University of Parma

Industrial Engineering Department

HTTP://ied.unipr.it

AcMus

I Seminário

Música Ciência e Tecnologia

3 a 5 de novembro 2004

ACÚSTICA MUSICAL

# “Listening Tests Performed Inside a Virtual Room Acoustic Simulator”

Authors: A. Farina, P. Martignon, A. Azzali, A. Capra

# GOALS

This paper describes the processing required for obtaining a **realistic audible sound reproduction** and how to perform **good listening tests**.

The basic tool is a convolution procedure between “dry” music and the impulse responses measured (or numerically simulated) of a theatre

- The overall process is usually known as “**Auralization**”, and traditionally it was performed through the binaural technology (headphones reproduction)

- Here the process is generalized to many more loudspeaker-based reproduction systems:
  Stereo, Binaural, Stereo Dipole, Dual Stereo Dipole, Ambisonics, Ambiophonics.

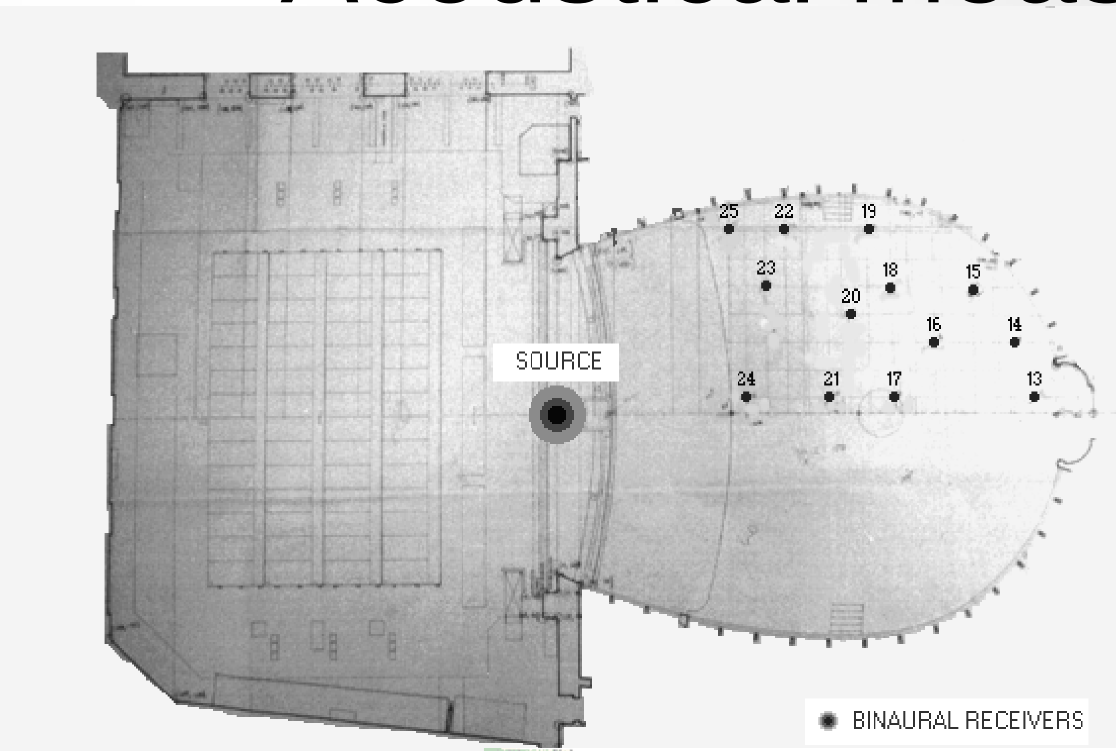
# Theatre la Fenice, Venice

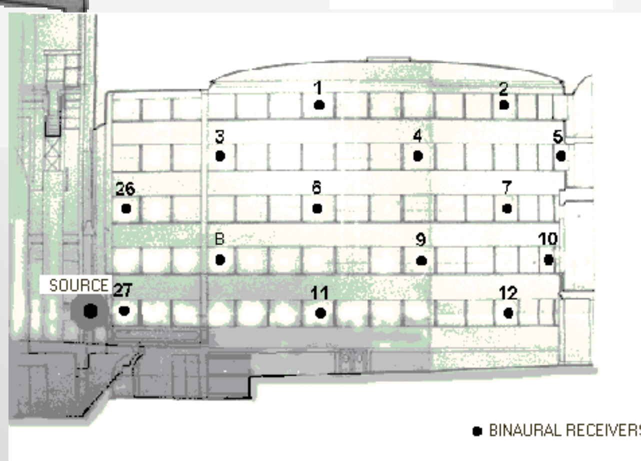
- The first theatre was realised in 1792 by Gian Antonio Selva, after the burning of Teatro San Benedetto
- In December 1836 the theatre burned down again and was rebuilt by G. and T. Meduna the year after
- The theatre was closed in 1995 for maintainance; it had to open again in February 1, 1996, but it burned two days before (January 29, 1996)
- A few weeks before the fire, L.Tronchin measured binaural impulse responses

# Acoustical measurements

- Measurements were performed by L.Tronchin in November-December 1995 with a modified gun (with omnidirectional diffuser) and binaural microphones
- The goal was to analyze some acoustical problems about intercommunication between orchestra pit and stage and to gather information for designing the orchestra shell
- The data were processed with the software Aurora, which had been developed just 2 months before.

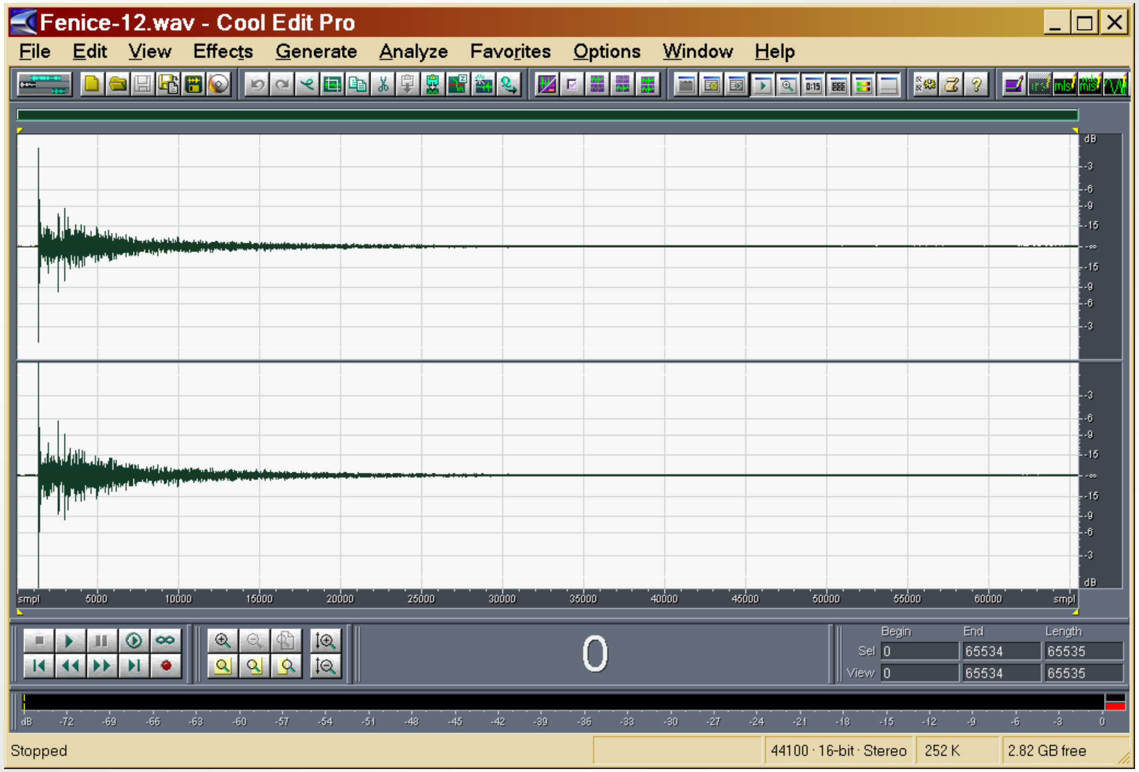
# Acoustical measurements




- In 27 positions a series of binaural impulse responses (with gun shots) was recorded
- Each recording is consequently a stereo file at 16 bits, 48 kHz
- During the measurements the room was perfectly fitted, whilst the stage was empty (no scenery)

# Impulse Responses of La Fenice




Point n. 12

# Auralization Examples – La Fenice

Overture alle Nozze di Figaro di Mozart

- Dry music
- Convolution with experimental I.R. (pt. 12)
- Convolution with computer-simulated IR

Preludio al primo atto della Traviata di G.Verdi

- Dry music
- Convolution with experimental I.R. (pt. 12)
- Convolution with computer-simulated IR

Stop

# The Past

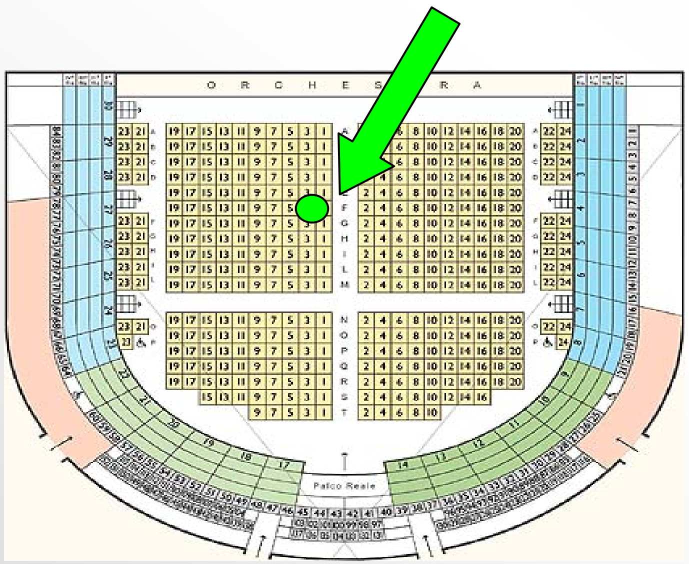
# A PRELIMINARY SUBJECTIVE TEST

Test: questionnaires compiled during headphones listening of several tracks played in 5 different theatres

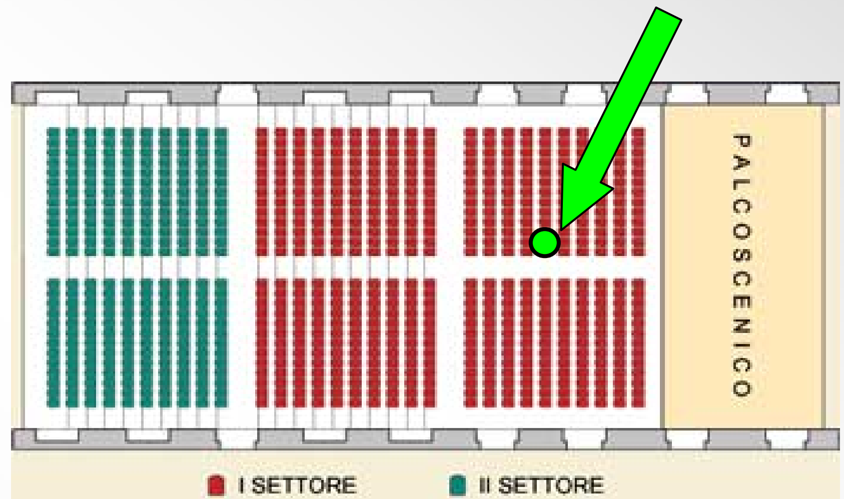
Italian theatres choosen:

- Opera theatres
  - ***Teatro Regio*** (Parma)
  - ***Teatro Valli***(Reggio E.)
- Auditoria
  - ***Paganini*** (Parma)
  - **Sala 700** (Roma)
- Historical theatres
  - ***Teatro Olimpico*** (Vicenza)

<u>Track used for the test</u>: anechoic tracks auralized with binaural Impulse Response of 5 theatres. Acquisition of Impulse Response made with source on the stage and receiver placed between 5th and 6th of every room.



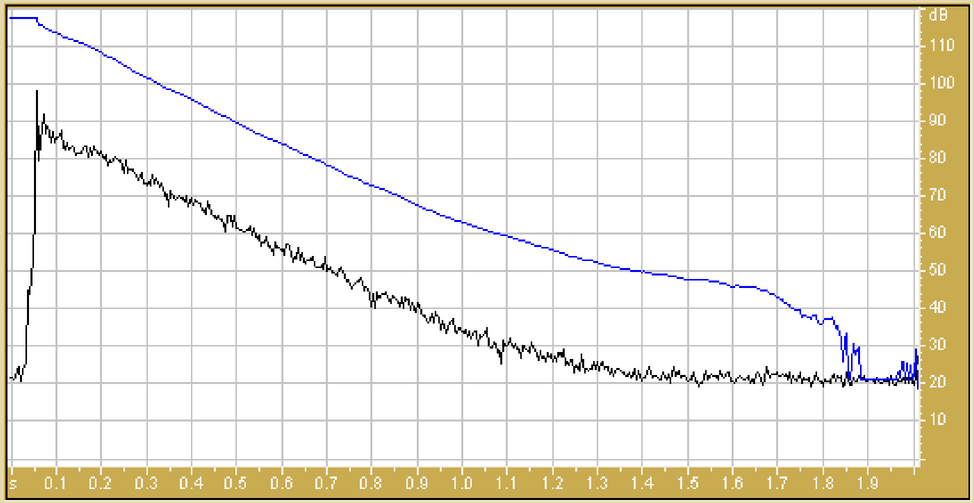

**Teatro Regio (Parma)**




**Auditorium Paganini (Parma)**

# Teatro Regio in Parma (Italy)

Acoustical Parameters according to ISO3382



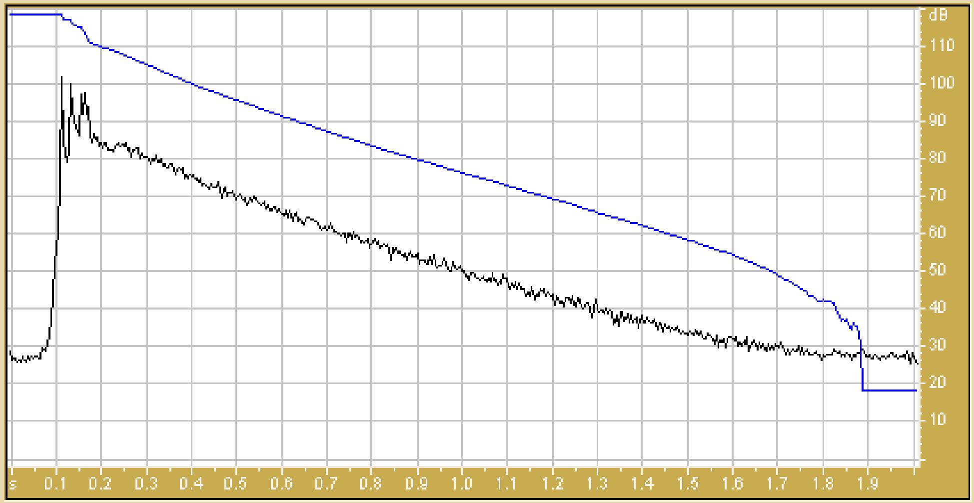

OK
Help
Save to File...
Copy to Clipboard
Store Reference Signal
Switch to Right Channel
Tuser limits:
(-5. dB, -15. dB)
Left Channel

| Freq. (Hz) | 31.5 | 63 | 125 | 250 | 500 | 1k | 2k | 4k | 8k | 16k | Lin | A |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Signal (dB) | 48.60 | 53.48 | 57.25 | 60.89 | 66.81 | 69.76 | 72.43 | 71.61 | 71.65 | 69.16 | 78.55 | 78.63 |
| Noise (dB) | 25.21 | 18.88 | 13.32 | 19.41 | 19.37 | 16.96 | 14.38 | 13.67 | 15.59 | 21.50 | 28.76 | 24.01 |
| strenGth(dB) | -20.40 | -15.52 | -11.75 | -8.11 | -2.19 | 0.76 | 3.43 | 2.61 | 2.65 | 0.16 | -0.45 | 1.63 |
| C50 (dB) | -5.46 | 1.37 | 0.23 | 0.25 | 0.77 | 0.46 | 1.29 | 2.25 | 6.99 | 11.14 | 2.57 | 2.41 |
| C80 (dB) | -3.70 | 2.80 | 3.48 | 1.82 | 4.26 | 2.61 | 3.24 | 4.88 | 10.51 | 15.60 | 4.93 | 4.69 |
| D50 (%) | 22.15 | 57.81 | 51.30 | 51.45 | 54.44 | 52.67 | 57.36 | 62.67 | 83.33 | 92.86 | 64.40 | 63.55 |
| Ts (ms) | 202.12 | 123.81 | 102.50 | 86.64 | 67.18 | 76.24 | 67.54 | 52.41 | 20.87 | 10.71 | 53.18 | 54.42 |
| EDT (s) | 2.26 | 1.79 | 1.64 | 1.18 | 1.09 | 1.20 | 1.16 | 1.16 | 0.69 | 0.38 | 1.12 | 1.13 |
| T20 (s) | 2.94 | 2.51 | 1.56 | 1.21 | 1.11 | 1.07 | 0.98 | 0.87 | 0.68 | 0.47 | 1.01 | 0.97 |
| r T20 | 0.93 | 0.99 | 0.98 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 1.00 | 1.00 |
| T30 (s) | 2.88 | 2.17 | 1.65 | 1.39 | 1.18 | 1.05 | 1.00 | 0.87 | 0.65 | 0.42 | 1.07 | 0.99 |
| r T30 | 0.96 | 0.98 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 1.00 | 1.00 |
| Tuser (s) | 1.95 | 2.45 | 1.36 | 0.97 | 1.17 | 1.05 | 0.94 | 0.95 | 0.85 | 0.40 | 1.02 | 0.99 |
| r Tuser | 0.96 | 0.96 | 0.89 | 0.99 | 1.00 | 1.00 | 0.99 | 0.98 | 0.98 | 0.98 | 0.99 | 0.99 |

$T_{20}$ = 1.11 s

# Teatro Valli, Reggio Emilia, Italy

**Acoustical Parameters according to ISO3382**




OK | Help

Save to File...

Copy to Clipboard

Store Reference Signal

Switch to Right Channel

Tuser limits:
(-5. dB, -15. dB)

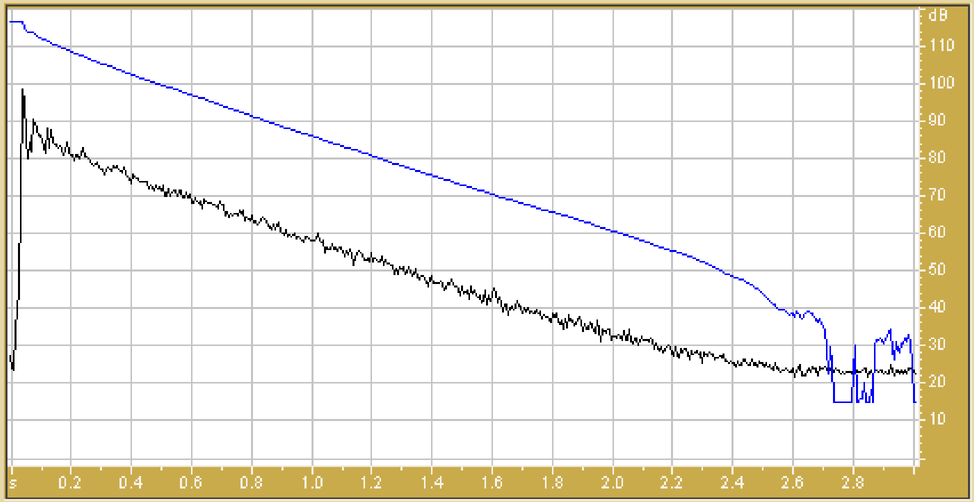
Left Channel

| Freq. (Hz) | 31.5 | 63 | 125 | 250 | 500 | 1k | 2k | 4k | 8k | 16k | Lin | A |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Signal (dB) | 33.24 | 50.79 | 63.47 | 66.10 | 68.38 | 67.96 | 73.18 | 76.32 | 79.64 | 82.10 | 83.65 | 83.26 |
| Noise (dB) | 13.25 | 32.82 | 24.81 | 23.47 | 23.70 | 21.12 | 21.84 | 21.52 | 23.02 | 29.94 | 35.88 | 30.56 |
| strenGth(dB) | -35.76 | -18.21 | -5.53 | -2.90 | -0.62 | -1.04 | 4.18 | 7.32 | 10.64 | 13.10 | 4.65 | 6.26 |
| C50 (dB) | 1.55 | 1.19 | -5.72 | -1.74 | 0.52 | 2.71 | 5.06 | 1.40 | 1.62 | 10.18 | 2.94 | 2.63 |
| C80 (dB) | 3.61 | 1.97 | 5.83 | 7.28 | 7.97 | 5.57 | 7.58 | 6.06 | 8.20 | 17.01 | 8.46 | 7.78 |
| D50 (%) | 58.81 | 56.81 | 21.12 | 40.11 | 53.00 | 65.09 | 76.24 | 57.99 | 59.23 | 91.24 | 66.31 | 64.72 |
| Ts (ms) | 147.19 | 136.86 | 101.13 | 73.80 | 56.28 | 56.96 | 44.84 | 62.36 | 49.52 | 17.35 | 45.61 | 49.03 |
| EDT (s) | -- | 2.08 | 1.01 | 0.72 | 0.71 | 1.37 | 0.94 | 0.92 | 0.57 | -- | 0.60 | 0.71 |
| T20 (s) | 2.70 | 2.42 | 1.96 | 1.91 | 1.55 | 1.54 | 1.47 | 1.30 | 0.98 | 0.50 | 1.25 | 1.24 |
| r T20 | 0.95 | 0.94 | 0.98 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.93 | 1.00 | 1.00 |
| T30 (s) | -- | 1.94 | 2.16 | 1.92 | 1.68 | 1.58 | 1.55 | 1.37 | 1.03 | 0.57 | 1.38 | 1.34 |
| r T30 | -- | 0.95 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.97 | 1.00 | 1.00 |
| Tuser (s) | 4.12 | 3.41 | 1.81 | 1.64 | 1.46 | 1.47 | 1.35 | 1.30 | 1.07 | -- | 1.22 | 1.23 |
| r Tuser | 0.89 | 0.98 | 0.95 | 0.99 | 0.98 | 1.00 | 1.00 | 0.99 | 0.98 | -- | 0.98 | 0.99 |

$T_{20} = 1.55$ s

# Paganini Auditorium, Parma, Italy

**Acoustical Parameters according to ISO3382**




OK | Help | Save to File... | Copy to Clipboard | Store Reference Signal | Switch to Right Channel

Tuser limits: (-5. dB, -15. dB)

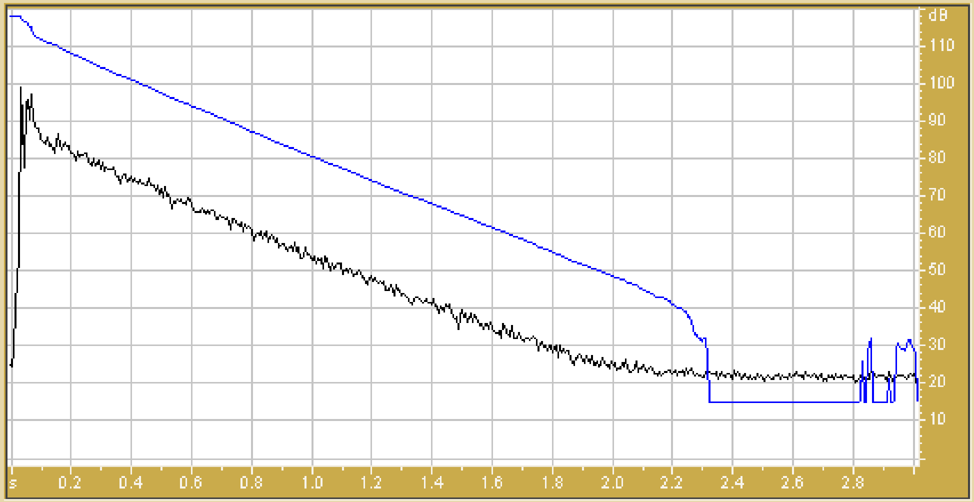
Left Channel

| Freq. (Hz) | 31.5 | 63 | 125 | 250 | 500 | 1k | 2k | 4k | 8k | 16k | Lin | A |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Signal (dB) | 40.84 | 57.24 | 68.80 | 66.26 | 67.20 | 70.03 | 73.41 | 74.09 | 77.50 | 75.02 | 82.52 | 81.08 |
| Noise (dB) | 20.40 | 26.31 | 23.57 | 24.89 | 22.64 | 20.25 | 19.91 | 17.27 | 23.54 | 20.46 | 33.16 | 28.11 |
| strenGth(dB) | -28.16 | -11.76 | -0.20 | -2.74 | -1.80 | 1.03 | 4.41 | 5.09 | 8.50 | 6.02 | 3.52 | 4.08 |
| C50 (dB) | 2.38 | -2.31 | -5.27 | -5.18 | -1.30 | 1.21 | 1.29 | 0.09 | 4.29 | 8.19 | 2.33 | 2.04 |
| C80 (dB) | 4.35 | -0.33 | -0.29 | -0.84 | 0.44 | 2.20 | 2.45 | 1.47 | 6.37 | 9.33 | 3.97 | 3.49 |
| D50 (%) | 63.36 | 37.01 | 22.92 | 23.29 | 42.57 | 56.89 | 57.36 | 50.51 | 72.87 | 86.82 | 63.12 | 61.53 |
| Ts (ms) | 145.30 | 159.46 | 155.87 | 138.71 | 122.80 | 98.06 | 94.93 | 100.07 | 44.10 | 15.99 | 69.01 | 75.79 |
| EDT (s) | 2.25 | 1.84 | 1.98 | 1.64 | 2.19 | 1.95 | 2.06 | 1.81 | 1.01 | -- | 1.49 | 1.61 |
| T20 (s) | 3.12 | 2.76 | 2.75 | 2.48 | 2.08 | 2.37 | 2.35 | 1.98 | 1.20 | 0.56 | 2.05 | 2.08 |
| r T20 | 0.99 | 0.99 | 0.97 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.96 | 1.00 | 1.00 |
| T30 (s) | 5.53 | 2.77 | 2.66 | 2.66 | 2.08 | 2.42 | 2.32 | 1.99 | 1.24 | 0.60 | 2.20 | 2.14 |
| r T30 | 0.91 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.98 | 1.00 | 1.00 |
| Tuser (s) | 3.36 | 2.14 | 1.98 | 2.16 | 2.00 | 2.30 | 2.40 | 1.90 | 1.17 | 1.18 | 1.81 | 1.91 |
| r Tuser | 0.98 | 0.98 | 0.97 | 0.99 | 0.98 | 0.99 | 1.00 | 1.00 | 1.00 | 0.93 | 0.99 | 1.00 |

$T_{20} = 2.08$ s

# S.Cecilia Auditorium 700, Rome, Italy

**Acoustical Parameters according to ISO3382**




OK

Help

Save to File...

Copy to Clipboard

Store Reference Signal

Switch to Right Channel

Tuser limits:

(-5. dB, -15. dB)

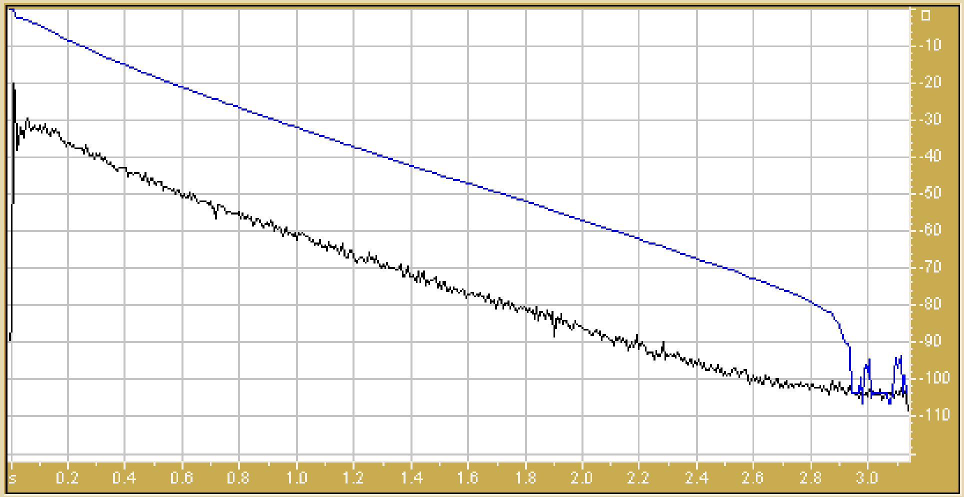
Left Channel

| Freq. (Hz) | 31.5 | 63 | 125 | 250 | 500 | 1k | 2k | 4k | 8k | 16k | Lin | A |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Signal (dB) | 37.50 | 58.10 | 69.58 | 67.65 | 69.29 | 70.45 | 74.74 | 78.03 | 76.63 | 74.34 | 83.48 | 82.62 |
| Noise (dB) | 17.10 | 19.19 | 23.26 | 19.21 | 18.71 | 16.64 | 18.07 | 20.72 | 21.58 | 19.77 | 30.42 | 26.83 |
| strenGth(dB) | -31.50 | -10.90 | 0.58 | -1.35 | 0.29 | 1.45 | 5.74 | 9.03 | 7.63 | 5.34 | 4.48 | 5.62 |
| C50 (dB) | -0.90 | 3.27 | 0.60 | 2.07 | 0.19 | 0.67 | 2.73 | 5.41 | 5.29 | 12.80 | 4.45 | 4.32 |
| C80 (dB) | 0.61 | 3.81 | 2.54 | 3.12 | 2.19 | 2.20 | 3.94 | 6.93 | 7.59 | 17.00 | 6.04 | 5.85 |
| D50 (%) | 44.81 | 67.98 | 53.45 | 61.70 | 51.11 | 53.86 | 65.24 | 77.67 | 77.16 | 95.02 | 73.58 | 72.99 |
| Ts (ms) | 181.27 | 107.51 | 111.52 | 100.03 | 108.51 | 104.34 | 81.09 | 48.46 | 39.77 | 15.22 | 57.02 | 58.14 |
| EDT (s) | 2.63 | 2.43 | 1.62 | 1.74 | 1.96 | 1.95 | 1.51 | 0.98 | 0.74 | 0.35 | 1.11 | 1.14 |
| T20 (s) | 3.16 | 2.36 | 1.75 | 2.00 | 2.04 | 1.90 | 1.85 | 1.49 | 0.93 | 0.44 | 1.72 | 1.68 |
| r T20 | 0.97 | 0.98 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.96 | 1.00 | 1.00 |
| T30 (s) | 3.18 | 2.44 | 1.74 | 1.99 | 2.06 | 1.91 | 1.82 | 1.52 | 0.94 | 0.47 | 1.75 | 1.72 |
| r T30 | 0.98 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 |
| Tuser (s) | 2.91 | 2.77 | 2.00 | 2.14 | 2.05 | 2.15 | 1.82 | 1.49 | 0.93 | 0.23 | 1.64 | 1.60 |
| r Tuser | 0.96 | 0.99 | 0.97 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 0.97 | 1.00 | 1.00 |

$T_{20} = 2.04$ s

# Teatro Olimpico, Vicenza, Italy

**Acoustical Parameters according to ISO3382**




OK | Help

Save to File...

Copy to Clipboard

Store Reference Signal

Switch to Left Channel

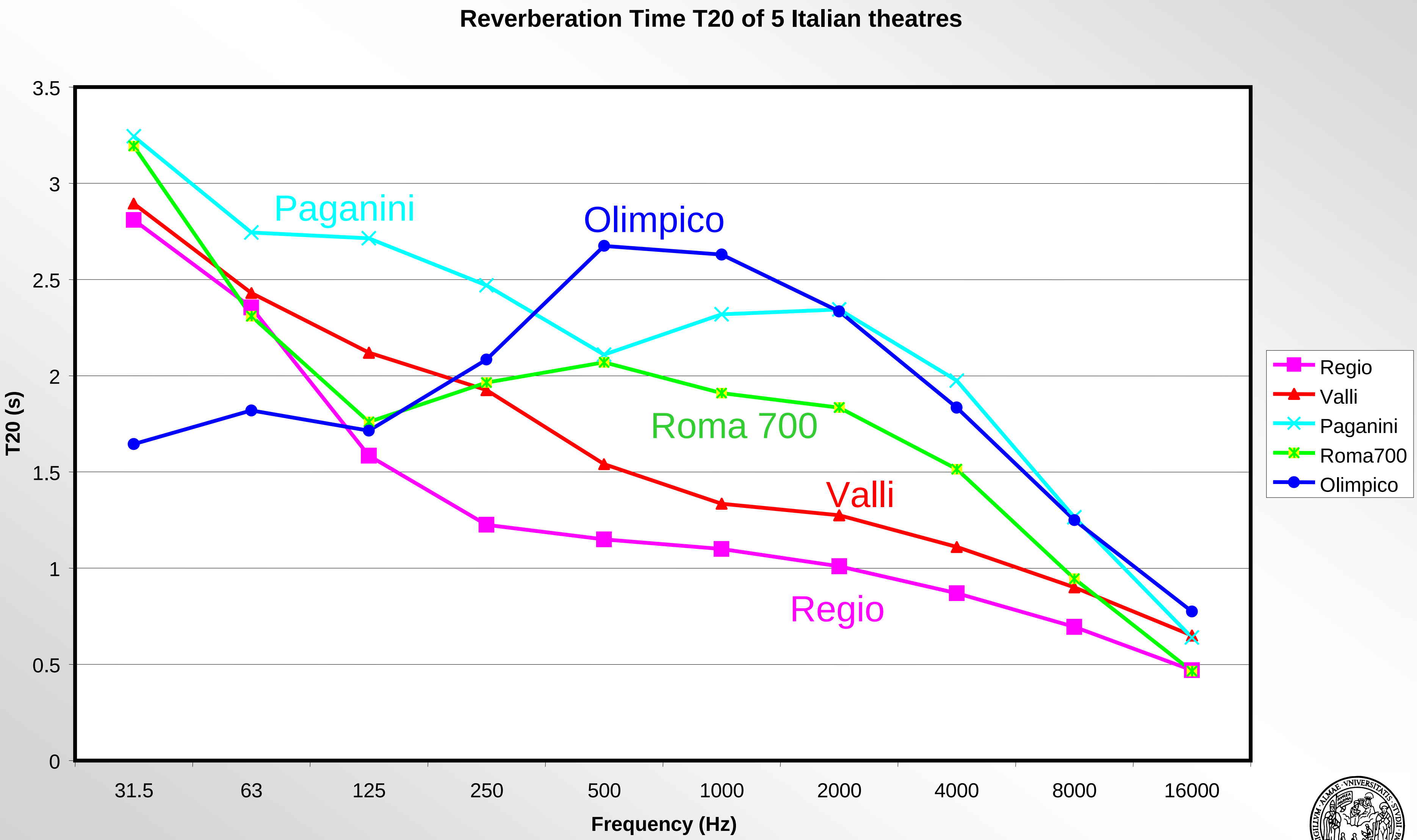
Tuser limits:
(-5. dB, -15. dB)

Right Channel

| Freq. (Hz) | 31.5 | 63 | 125 | 250 | 500 | 1k | 2k | 4k | 8k | 16k | Lin | A |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Signal (dB) | -102.98 | -67.92 | -64.66 | -55.81 | -54.84 | -52.48 | -46.50 | -41.06 | -41.43 | -51.80 | -36.76 | -37.05 |
| Noise (dB) | -128.23 | -109.58 | -109.43 | -105.18 | -103.86 | -106.06 | -104.11 | -109.57 | -111.06 | -102.07 | -95.95 | -98.74 |
| strenGth(dB) | -171.98 | -136.92 | -133.66 | -124.81 | -123.84 | -121.48 | -115.50 | -110.06 | -110.43 | -120.80 | -115.76 | -114.05 |
| C50 (dB) | 5.53 | -2.97 | 1.02 | 2.65 | -1.87 | -2.16 | 0.02 | 0.30 | -2.50 | 1.53 | -0.77 | -0.69 |
| C80 (dB) | 6.77 | 3.28 | 3.65 | 3.83 | 0.71 | -0.66 | 1.31 | 1.93 | 0.07 | 4.44 | 1.18 | 1.18 |
| D50 (%) | 78.14 | 33.51 | 55.85 | 64.82 | 39.40 | 37.80 | 50.12 | 51.72 | 35.97 | 58.69 | 45.58 | 46.05 |
| Ts (ms) | 91.66 | 109.85 | 94.44 | 93.63 | 139.20 | 147.28 | 118.12 | 95.79 | 101.26 | 56.53 | 100.65 | 102.32 |
| EDT (s) | 1.28 | 0.81 | 1.47 | 2.15 | 2.42 | 2.29 | 2.31 | 1.76 | 1.26 | 0.89 | 1.58 | 1.66 |
| T20 (s) | 1.66 | 1.75 | 1.78 | 2.09 | 2.63 | 2.66 | 2.30 | 1.84 | 1.25 | 0.74 | 1.85 | 1.90 |
| r T20 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| T30 (s) | 1.55 | 1.65 | 1.83 | 2.11 | 2.59 | 2.65 | 2.38 | 1.90 | 1.29 | 0.77 | 2.03 | 2.05 |
| r T30 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Tuser (s) | 1.66 | 1.62 | 1.53 | 2.02 | 2.48 | 2.30 | 2.14 | 1.77 | 1.15 | 0.71 | 1.62 | 1.70 |
| r Tuser | 0.97 | 0.96 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 |

$T_{20}$ = 2.63 s

# Acoustical properties of the theatres



# Acoustical properties of the theatres

## According to ISO 3382 standard

| Param. | Regio | Valli | Paganini | Roma-700 | Olimpico |
|---|---|---|---|---|---|
| C50 [dB] | 1.82 | 7.48 | -1.47 | -2.45 | 0.03 |
| C80 [dB] | 4.93 | 9.28 | 0.81 | 0.83 | 1.20 |
| D50 [%] | 60 | 84 | 42 | 37 | 50 |
| Ts [ms] | 48 | 28 | 115 | 144 | 110 |
| EDT [s] | 1.08 | 1.26 | 2.09 | 1.98 | 2.43 |
| T20 [s] | 1.10 | 1.44 | 2.22 | 1.99 | 2.65 |
| T30 [s] | 1.11 | 1.55 | 2.24 | 1.99 | 2.64 |
| LF | 0.10 | 0.12 | 0.22 | 0.12 | 0.18 |
| IACC (Early) | 0.71 | 0.88 | 0.60 | 0.70 | 0.81 |
| TB | 1.47 | 1.70 | 1.20 | 1.11 | 0.91 |
| BR | 1.27 | 1.41 | 1.17 | 0.94 | 0.72 |

# Sound Samples

Chosen in order to evidence the difference between rooms designed for opera and rooms for purely orchestral music:

**ORCHESTRAL:**

- G.Verdi, Preludio al Primo Atto de “*la Traviata*”
- W.A. Mozart, Overture de “*le Nozze di Figaro*”
- Strauss, “*Pizzicate Polka*”

**VOCAL:**

- Mozart, *“Cosi’ fan tutte”* (voice and piano)
- Tosti, “*Non t’amo più*” (voice and piano)
- “*My Funny Valentine*” (jazz,voice solo)

# Listening seat in **Casa della Musica** (Parma-Italy):

<u>Instrumentation</u>:

- A liquid cooled Computer (*Futureclient*)

- Open dynamic headphones *Sennheiser HD 580 Precision*

- Audio-pro Subwoofer for very low frequencies (18-50 Hz)

# Headphones

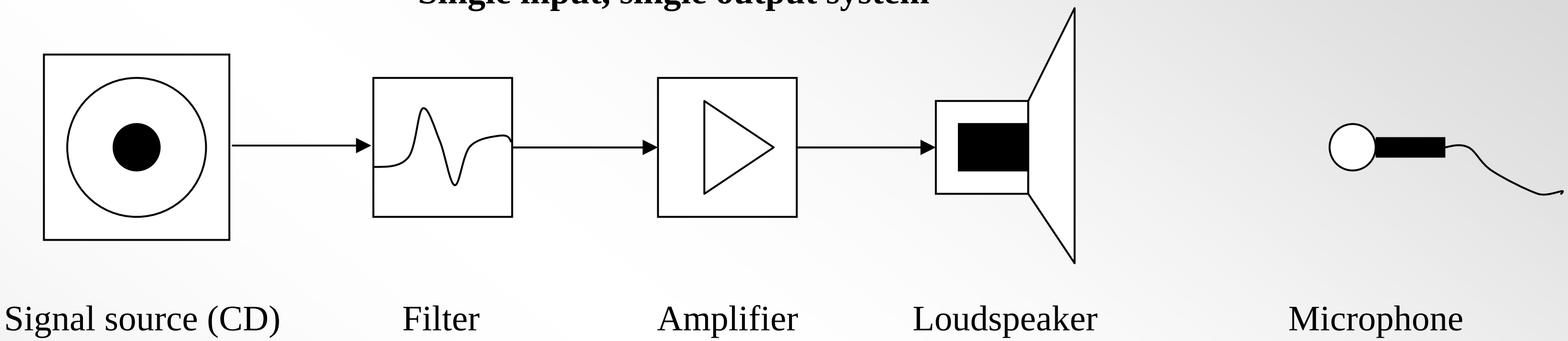
- Open dynamic headphones: *Sennheiser HD 580 Precision*

- Digital equalizing filters are employed for compensating the frequency response of the hedphone+dummy head

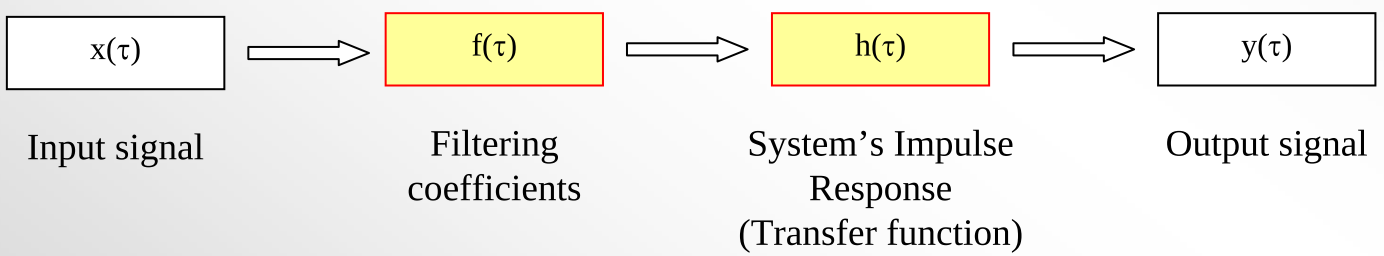
- The computation of these inverse filters revealed to be very important for the "transparency" of the reproduction chain

# Theory of inverse filters

**Single input, single output system**




**Block diagram**



# Combined transfer functiom

**As all the stages are linear:**

$$y(i) = x(i) \otimes f(j) \otimes h(l)$$

**The gola of the filter is to "equalize" – this means to make the output y(i) to be equal to the input x(i). This is obtained if:**

$$f(j) \otimes h(l) \Rightarrow \delta(i)$$

**In which δ(i) is the Dirac's Delta function (a single sample having unit value, peceded and followed by thousands of zeroes) – this way the total effect of the filter+system is simply a delay of a few millisecond, with no other evident alteration**

# Possible designe strategies for an equalizing filter

- Mourjopoulos – Least-squares recursive method in time domain – the whole frequency range is always completely inverted.
- Neely & Allen – the filter is designed in the frequency domain – only the magnitude of the transfer function is inverted, so the equalized system will have flat frequency response, but it will still be “smeared” in time.
- Nelson & Kirkeby – again in the frequency domain, but the whole complex value is inverted, adding a small regularization quantity at the denominator for avoiding instabilities and ensuring a finite lenght of the inverse filter

# Theory of Kirkeby inversion

- **Step 1 – pass to frequency domain through FFT**

$$H(\omega) = \mathrm{FFT}\left[h(\tau)\right]$$

- **Step 2 – make the complex reciprocal at each frequency:**

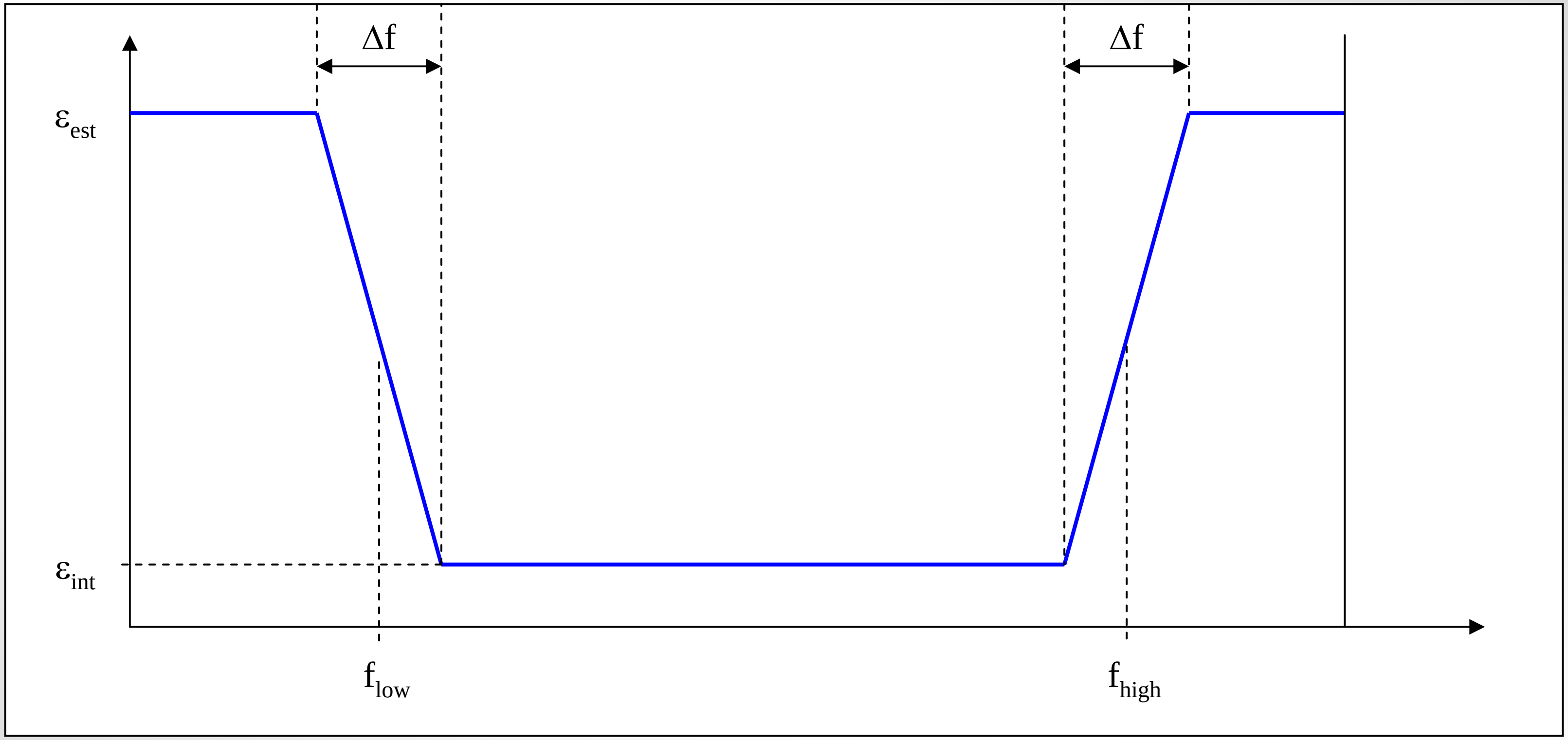
$$F(\omega) = \frac{\mathrm{Conj}\left[H(\omega)\right]}{\mathrm{Conj}\left[H(\omega)\right] \cdot H(\omega) + \varepsilon(\omega)}$$

- **Step 3 – go back to time domain through an IFFT:**

$$f(\tau) = \mathrm{IFFT}\left[F(\omega)\right]$$

**Parametro di regolarizzazione**

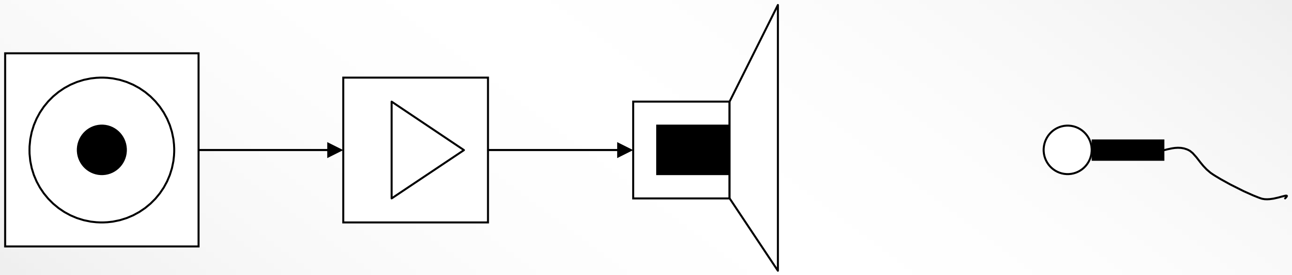
# Regularization parameter $\varepsilon(\omega)$ variable with frequency

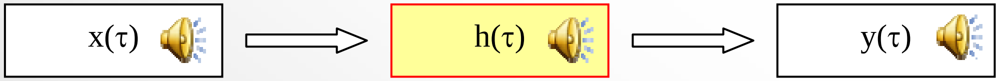



**Changing the value of the regularization parameter allows for a very accurate filter at central frequency, and progressively a less aggressive filtering at very low or very high frequencies.**

# Example

- A loudspeaker+microphone system was measured:





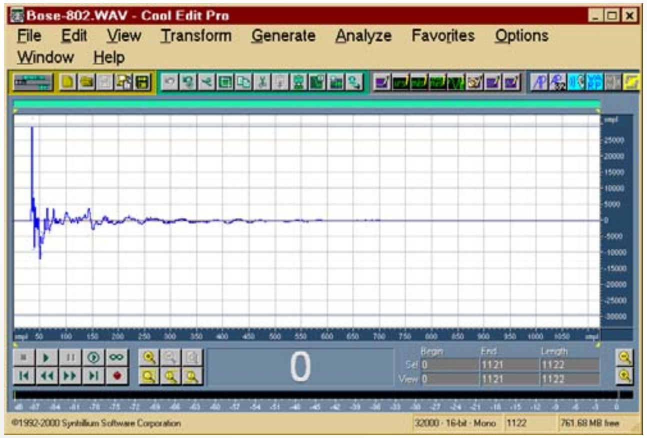

Input signal

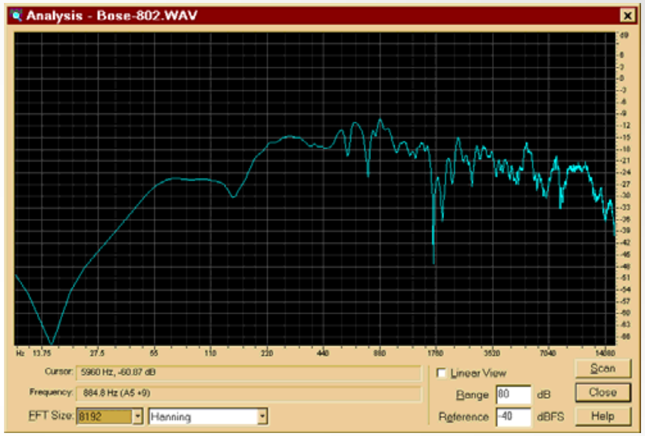
Impulse response
(Transfer function)

Output signal

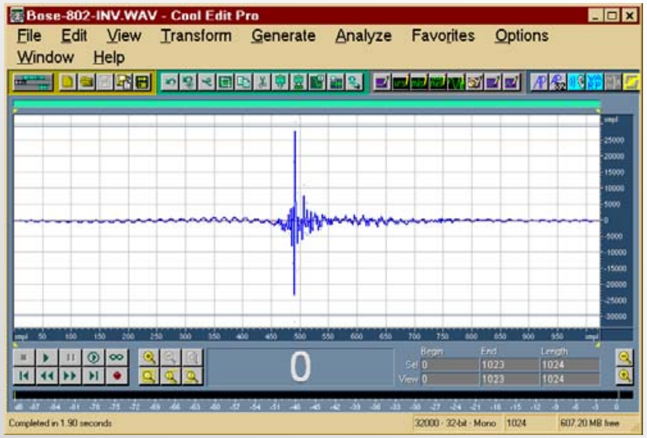
# Inverse filter example

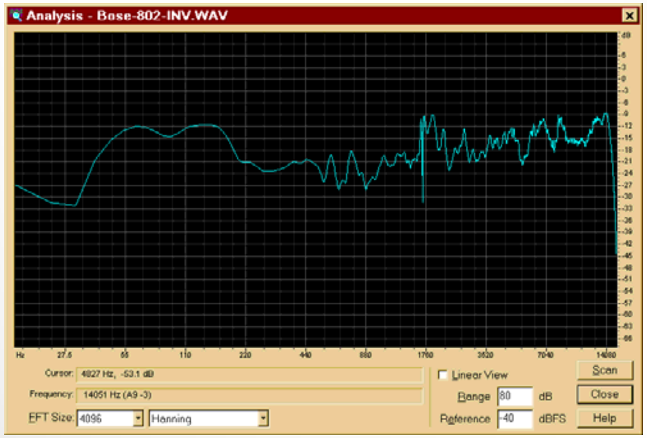
- **System's impulse response**



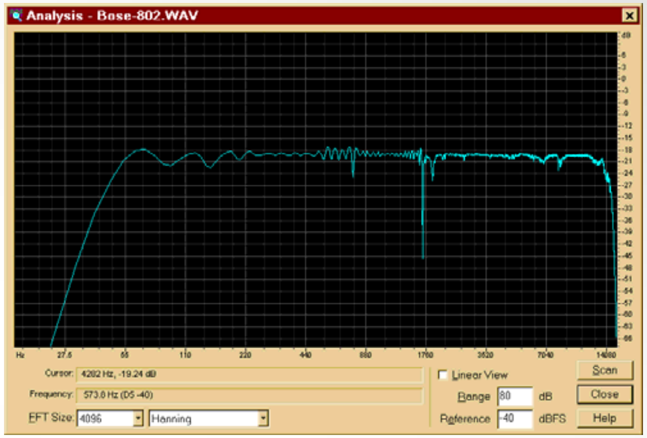

❑ **Filtro inverso**



# Inverse filter example

- **Convolution of inverse filter with the system's impulse response**


**Not filtered system**

| x(τ) | → | h(τ) | → | y(τ) |
|---|---|---|---|---|
| Input signal | | System | | Output signal |

**System with filter**

| x(τ) | → | f(τ) | → | z(τ) | → | h(τ) | → | y(τ) |
|---|---|---|---|---|---|---|---|---|
| Input signal | | Filtering coefficients | | Filtered signal | | System | | Output signal |

# The software

- Real-time control over the theatre (ABCDE) and the music piece (123456)

- Collection of the questionnaires with a graphical user's interface

- The user is allowed to switch at will the sound samples and theatres, and to change the responses

# Statistical Analysis of the results

**Aim**: possible correlation between objective parameters measured inside the theatres and subjective descriptors.

Objective parameters choosen:

- **Monophonic parameters**: reverberation times **T10,T20, T30, EDT**; clarity **C50** e **C80**; center time **Ts**; **D50**.
- **Spatial parameters**: *Lateral Fraction*(**Lf**) and *Inter Aural Cross Correlation* (**IACC**)
- **Tonal parameters**: *Tonal Balance* (**TB**) and *Bass Ratio* (**BR**)

**Method of analysis:** *linear regression* (multiple regression to manage the whole matrix of subjective/objective data).

Defining r as coefficient of correlation:

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

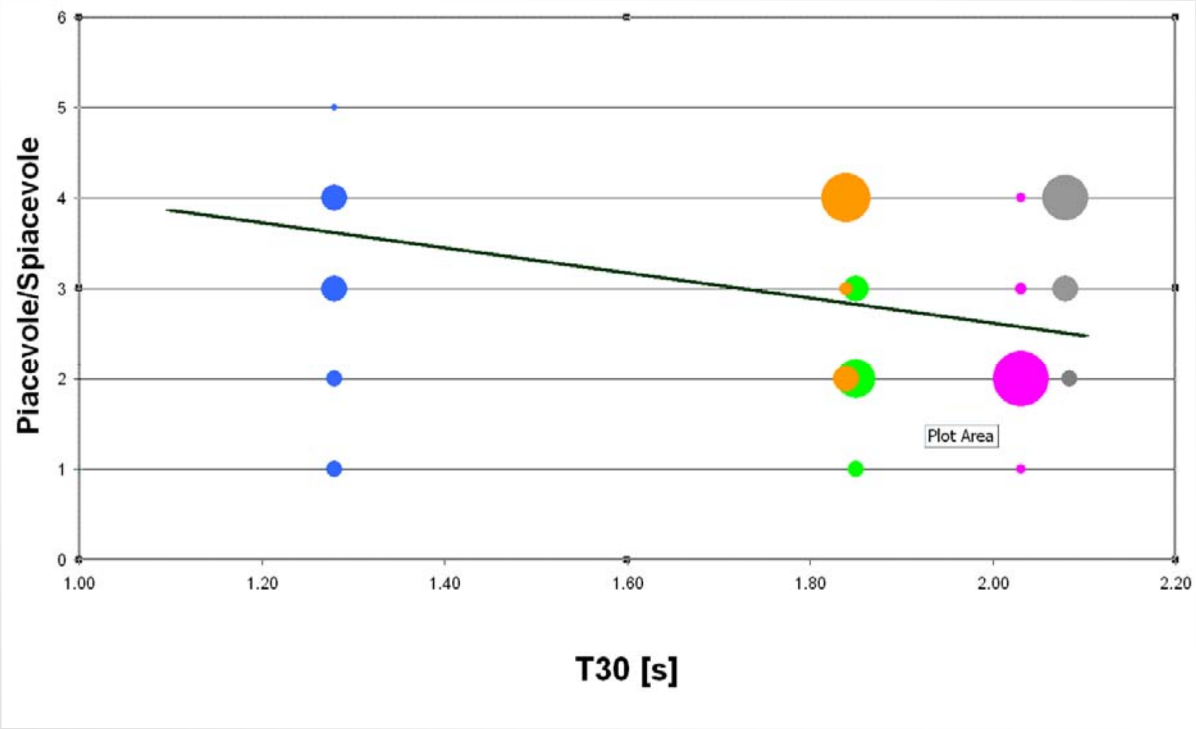
Objective parameter $\xleftrightarrow[|r|>0,3]{r}$ *Subjective parameter*

# Matrix 9x11 of “**r**”: objective parameters on the abscissa and subjective parameters on the ordinate.

| Coeff. Di Regressione Lineare ORCHESTRALI | C50 | C80 | D50 | Ts | EDT | T20 | T30 | LF | IACC | TB | BR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Piacevole-Spiacevole** | 0,14 | 0,19 | 0,19 | -0,24 | -0,19 | -0,20 | -0,12 | -0,21 | 0,16 | -0,07 | 0,09 |
| **Rotondo-Spigoloso** | **0,35** | **0,42** | **0,46** | **-0,51** | **-0,35** | **-0,37** | -0,22 | **-0,51** | **0,32** | -0,22 | 0,15 |
| **Morbido-Duro** | 0,17 | **0,33** | 0,28 | **-0,45** | **-0,40** | **-0,39** | -0,19 | **-0,35** | 0,28 | -0,11 | 0,13 |
| **Diffuso-Localizzabile** | 0,17 | **0,30** | 0,22 | **-0,36** | **-0,35** | **-0,37** | -0,26 | -0,25 | 0,26 | 0,02 | 0,21 |
| **Distaccato-Avvolgente** | -0,17 | -0,24 | -0,19 | 0,23 | 0,16 | 0,19 | 0,09 | 0,21 | -0,21 | -0,02 | -0,17 |
| **Secco-Rimbombante** | -0,24 | **-0,42** | **-0,32** | **0,50** | **0,48** | **0,50** | **0,34** | **0,37** | **-0,36** | 0,01 | -0,26 |
| **Acuti Accentuati-Acuti Ridotti** | -0,19 | -0,28 | **-0,35** | **0,45** | **0,31** | 0,26 | 0,04 | **0,44** | -0,22 | **0,31** | 0,02 |
| **Bassi Accentuati-Bassi Ridotti** | 0,22 | 0,29 | **0,38** | **-0,48** | **-0,36** | **-0,32** | -0,18 | **-0,46** | 0,20 | **-0,34** | -0,04 |
| **Sommesso-Sonoro** | -0,08 | -0,01 | 0,00 | -0,11 | -0,15 | -0,09 | -0,02 | -0,05 | -0,01 | -0,12 | -0,11 |

# Results

Bad correlation between objective and subjective parameters.
Example: pleasant vs. T30 gives a correlation r=-0.20




● T.Regio ● Roma 700 ● T.Olimpico ● A.Paganini ● T.Valli

# Results

**IACC** $\xleftrightarrow{0.42}$ **diffuse-localisable**

**Tonal Balance** $\xleftrightarrow{0.31}$ **treble boosted-treble reduced**

**Tonal Balance** $\xleftrightarrow{-0.34}$ **bass boosted-bass reduced**

**Lateral Fraction** $\xleftrightarrow[-0.51]{}$ **round-sharp**

Although the results were inherently bad, they were useful for improving the methodology, along these findings:

- Future listening tests

Only one music sample for test

Short explanation of the attributes

- Statistical analysis of the results

Advanced statistical methods: *factor analysis* or *multivariate regression.*

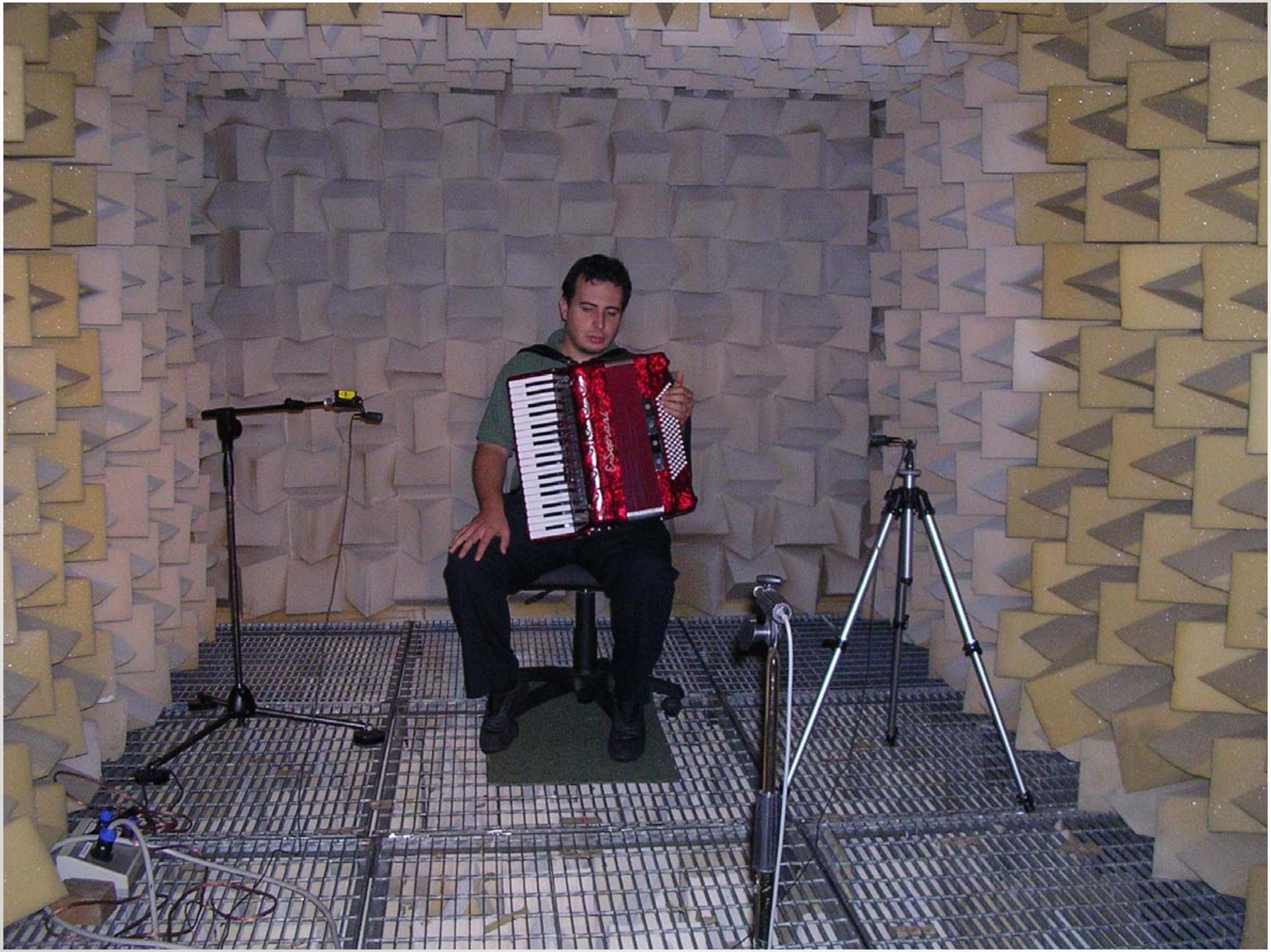
# The Present

# THE NEW LISTENING TEST WITH 4 DIFFERENT SOUND SYSTEMS

- 4 reproduction sound systems being compared:
    1. Headphones
    2. Traditional normal stereo
    3. Stereo Dipole
    4. Dual Stereo Dipole
- Monophonic sound source (accordion) – no multiple soundtracks
- The same 5 virtual rooms as in the preliminary experiment
- Different source-receiver distances in each room

# Recording of the anechoic music piece

(ASK Industries, Reggio Emilia)



# THE NEW LISTENING ROOM

("*Casa della Musica*" – Parma)

To reduce the reverberation:

- <u>High Frequencies</u>: traditional absorbing panel (glass wool, pyramid, ecc... )

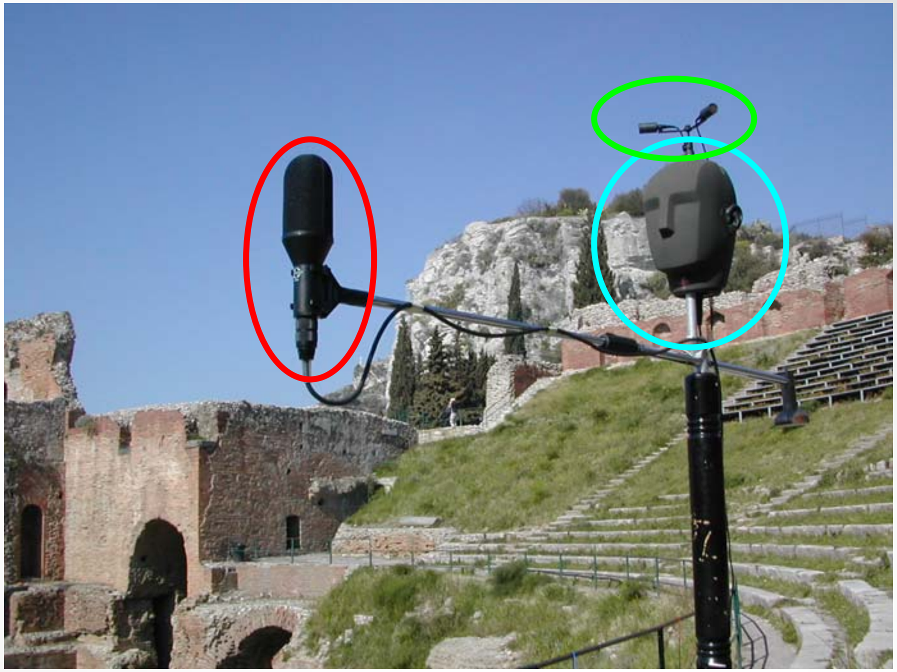
- <u>Low Frequencies</u>: resonant open cavities (cartoon boxes and tube traps), double side rigid and vertical absorbing panels.

# Modern multiple-format microphonic systems

- Microphones *Neumann KS-140* in configuration ORTF (angle of 110°)

- Dummy head *Neumann KU-100* with binaural microphones

- Soundfield ST-250 microphone probe



- Turn table for recordings to different angles

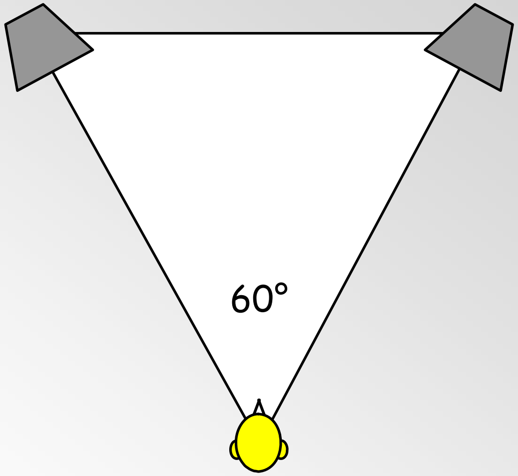
# Binaural / HEADPHONES

- *Sennheiser HD 580 Precision*

-In principle, they should put the right pressure exactly where it was recorded, at the ears, maintaining perfect separation and not being affected by the room response.

- Each signal is passed trough an inverse FIR filter, which is computed after a measurement performed with the headphones over the dummy head

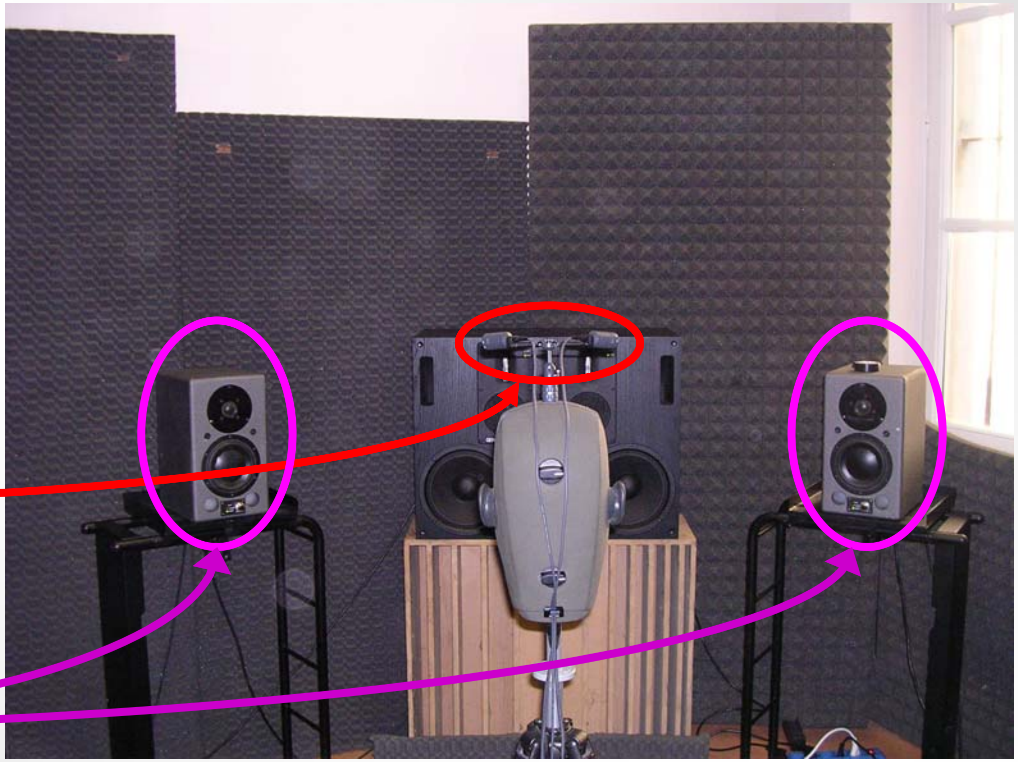
# THE NORMAL STEREO



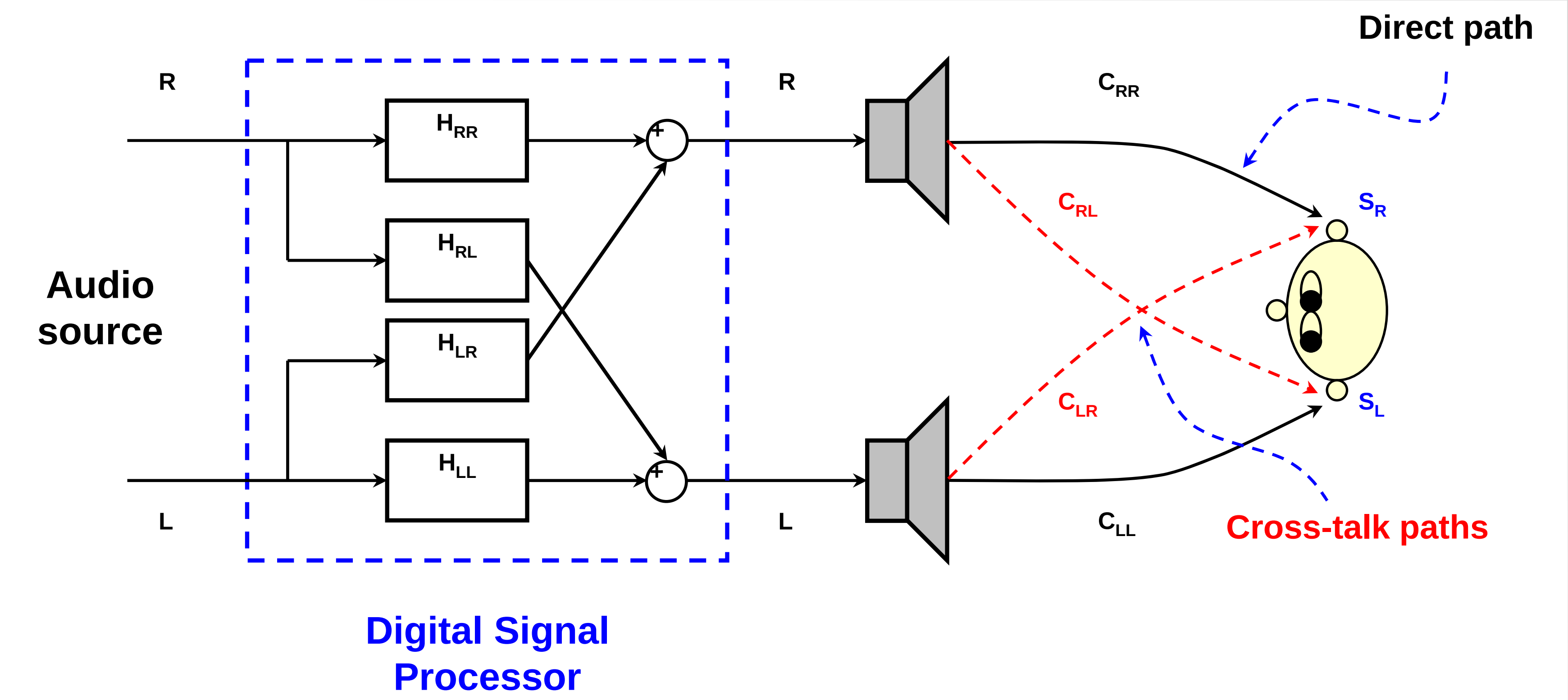

- *Dynaudio* self-powered studio monitors

-The sound is picked up by the two cardioid microphones

- Each of the two signals is passed through a proper inverse filter, computed after a measurement performed placing the ORTF microphones at the listening position, in front of the loudspeakers

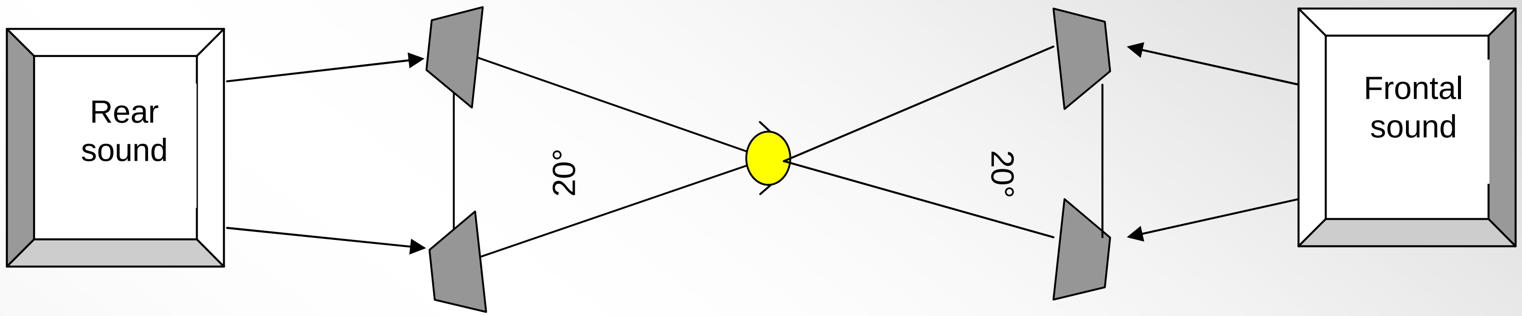


# THE STEREO DIPOLE




The cross-talk cancellation allows for the replica of the recorded signals at the ears of the listener

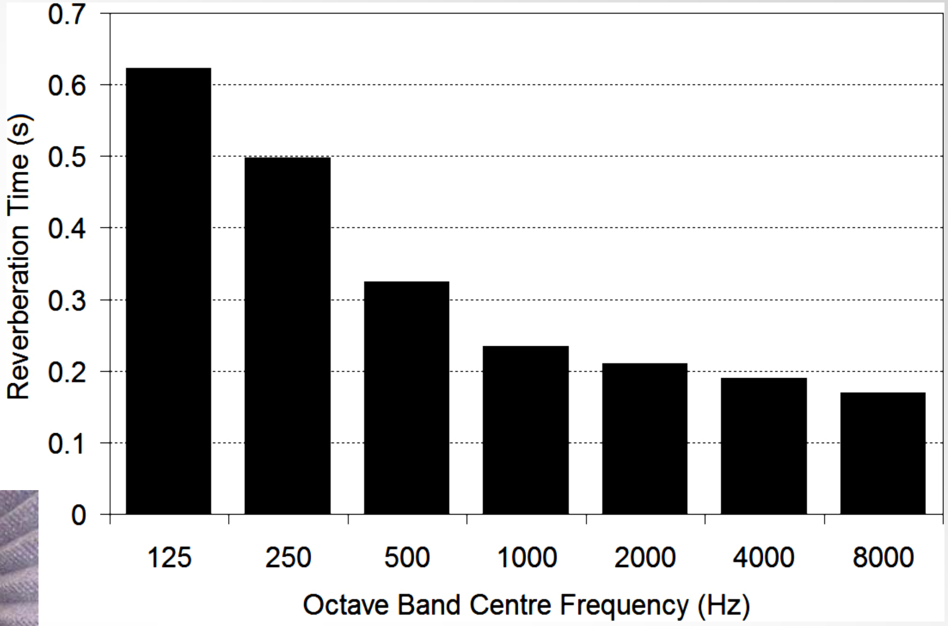
# THE DUAL STEREO DIPOLE




Reproduction over a Dual-Stereo-Dipole loudspeaker rig

It is a four-channel system, in which a frontal stereo dipole is employed for reproducing the sound coming from directions located in the frontal hemispace, and the rear stereo dipole reproduces the sound coming from the rear hemispace.

# Reverberation Time of the listening room

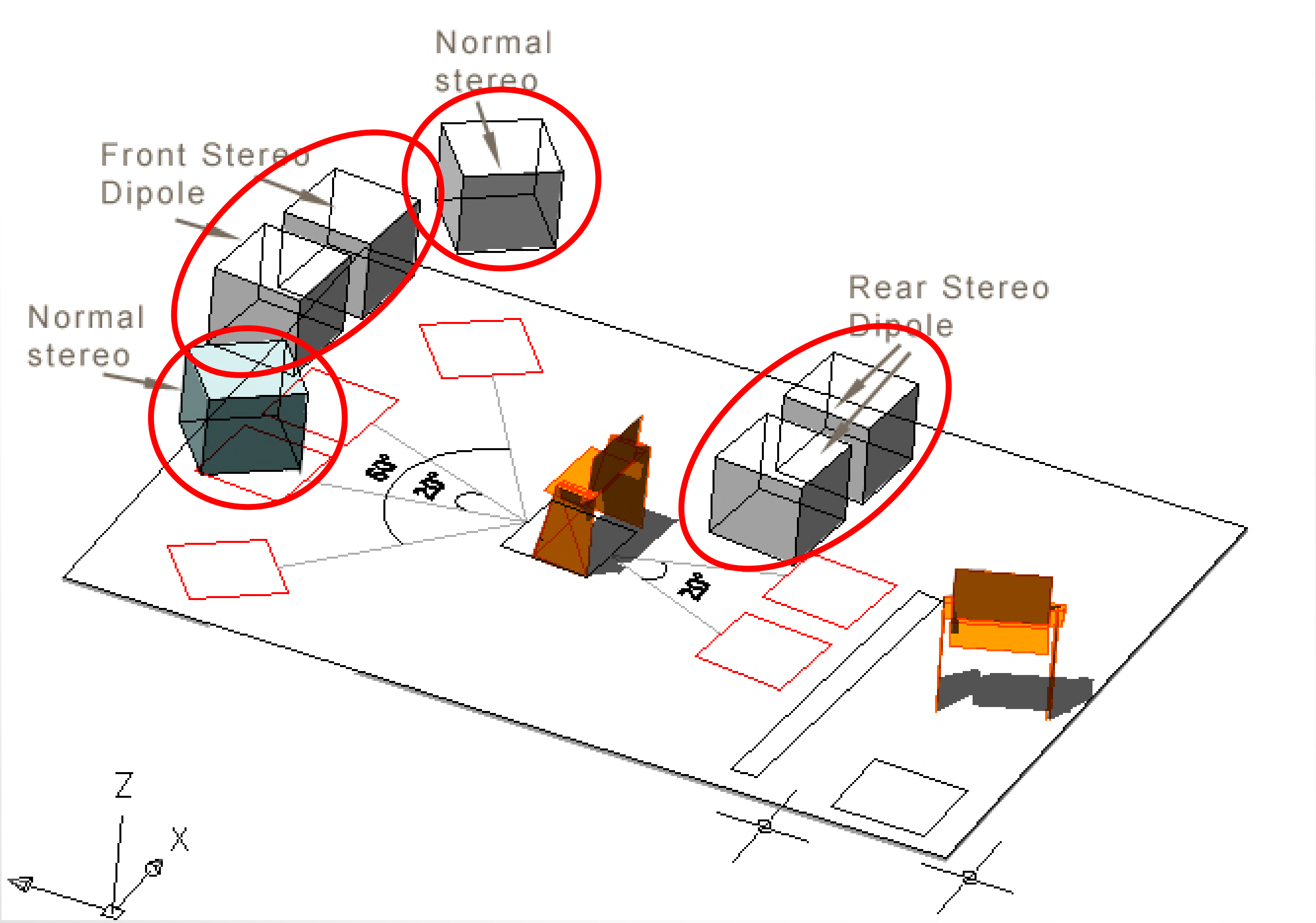



# Speakers positioning

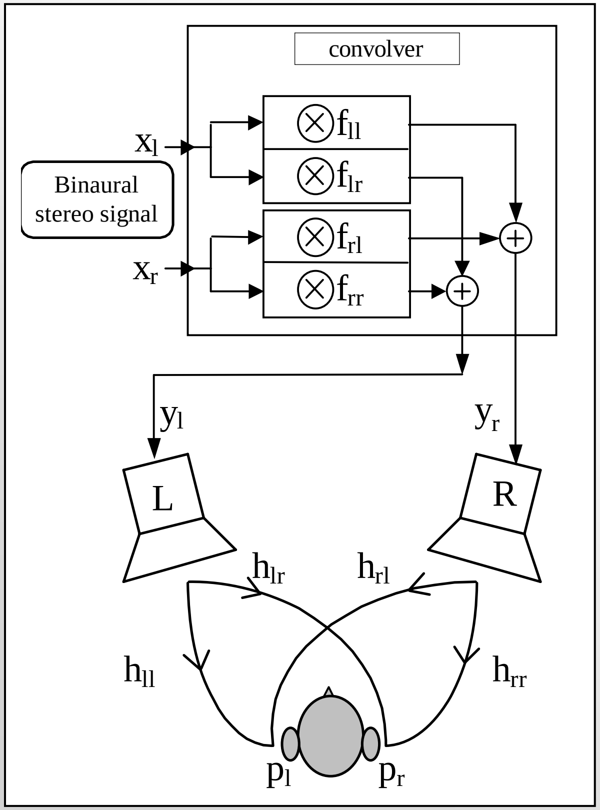
Genelec S30D

Dynaudio

QSC AD-S82H



# Design of cross-talk canceling filters




- First, a binaural measurement is made in front of the Stereo Dipole loudspeakers

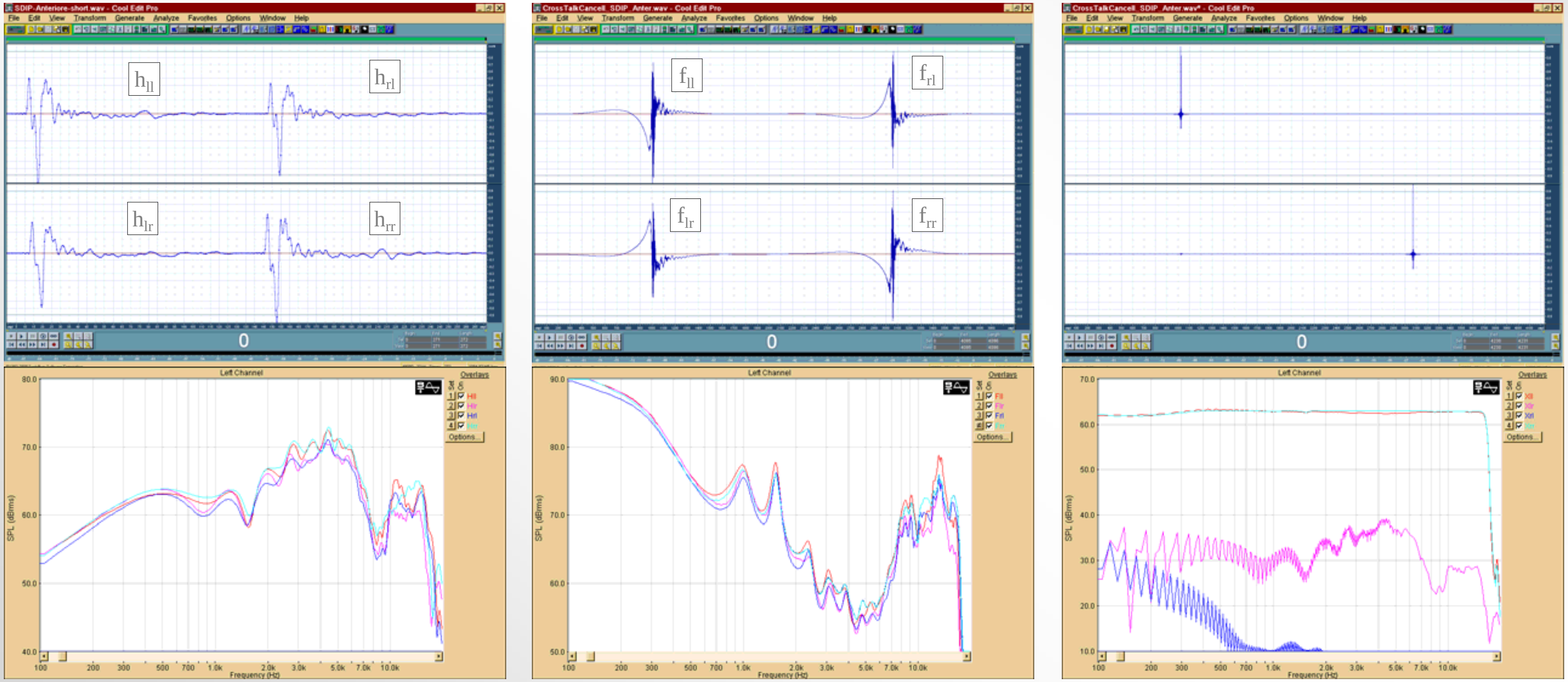
- Then, the cross-talk cancelling filters are computed, so that their convolution with the measured impulse responses reduces to the identity matrix

# Stereo Dipole inverse filters

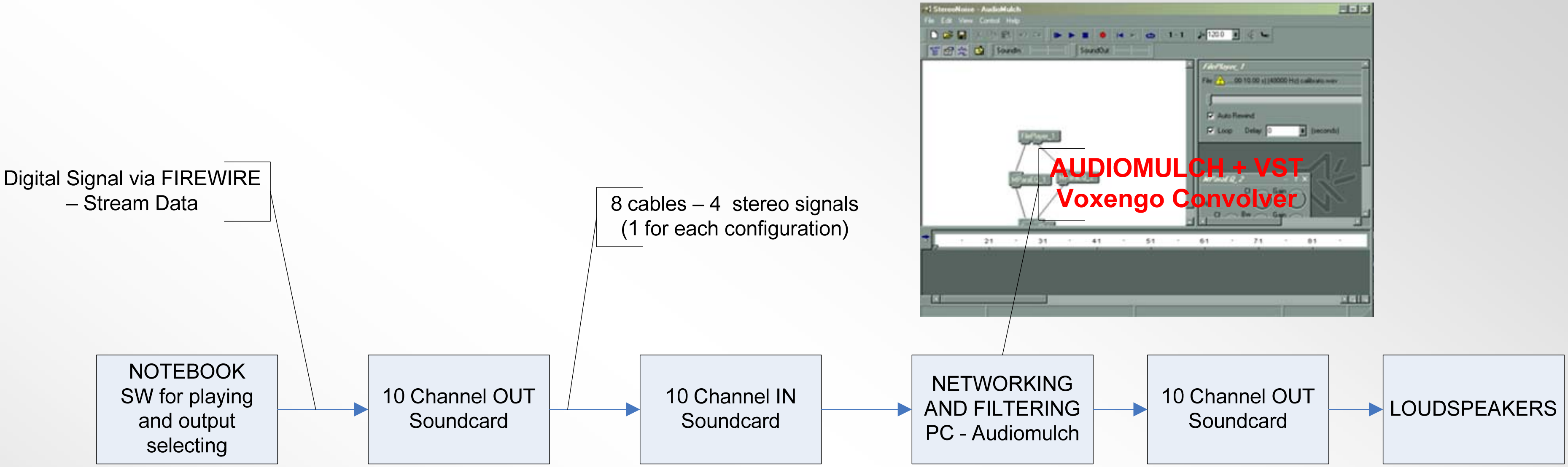
$$\begin{cases} f_{ll} = (h_{rr}) \otimes InvDen \\ f_{lr} = (-h_{lr}) \otimes InvDen \\ f_{rl} = (-h_{rl}) \otimes InvDen \\ f_{rr} = (h_{ll}) \otimes InvDen \\ InvDen = InvFilter(h_{ll} \otimes h_{rr} - h_{lr} \otimes h_{rl}) \end{cases}$$

$$C(\omega) = FFT(h_{ll}) \cdot FFT(h_{rr}) - FFT(h_{lr}) \cdot FFT(h_{rl})$$

$$InvDen(\omega) = \frac{Conj[C(\omega)]}{Conj[C(\omega)] \cdot C(\omega) + \varepsilon(\omega)}$$



# Hardware and Software setup

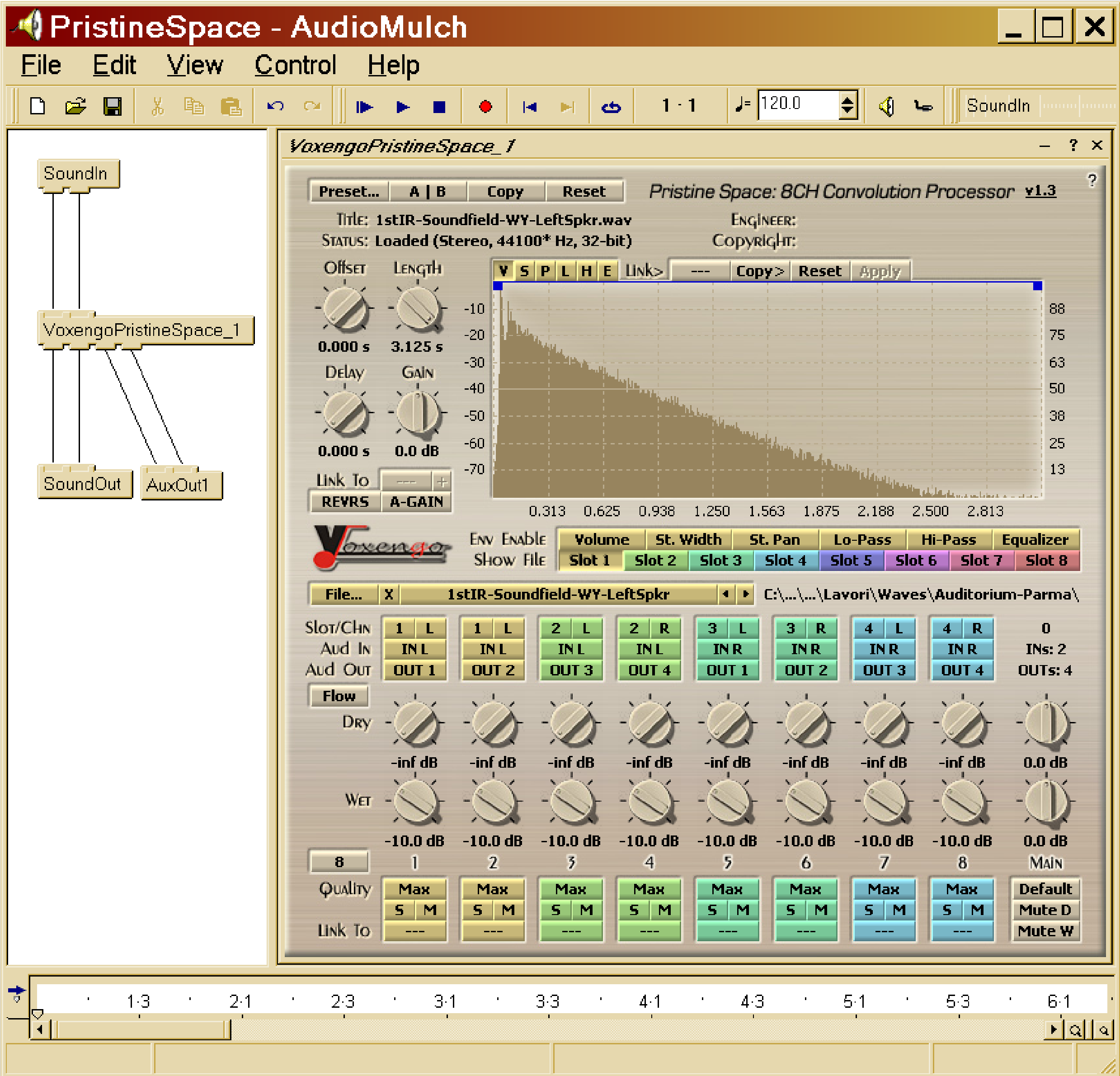



- Anechoic tracks preconvolved with IRs of the theatres at the beginning of the chain

- Filtering through Audiomulch in real time

- Voxengo Pristine Space VST plugin employed as multichannel convolver

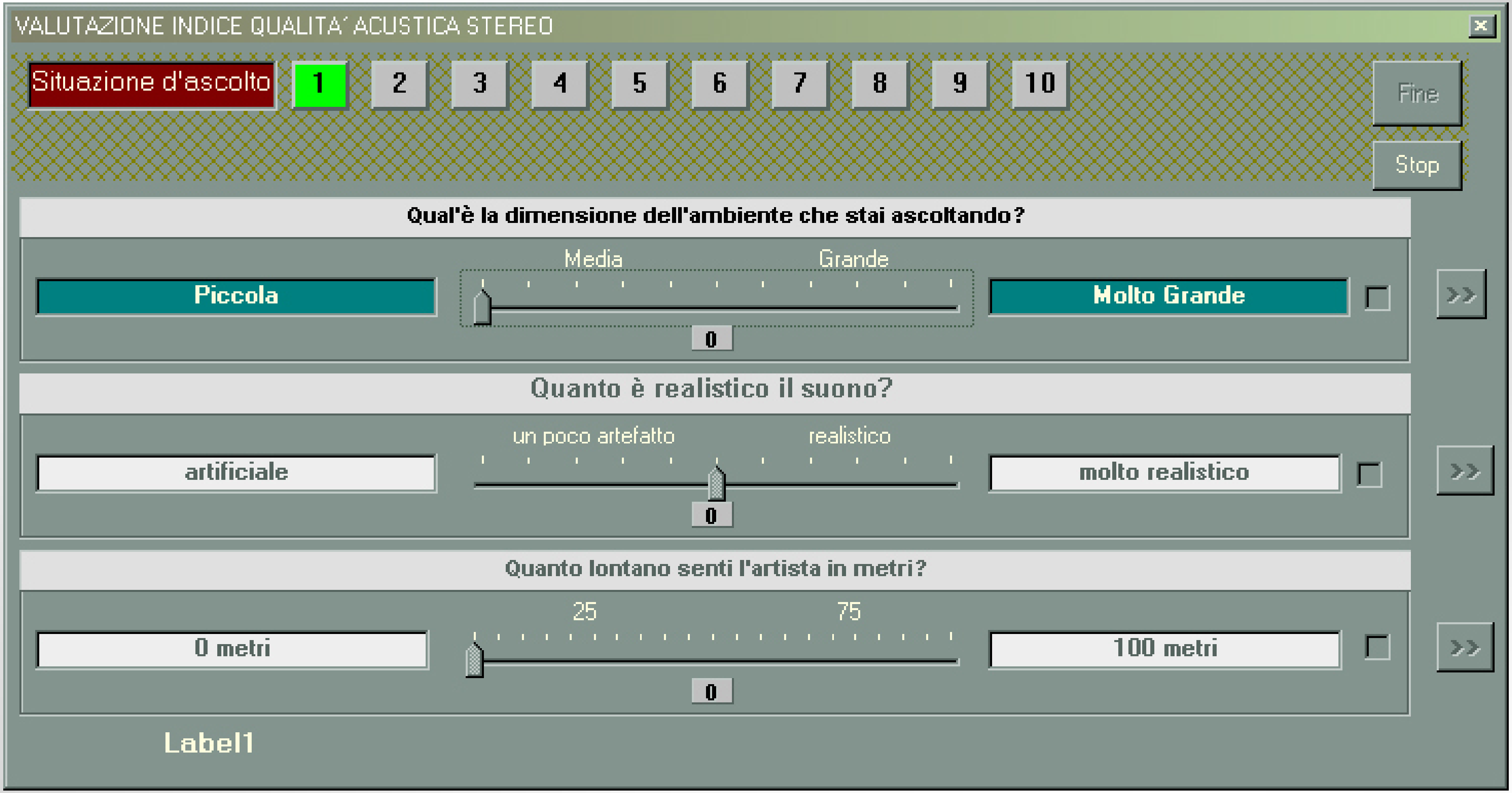
# The Voxengo Pristine Space multichannel convolver




- Filtering through Audiomulch in real time

- Voxengo Pristine Space VST plugin employed as multichannel convolver

# Software for collecting the answers




- Listening and simultaneous evaluation of 10 sounds, differing for position inside the theatre and system of reproduction

# GOALS of the new subjective experiment

- Evaluation of spatiality and distance in auditoria through auralized tracks and different reproduction systems

- Evaluation of the best system for a realistic reproduction

- The subjective tests are undergoing in these months, the results will be published in 2005

# Conclusions (preliminary yet)

• The loudspeakers are generally judged more natural than the headphones

• The frontal (single) stereo dipole is the system providing the better result/effort ratio: it works with a normal stereo system, it does not require a lot of computational power for running the inverse filters in real time, and is generally very natural and provides good localization

• Adding the second stereo dipole gives some advantages, but generally they are not worth the extra effort required (4-channels processing, etc.)

• The normal stereo gives very poor localization, and the perceived spectrum changes with the direction of the sound source, so the "colour" of the sound is not preserved

• It is very important to provide a confortable user's interface: large LCD screen, wireless mouse, comfortable seat, proper lighting, thermohygrometric confort.

# The Future

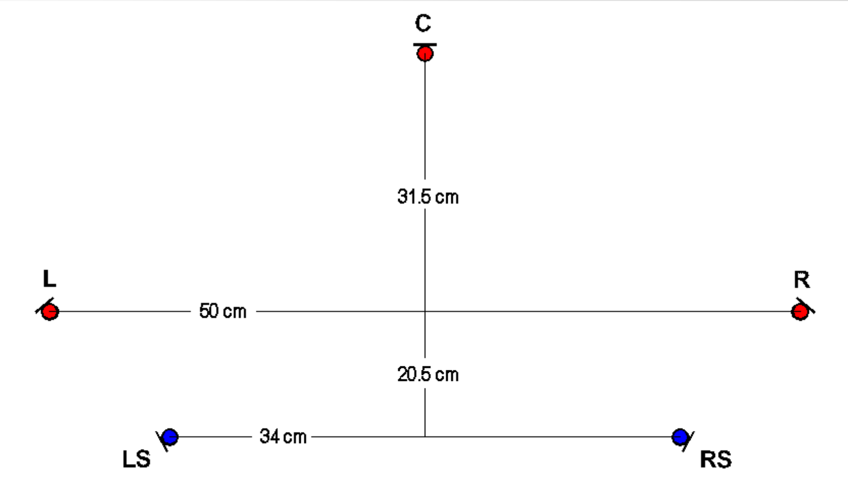
# Future improvement of the listening room

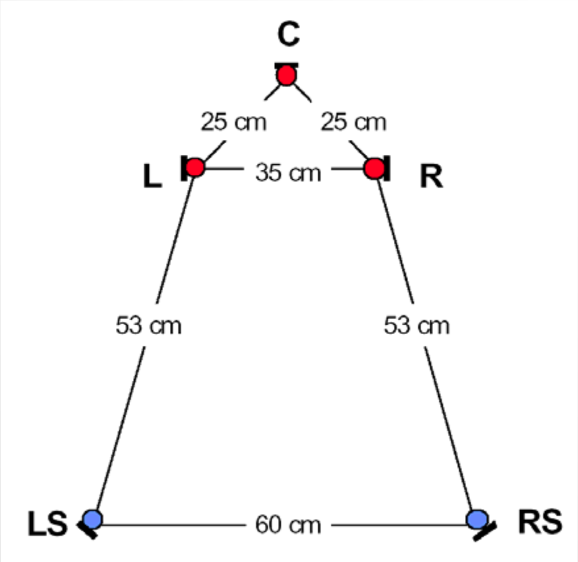
- Multichannel (surround) sound systems, including:
  - ITU 5.1 horizontal surround (2D) through Dolby AC3 or DTS
  - Advanced 7.1 horizontal surround (2D) through DTS-ES or Microsoft WMA
  - 1°-order periphonic (3D) Ambisonics (8 loudspeakers)
  - Hybrid Stereo Dipole + Ambisonics = Ambiophonics (3D)
- A single computer running both the playback tool and the filtering tool
- Large LCD widescreen display (also good for films!)
- Fanless, completely silent liquid-cooled computer
- 16 playback channels through an Apogee DAC 16 converter and an RME Hammerfall soundcard with 2 ADAT optical outputs

# ITU 5.1 surround

- Williams MMA

**Schematic of the setup**
C : Cardioid, 0°
L, R : Cardioid, ± 40°
LS, RS : Cardioid, ± 120°




- INA-5

**Schematic of the setup**
C : Cardioid, 0°
L, R : Cardioid, ± 90°
LS, RS : Cardioid, ± 150°

# ITU 5.1 surround

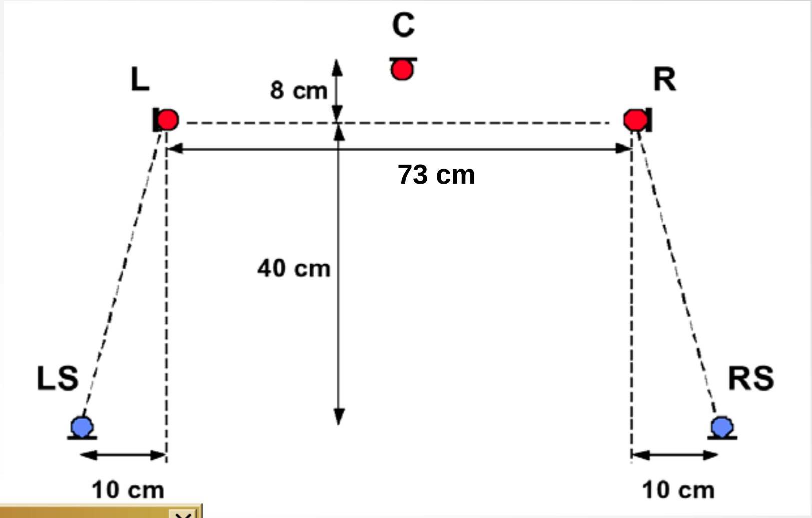
## OCT

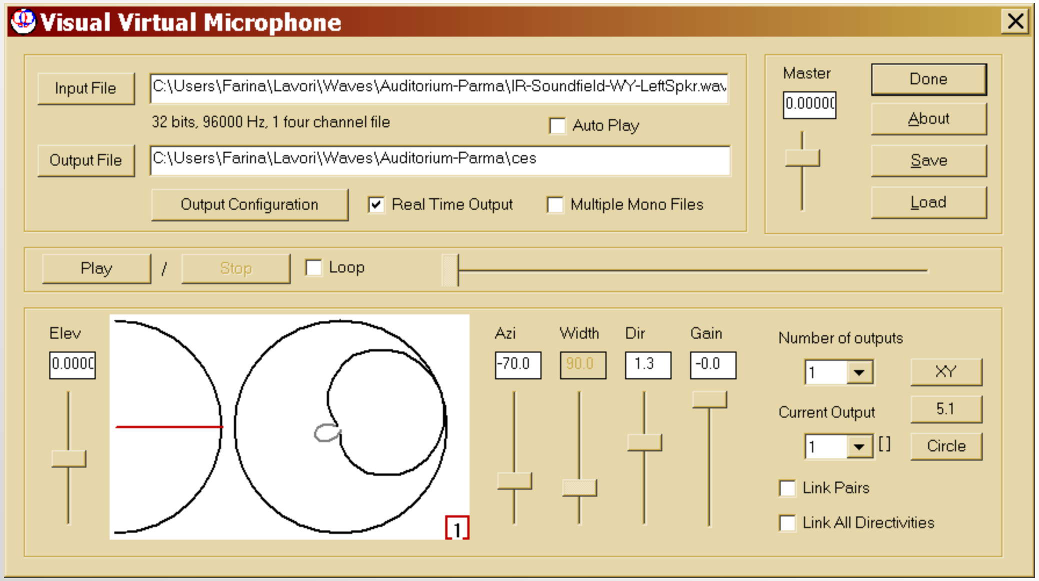
**Schematic of the setup**

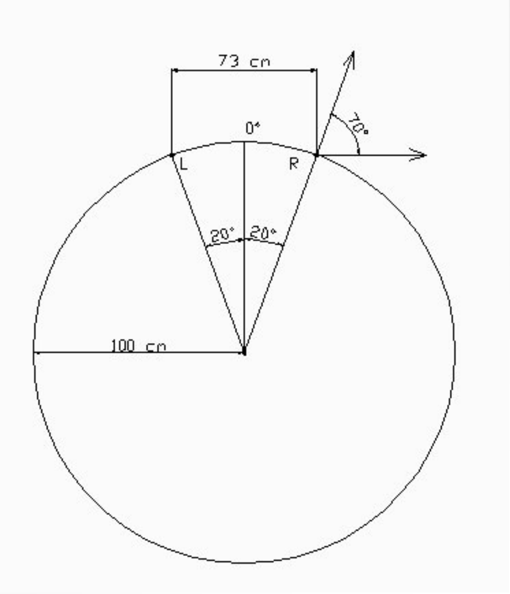
C : Cardioid, 0°

L, R : Super Cardioid, ± 90°

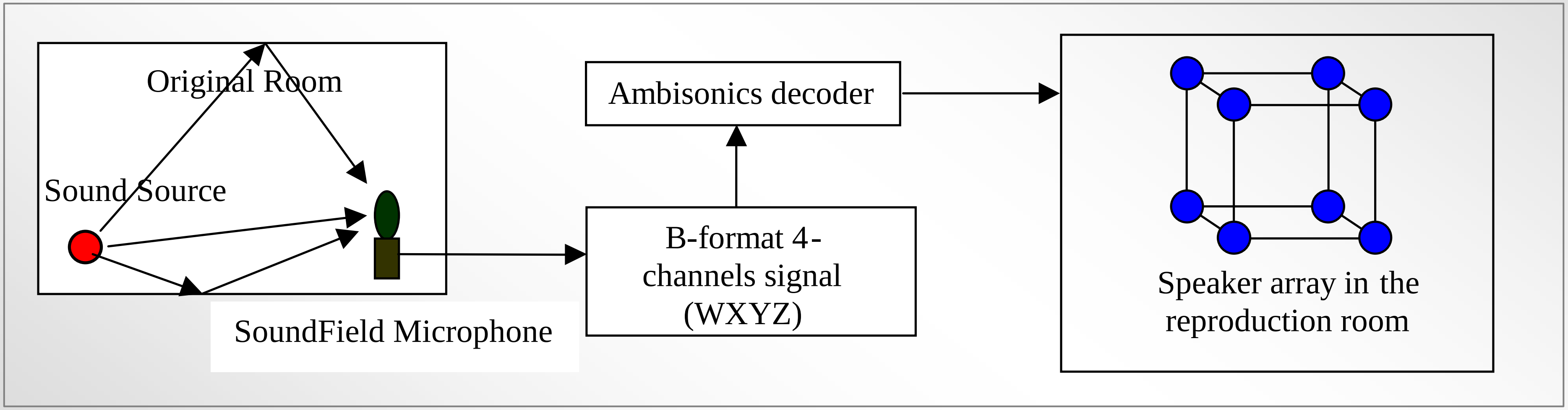
LS, RS : Cardioid, ± 180°



# Ambisonics 3D 1st order

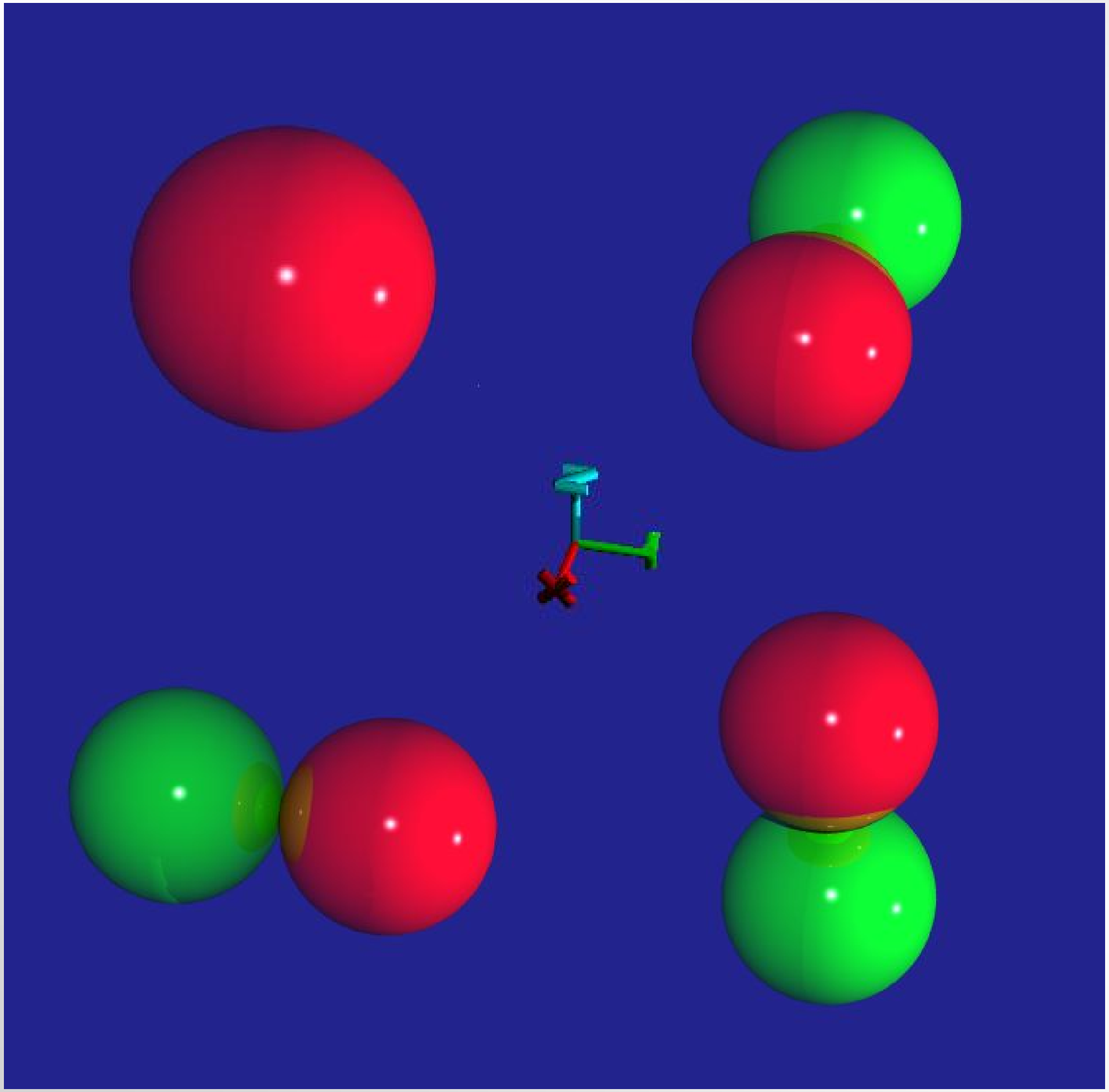



Reproduction occurs over an array of 8-24 loudspeakers, through an Ambisonics decoder

# The Soundfield microphone

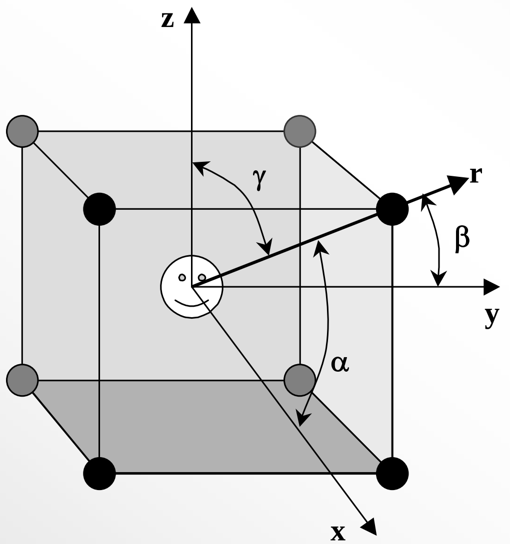
- This microphone is equipped with 4 subcardioid capsules, placed on the faces of a thetraedron
- The signal are analogically processed in its own special control box, which derives 4 “B-format” signals
- These signals are:
  - W : omnidirectional
  - X,Y,Z : the three figure-of-eight microphones aligned with the ISO cartesian reference system

# The B-format components

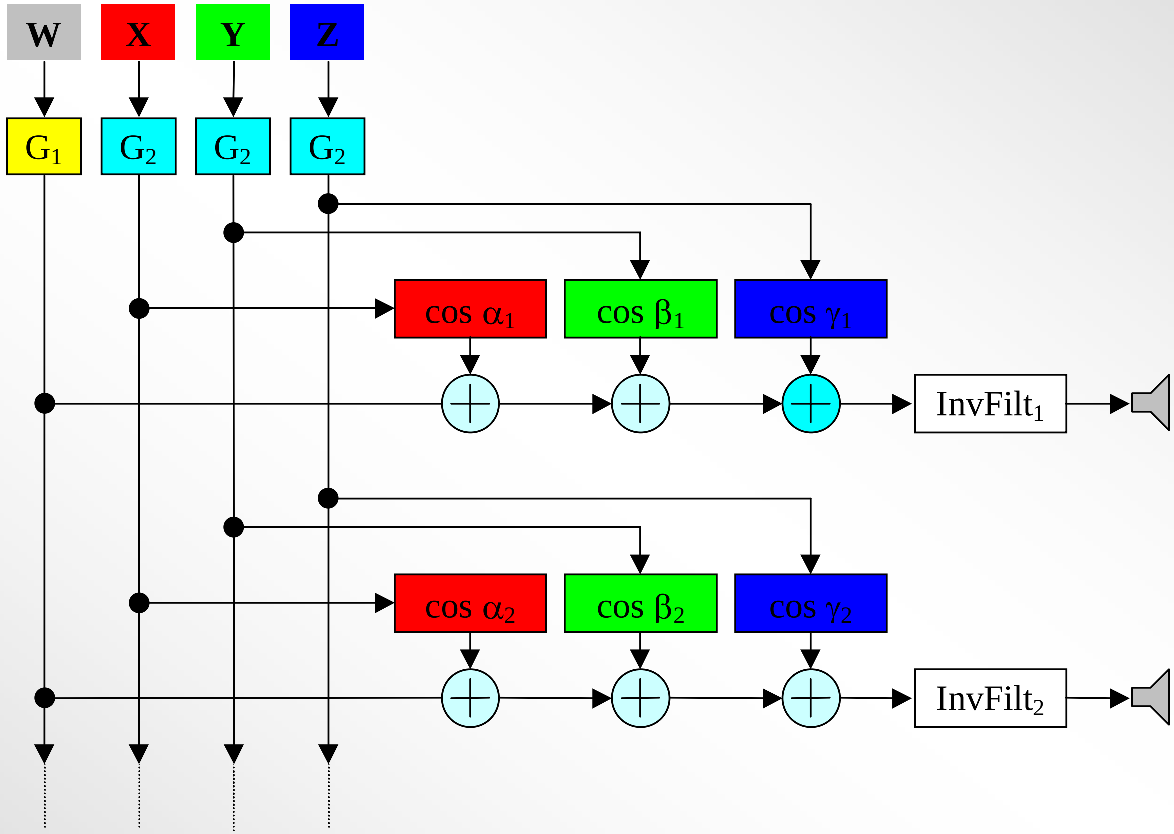



- Physically, W is a signal proportional to the pressure, XYZ are signals proportional to the three Cartesian components of the particle velocity
- when a sound wave impinges over the microphone from the “negative” direction of the x-axis, the signal on the X output will have polarity reversed with respect to the W signal

# Ambisonics decoding



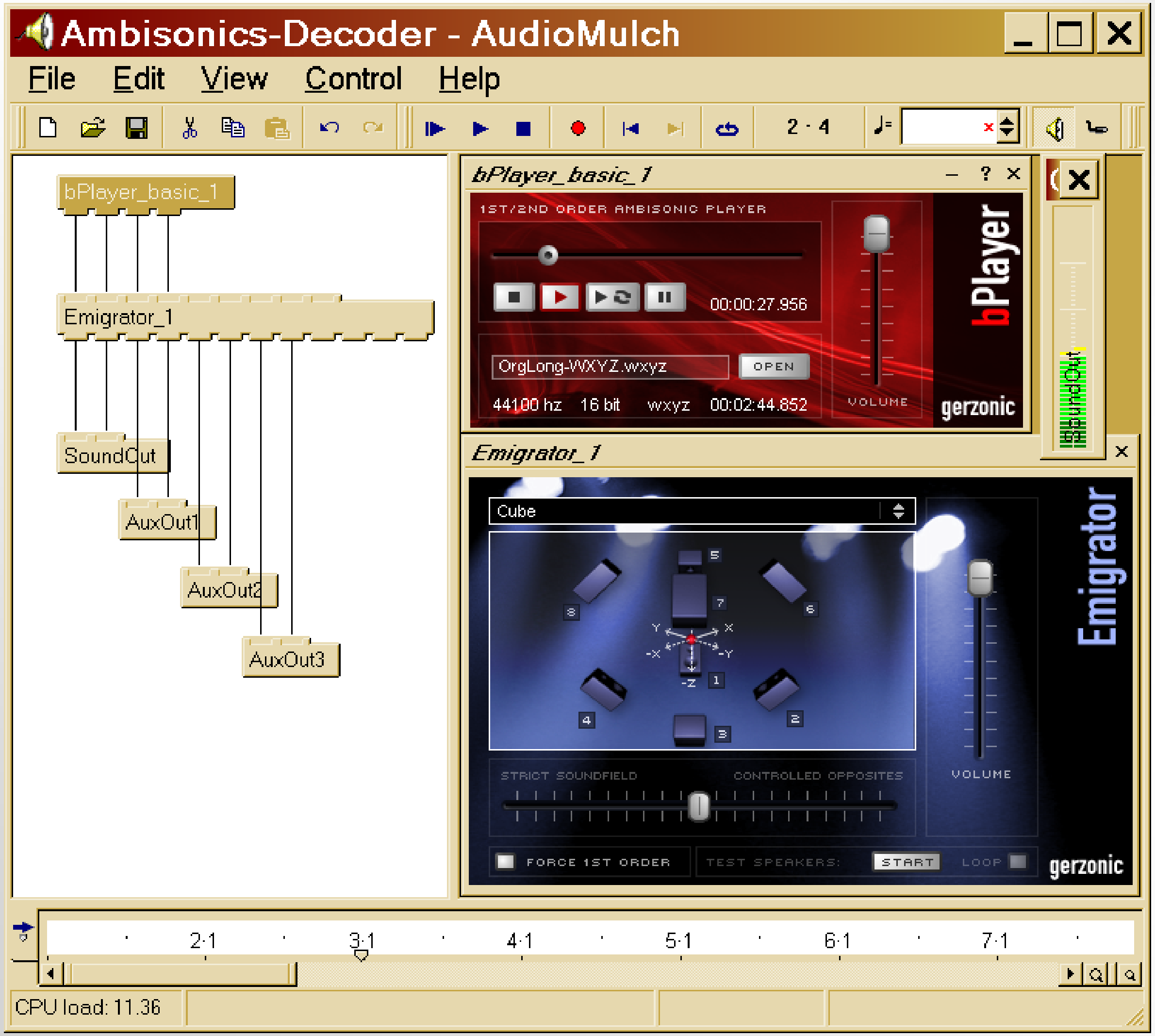

$$F_i = \frac{1}{2} \cdot \left[ G_1 \cdot W + G_2 \cdot \left( X \cdot \cos(\alpha) + Y \cdot \cos(\beta) + Z \cdot \cos(\gamma) \right) \right]$$




Each speaker feed is simply a weighted sum of the 4 B-format signals.

The weighting coefficients are computed by the cosines of the angles between the loudspeaker and the three Cartesian axes

# A software Ambisonics decoder




Audiomulch VST host

Gerzonic bPlayer

Gerzonic Emigrator

# Rooms for Ambisonics 3D 1st order

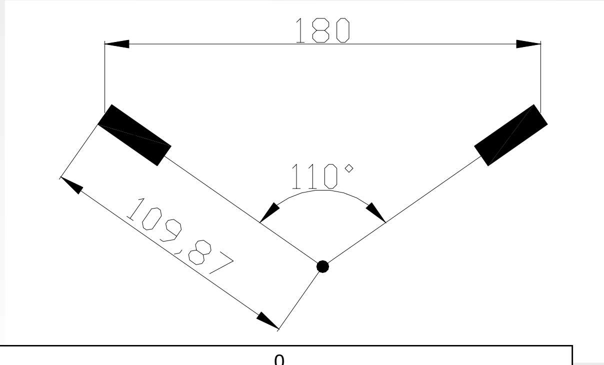
ASK – Reggio Emilia

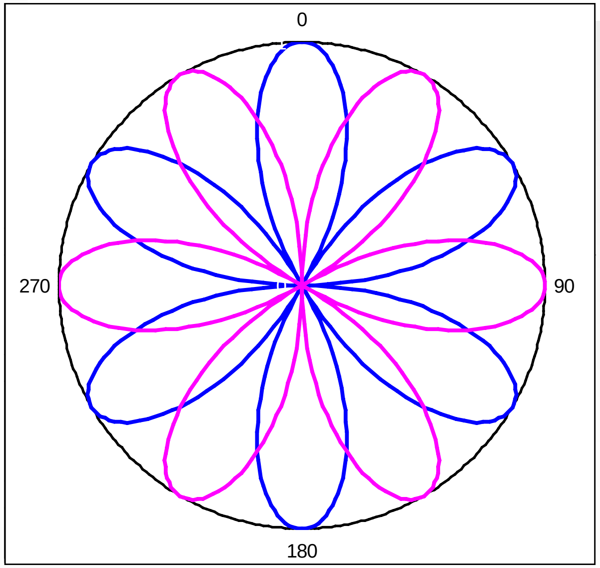
University of Bologna

University of Ferrara

# Virtual high-order microphones (M. Poletti)

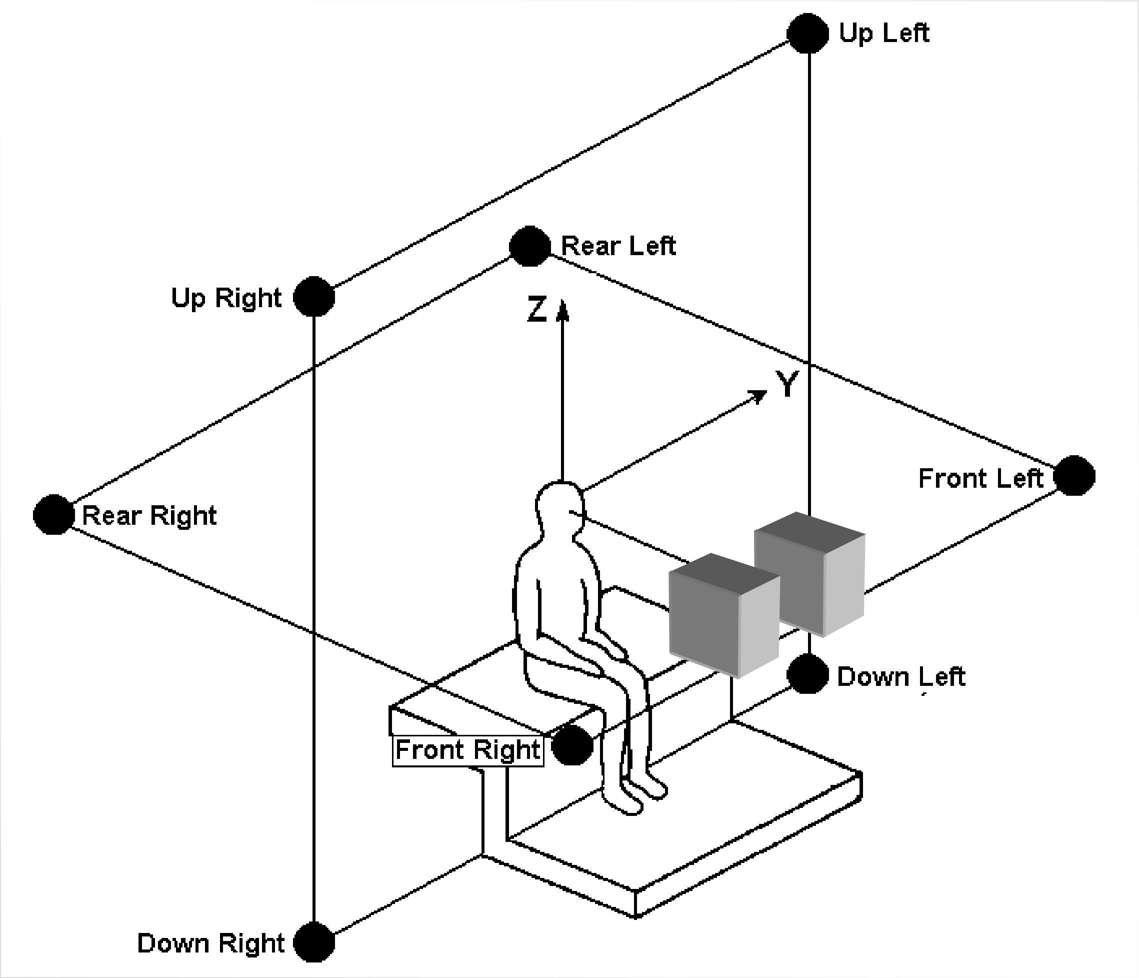
One of the two ORTF cardioid is employed, which samples 36 positions along a 110 mm-radius circumference



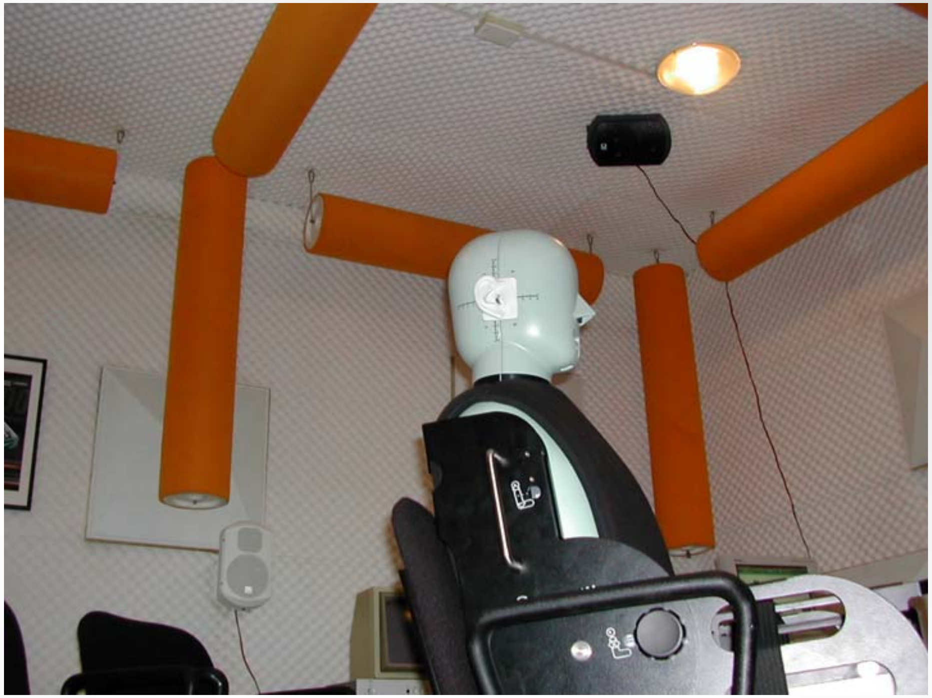




From these 36 impulse responses it is possible to derive the response of cylindrical harmonics microphones (2D Ambisonics) up to 5th order.

# Ambiophonics 3D (10 loudspeakers):

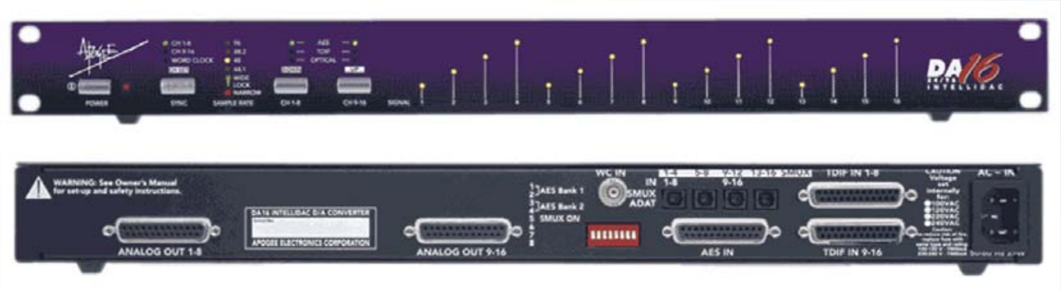


# Ambiophonics Room at ASK



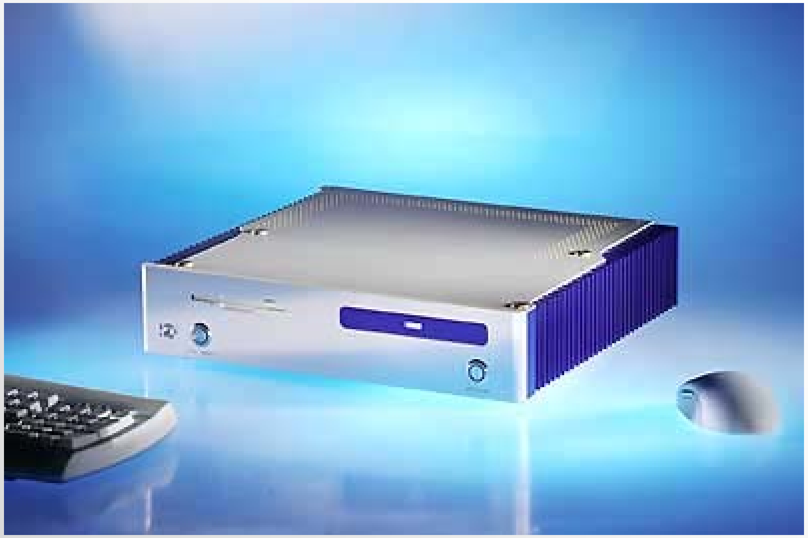
# Ralph Glasgal’s Ambiophonics Room at the Ambiophonics Institute

# New hardware tools



Apogee DAC 16

RME Hammerfall soundcard



Liquid-cooled FutureClient PC

Toshiba 23” LCD TV/Monitor

# Internet resources

All the papers previously published by Angelo Farina can be downloaded from his personal web site:

**www.angelofarina.it**

The software system employed for this research is based on the following modules:

Adobe Audition (**www.adobe.com**)

Aurora Plugins (**www.aurora-plugins.com**)

Audio Mulch (**www.audiomulch.com**)

Voxengo Pristine Space (**www.voxengo.com**)