UNTREF UNIVERSIDAD NACIONAL DE TRES DE FEBRERO

Room Acoustic Parameters

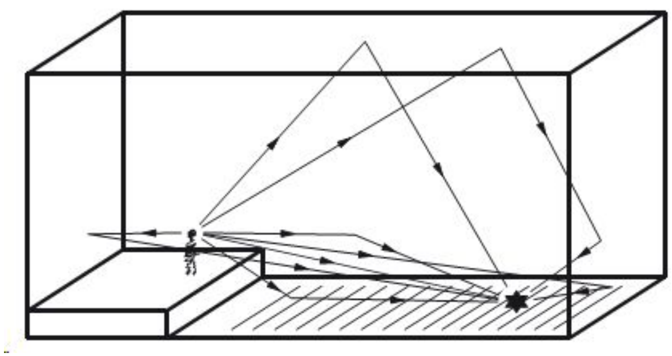
Acoustics and Psychoacoustics II (Module I)

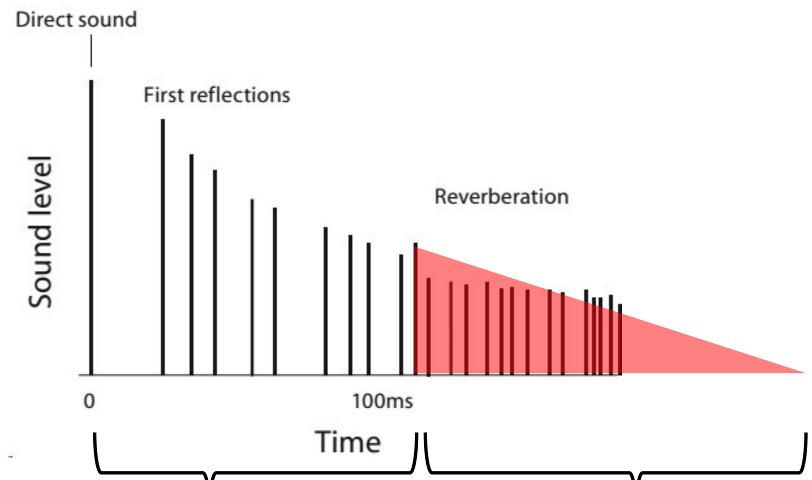
18 April 2022

SATO Shin-ichi (ssato@untref.edu.ar)

Acoustic simulation

- Sound ray-tracing to detect the reflection paths from sound source to receivers





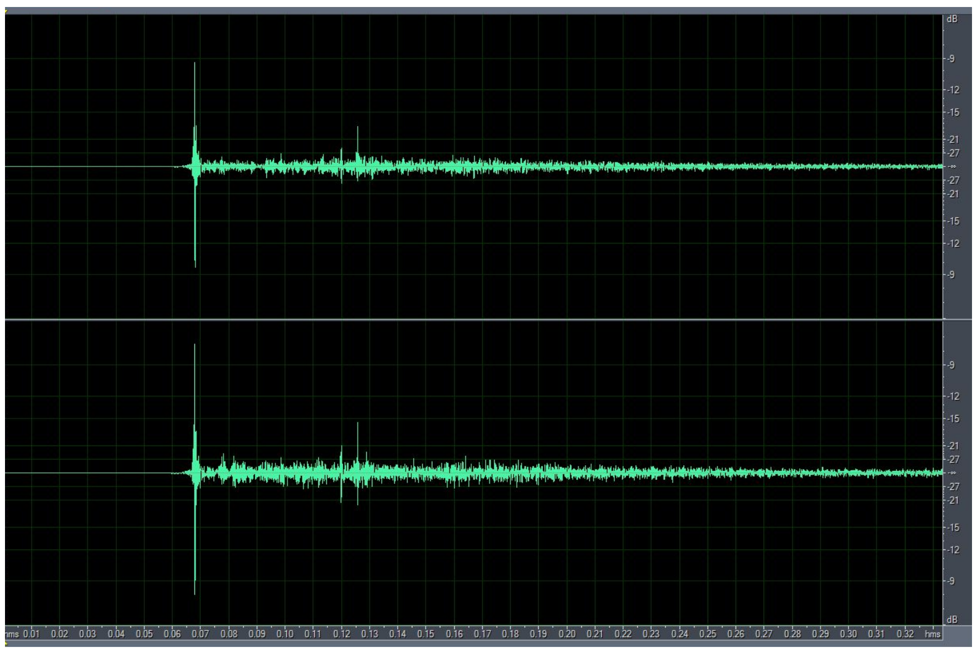
Hybrid method

Early reflections
Ray tracing
- Amplitude
- Delay time
- Direction

Late reverberation
Sabine equation

Acoustic measurement

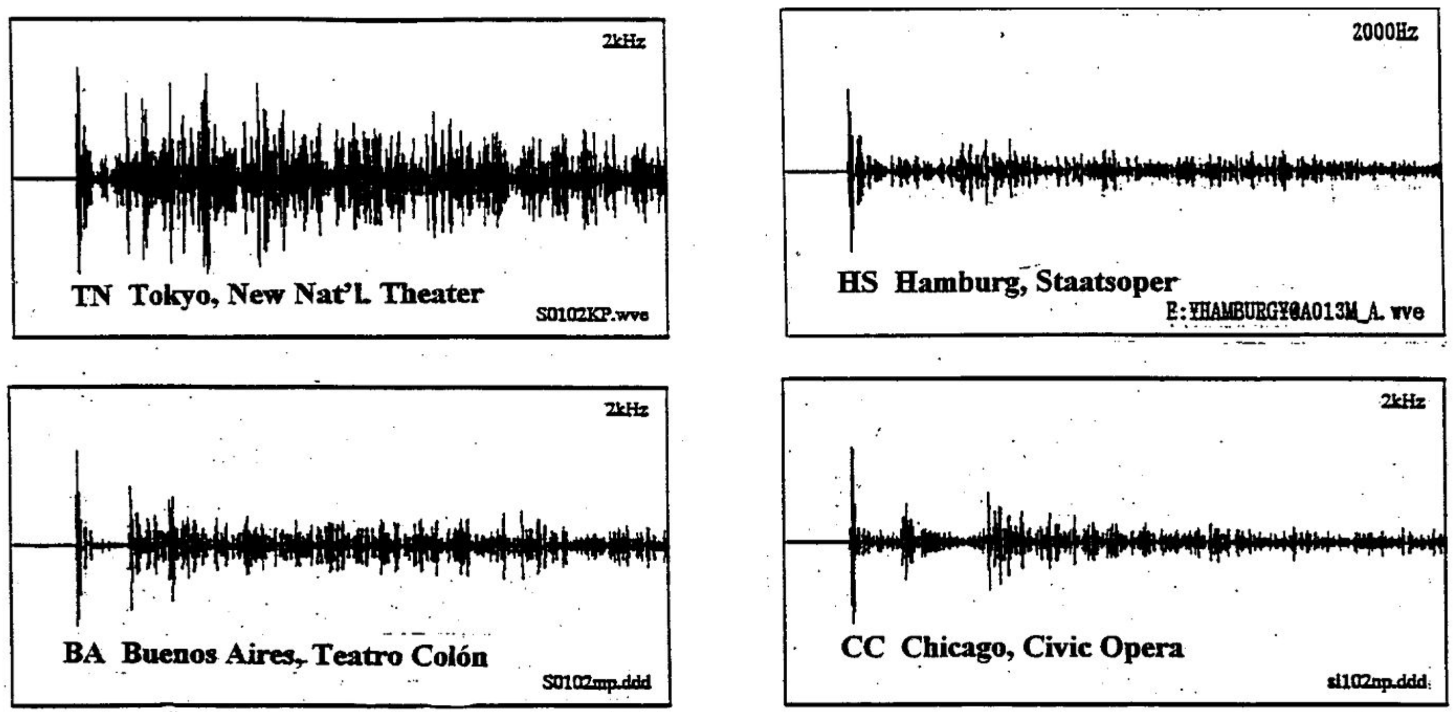
- (Binaural) Impulse response
  - Observation of (early) reflection pattern
  - Calculation of acoustic parameters



UNTREF
UNIVERSIDAD NACIONAL
DE TRES DE FEBRERO

Evaluation of acoustics

- Useful information from impulse response:
  - Initial time delay gap (ITDG)
  - (Early) reflection density
  - Long-delay echo



Evaluation of acoustics

- What is good acoustics for auditoriums?
  - Loudness
  - Clarity (intelligibility)
  - Reverberance
  - Spatial impression
  - Balance (spectral, performers and instruments)

  etc.

Loudness

- Subjective perception of the sound level/volume
- "**Strength of sound**" measured in dB, to formulate loudness
- We need a reference (level) because the level at each listener position depends on the sound power level of the source
- Acoustical parameter: "**Strength (G)**"

Strength (G)

- Definition

$$G = 10\log_{10}\left(\int_0^\infty p^2(t)dt \Big/ \int_0^\infty p_{ref}{}^2(t)dt\right)$$

$p(t)$: Impulse response at a receiver point

$p_{ref}(t)$: Reference impulse response, measured at **10 m** from the source in a **free field**

- 10 m is usually the closest source-receiver distance in a hall
- G value usually takes the positive value due to the reflections arriving at each listener position

Intelligibility/clarity: Early to late energy ratio

- Limited ability of our hearing to distinguish all the countless repeated sound signals
- Definition (Meyer and Thiele, 1956)

$$D = \left[\int_0^{50ms} p(t)^2 \, dt \Big/ \int_0^{\infty} p(t)^2 \, dt \right] \cdot 100 \quad [\%]$$

  - Reliable measure of speech (syllable) intelligibility
- Clarity (Reichardt et al., 1974)

$$C = 10 \log_{10} \left[\int_0^{80ms} p(t)^2 \, dt \Big/ \int_{80ms}^{\infty} p(t)^2 \, dt \right] \quad [\mathrm{dB}]$$

  - Transparency of music in a concert hall
- These parameters are calculated from room impulse response, not music signals

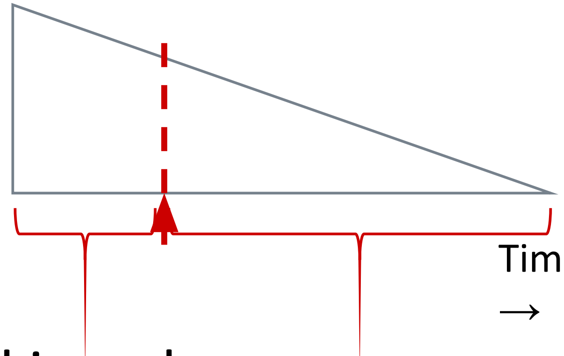
Clarity (of music)

- "Clarity" is the degree to which a listener can distinguish sounds in a musical performance
- Two forms of Clarity:
  - **Horizontal**: Related to played in succession
  - **Vertical**: Related to tones played simultaneously

- (Horizontal) clarity increases
  - as the reverberation time decreases
  - as the early to reverberant sound energy ratio increases

Center time

- In critical cases, there is unfavorable effect that a small change in the arrival time of a strong reflection may result in a significant change in C_{80}



- Center time (Kürer, 1969)

$$t_s = \int_0^\infty t \cdot p(t)^2 dt \,/ \int_0^\infty p(t)^2 dt$$

  - The first moment of the squared impulse response
  - No delay limit

Sound level in a room

- The sound level in a room is determined by:
  - Sound energy being generated
  - Acoustic characteristics of the room
    - Distance from the source
    - Total absorption (A)

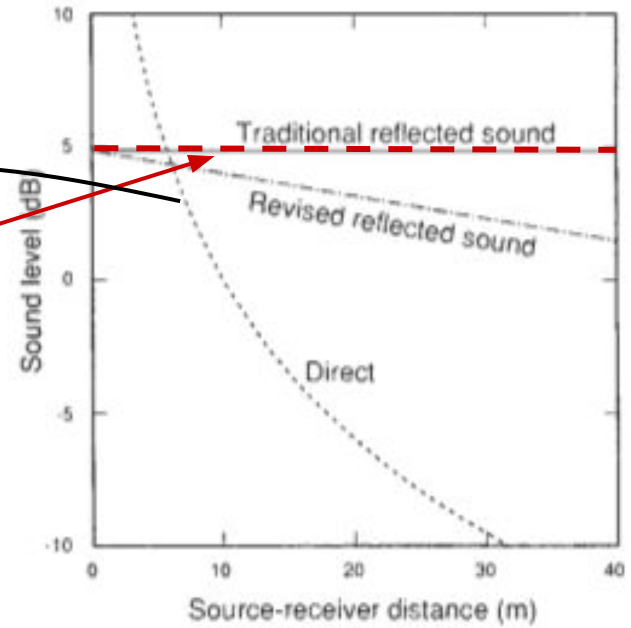
- The total sound is divided into:
  - The direct sound
  - The reflected component

Sound level in a room

- The direct sound decreases 6 dB for every doubling of distance (if the floor has appropriate slope)
- Traditionally, the reflected component was assumed constant throughout the space
(= the space is **totally diffusive**)

$$L = 10\log\left[\left(\frac{W}{4\pi r^2} + \frac{4W}{A}\right)\right]$$

W: Sound power level
r: Source-receiver distance
A: Total absorption

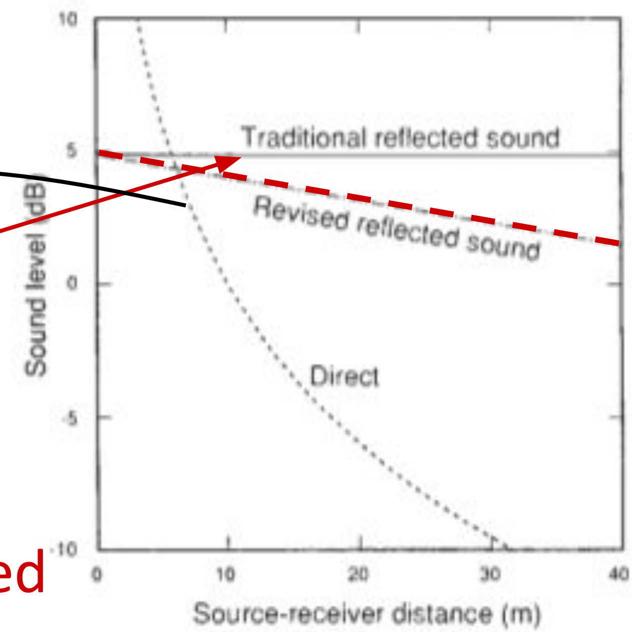


Barron’s revised theory

- However, the assumption of constant reflected sound level is not valid in actual auditoriums
- The reflected sound level decreases with distance from the source (→ Barron’s revised theory, 1988)

$$L = 10\log\left[\left(\frac{W}{4\pi r^2} + \frac{4W}{A}\right)\right]$$

This term
should be revised



Barron's revised theory

- Sound level at 10 m from the source (reference)

$$L_0 = 10\log(W/400\pi)$$

- Total sound pressure level

$$L - L_0 = 10\log(100/r^2 + 31200T/V)$$

- Direct sound

$$d = 100/r^2$$

- Early reflections

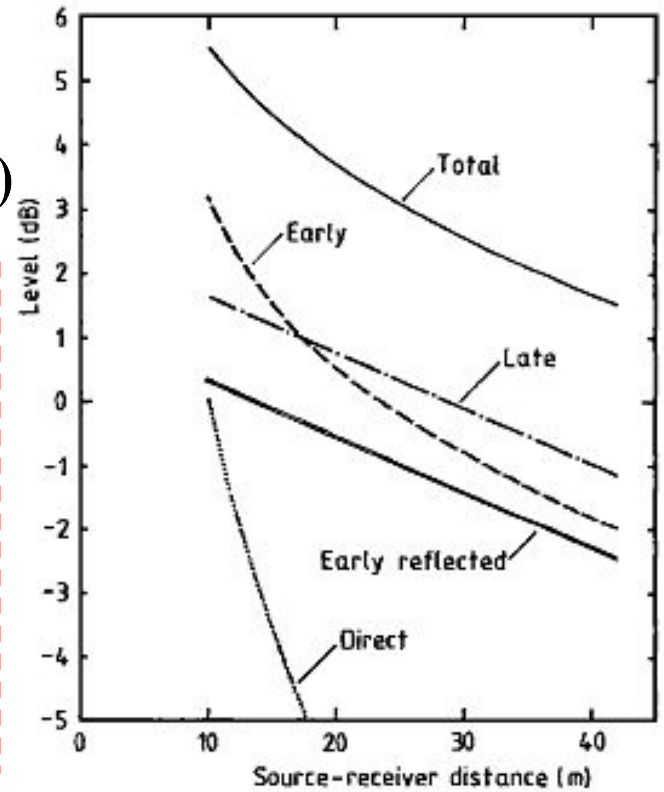
$$e = (31200T/V)e^{-0.04r/T}(1 - e^{-1.11/T})$$

- Late reflections

$$l = (31200T/V)e^{-0.04r/T}e^{-1.11/T}$$

- Total sound level (G value)

$$L - L_0 = 10\log(d + e + l)$$



Barron’s revised theory

- The revised theory can provide a value for what is usual
- Some design features cause unusually quiet or loud sound
  - **Lower** strength
    - Lack of early reflections
    - Excess absorptions
  - **Higher** strength (at particular positions)
    - Focusing
    - Strong specular (non-diffused) reflections

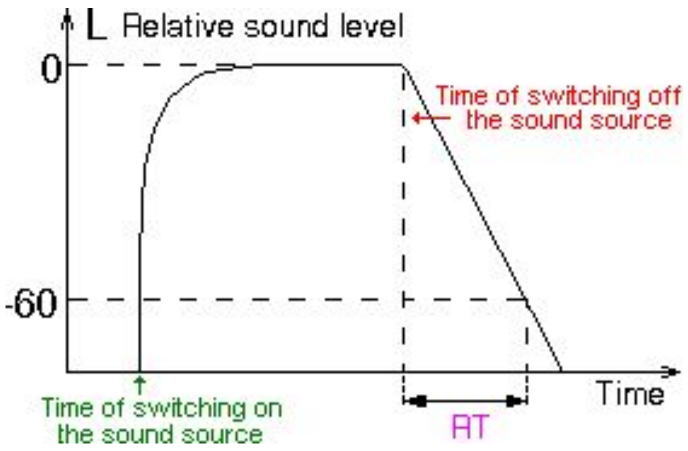
Reverberation time

- Definition of reverberation time T_{60}
  - Time required for the sound pressure level to decrease by 60 dB from its initial level once the source is shut off
  - Sabine's equation

$$T_{60} = 0.161\frac{V}{A}$$

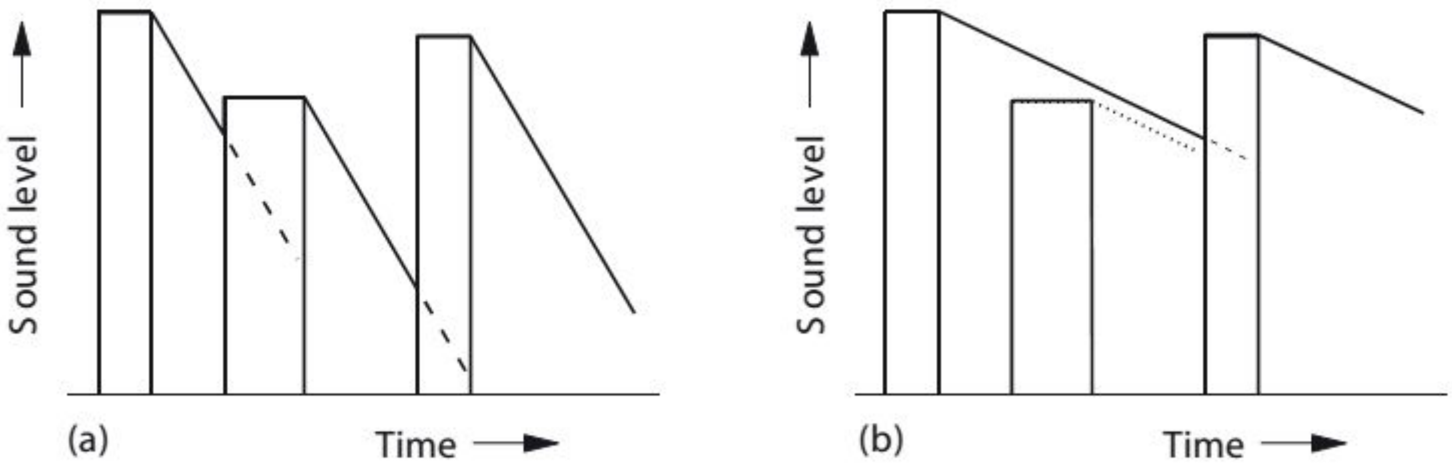
V: Room volume [m^3]

A: Total absorption area [m^2]



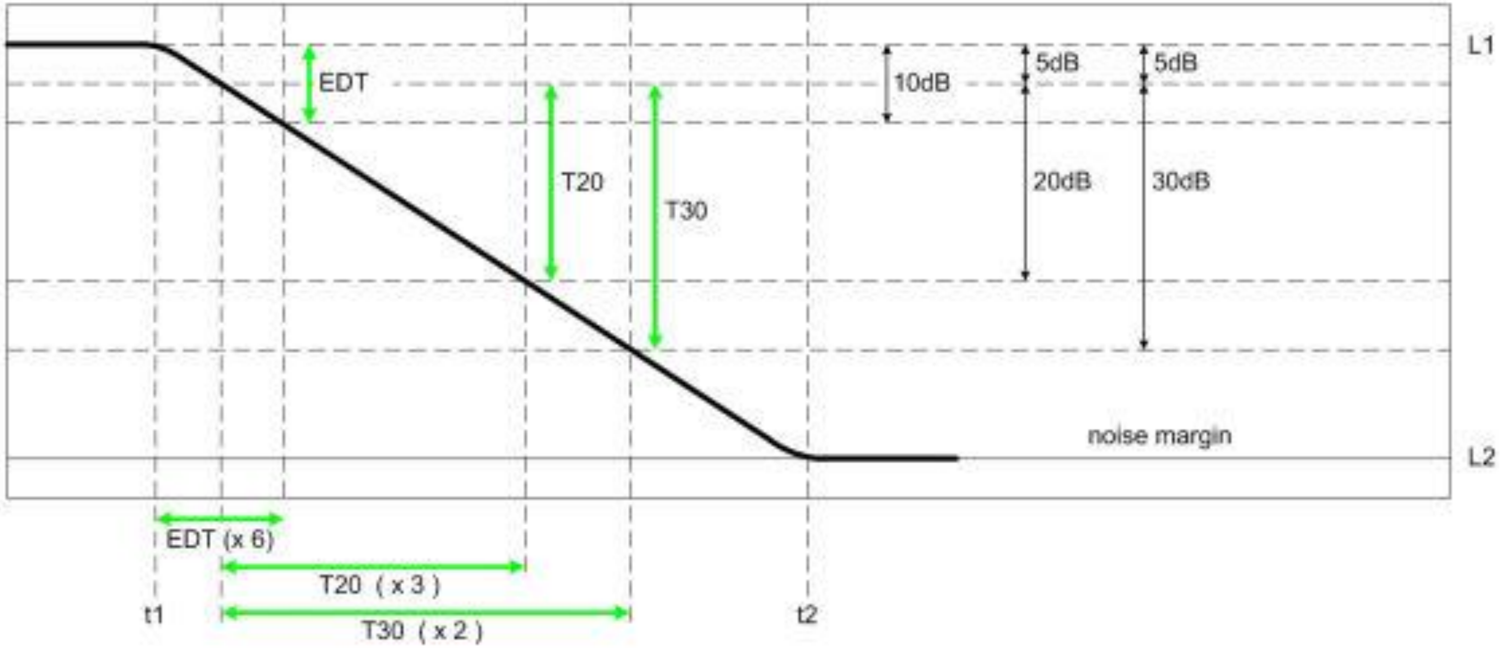
Early decay time (EDT)

- When a musician or ensemble plays rapidly, only the early part of the sound decay process remains audible between successive notes.



Early decay time (EDT)

- Early decay time (EDT) designates that initial part of sound decay
- EDT is the time in which the first 10 dB fall of a decay process occurs, multiplied by a factor 6. It allows a direct comparison between EDT and RT.



Early decay time (EDT)

- EDT is mostly **shorter** than the reverberation time
- EDT corresponds more to the perceived **reverberance** than RT
- **RT** shows **less variations** with room shape because the decay process as a whole is made up of numerous reflections with different delays, strengths and wall portions where they originated
- On the contrary, **EDT** is strongly influenced by early reflections, and therefore **depends on the position**; furthermore, it is sensitive to details of the room's geometry.

Early decay time (EDT)

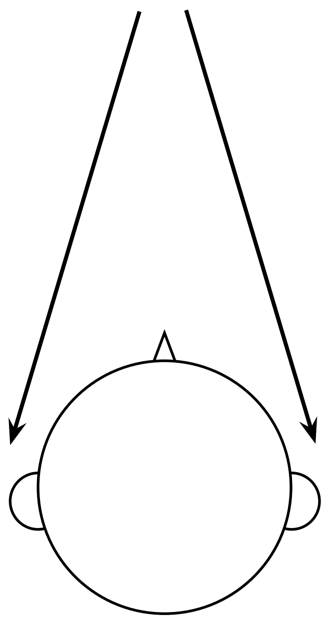
- Cliff-type decay curve (EDT < RT)
  - The direct sound or the early reflections are very strong

- Plateau-type decay curve (EDT > RT)
  - Lack of early reflections

- Sagging decay curves (EDT < RT)
  - Two-dimensional reverberation by absorption concentrated on the floor and with vertical walls
  - **Diffusion** on the side walls is needed for a linear decay curve

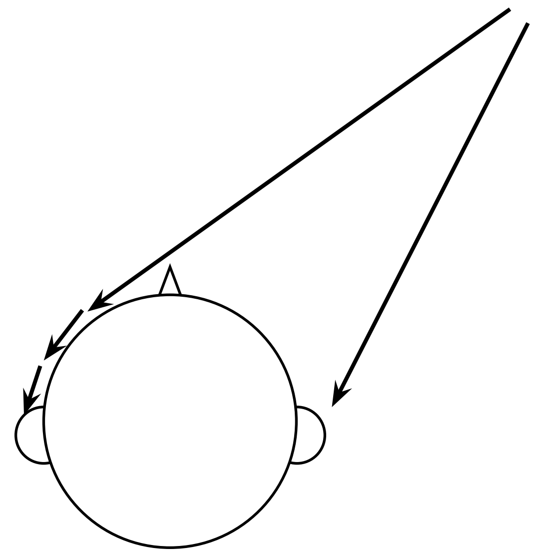
Lateral reflection and stereo effect

- Reflection from front, above, and back



Left and right ear signals are
(almost) identical

- Reflection from sides



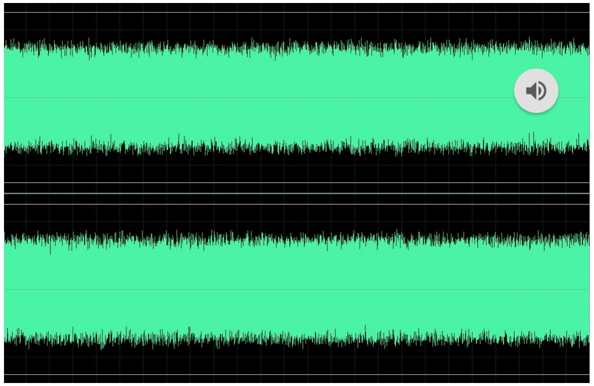
Dissimilarity of sounds between left and right ears signals
Interaural difference of:
Delay time/Sound level/Spectrum

Controlling IACC

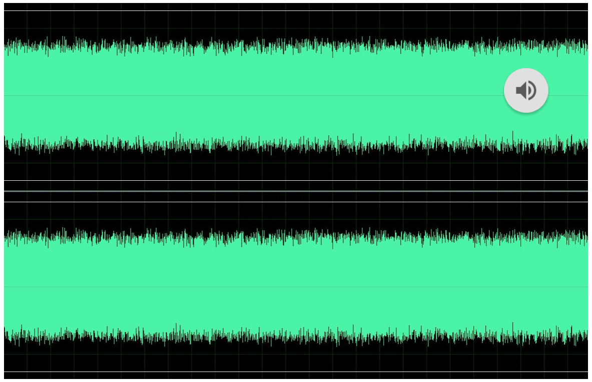
- Gabriel and Colburn (1981)
  - Uncorrelated noise n_1 and n_2 and correlation ρ

$$n_L(t) = n_1(t),$$

$$n_R(t) = \rho n_1(t) + (1-\rho^2)^{1/2} n_2(t).$$



0.5 0.8

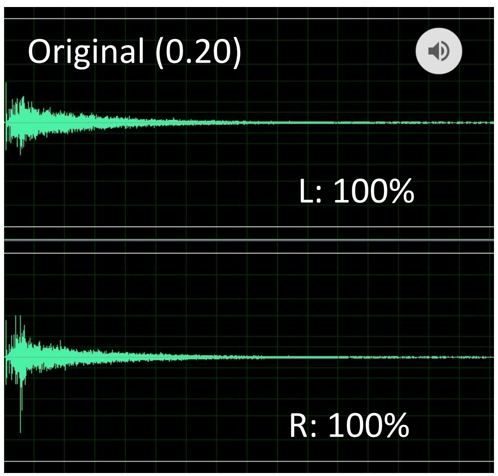


Independently generated broadband noises
→ Correlation is 0.0

The same broadband noise to both signals
→ Correlation is 1.0

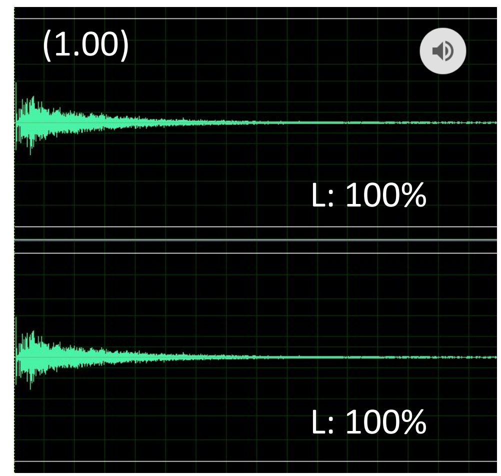
Controlling IACC

- Impulse response:
  - Left channel was mixed with the right channel at different percentages

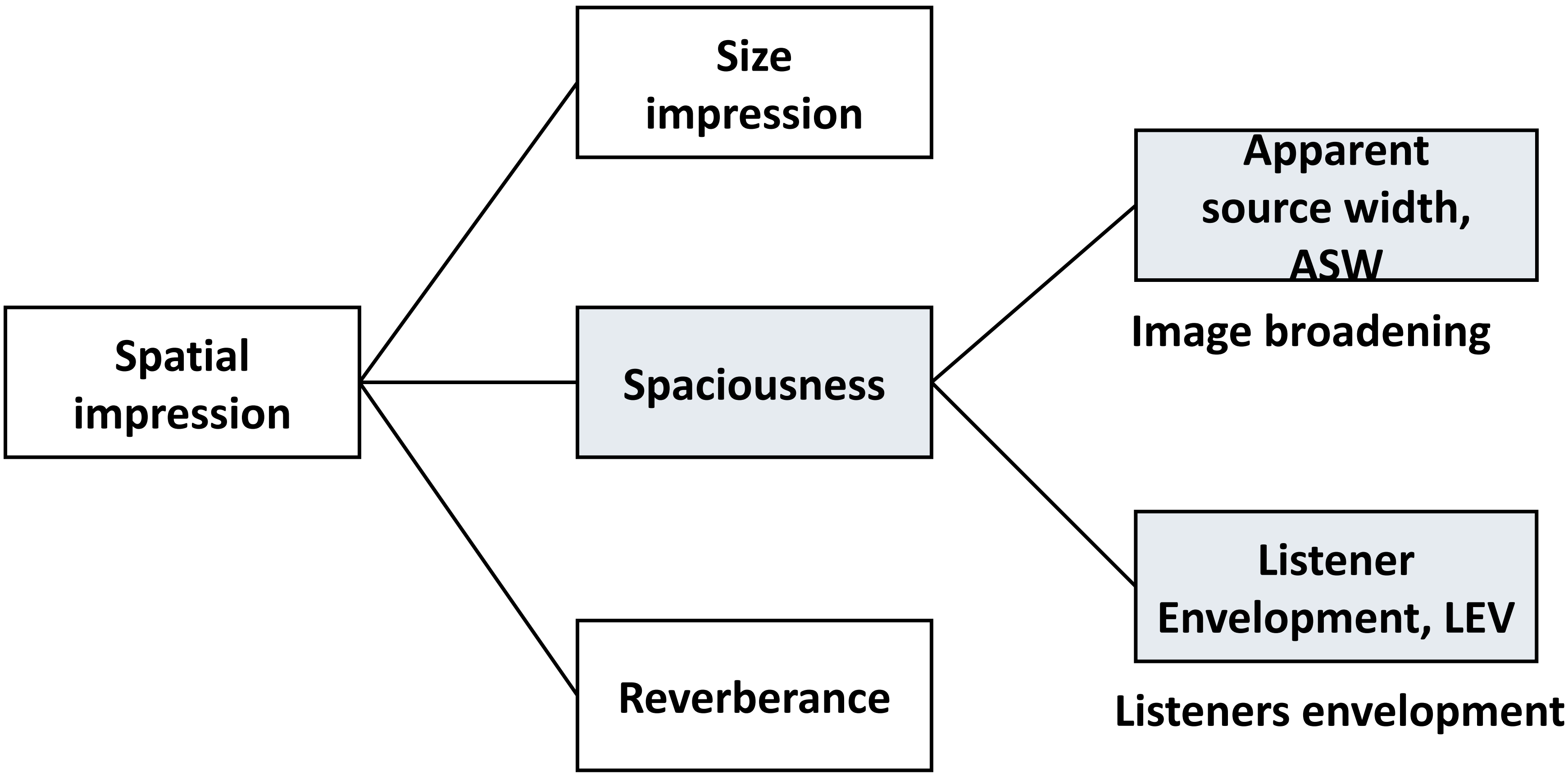


0.5

0.7



Spatial impression (in auditorium)



Two components of spatial impression

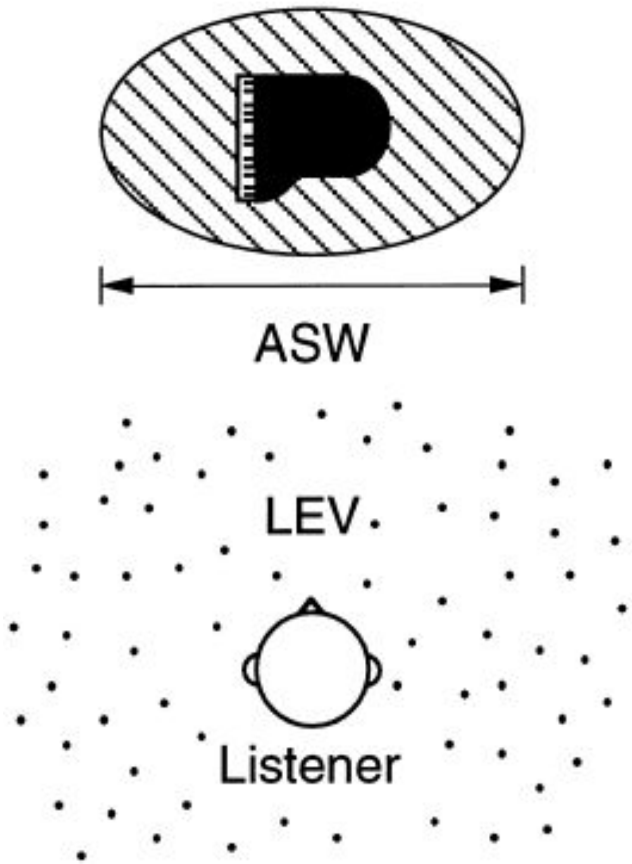
- Many terms had been used:

  Binaural similarity, Interaural coherence, Spatial impression, Subjective diffuseness, Broadening, Auditory spaciousness, etc.

- **Two components of spatial impression**

  (Morimoto & Maekawa, 1989)

  - Apparent source width (ASW)
  - Listener envelopment (LEV)



Two components of spatial impression

- **Apparent source width (ASW)**
  - Width of a sound image fused temporally and spatially with the direct sound
  - Listener perceives the sound substantially wider than the actual width of the source
  - (Early) lateral reflections from sidewalls and side-balcony fronts

Two components of spatial impression

- **Listener envelopment (LEV)**
  - Degree of fullness of sound images around the listener
  - Degree to which the reverberant sound seems to surround the listener—to come from all directions
  - Reverberation from the entire upper-hall space—above, ahead, and behind the listener
  - Surface irregularities and ornamentation help spread the reflections to various directions (= diffusion)
  - Listener seated under a deep balcony perceives the reverberation as coming only from the front

Investigation on good acoustics in concert halls

1960's

The importance of the spatial attributes in terms of the shape (cross-section) of the concert halls

- **West** (1966): 2H/W correlates to the subjective evaluation of halls (H: height; W: width of a hall)
- **Marshall** (1967): Importance of room cross-section ratio (W/H ratio) in concert halls
- **Marshall** (1968): The reflection from the side walls should arrive faster than that from the ceiling
  - Importance of **lateral reflection**
- Possible objective measure
  - Interaural cross-correlation and lateral fraction

Interaural cross-correlation function

- Interaural cross-correlation function between the binaural impulse responses at both ears $h_l(t)$ and $h_r(t)$ is defined by:

$$\Phi_{lr}(\tau) = \int_{-\infty}^{\infty} h_l(t) h_r(t+\tau) dt$$

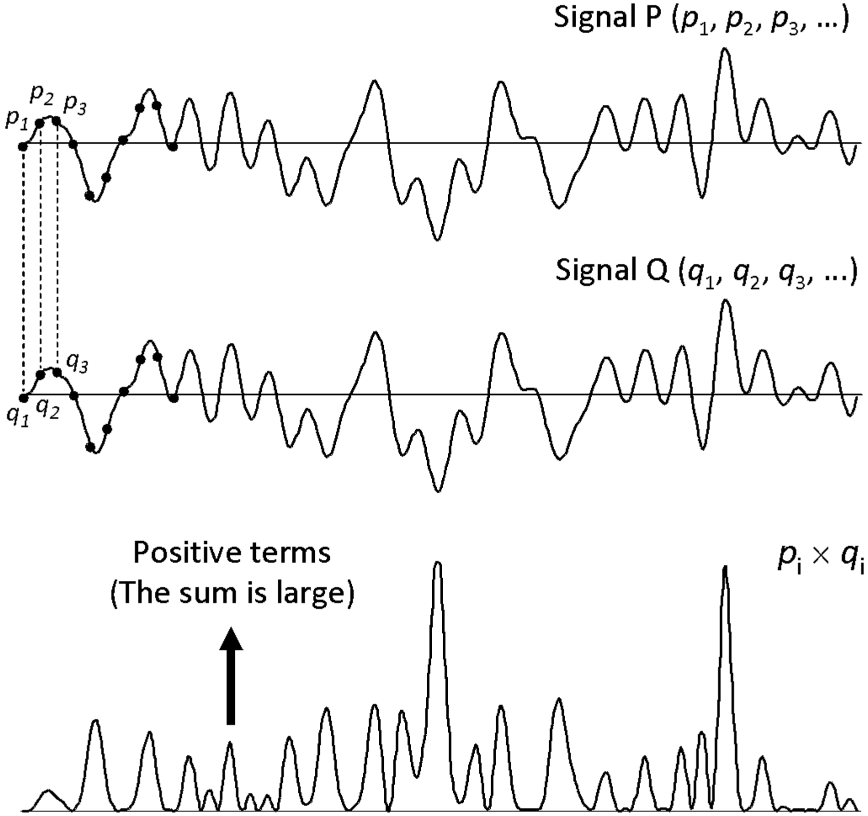
- Similarity between two waveforms is measured by multiplying them together, ordinate by ordinate, and to add the products over the duration of the waveforms
- To assess the similarity between waveforms $P(p_1, p_2, p_3, \ldots)$ and $Q(q_1, q_2, q_3, \ldots)$, we multiply ordinate p_1 by ordinate q_1, p_2 by q_2, p_3 by q_3, …, and then we add these products to obtain a single number, to measure the similarity

Basic propositions

Waveforms P and Q (identical)

- Every ordinate, positive or negative, contributes a positive term to the sum
- The sum is large (similar)

$p_1 * q_1 + p_2 * q_2 + p_3 * q_3 + \ldots$

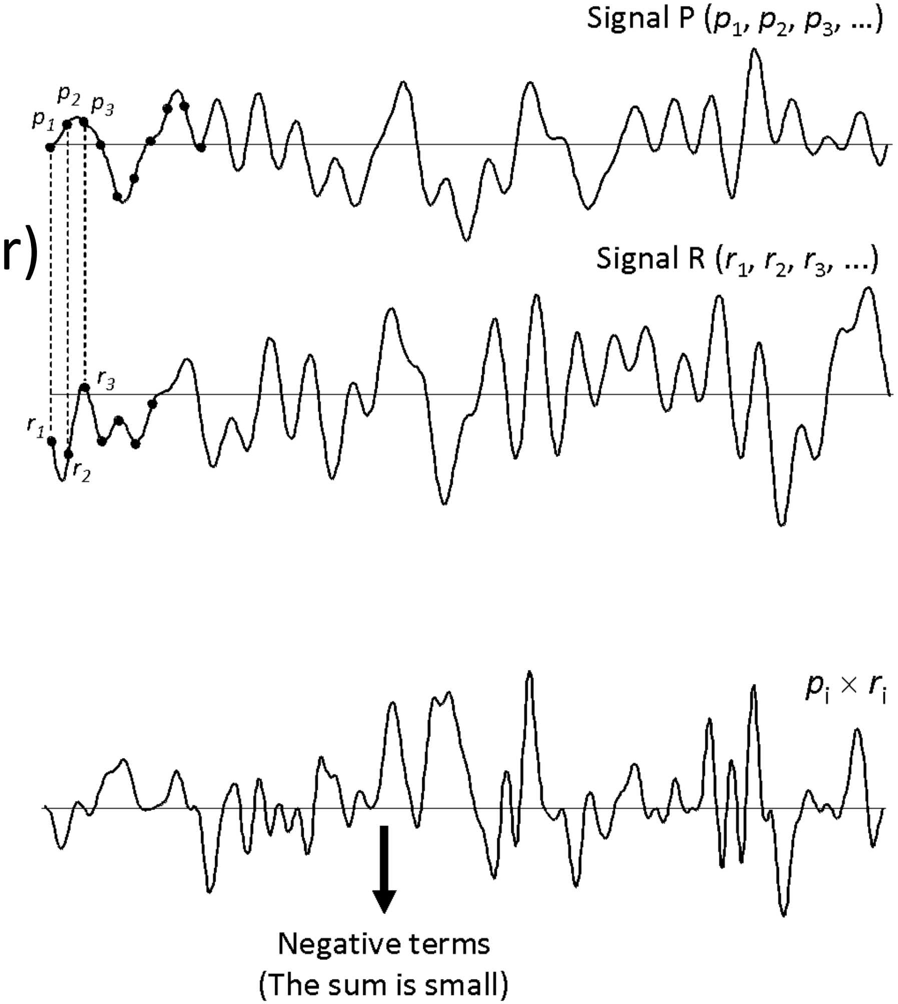


Basic propositions

Waveforms P and R (different)

- Each positive is offset by another negative product
- The sum is small (dissimilar)

$$p_1 * r_1 + p_2 * r_2 + p_3 * r_3 + \dots$$



Interaural cross-correlation function

UNTREF
UNIVERSIDAD NACIONAL
DE TRES DE FEBRERO

- Interaural cross-correlation function between the binaural impulse responses at both ears $h_l(t)$ and $h_r(t)$ is defined by:

$$\Phi_{lr}(\tau) = \int_{-\infty}^{\infty} h_l(t) h_r(t+\tau) dt$$

- The normalized interaural cross-correlation function is defined by

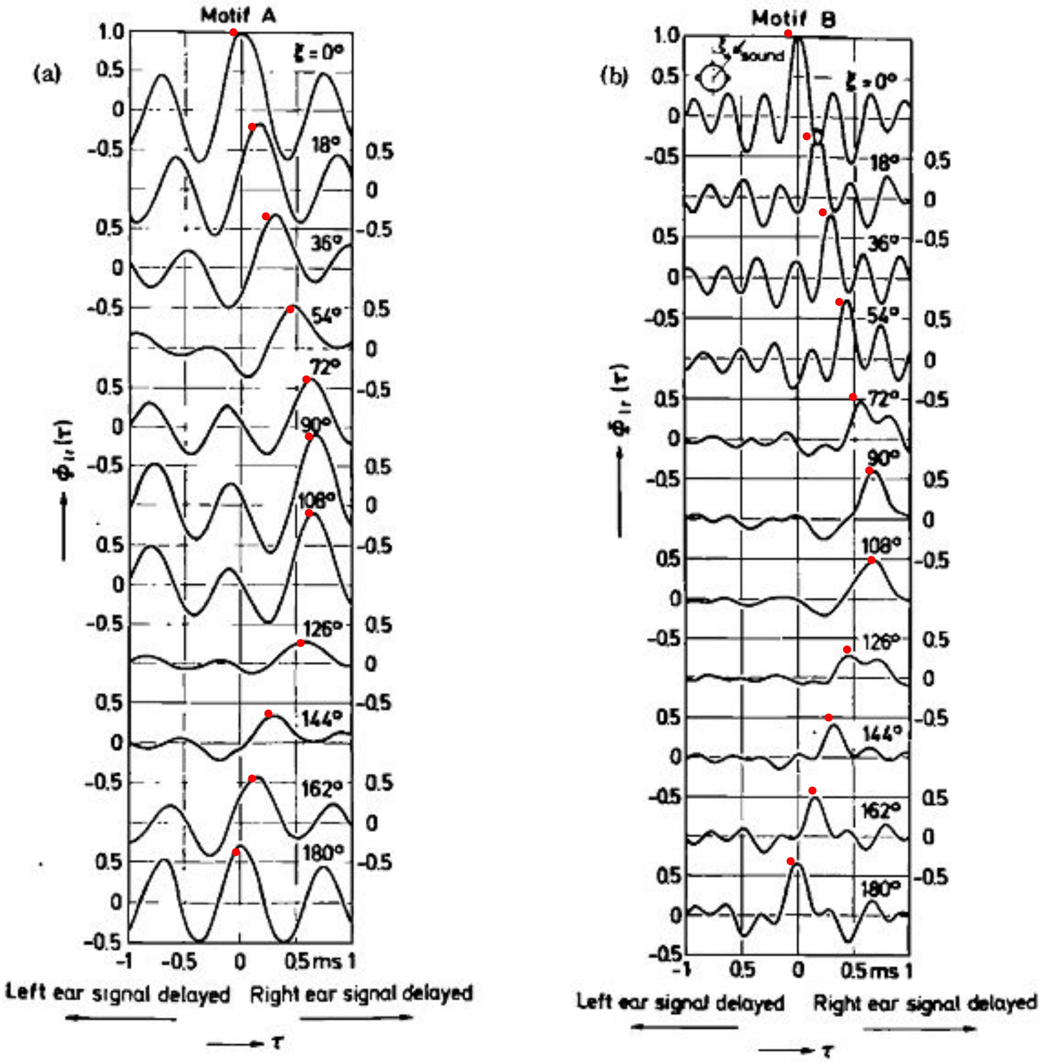
$$\varphi_{lr}(\tau) = \frac{\Phi_{lr}(\tau)}{\sqrt{\Phi_{ll}(0)\Phi_{rr}(0)}}$$

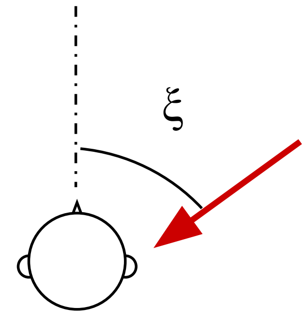
where $\Phi_{ll}(0)$ and $\Phi_{rr}(0)$ are the sound energies at both eras

- Interaural cross-correlation coefficient (IACC) is defined by

$$\mathrm{IACC} = \left|\varphi_{lr}(\tau)\right|_{\max}, \quad |\tau| \le 1ms$$

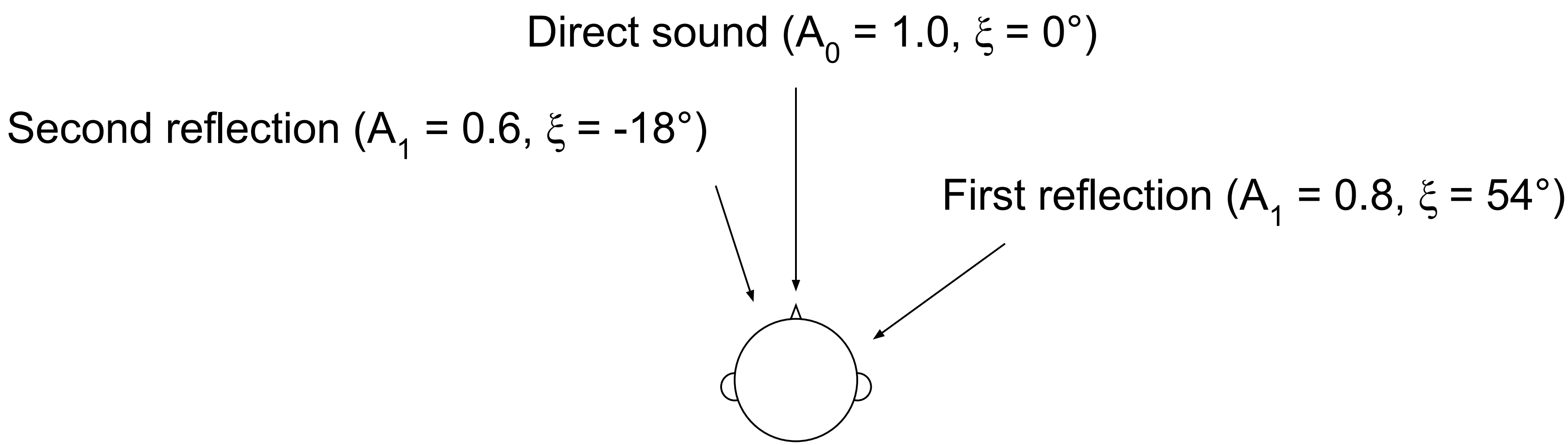
Interaural cross-correlation function of music signal





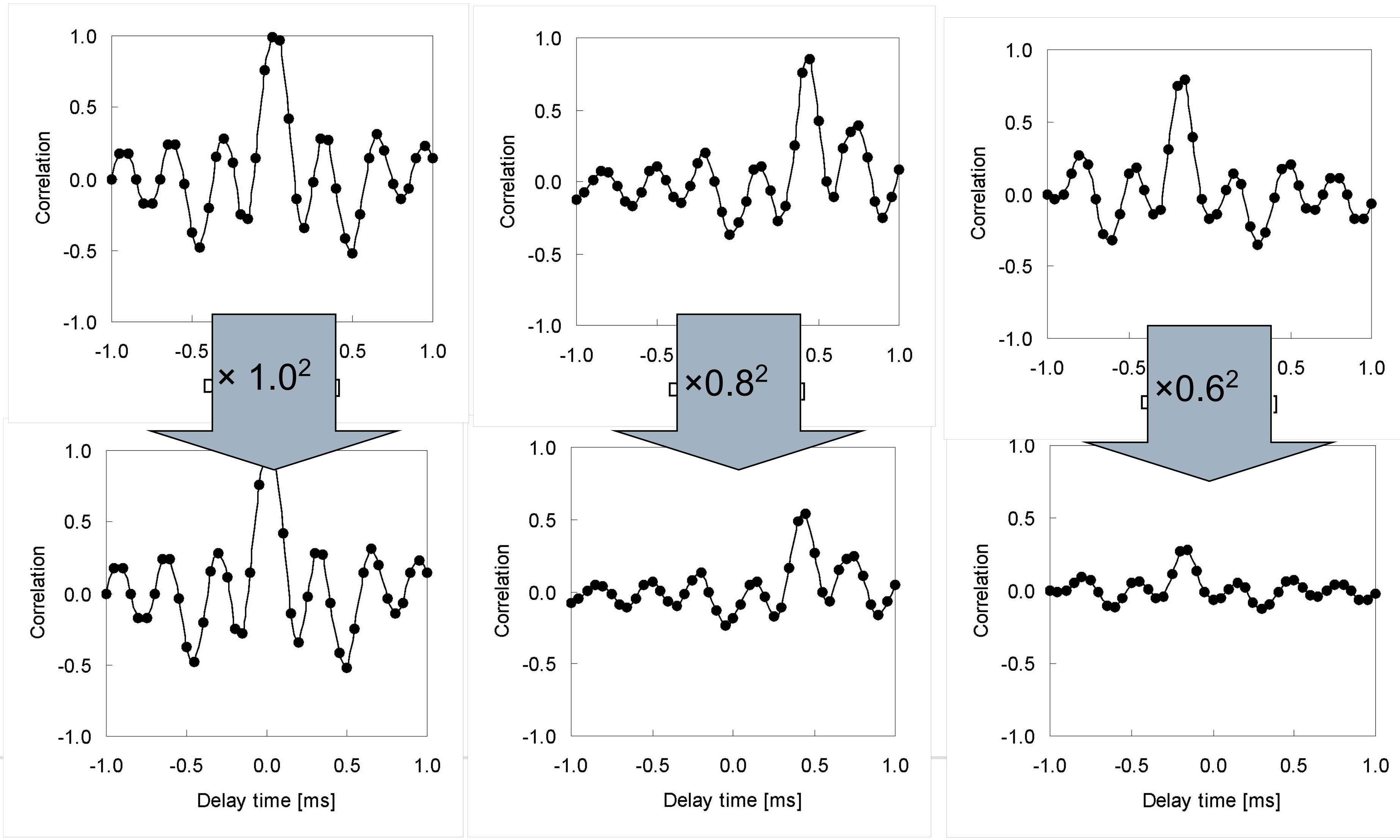
IACF depends on the frequency contents of source signal

Calculation of IACF (Example)



Calculation of IACF (Example)

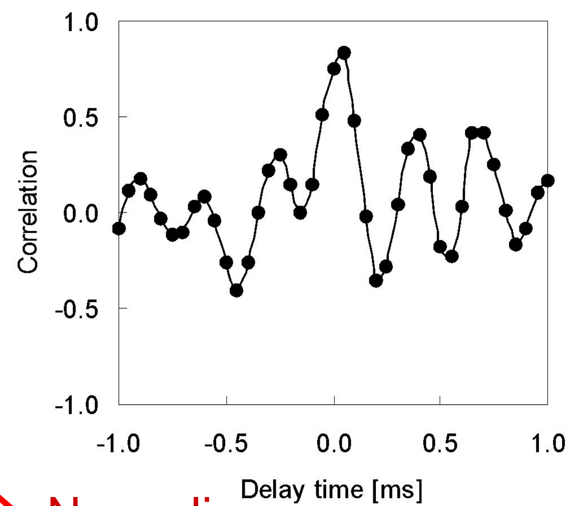
Direct sound (ξ = 0°) First reflection (ξ = 54°) Second reflection (ξ = -18°)



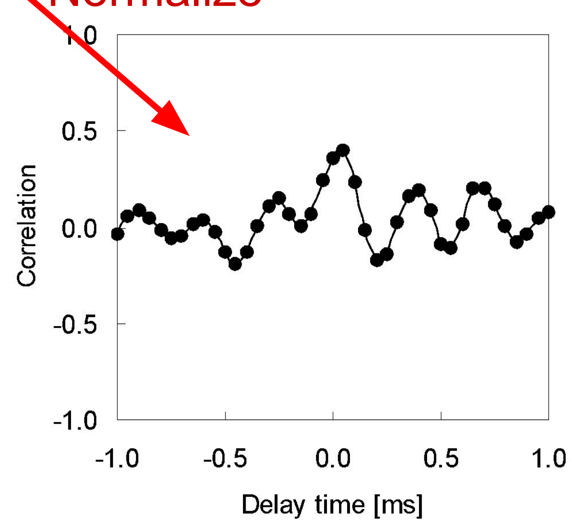
Calculation of IACF (Example)

Summation of the function for each sound

$$\phi_{lr}(\tau) \approx \frac{\sum_{n=0}^{N} A_n{}^2 \Phi_{lr}(\tau)}{\sqrt{\sum_{n=0}^{N} A_n{}^2 \Phi_{ll}(0) \sum_{n=0}^{N} A_n{}^2 \Phi_{rr}(0)}}$$

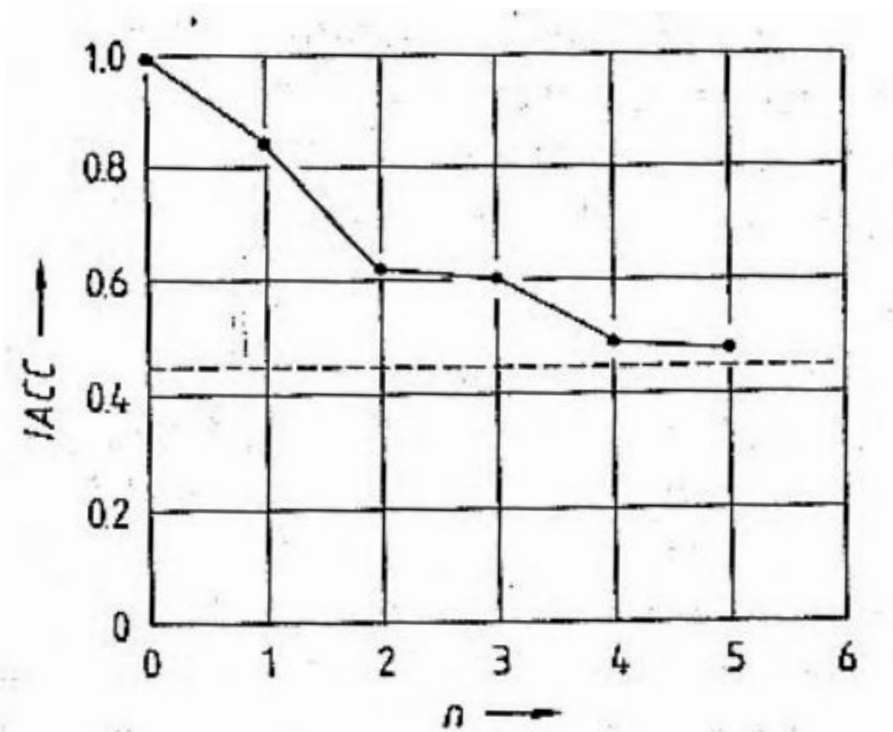


Normalize



Calculation of IACF (Example)

- The first 4 or 5 reflections determine the IACC value (Ando, 1983)



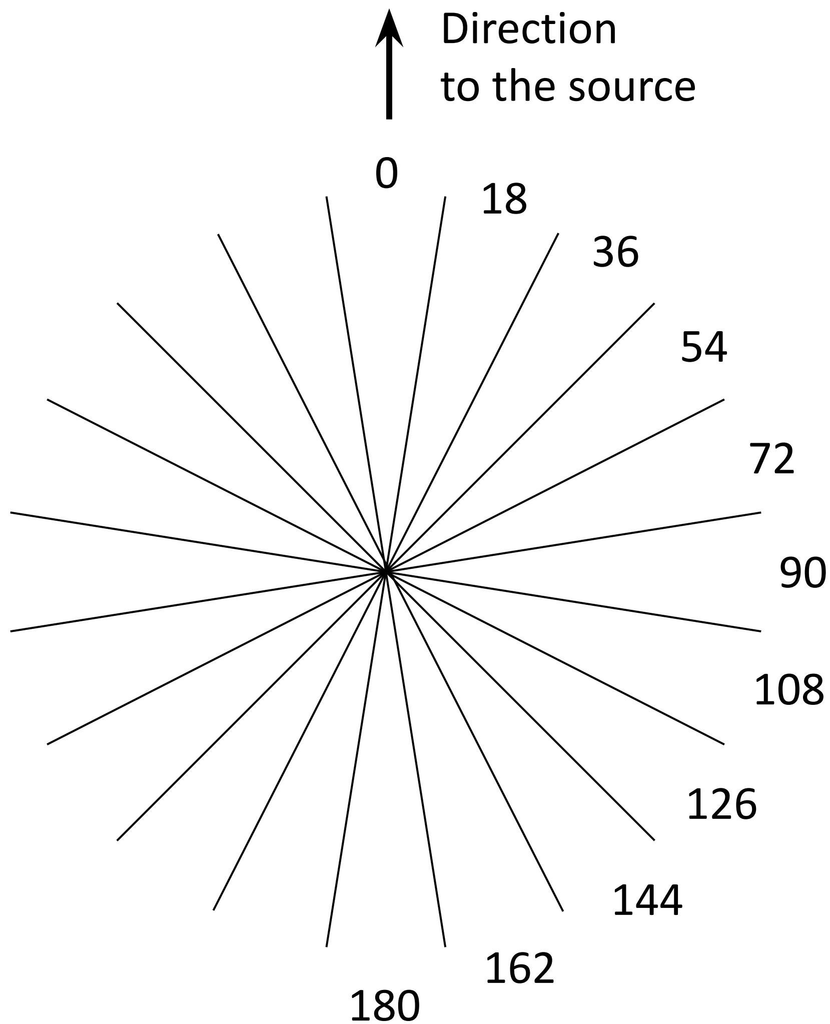
Number of reflection

IACC data (18° step)

UNTREF
UNIVERSIDAD NACIONAL
DE TRES DE FEBRERO

	Gibbons			Arnold			Hyden			Wagner			Mozart		
	Music A			**Music B**			**Music C**			**Music D**			**Music E**		
	Φ_{lr}	Φ_{ll}	Φ_{rr}	Φ_{lr}	Φ_{ll}	Φ_{rr}	Φ_{lr}	Φ_{ll}	Φ_{rr}	Φ_{lr}	Φ_{ll}	Φ_{rr}	Φ_{lr}	Φ_{ll}	Φ_{rr}
0	0.99	1.00	1.00	0.99	1.00	1.00	0.97	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00
18	0.30	0.71	1.12	-0.17	0.54	1.38	0.18	0.66	1.17	-0.11	0.65	1.40	0.21	0.73	1.31
36	-0.32	0.42	1.31	0.18	0.39	1.73	-0.06	0.42	1.34	0.02	0.46	1.66	-0.19	0.31	1.55
54	-0.32	0.32	1.42	-0.28	0.35	2.06	-0.28	0.40	1.50	-0.23	0.20	1.66	-0.38	0.39	2.61
72	0.09	0.34	1.27	-0.04	0.28	1.42	-0.16	0.42	1.39	0.06	0.28	1.56	-0.42	0.44	2.51
90	0.13	0.65	1.51	0.04	0.34	1.25	-0.14	0.60	1.39	0.15	0.36	1.38	-0.34	0.60	2.44
108	0.00	0.62	1.51	0.00	0.30	1.13	-0.16	0.57	1.30	0.10	0.28	1.14	-0.37	0.47	2.30
126	-0.07	0.19	0.75	-0.06	0.23	0.87	-0.12	0.26	0.87	-0.08	0.22	0.97	-0.46	0.35	1.90
144	-0.09	0.24	0.84	0.03	0.27	0.92	-0.04	0.30	0.85	-0.06	0.23	0.87	-0.26	0.28	1.28
162	0.30	0.52	0.75	0.00	0.39	0.75	0.22	0.45	0.74	-0.06	0.44	0.80	0.18	0.63	1.13
180	0.69	0.69	0.71	0.63	0.66	0.66	0.61	0.72	0.58	0.88	0.90	0.96	0.90	0.87	0.97

Approximate direction of sound incidence



3D sound

- For arbitrary angles of 3D incidence (horizontal ξ and vertical η), the angle given by

$$\sin^{-1}(\sin\xi\cos\eta)$$

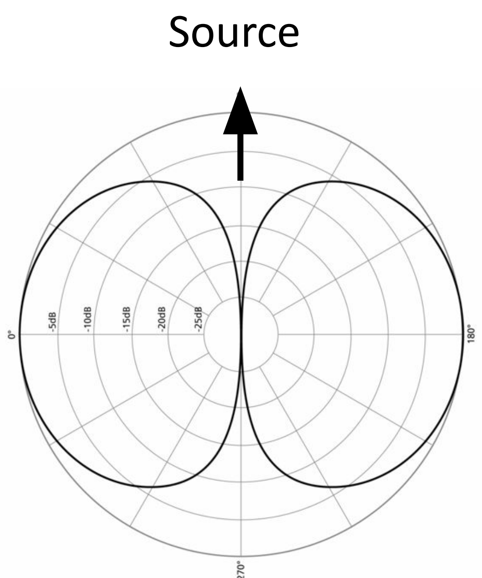
is substituted for the horizontal angle ξ

Lateral energy fraction

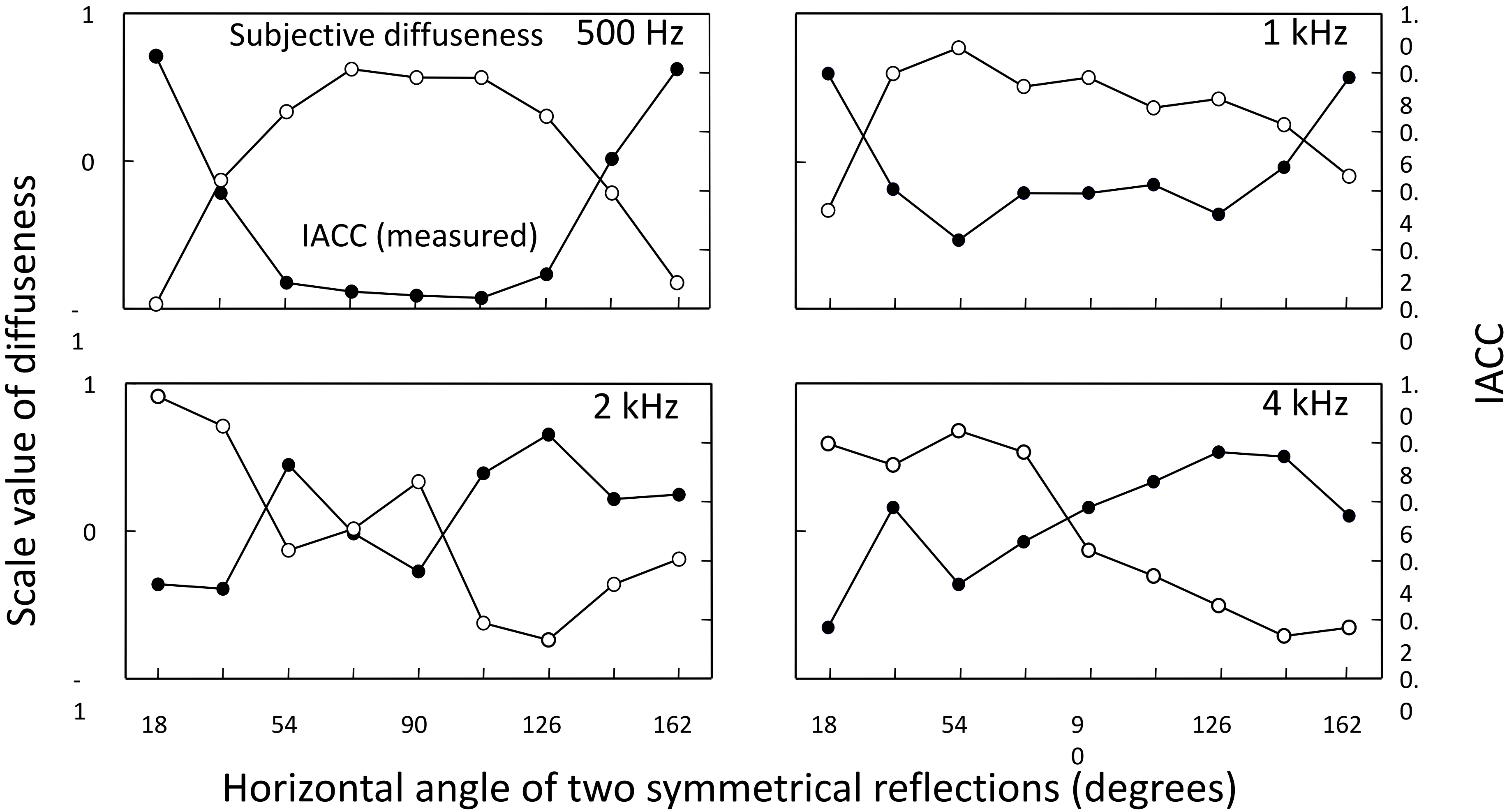
- Barron (1971)
  - The degree of spatial impression is related to the ratio of lateral (90°) to non-lateral sound arriving within 80 ms
- Barron and Marshall (1981)
  - Early lateral Energy Fraction

$$J_{LF} = \int_{0.005}^{0.08} p_L^{\,2}(t)dt \Big/ \int_{0}^{0.08} p^2(t)dt$$

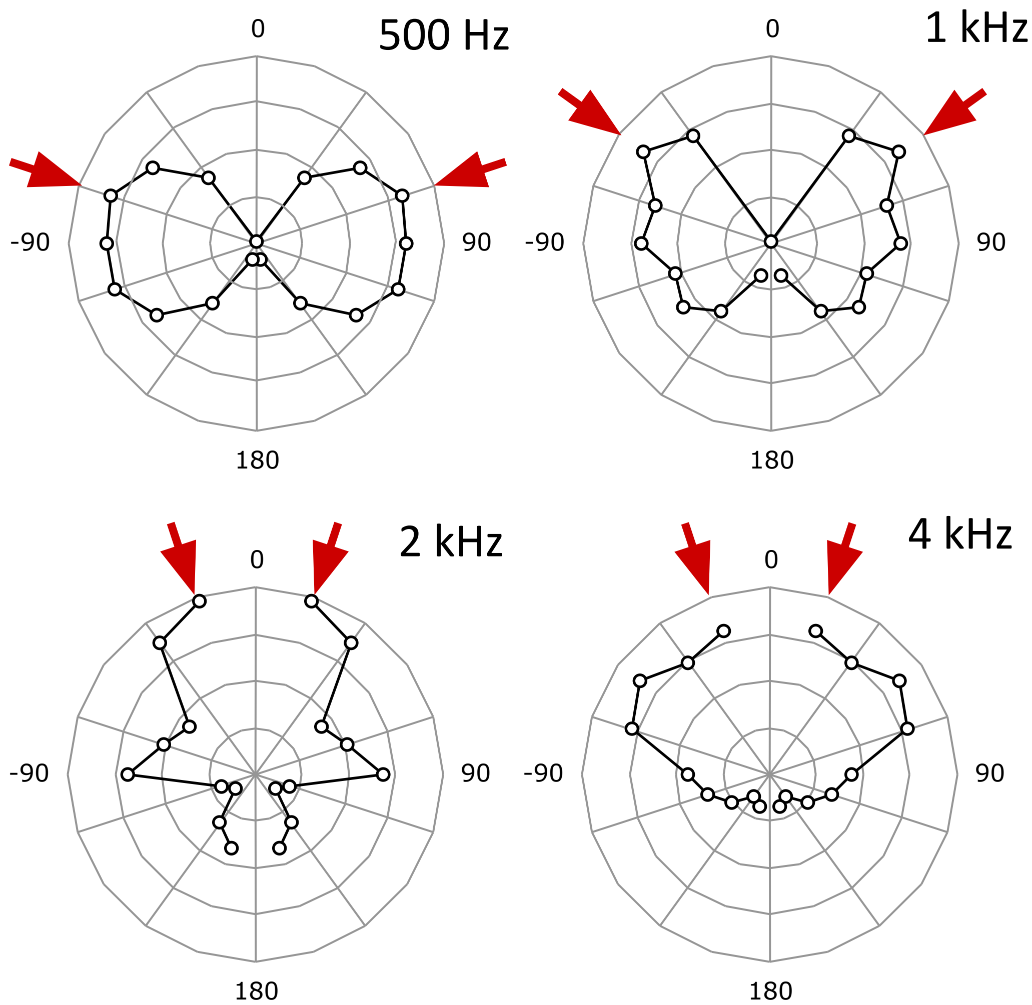
$p_L(t)$: impulse response measured with a figure-of-eight microphone
$p(t)$: impulse response measured with an omnidirectional microphone



Effective direction for a lower IACC (Bandpass noises)

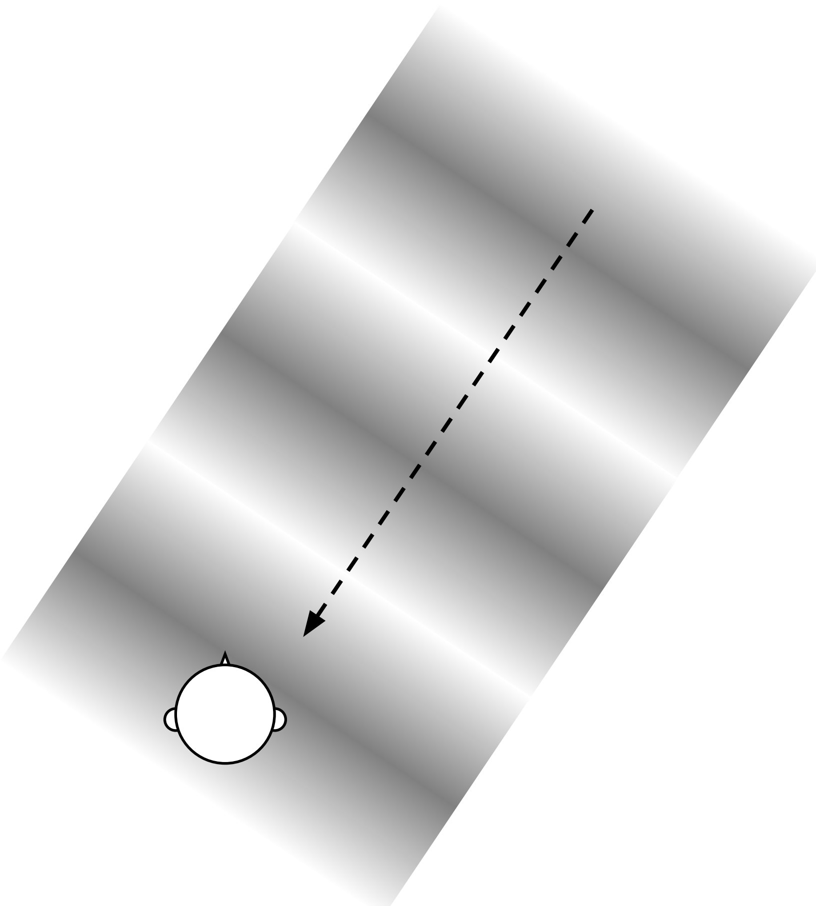


Effective direction for a lower IACC (Bandpass noises)



The most effective horizontal angles of reflections depend on the frequency of the source signal

Frequency dependence of IACC

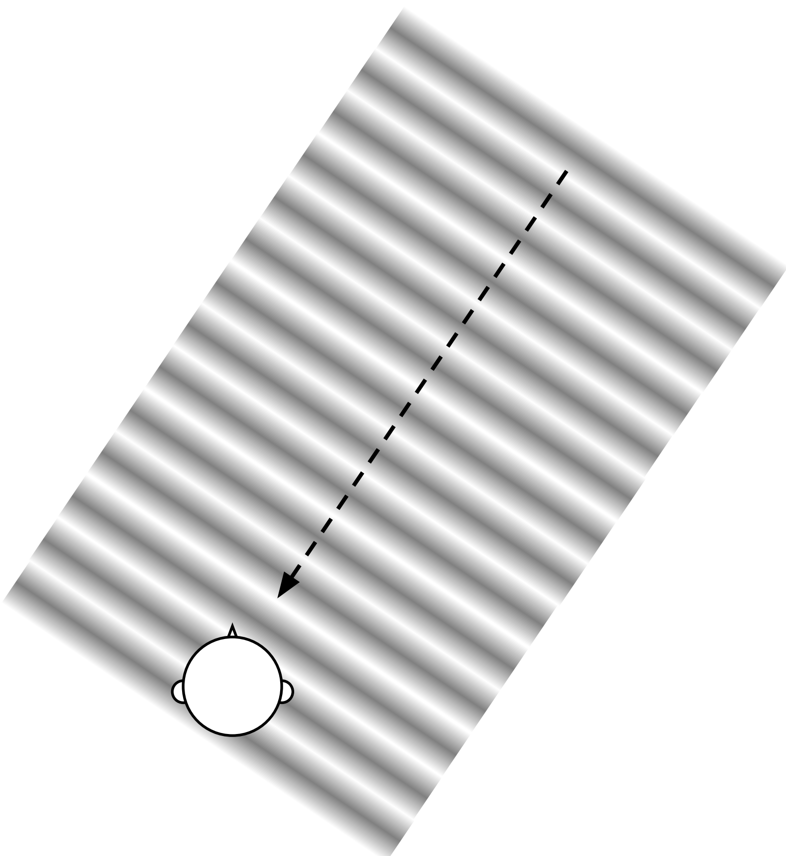


Lower frequency signal

Wavelength is larger than the dimension of the head. There is little phase difference between the two ears.

90° sound incidence maximizes the phase difference between the ears

Frequency dependence of IACC

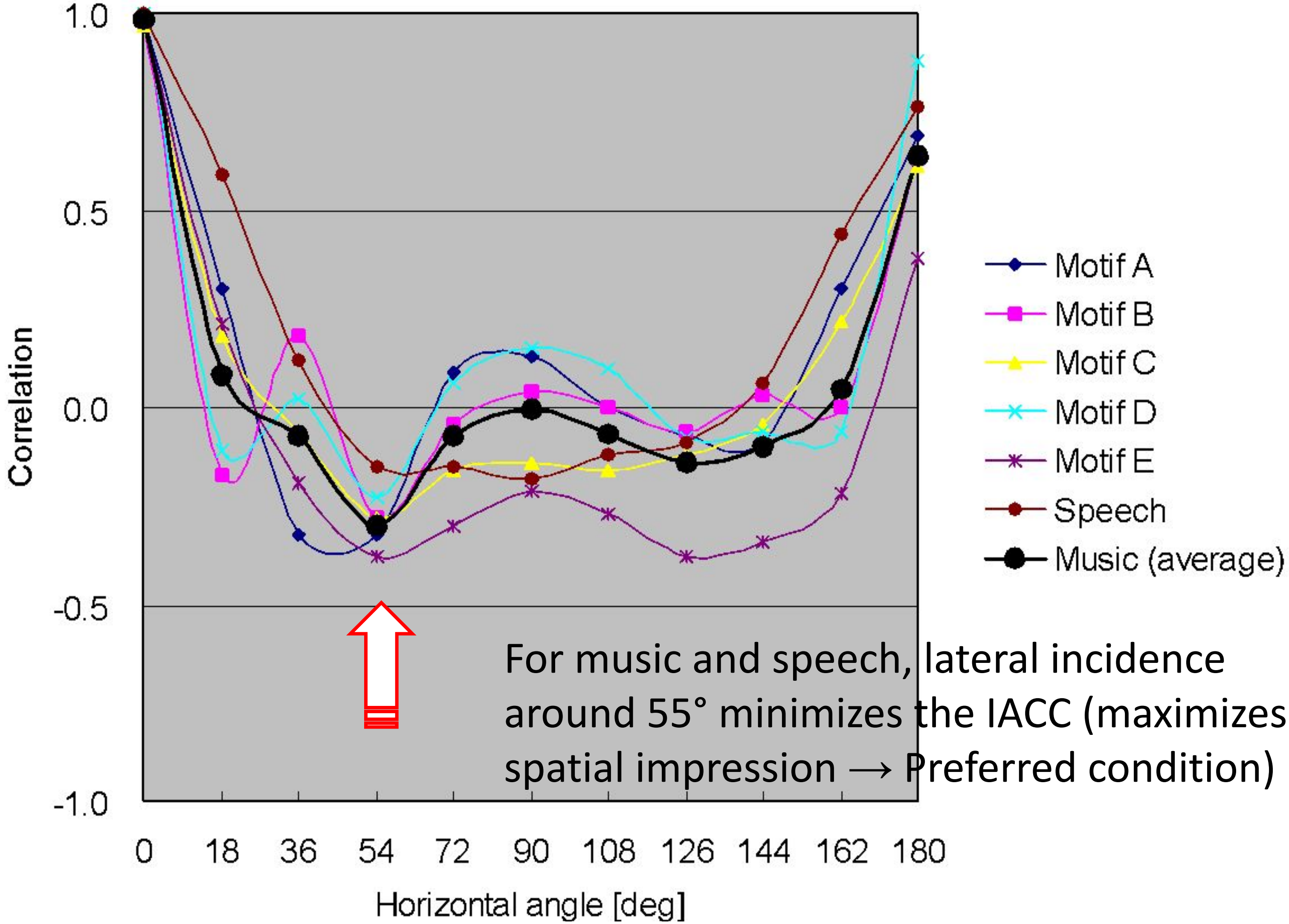


Higher frequency signal

Wavelength is smaller than the dimension of the head.

Not only the phase difference but also the diffraction and scattering by the head and pinna affect the interaural cross-correlation.

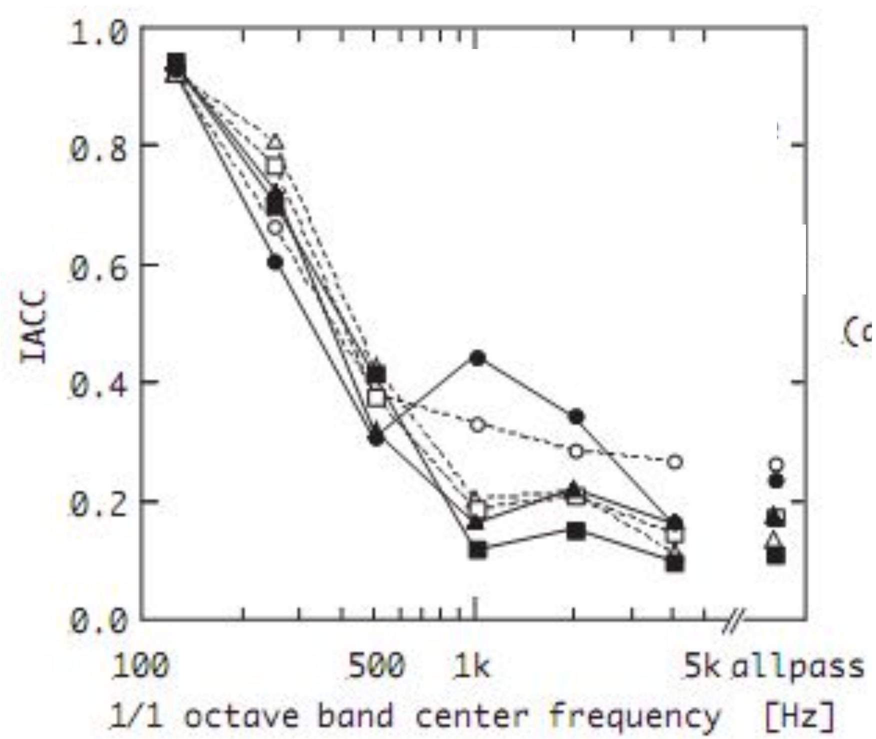
Effect of direction of a single reflection on IACC



For music and speech, lateral incidence around 55° minimizes the IACC (maximizes the spatial impression → Preferred condition)

IACC obtained from binaural impulse responses

- Usually, IACC values in octave bands show higher values for lower frequency and lower values for mid and high frequencies



Criteria for concert halls (recommendation)

UNTREF
UNIVERSIDAD NACIONAL DE TRES DE FEBRERO

- A. C. Gade, Springer Handbook of Acoustics (2007) Chap. 9

Parameter (Mid-frequency)	Chamber hall	Symphony hall
Reverberation time	1.5 s	2.0-2.4 s
Early decay time, EDT	1.4 s	2.2 s
Sound strength, G	10 dB	3 dB
Clarity, C80	3 dB	-1 dB
Lateral energy fraction, LF	0.15-0.20	0.2-0.25
Interaural cross-correlation, IACC	0.4	0.3

(Each researcher proposes different set of values)

Criteria for concert halls (probable values)

- ISO3382

Parameter	**Single number frequency averaging**	**Typical range**
Sound strength G (dB)	500, 1000 Hz	-2; 10 dB
Early decay time (s)	500, 1000 Hz	1.0; 3.0 s
Clarity, C80 (dB) Center time (ms)	500, 1000 Hz 500, 1000 Hz	-5; 5 dB 60; 260 ms
Early lateral energy fraction, LF	125, 250, 500, 1000 Hz	0.05; 0.35

(IACC in Annex due to the difficulty of standardization)

Checkpoints

1. Uniform loudness throughout the audience area
   - Enough diffusion, No focusing, No excess absorption
   - Barron’s revised theory
2. Diffusions
   - No huge flat and hard surfaces
   - Impulse responses, Reverberation decay
3. Plenty of early-lateral reflections
   - Boundaries facing to the audience area
   - ITDG, C_{80}, RT and EDT, and IACC
4. Reverberation time
   - Choice of materials (and their positions)
   - V/N (Room volume per person)