Computer simulation of Binaural, Stereo-Dipole, B-format and Ambiophonics impulse responses

Angelo Farina (1), Lamberto Tronchin (2)

(1) Industrial Engineering Dept., University of Parma, Via delle Scienze 181/A
43100 Parma, ITALY – **HTTP://pcfarina.eng.unipr.it**

(2) DIENCA-CIARM, University of Bologna, Via Risorgimento 42
40136 Bologna, Italy – **HTTP://ciarm.ing.unibo.it**

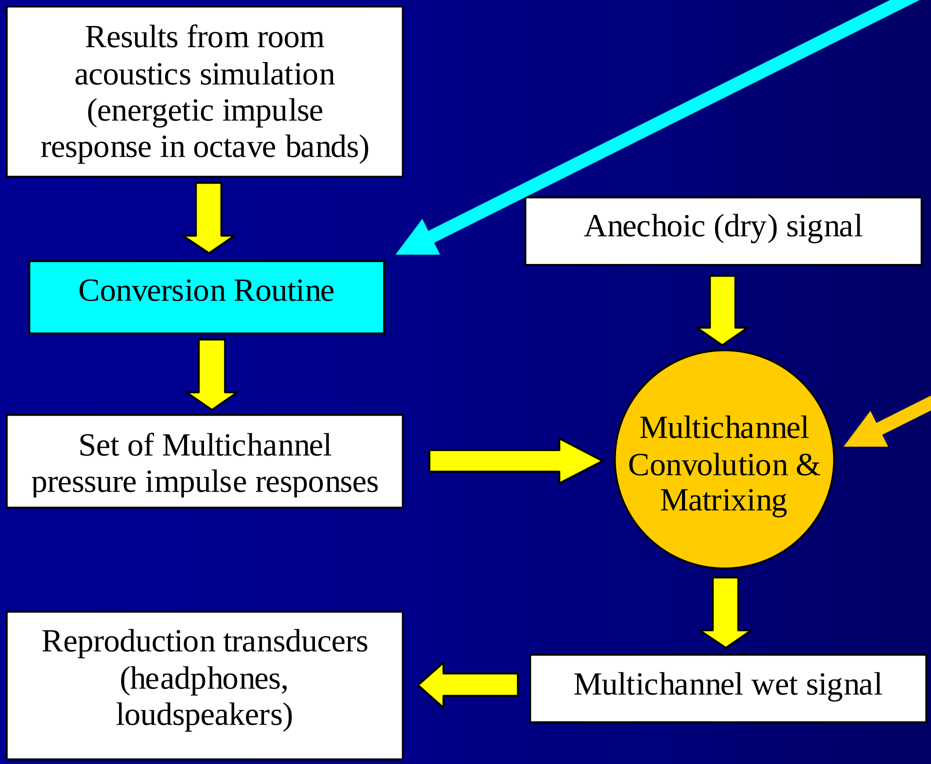
Goals

- This paper describes the processing required for obtaining a **realistic audible sound reproduction** from the results of a geometrical room acoustics program
- This overall process is usually known as “Auralization”, and traditionally is performed through the binaural technology (headphone reproduction)
- Here the process is generalized to many more reproduction systems: Mono, Binaural, Stereo Dipole, Ambisonics and Ambiophonics.
- The last three systems are loudspeaker-based

Methods

- In all these 5 cases, anyway, the auralization is obtained by means of real-time convolution of dry signals with properly derived impulse responses
- Convolution is now possible in real time, with many simultaneous channels, directly on a low cost PC without any added hardware, thanks to available free sofware, which outperforms traditional DSP-based convolvers
- Consequently, this paper focuses on the derivation of the proper sets of impulse responses for each reproduction method, starting from the results of the geometrical room acoustic program.

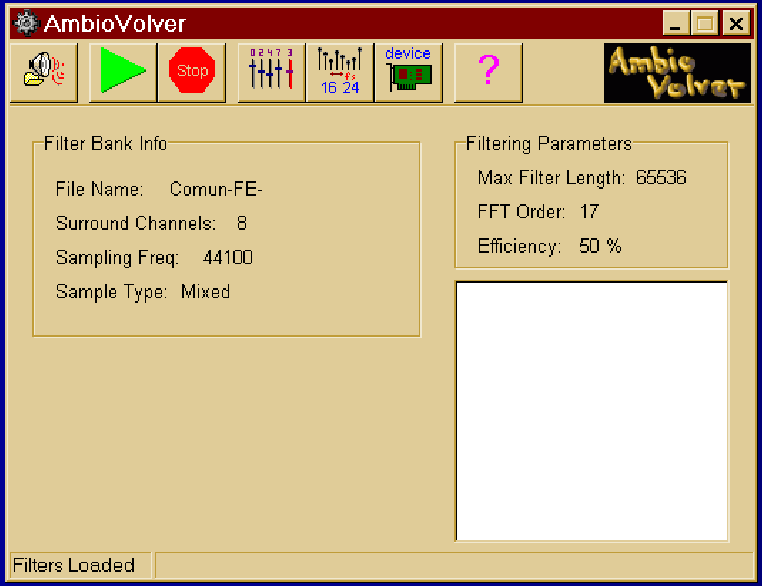
Processing path of the results of a geometrical room acoustic simulation



- The Conversion Routine is the main topic of this presentation
- The convolution engine is nowadays based on free software running on the PC platform

Software convolvers

- The multichannel convolution can be done for free on a low cost PC nowadays. Two solutions are currently available:



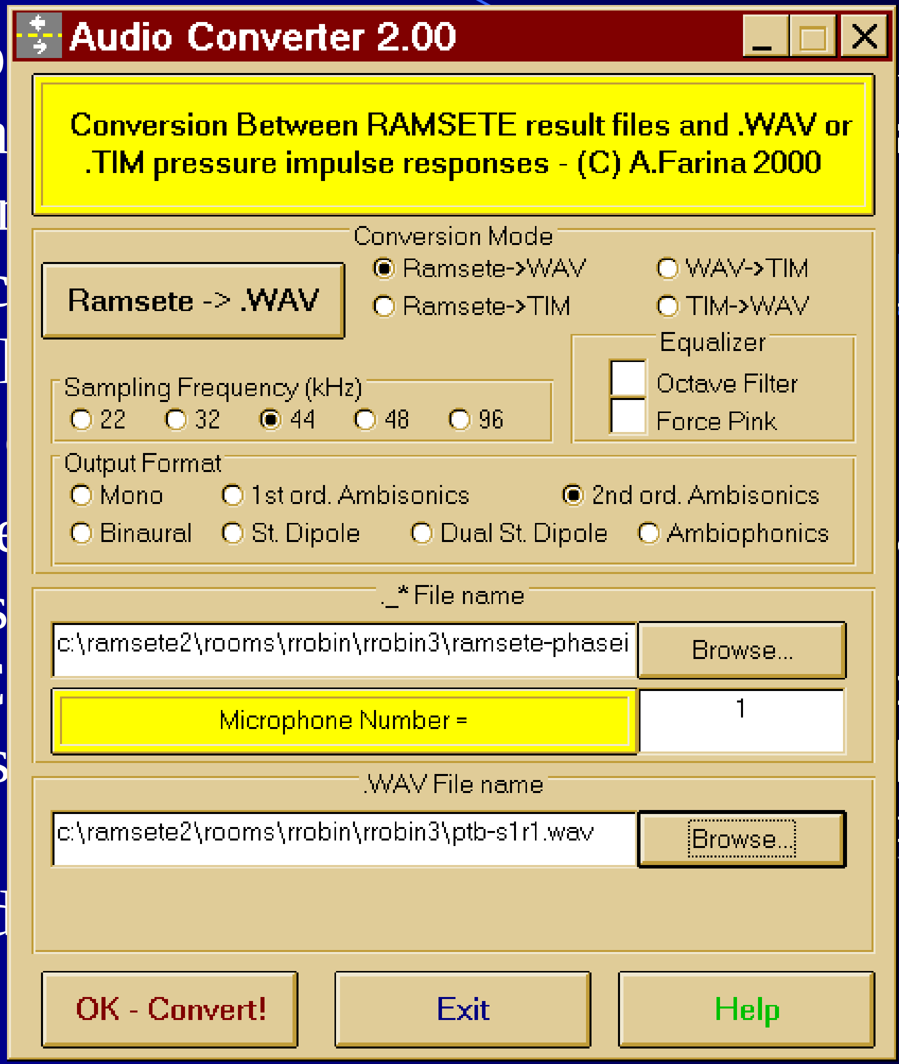
Ambiovolver is a program running on any Windows PC equipped with a multichannel soundboard

BruteFIR is an highly optimized Linux program running on a customized, noiseless computer with a digital RME audio interface

Internet links to these softwares are in the last slide

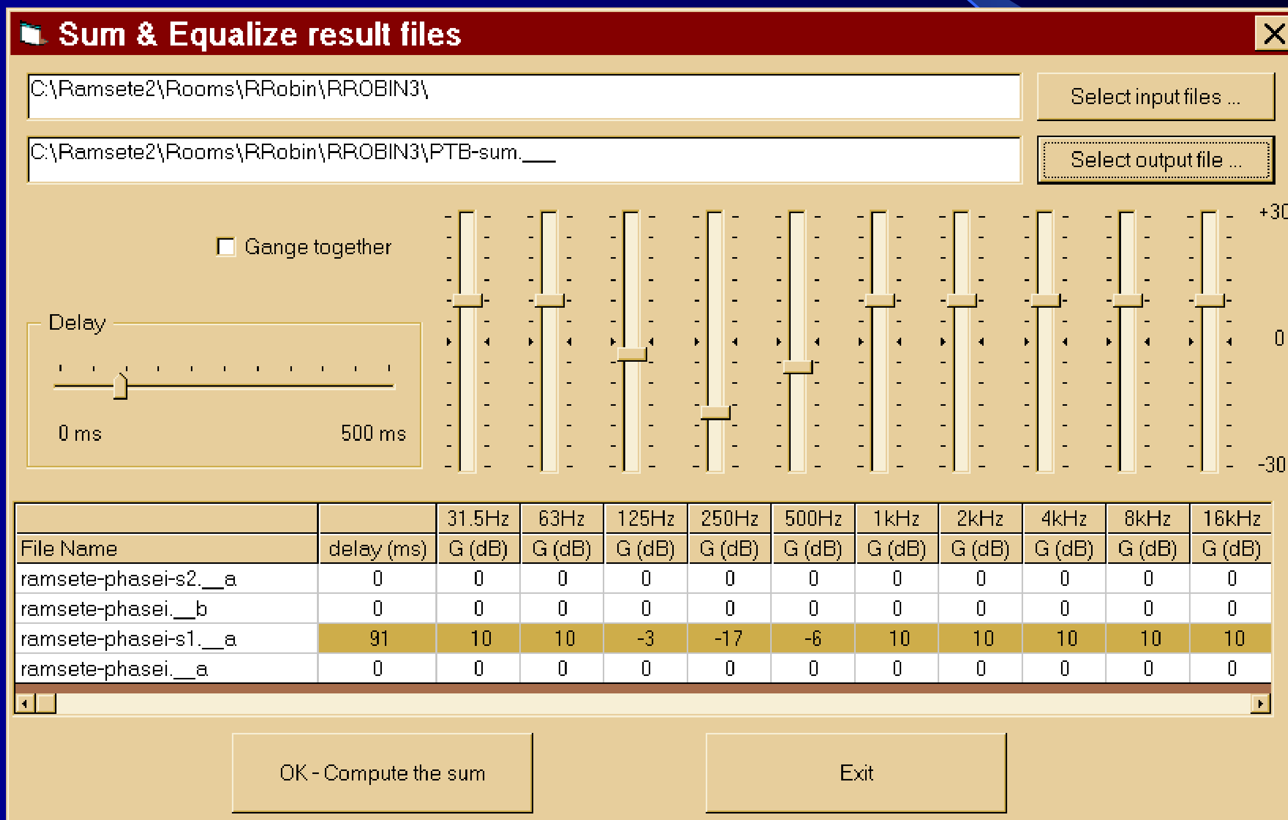
The conversion routine

- This piece o… les produced by the room a… amsete 2), and process the… set of impulse responses, c… uld be recorded experimental… a multichannel microphone …
- Actually the… veloped for the release vers… se microphonic standards: C… ono, 1st and 2nd order Ambis… Dipole (Kemar), Dual Stereo … Pinnaless sphere dummy head…



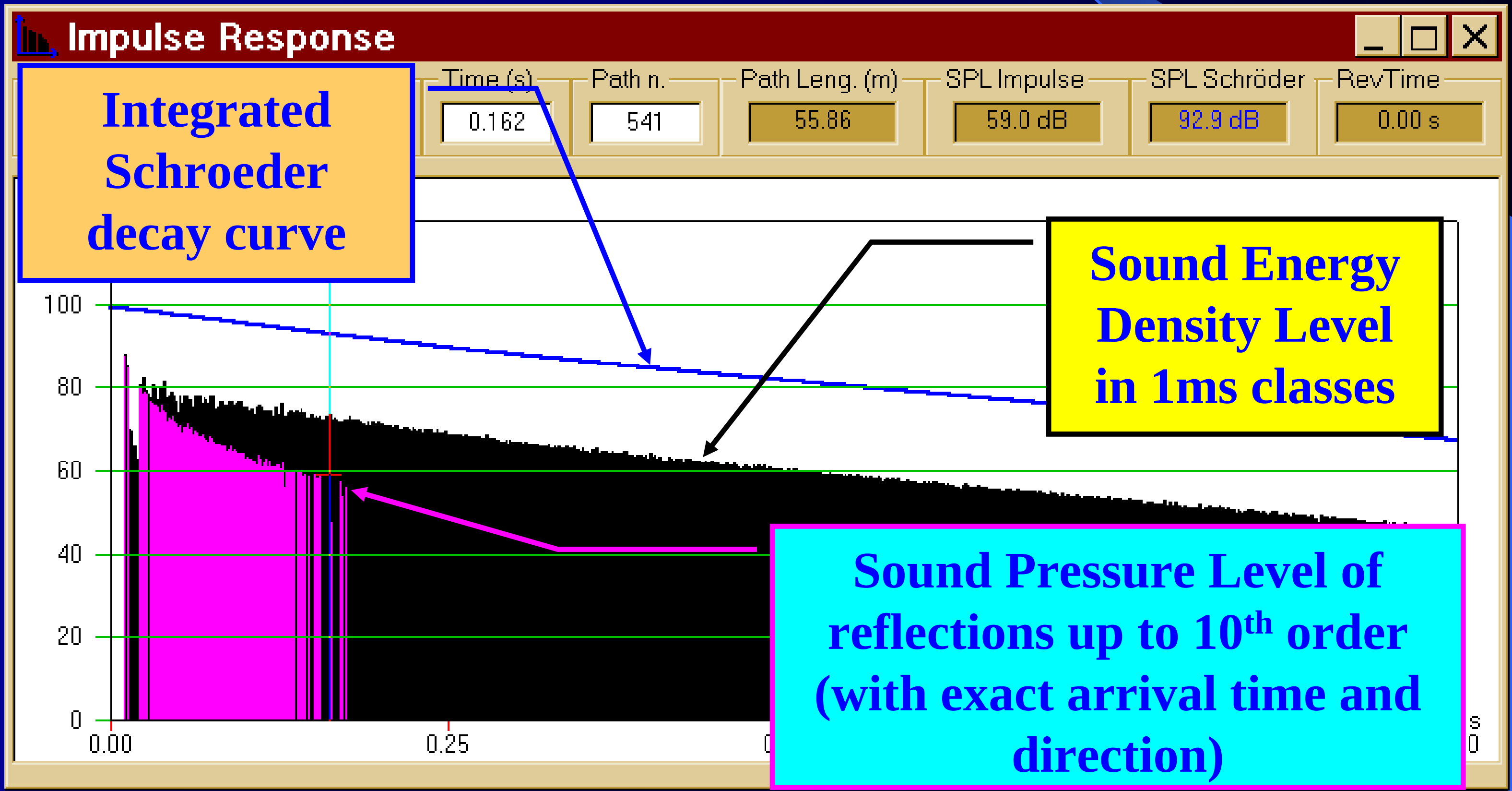
Combination of multiple sources

- Ramsete can combine the results of several sound sources emitting the same signal, with optional delay and equalization (multi-source sound reinforcement systems):



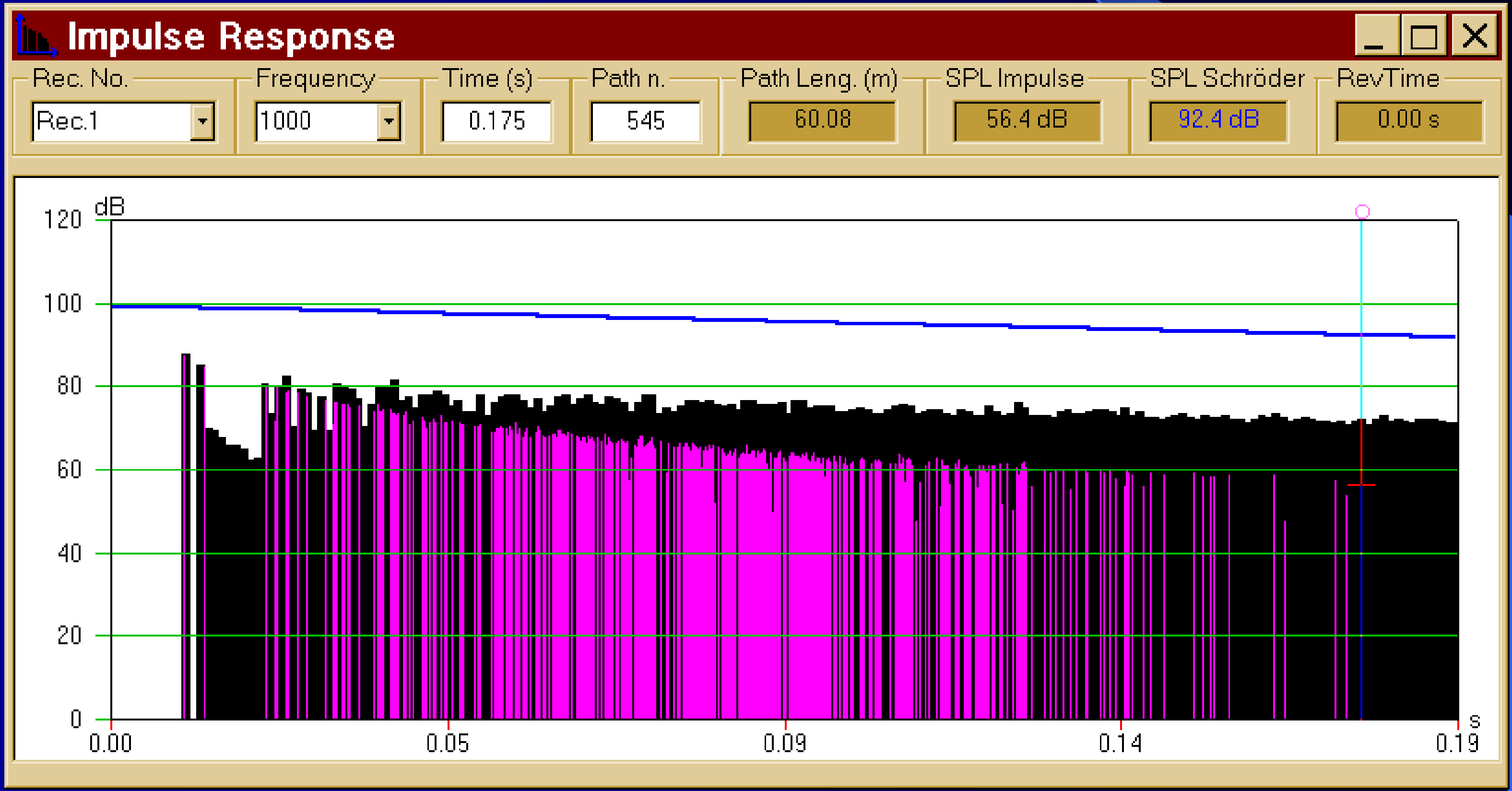
Results from the Ramsete program

- In each receiver point, Ramsete computes an energetic impulse response in ten octave bands for each sound source (or combination of multiple sources):



Detail of the initial part of the IR

- The temporal density of the reflections initially grows (theoretically with the square of time), but later it starts reducing and eventually vanishes.



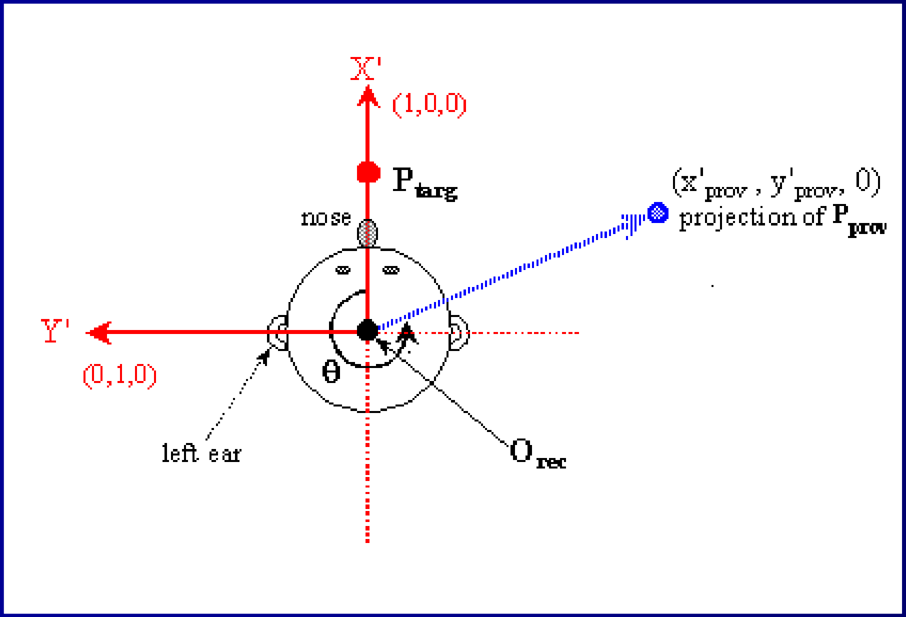
Rationale of the conversion process

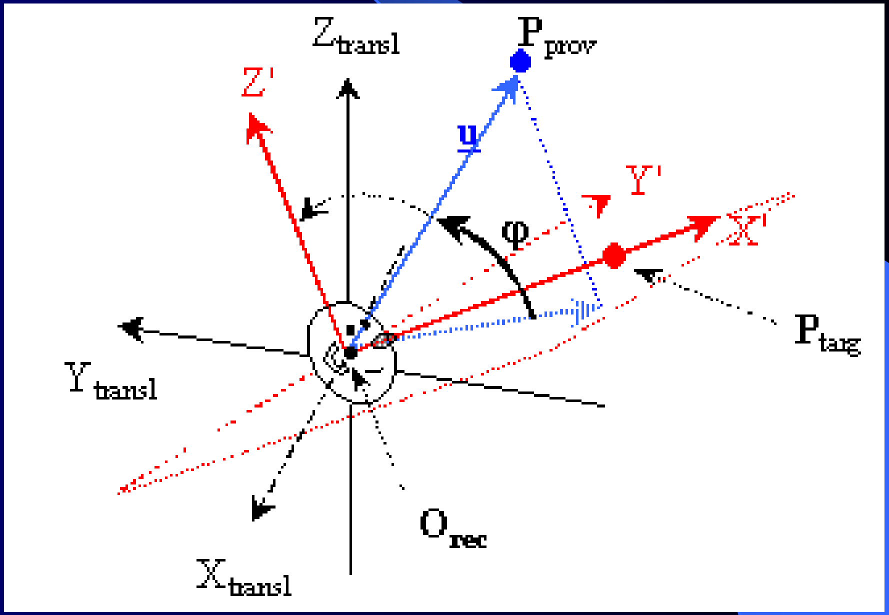
- First the discrete early reflections are processed, taking into account their known arrival direction and exact timing.
- For each discrete arrival, a Dirac's delta is generated at the exact arrival time,
- then it is convolved with the impulse response of an octave-band equalizer which imposes the proper SPL value in the 10 octave bands,
- and finally it is convolved with the multichannel impulse response of the selected type of microphone, chosen depending on the direction of arrival.

- Second, the subsequent reverberant tail is added, based on the whole energetic impulse response data (which do not contain any directional information), after subtraction of the energy of the discrete reflections (already processed).
- An independent sample of white noise for each channel of the virtual microphone is generated
- The white noise sample is splitted in ten octave-band filtered versions
- Each of them is amplitude-modulated with the square root of the energetic IR in the corresponding octave band
- Finally the wide-band response is obtained simply summing together the ten filtered signals

Reference system definition

- A local cartesian reference system is assumed solidal with the listener head; X’ axis is pointing forward (nose), Y’axis is pointing on the left ear, and the Z’ axis towards the top of the head.





We are searching for the local arrival angles of the incoming ray: azimuth (θ) and elevation (φ)

Conversion formulas between absolute points and local angles

- The Ramsete program saves the coordinates of three points in the absolute reference system:

– $\mathbf{P}_{\mathrm{prov}} \equiv (\mathbf{x}_{\mathbf{prov}}, \mathbf{x}_{\mathbf{prov}}, \mathbf{x}_{\mathbf{prov}})$ = provenience point of the ray;

– $\mathbf{O}_{\mathrm{rec}} \equiv (\mathbf{x}_{\mathbf{rec}}, \mathbf{y}_{\mathbf{rec}}, \mathbf{z}_{\mathbf{rec}})$ = receiver origin;

– $\mathbf{P}_{\mathrm{targ}} \equiv (\mathbf{x}_{\mathbf{t}}, \mathbf{y}_{\mathbf{t}}, \mathbf{z}_{\mathbf{t}})$ = receiver target point.

First of all, the position of the provenience point of the ray is recomputed in the local reference system:

$\mathbf{P}_{\mathrm{prov}} \equiv> \mathbf{P'}_{\mathrm{prov}} (\mathbf{x'}_{\mathbf{prov}}, \mathbf{y'}_{\mathbf{prov}}, \mathbf{z'}_{\mathbf{prov}})$

Conversion formulas between absolute points and local angles

- Elevation angle φ :

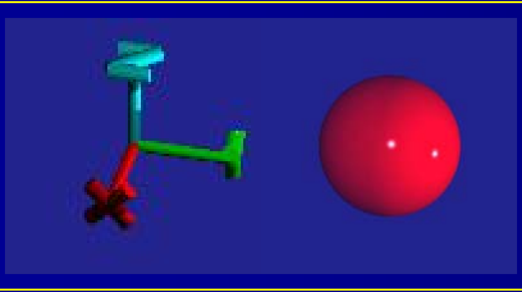
$$\varphi = 90^\circ - \arccos\left[z'_{prov} \Big/ \sqrt{\left(x'_{prov}\right)^2 + \left(y'_{prov}\right)^2 + \left(z'_{prov}\right)^2} \right]$$

- Azimuth angle θ :

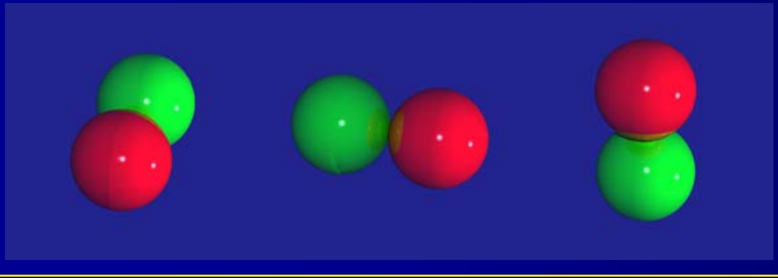
$$\theta = \arccos\left[x'_{prov} \Big/ \sqrt{\left(x'_{prov}\right)^2 + \left(y'_{prov}\right)^2} \right]$$

if $\mathbf{y'_{prov} < 0} \quad \Rightarrow \quad \boldsymbol{\theta = 360^\circ - \theta}$

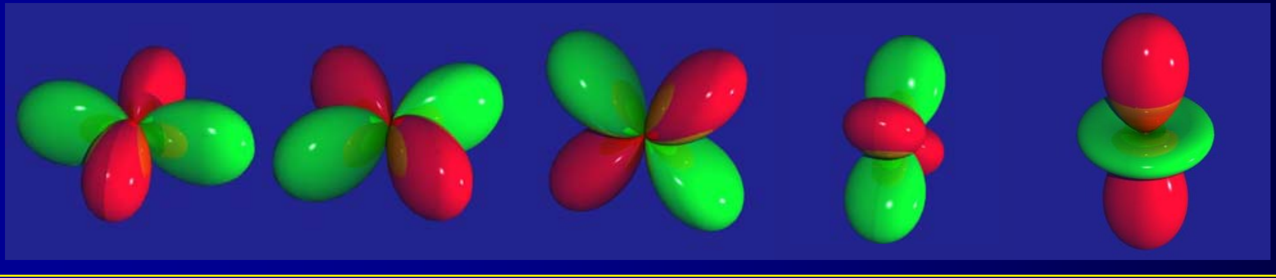
Ambisonics Impulse Responses



0th Order Ambisonics (W omnidirectional microphone, single-channel IR)



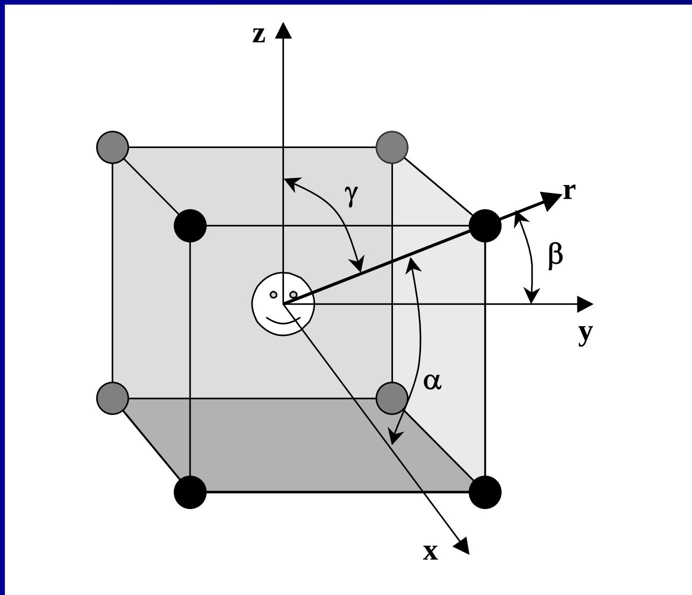
1th Order Ambisonics (XYZ pressure-gradient microphones, 4-channels IR)



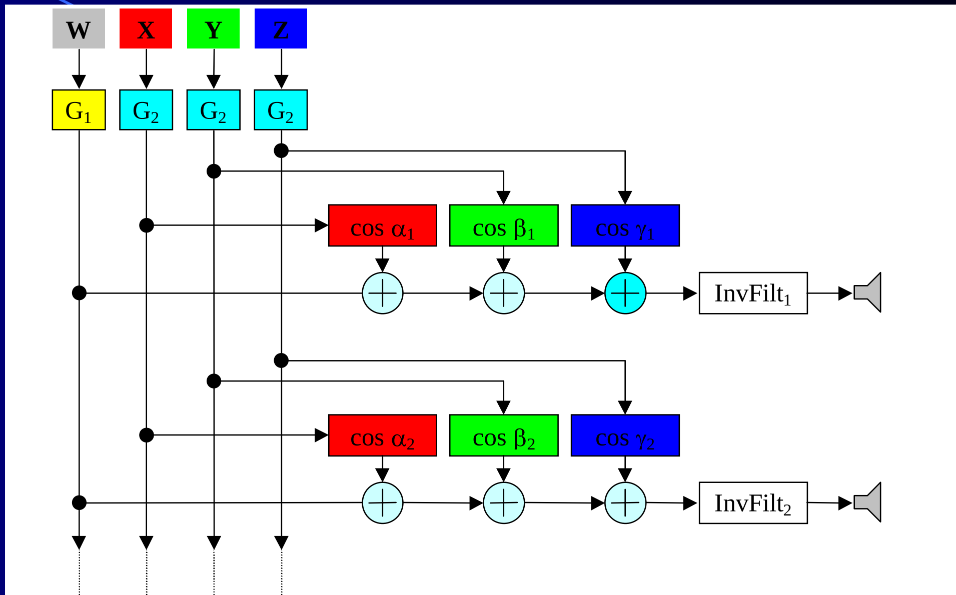
2nd Order Ambisonics (UVRST mikes)

- 0th Order can be listened to through a single loudspeaker
- 1st and 2nd orders must be properly matrixed, for driving an Ambisonics array of loudspeakers surrounding the listener

Ambisonics decoding



Geometry of the loudspeaker array



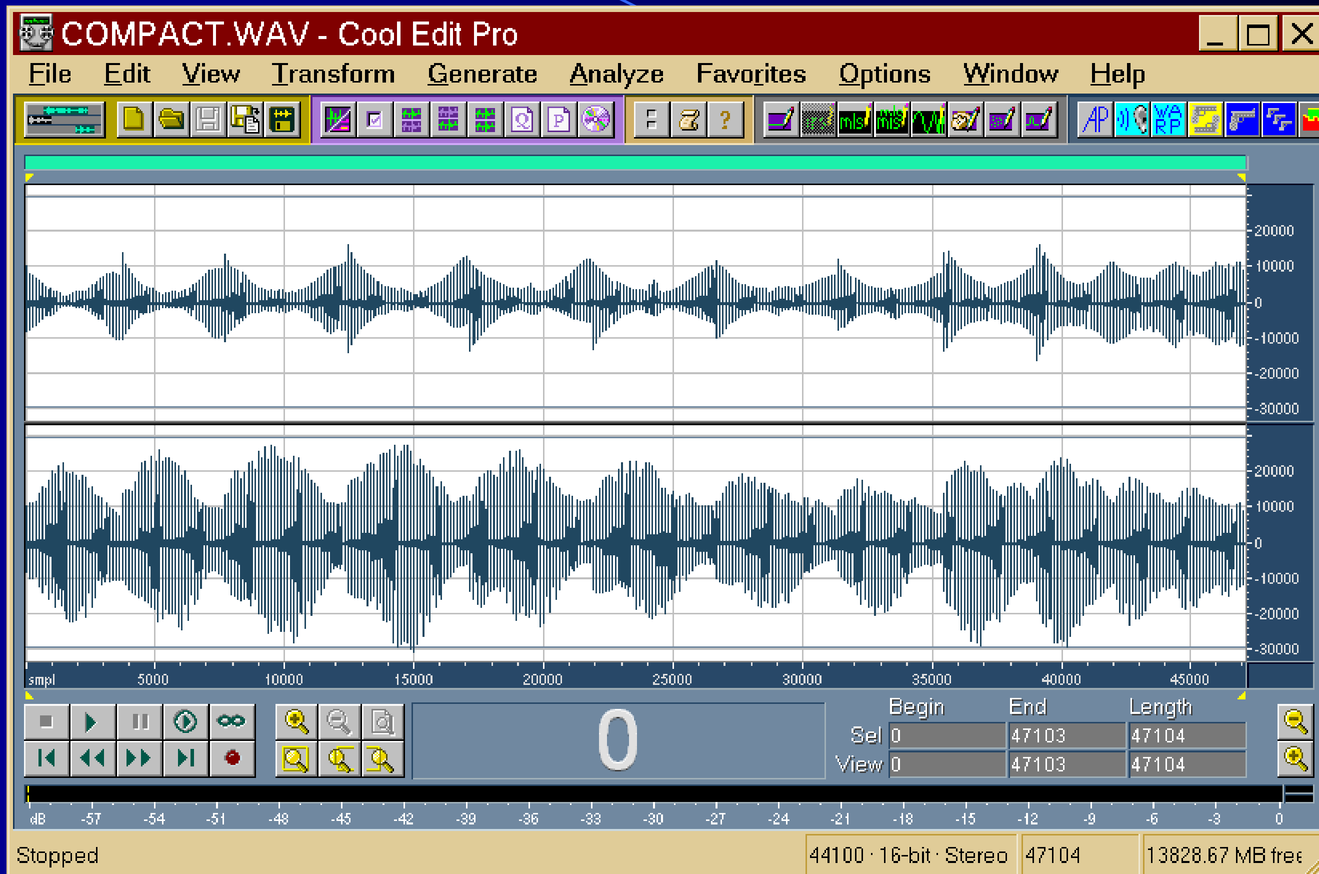
Flow diagram of the 1st order decoder

- Hardware decoders are obsolete nowadays.
- Free software is available for 1st and 2nd order real-time decoding, driving a multichannel sound board (thanks to Richard Furse)
- The decoding structure can be realized also by means of a multichannel convolution software (BruteFir, Ambiovolver, FIRreverb)

Binaural impulse responses

- In Ambisonics, the directive microphones IRs are simply matter of changing the gain (and perhaps the polarity) of a Dirac’s delta function
- In the binaural case, instead, a complex stereo IR is required for any direction-of-arrival
- A large data base of stereo binaural IRs (HRTF) is available from MIT-Medialab
- A proper routine has been developed for interpolating the required IR starting from the knowledge of the direction of arrival of each discrete early reflection.

Data-base of HRTF impulse responses



- The complete set of binaural IRs measured on the Kemar dummy head at MIT-Medialab

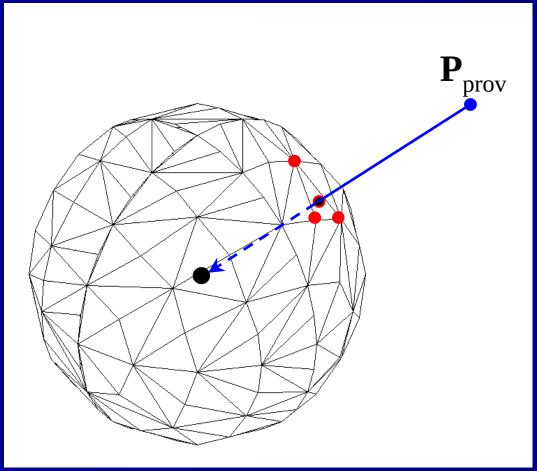
Geometry of the HRTF data-base

- Number of measurements at each elevation

Elevation φ (°)	Number of Measurements	Azimuth Increment (°)
± 90	1	***
± 80	12	30
± 70	24	15
± 60	36	10
± 50	45	8
± 40	56	6.43
± 30	60	6
± 20	72	5
± 10	72	5
0	72	5

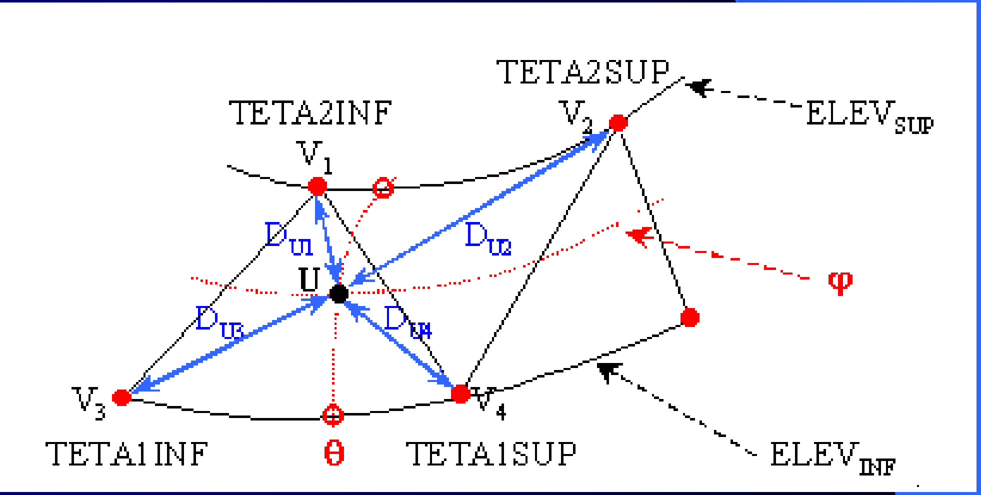
Thus: complex interpolation required on a spherical surface

Interpolation of the three nearest HRTFs



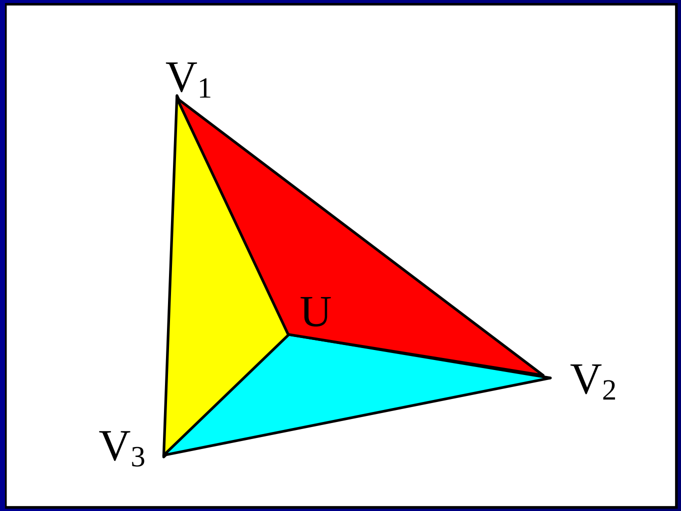
First, 4 possible nearest measurement points are located, choosing the values of φ and θ which are just lower and higher than the theoretical ones.

Then, one of the point is discarded, so that a triangle surrounding the arrival direction is defined



Evaluation of the weights

- the weight $\mathbf{P_i}$ (i = 1,2,3) relative to each HRTF is obtained calculating the opposite triangle area and dividing it for the total initial triangle area; so we have $\mathbf{P_1+P_2+P_3=}1$.



$$P_1 = \text{Area}(V_3,U,V_2)/A_{tot}$$

$$P_2 = \text{Area}(V_3,U,V_1)\ /A_{tot}$$

$$P_3 = \text{Area}(V_1,U,V_2)\ /A_{tot}$$

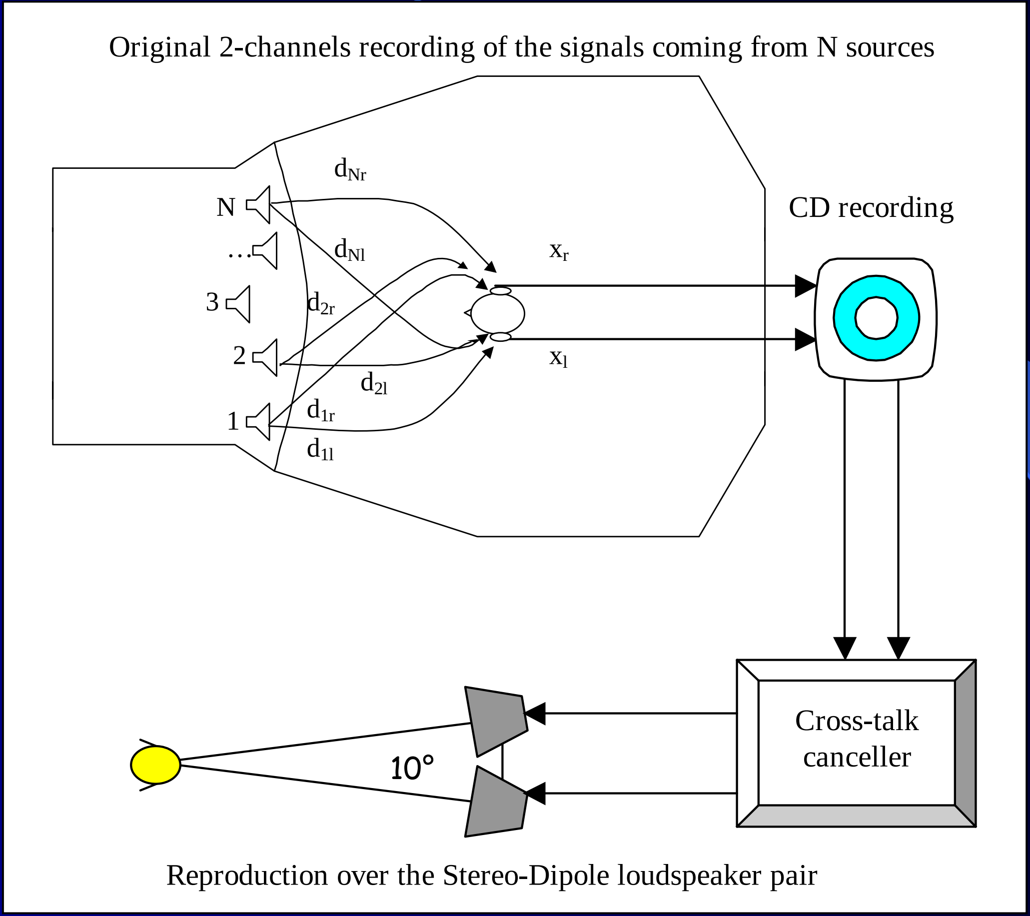
Computation of the binaural IR

- At the exact arrival time of the received ray, an averaged binaural IR is added to the global impulse response.
- The averaged IR is obtained by a frequency domain interpolation between the three HRTF complex spectra:

$$\mathrm{HRTF}(\varphi,\theta) = P_1 * \mathrm{HRTF}(\varphi_1,\theta_1) + P_2 * \mathrm{HRTF}(\varphi_2,\theta_2) + P_3 * \mathrm{HRTF}(\varphi_3,\theta_3)$$

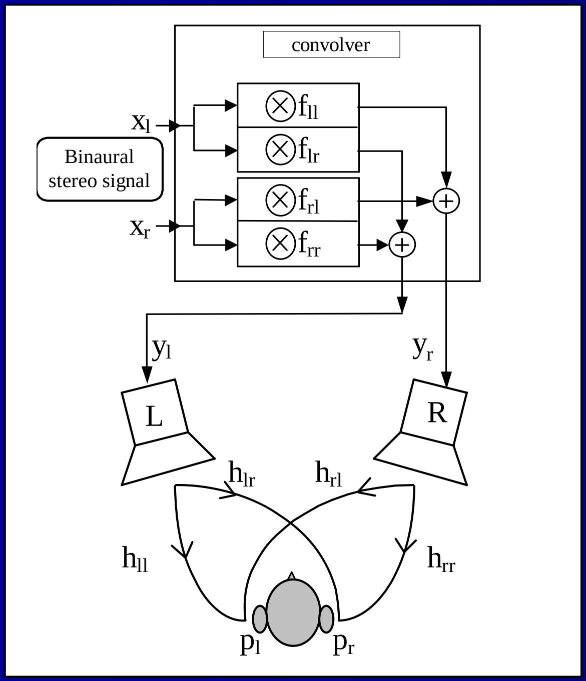
- The interpolation is actually done in the frequency domain, on the magnitude and unwrapped phase of the three HRTFs
- The octave-band amplitude equalization corresponding to the room transfer function is also applied in the frequency domain
- FFT and IFFT are used for converting between time domain and frequency domain

The Stereo Dipole



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

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

Theory of cross-talk canceling filters

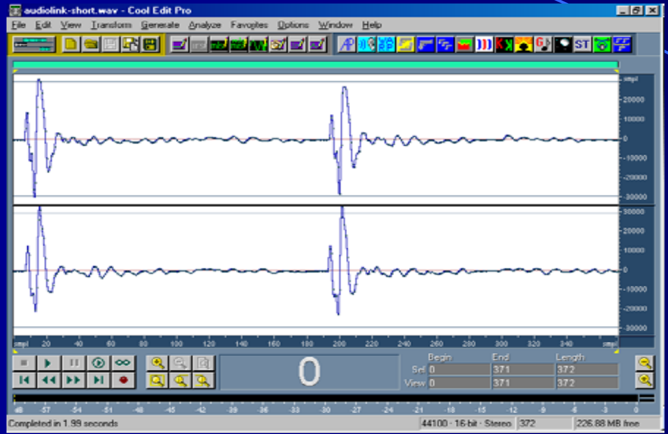
$$\begin{cases} f_{ll} = \left(h_{rr}\right) \otimes InvDen \\ f_{lr} = \left(-h_{lr}\right) \otimes InvDen \\ f_{rl} = \left(-h_{rl}\right) \otimes InvDen \\ f_{rr} = \left(h_{ll}\right) \otimes InvDen \\ InvDen = InvFilter\left(h_{ll} \otimes h_{rr} - h_{lr} \otimes h_{rl}\right) \end{cases}$$

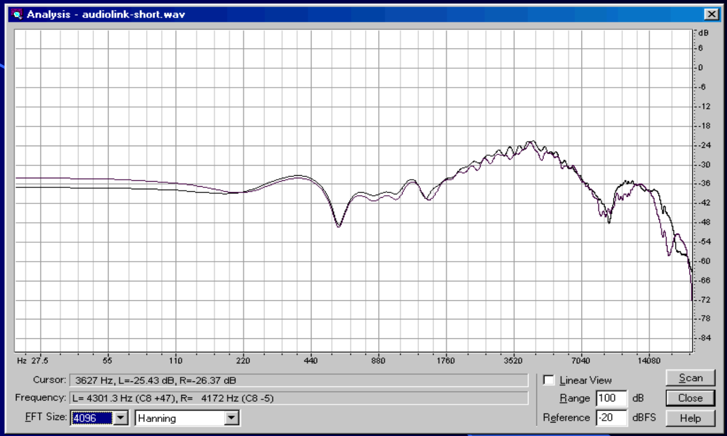
$$C(\omega) = FFT(h_{ll}) \cdot FFT(h_{rr}) - FFT(h_{lr}) \cdot FFT(h_{rl})$$

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

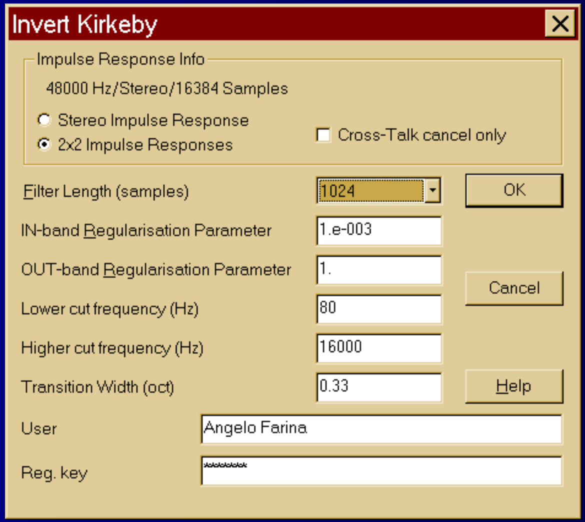
- The regularization parameter, ε, has to be adjusted by trials

Example



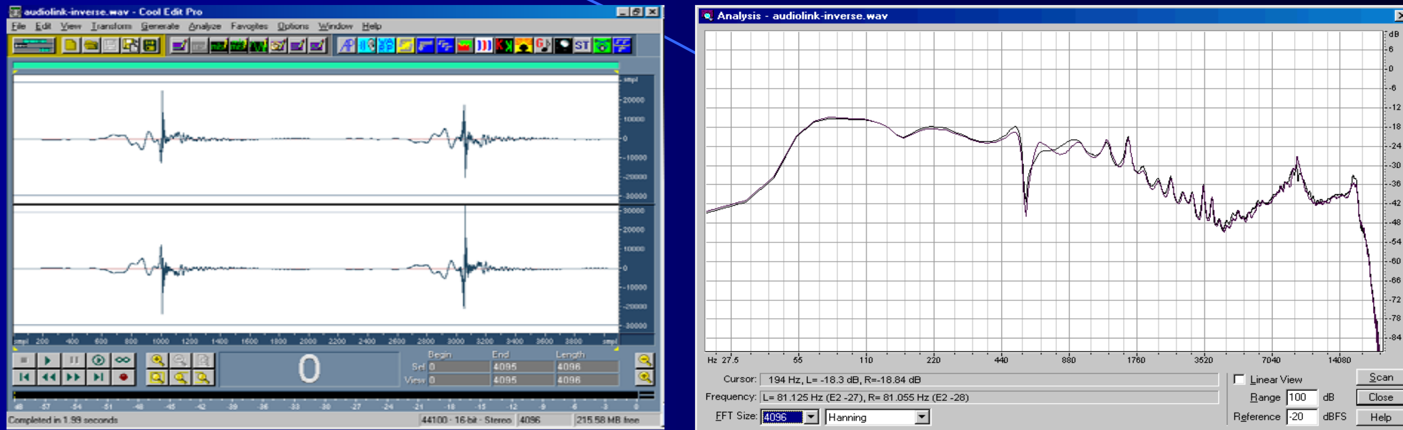


Measured impulse responses h

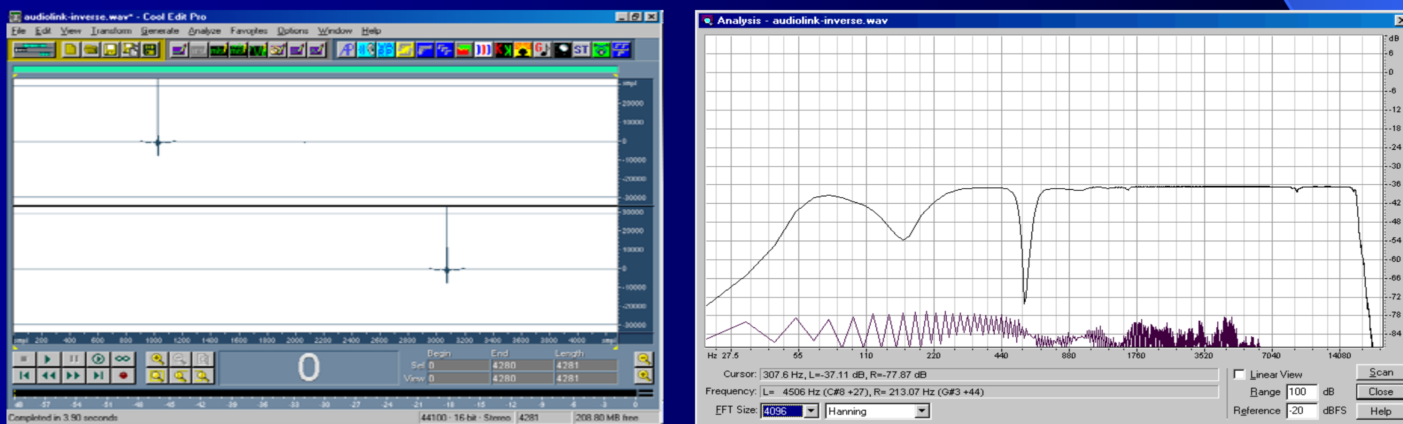


Plugin for CoolEdit which computes the inverse filters f

Example

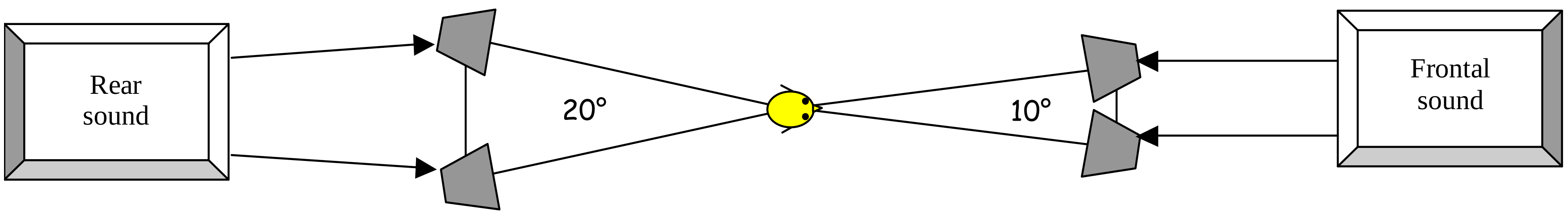


Computed long-FIR inverse filters f



Verification of the cross-talk cancellation

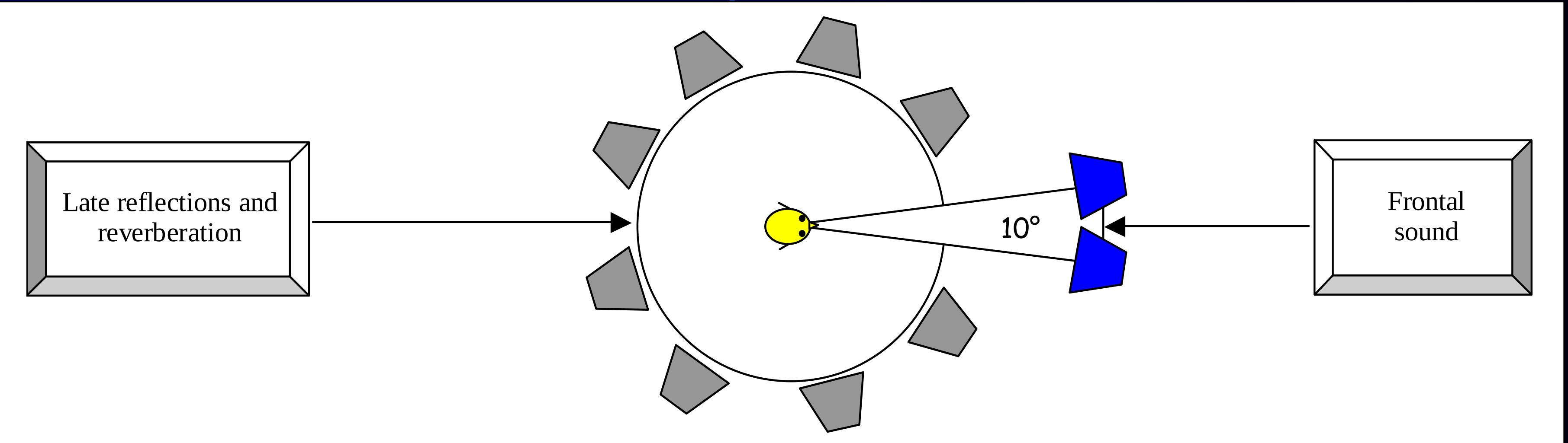
The Dual Stereo Dipole



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
- Usually the rear loudspeaker pair requires a larger angle than the frontal one

The Ambiophonics method



Reproduction over an Ambiophonics rig

- The system is based on two indipendently designed groups of loudspeakers:
- a Stereo Dipole, responsible for the reproduction only of the direct sound and early reflections coming from the stage,
- a surround periphonic array, driven by real-time convolution with room impulse responses (it can be quite irregularly shaped)

Synthesis of the IRs for surround convolution

- For any loudspeaker in the surround array, the WXYZ channels of a B-format IR can be processed, extracting a single (mono) response of a virtual microphone pointing along a given versor ***r*** *(r_x, r_y, r_z)*:

$$V(\vec{r}) = \frac{1}{2} \cdot \left[(2-D) \cdot W + D \cdot \left(r_x \cdot X + r_y \cdot Y + r_z \cdot Z \right) \right]$$

The directivity factor D can assume the following values:

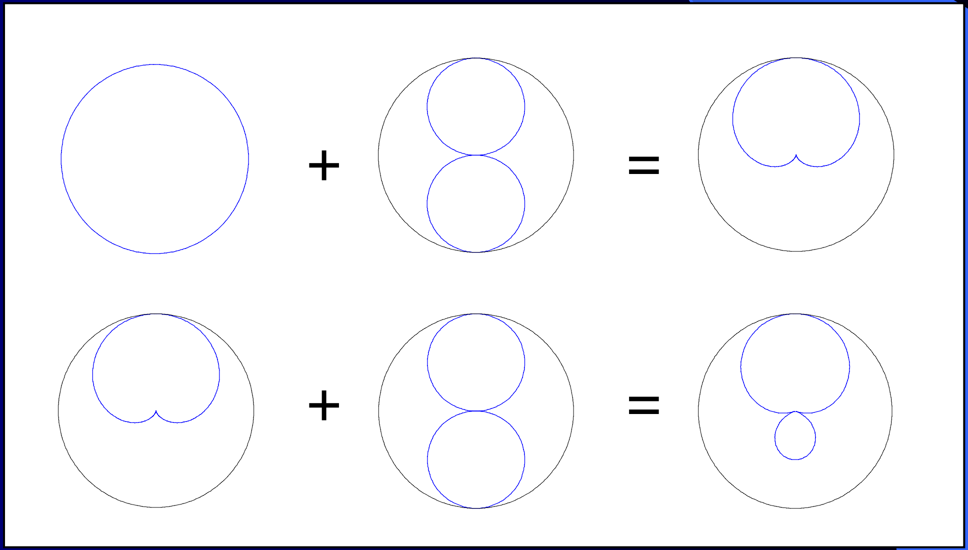
D=0.0 → omnidirectional

D=0.5 → subcardioid

D=1.0 → cardioid

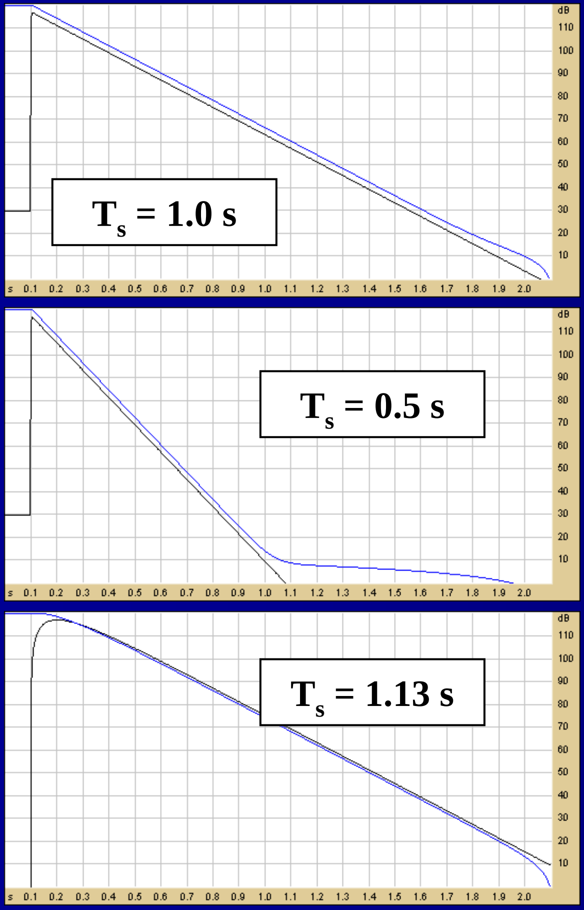
D=1.5 → hypercardioid

D=2.0 → figure-of-eight

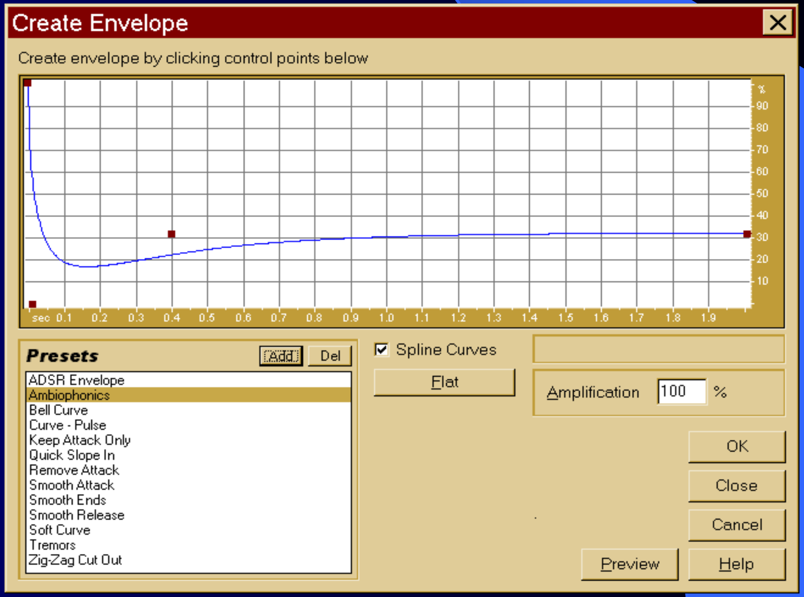


The Double-reverberation problem

When an impulse response is reproduced in another reverberant space, the resulting reverberant tail is the convolution of the two reverberant tails



The problem can be compensated for, by manual editing the first impulse response, applying (with CoolEdit) a time-varying amplitude shaping:



Conclusions

- Auralization can be done nowadays not only by the traditional binaural method (headphone listening), but also with modern methods of loudspeaker presentation
- Ambiophonics revealed to give significant advantages over the two surround systems which constitutes it (1st order Ambisonics and Stereo Dipole).
- In cases of limitation in number of reproduction channels, a Dual-Stereo-Dipole can also be very effective.
- When multiple listeners are required in a large listening area, the preferred method is 2nd order Ambisonics.
- The computational power required by these multichannel reproduction methods can be obtained cheaply by means of a modern PC running one of the available free convolution softwares

Internet Links

- The CoolEdit plugins, employed in the present work for measuring and processing the impulse responses, can be downloaded freely from **HTTP://www.ramsete.com/aurora**
- The BruteFIR convolver for Linux can be freely downloaded, with its source code, from **HTTP://www.ludd.luth.se/~torger/brutefir.html**
- The Ambiovolver convolver for Windows can be freely downloaded from **HTTP://www.dcomg.upv.es/jjlopez**
- Richard Furse's Ambisonics decoder can be freely downloaded from **http://www.muse.demon.co.uk/mn_index.html**