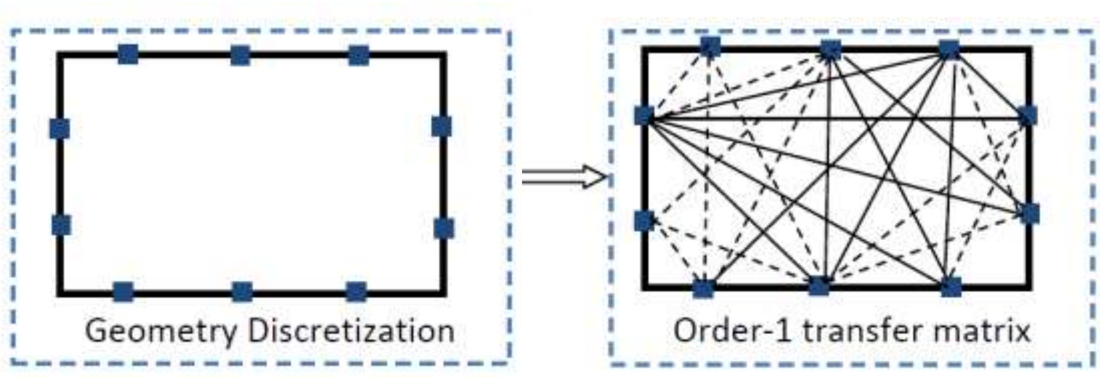


Radiance Transfer Method:
Digital simulation

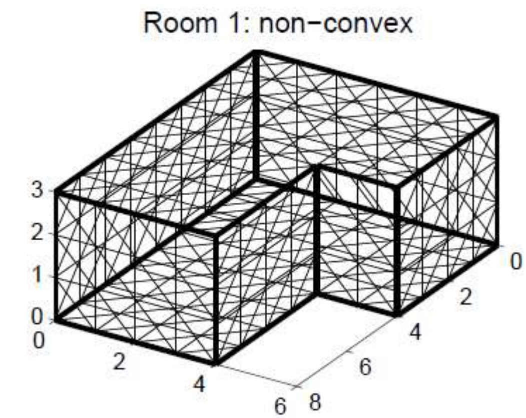


- **Room discretization**
  - Room is divised in patches
  - Iterative expression

$$I_i^{(n)}(t) = I_i^{(n-1)}(t) + \sum_{j=1, j\neq i}^{M} \boxed{F_{i,j}^{(1)}} I_j^{(n-1)}(t - \frac{r_{i,j}}{c}))$$

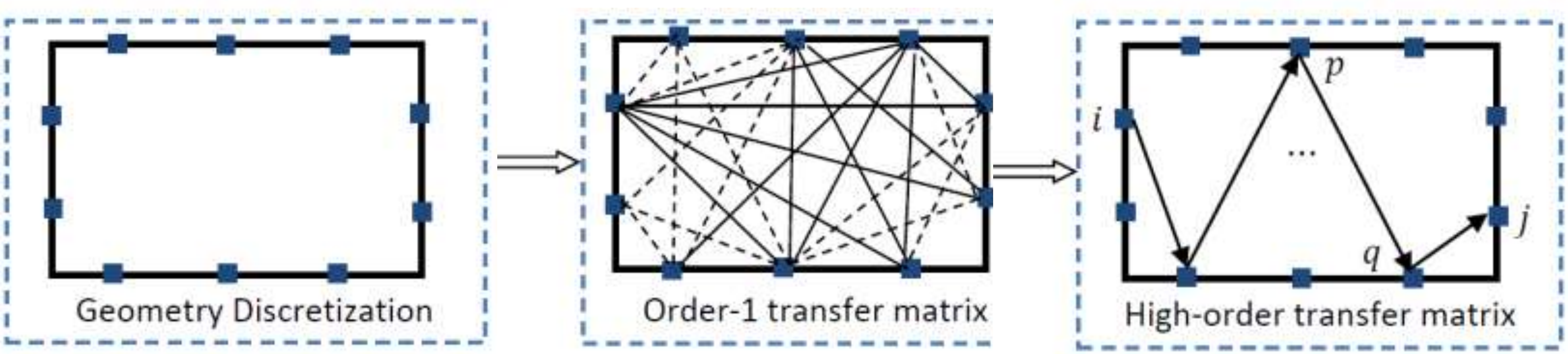


Radiance Transfer Method:
Digital simulation



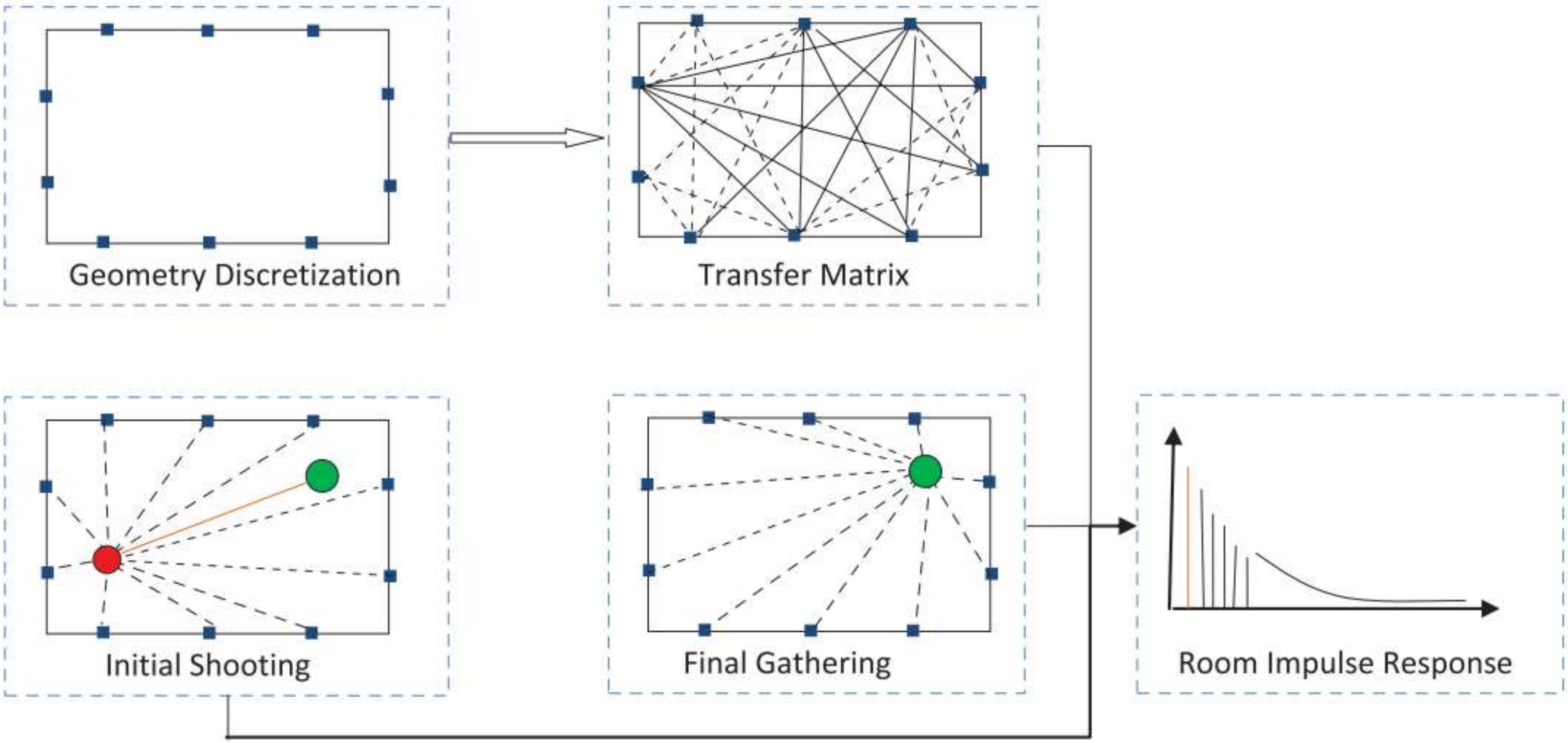
- **Room discretization**
  - Room is divised in patches
  - Iterative expression

$$I_i^{(n)}(t) = I_i^{\boxed{(0)}}(t) + \sum_{j=1, j\neq i}^{M} \boxed{F_{i,j}^{(n)}} I_j^{\boxed{(0)}}\left(t - \frac{r_{i,j}}{c}\right))$$

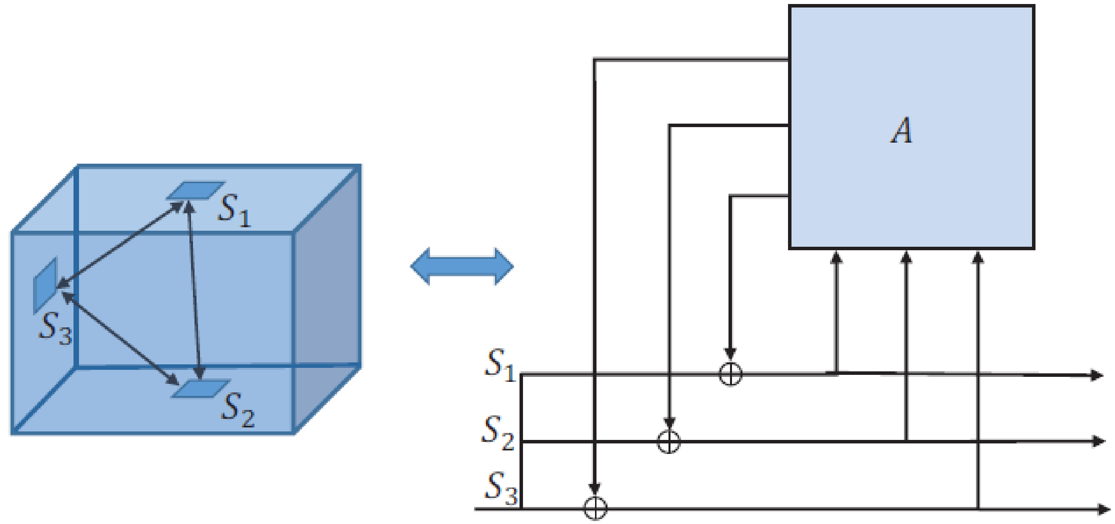


An alternative approach:
Radiance Transfer Method (RTM)

- **In summary**



Links between RTM and linear systems with reverberant filters



- The exchange of energy between patches of RTM can be linked to the recursive structure of the filter networks
- The exchange of energy of high order is equivalent to the infinite feedback loops of filter networks
- Brings efficient implementation of the RTM methods

Conclusions

- **Solutions for artificial reverberation do exist but:**
  - To exactly model a reverberant space calls for complex methods even with simple systems such as with unitary feedback matrix.
  - Most of the commercial systems probably used temporal variations to reduce coloration (but this is not well described in the literature).
  - The perception of tonal coloration is not well understood which explain why it is not clear why a given algorithm generate natural sounding reverberation and another does not.

A few references ….

- **Sound effects**
- V. Verfaille, U. Zölzer and D. Arfib, Adaptive Digital Audio Effects (A-DAFx): A New Class of Sound Transformations, IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING , VOL. 14, NO. 5, SEPTEMBER 2006 1817
- J. Laroche, J.M Jot, G. Richard, Effets et Réverbération artificielle, polycopié de cours Télécom ParisTech, 2005 (in French)

- **Reverberation**
- H. Kuttruff. Room acoustics. SPON Press, London, UK, 4th edition, 2009.
- E.-M. Nosal, M. Hodgson, and I. Ashdown. Improved algorithms and methods for room sound-field prediction by acoustical radiosity in arbitrary polyhedral rooms. The Journal of the Acoustical Society of America, 116(2):970–980, 2004.
- S. Siltanen, T. Lokki, S. Kiminki, and L. Savioja. The room acoustic rendering equation. The Journal of the Acoustical Society of America, 122(3):1624–1635, 2007.
- V. Valimaki, J. D. Parker, L. Savioja, J. O. Smith, and J. S. Abel. Fifty years of artificial reverberation. IEEE Transactions on Audio, Speech, and Language Processing, 20(5):1421–1448, 2012.
- J.-M. Jot. Etude et réalisation d'un spatialisateur de sons par modèles physiques et perceptifs. PhD thesis, Telecom ParisTech, 1992b.
- M. R. Schroeder. Natural sounding artificial reverberation. Journal of the Audio Engineering Society, 10(3): 219–223, 1962.
- H. Bai, « Modèles hybrides pour la réverbération tardive », PhD thesis, Telecom ParisTech, 2016.
- H. Bai, G. Richard, and L. Daudet. Late reverberation synthesis: From radiance transfer to feedback delay networks. Audio, Speech, and Language Processing, IEEE/ACM Transactions on, 23(12):2260–2271, 2015b.