EMULATION OF NOT-LINEAR, TIME-VARIANT DEVICES BY THE CONVOLUTION TECHNIQUE

Angelo Farina, Enrico Armelloni

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

AUDIO ENGINEERING SOCIETY
Italian Section

UNIVERSITÀ DEGLI STUDI DI PARMA

Goals for Auralization

- Transform the results of objective electroacoustics measurements to audible sound samples suitable for listening tests
- Traditional auralization is based on linear convolution: this does not replicates faithfully the nonlinear behaviour of most transducers
- The new method presented here overcomes to this strong limitation, providing a simplified treatment of memory-less distortion

Auralization by linear convolution

Convolving a suitable sound sample with the linear IR, the frequency response and temporal transient effects of the system can be simulated properly

Auralization by linear convolution

The beginnings: hardware DSP-based convolution units

Lake Technologies HURON

Yamaha SREV-1

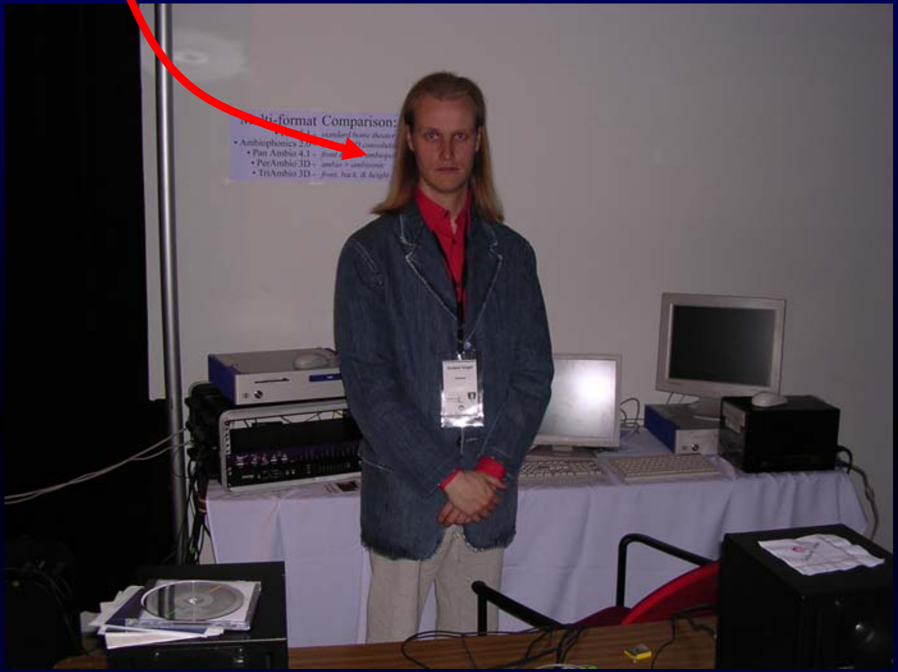
Sony DRES-777

The AMBIOPHONICS Institute: the home of convolution

Photos taken on
16 december 2002

Software Convolution: BruteFIR and AlmusVCU

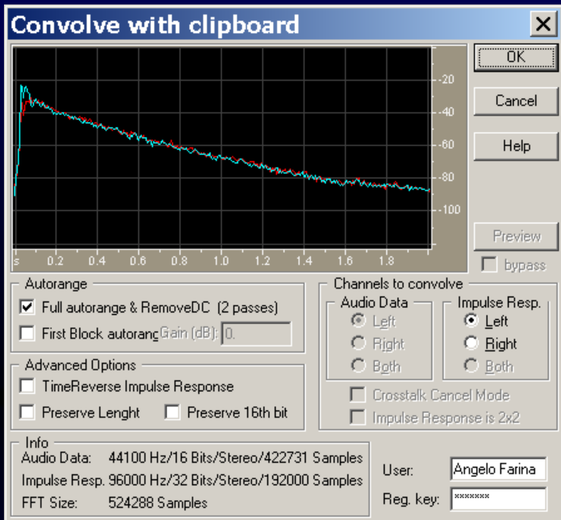
Open-source software for Linux by Anders Torger – AES 24° Conference

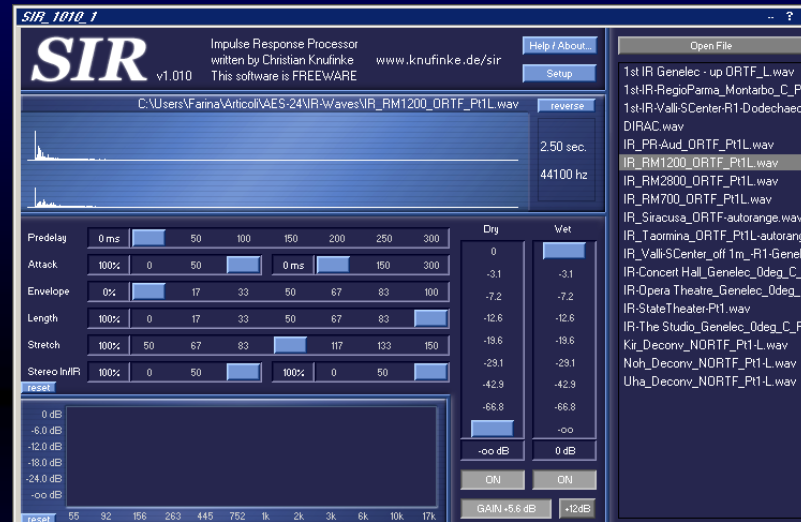


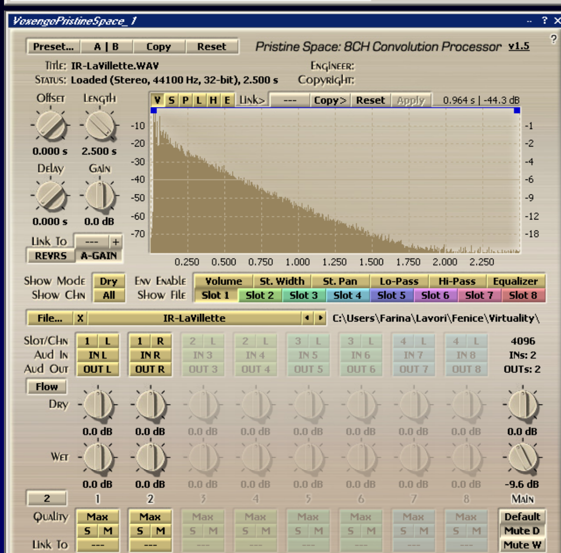
Performance: a fanless (silent) P-IV running at 2.5 GHz was capable of real time convolution of 2 inputs at 44.1 kHz, 24 bits, with 48 impulse responses, each 5s long, driving 24 Genelec loudspeakers (20 satellites + 4 subwoofers), employing 75% of the CPU time

Auralization by linear convolution

Nowadays many sytems or software plugins can perform satisfactorily the Linear Convolution operation, and are widely employed for audio processing







What’s missing in linear convolution ?

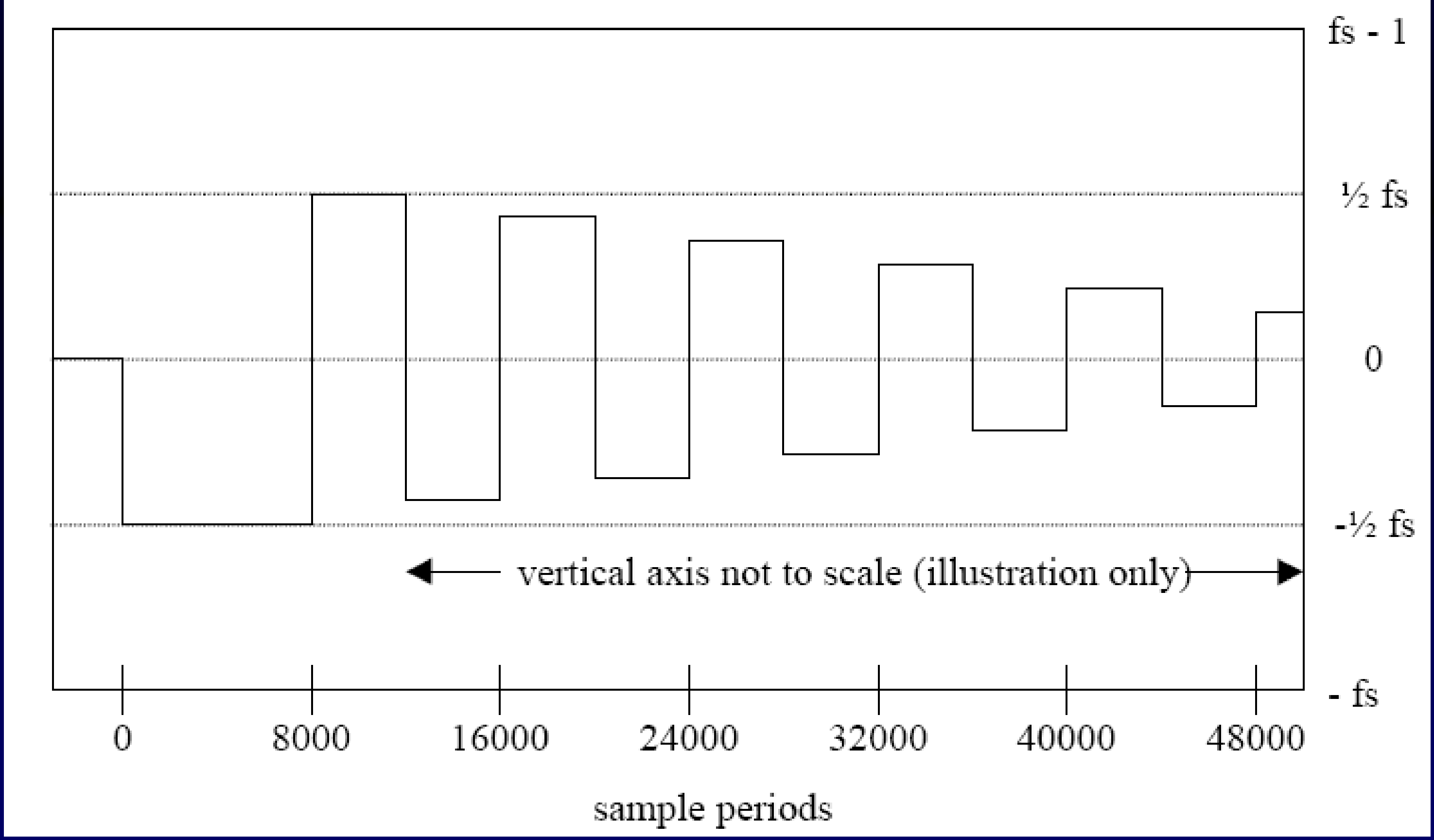
- No harmonic distortion, nor other nonlinear effects are being reproduced.
- From a perceptual point of view, the sound is judged “cold” and “innatural”
- A comparative test between a strongly nonlinear device and an almost linear one does not reveal any audible difference, because the nonlinear behavior is removed for both

Method 1 (IR switching)

- A very simple idea: a different IR is employed depending on the amplitude of each sample of the signal to be filtered
- The method is quite old: the first published papers are thoss of Bellini and Farina (1988) and Michael Kemp (1999)
- Several impulse responses are measured, employing test signals of different amplitudes, and stored for later usage.
- It is mandatory to implement the convolution as direct form in time domain, as each sample has to be processed with a different IR.

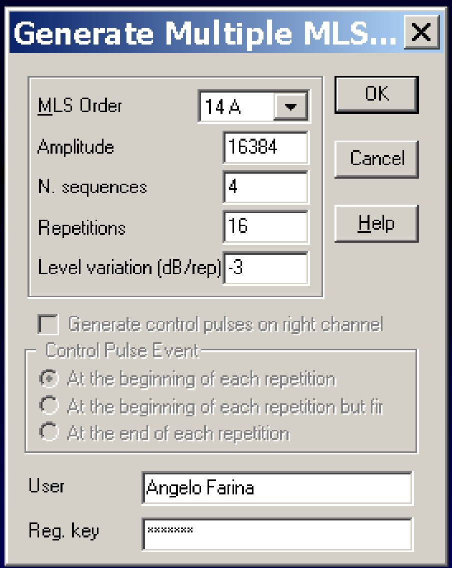
Measurement of multiple IRs

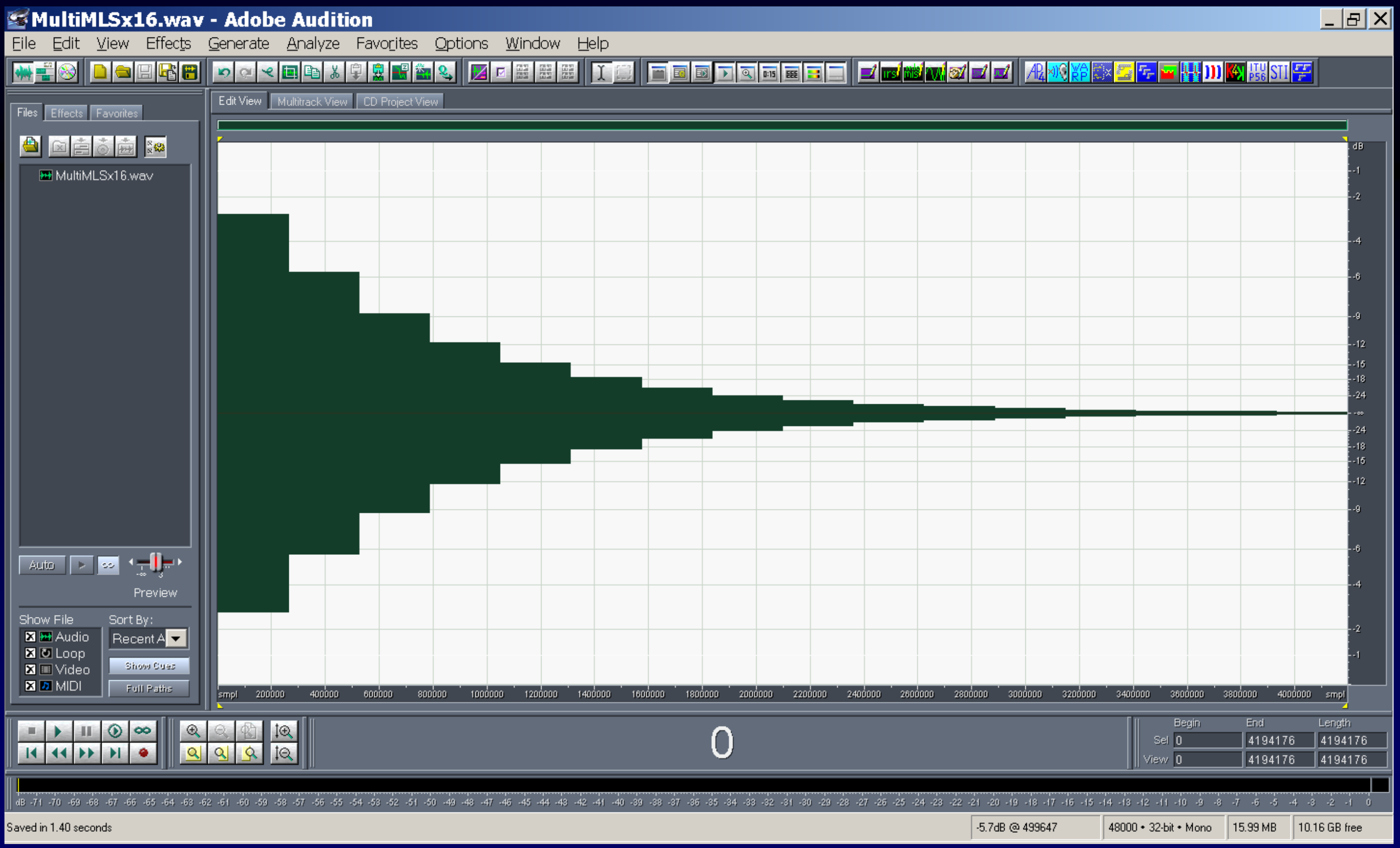
- Michael Kemp employed a step function, with several steps of decreasing amplitude



Misura delle IR multiple

- Farina e Bellini did employ a sequence of MLS repetitions, each with decreasing amplitude





Implementation (Michael J. Kemp)

- Focusrite did release recently Liquid Channel, the first “dynamic convolver” implementing the IR-switching technique

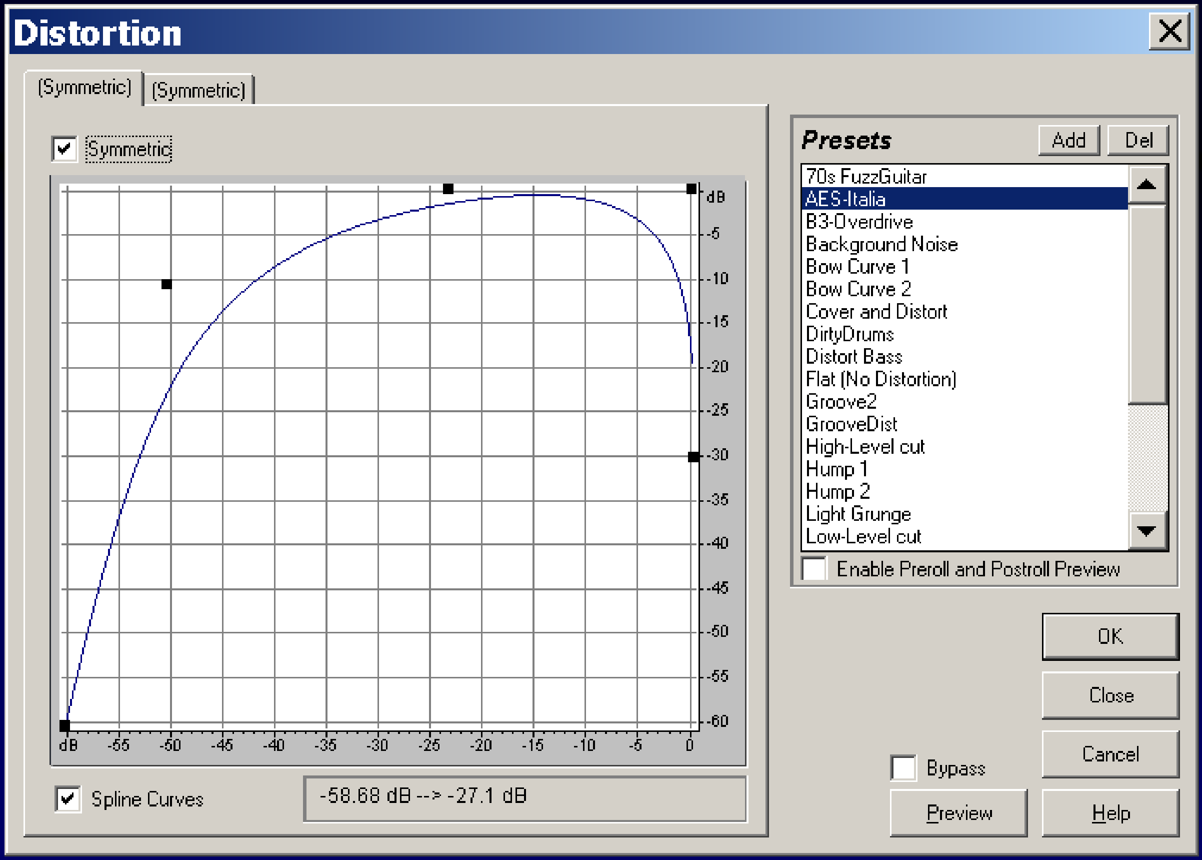
“The **Liquid Channel** is a revolutionary professional **channel** strip that can precisely replicate any classic mic-pre and compressor”

Implementation (Farina/Bellini)

- A FIR filtering algorithm, with the set of coefficients chosen depending on the sample amplitude, was implemented on a Sharc EZ-KIT 20161 board, and employed for car-audio applications

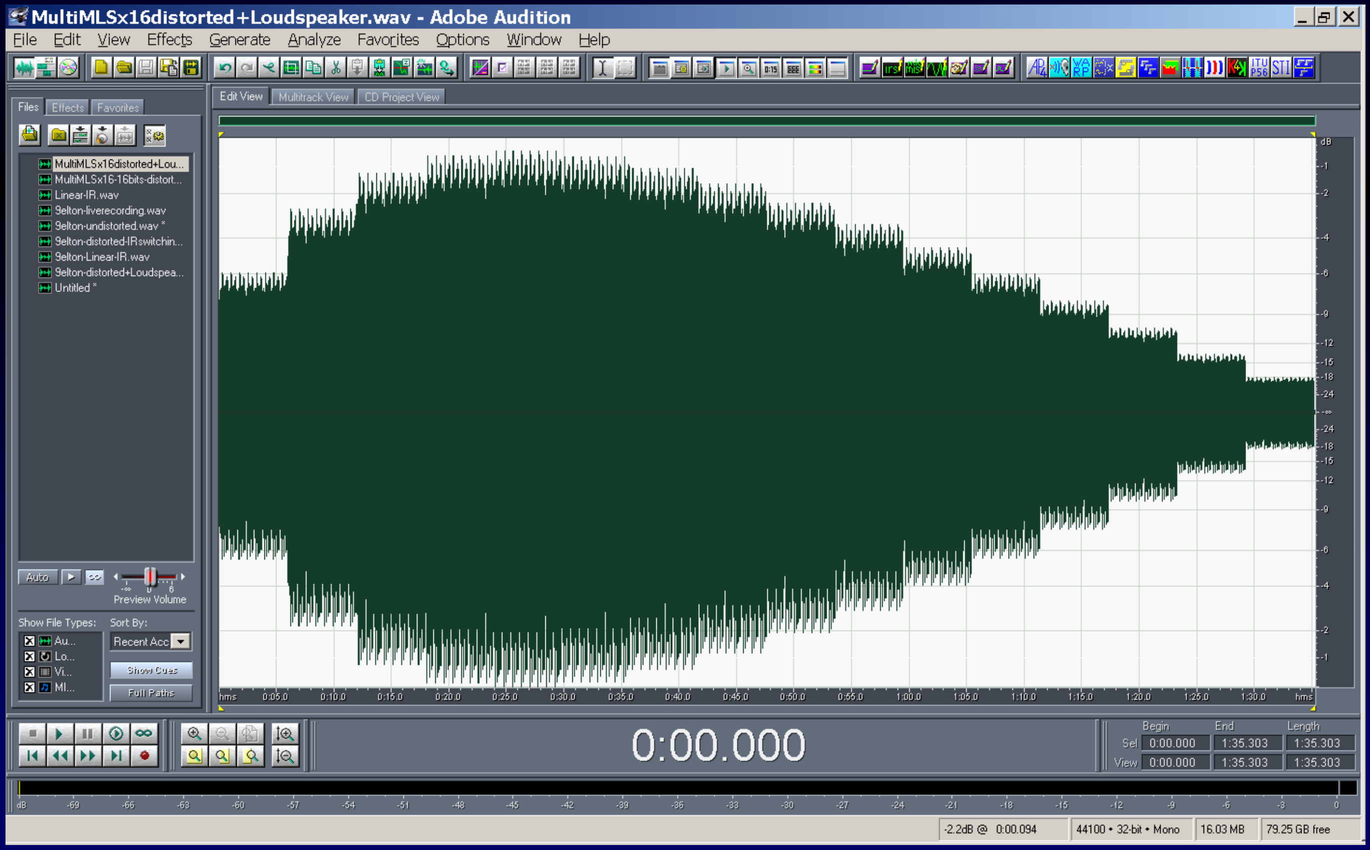
Example

- The “not linear device” is emulated by the DISTORTION plugin of Adobe Audition, followed by sound playback and simultaneous recording over the loudspeaker and microphone of a laptop PC



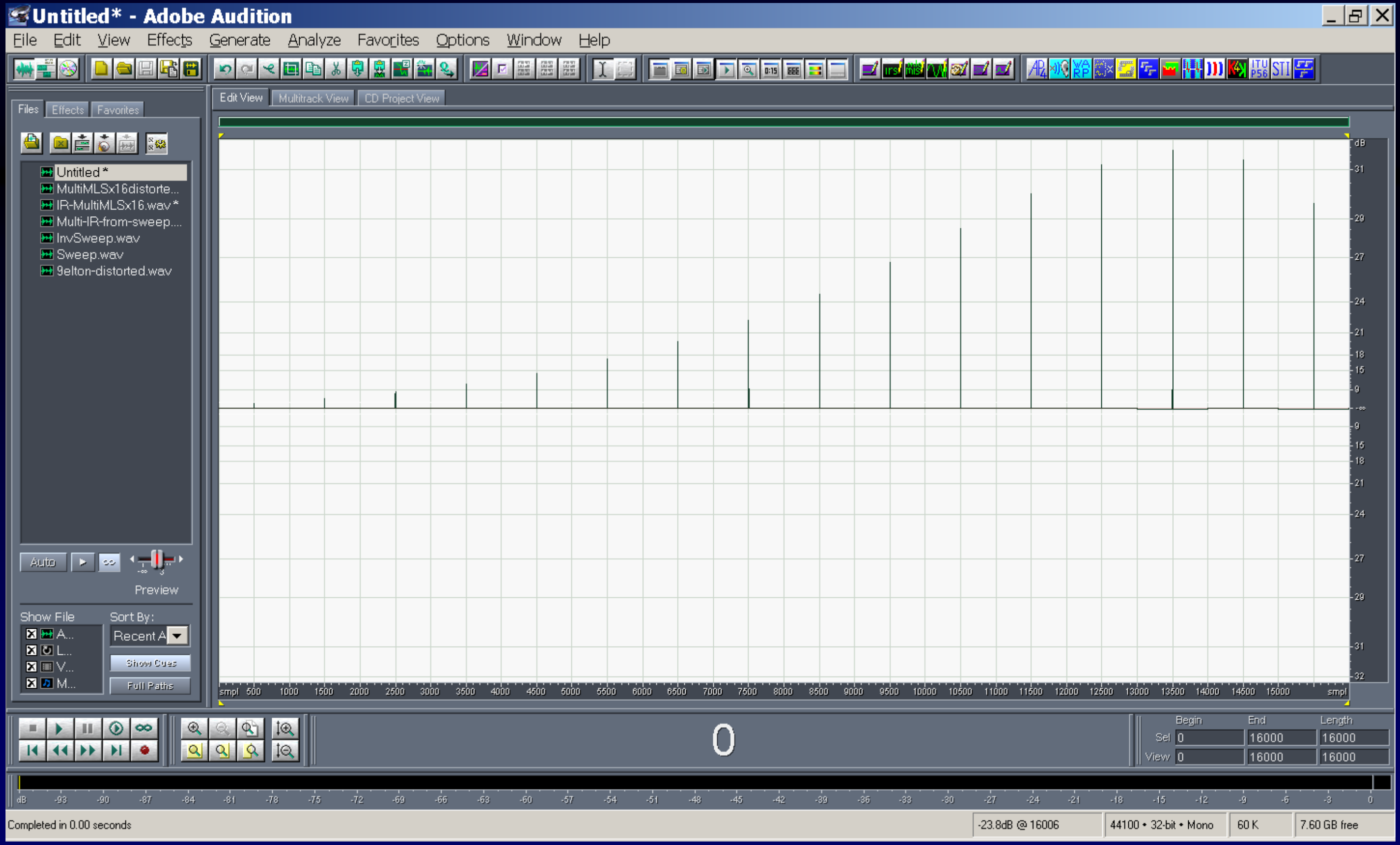
Example

- This is the multiple MLS signals after being processed through the not-linear device



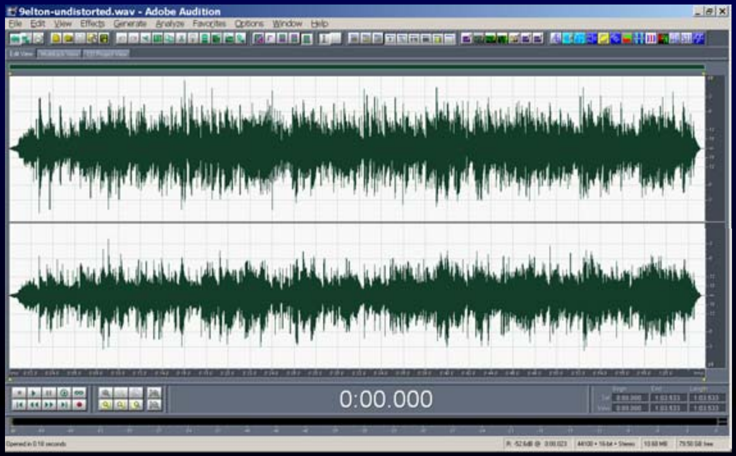
Example

- Here the 16 impulse responses measured with MLS of different amplitude (decreasing 3dB each from left to right) are shown

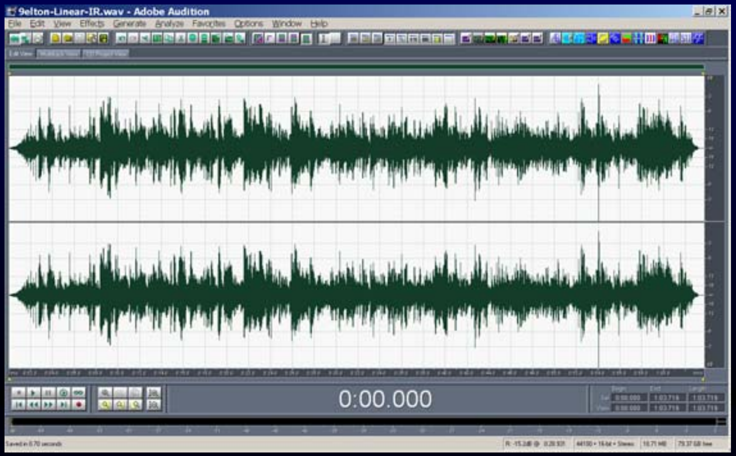


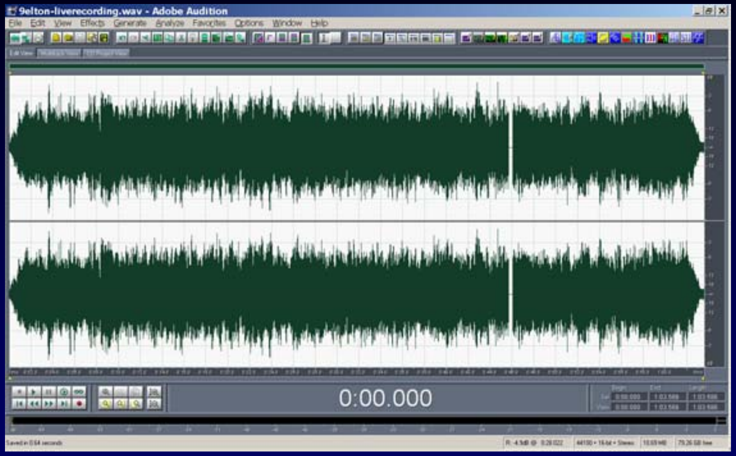
Audible evaluation of the performance

Original signal

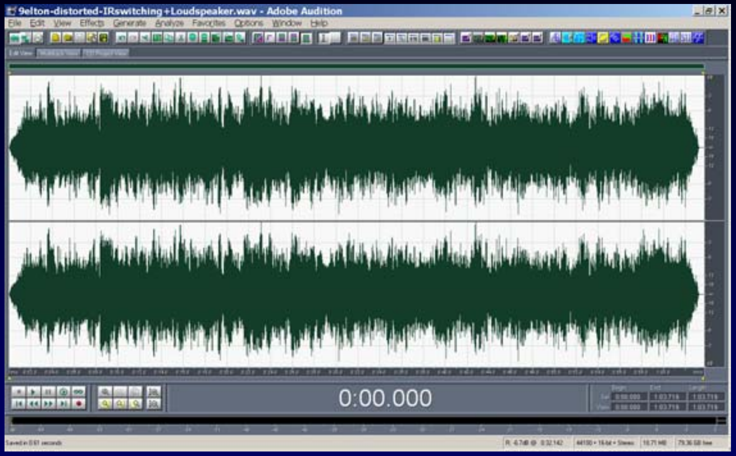


Linear convolution





Live recording



Non-linear (IR switching)

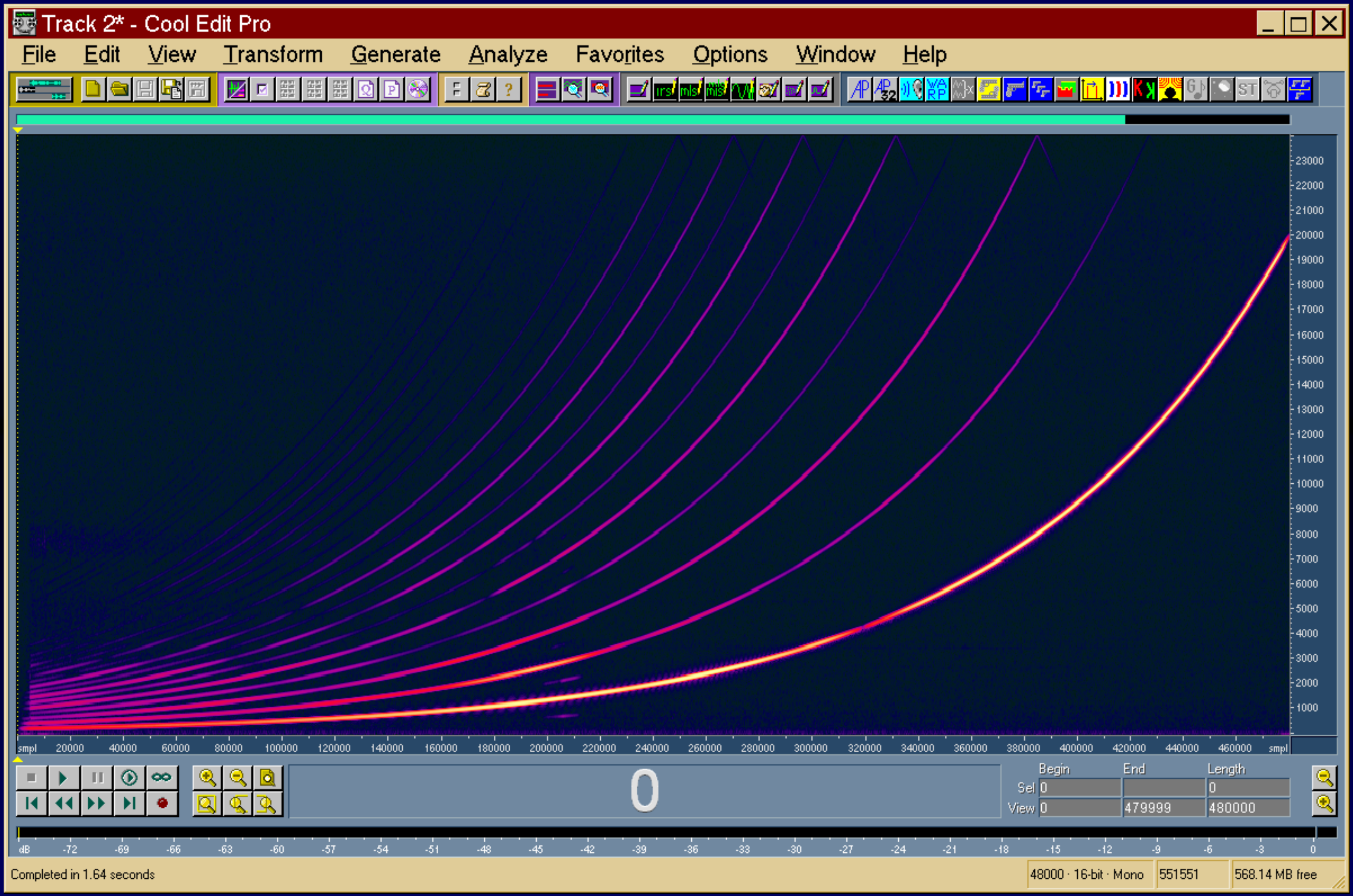
Method 2 – Diagonal Volterra Kernels

- We start from a measurement of the system based on exponential sine sweep (Farina, 108th AES, Paris 2000)
- Diagonal Volterra kernels are obtained by post-processing the measurement results
- These kernels are employed as FIR filters in a multiple-order convolution process (original signal, its square, its cube, and so on are convolved separately and the result is summed)

Exponential sweep measurement

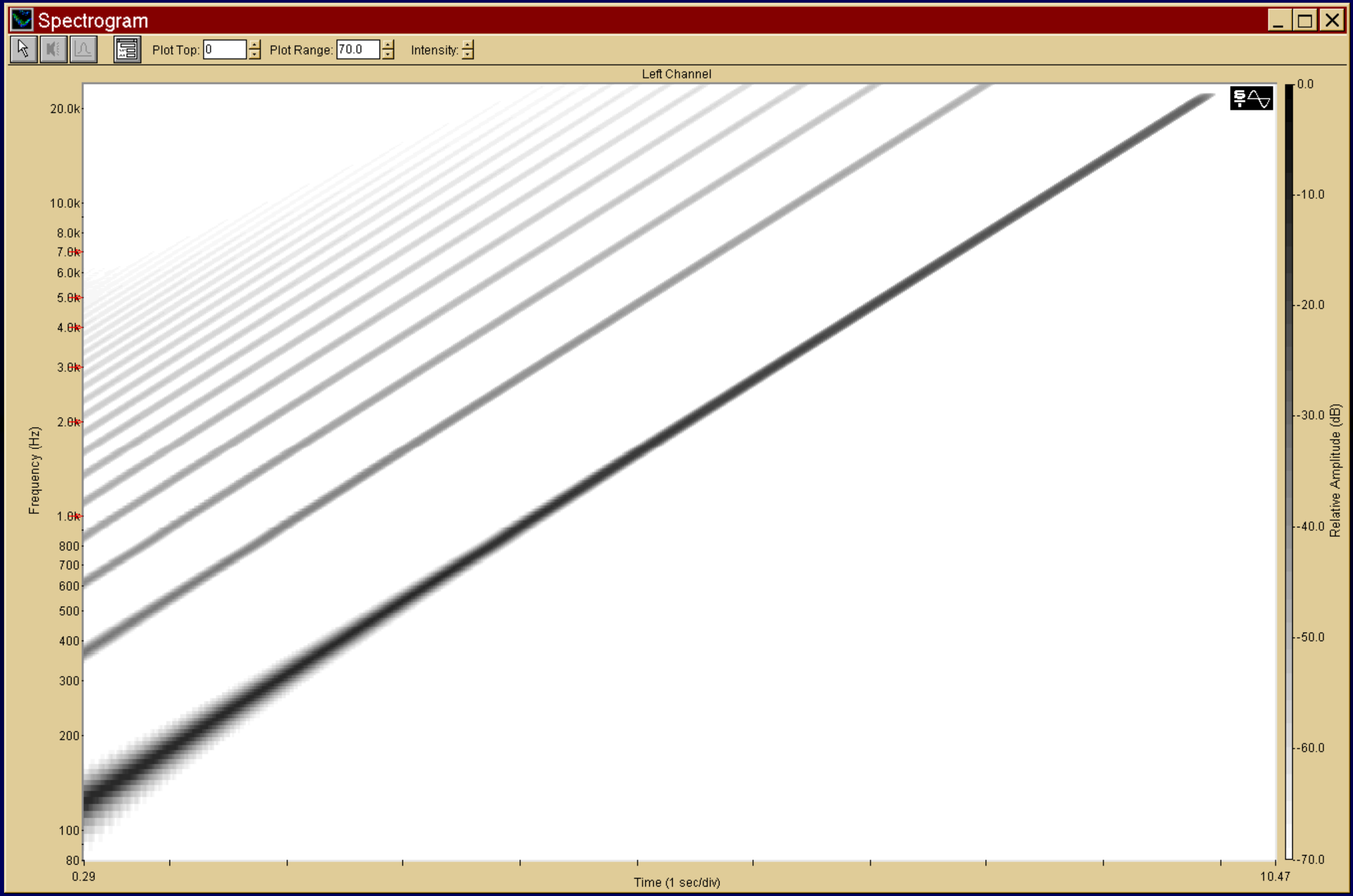
- The test signal is a sine sweep with constant amplitude and exponentially-increasing frequency

Raw response of the system



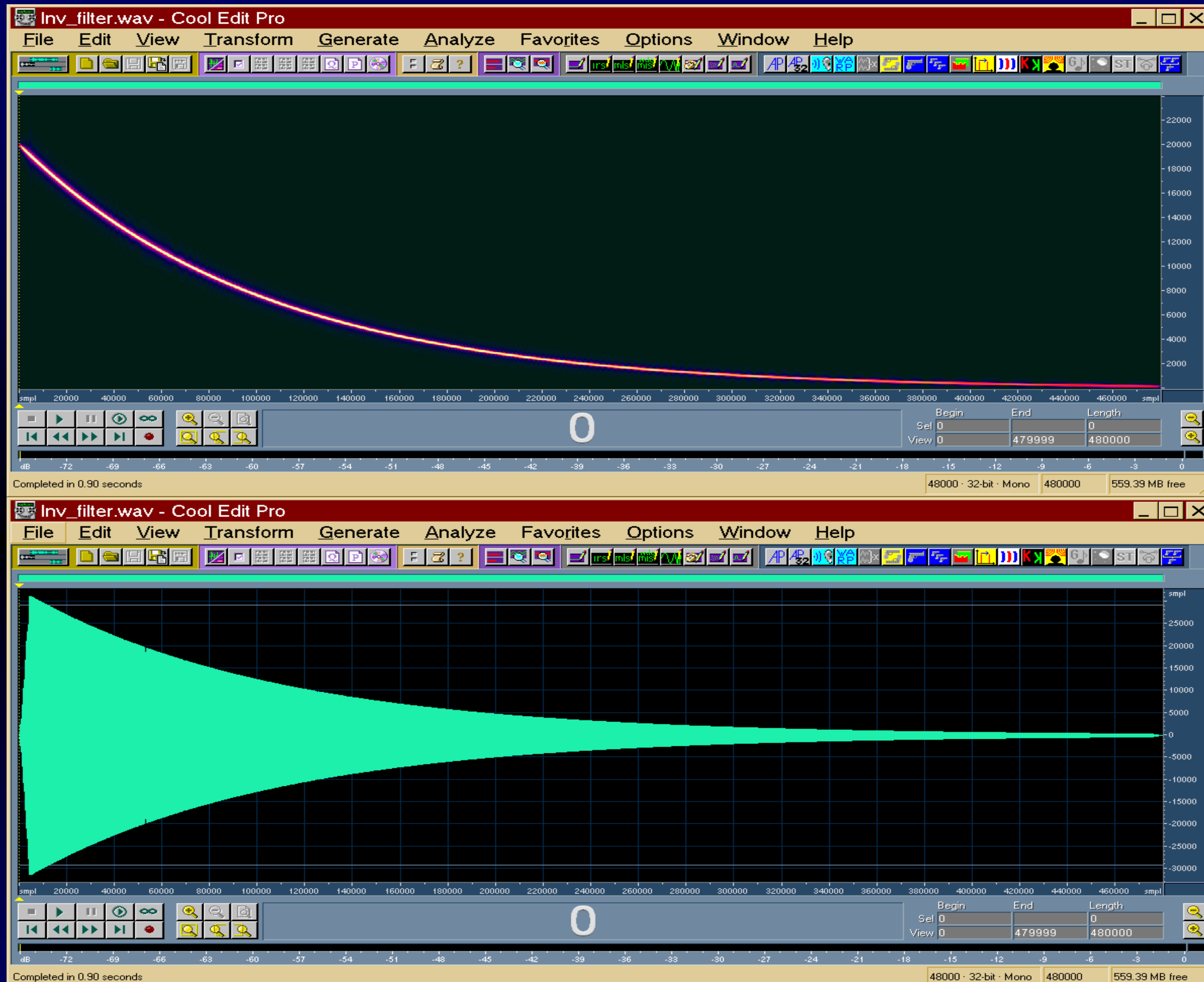
Many harmonic orders do appear as colour stripes

Raw response of the system



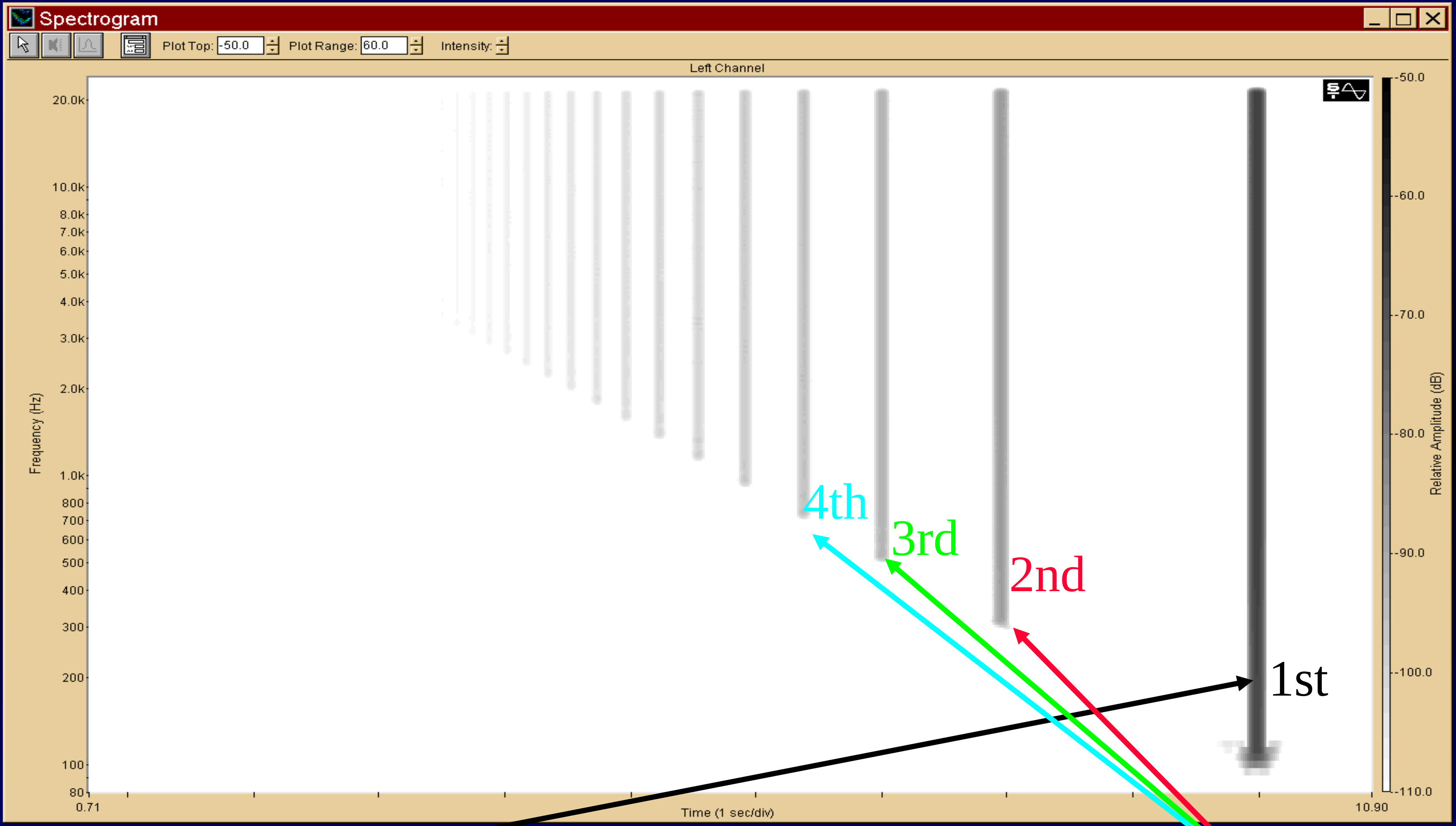
Many harmonic orders do appear as colour stripes

Deconvolution of system's impulse response



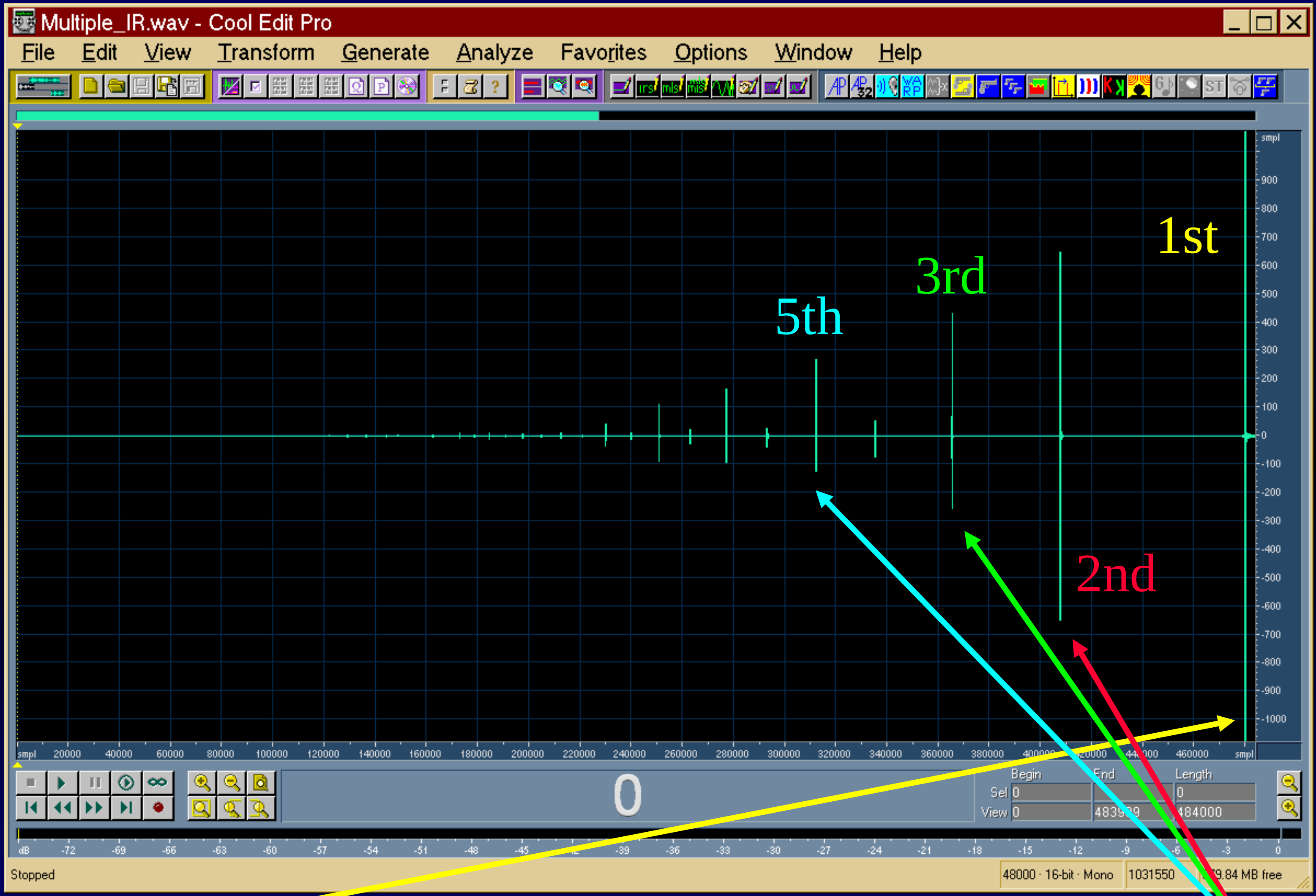
The deconvolution is obtained by convolving the raw response with a suitable inverse filter

Multiple impulse response obtained



The last peak is the linear impulse response, the preceding ones are the harmonic distortion orders

Multiple impulse response obtained



The last peak is the linear impulse response, the preceding ones are the harmonic distortion orders

Theory of nonlinear convolution

- The basic approach is to convolve separately, and then add the result, the linear IR, the second order IR, the third order IR, and so on.
- Each order IR is convolved with the input signal raised at the corresponding power:

$$y(n) = \sum_{i=0}^{M-1} h_1(i) \cdot x(n-i) + \sum_{i=0}^{M-1} h_2(i) \cdot x^2(n-i) + \sum_{i=0}^{M-1} h_3(i) \cdot x^3(n-i) + \ldots..$$

The problem is that the required multiple IRs **are not** the results of the measurements: they are instead the diagonal terms of Volterra kernels

Volterra kernels and simplification

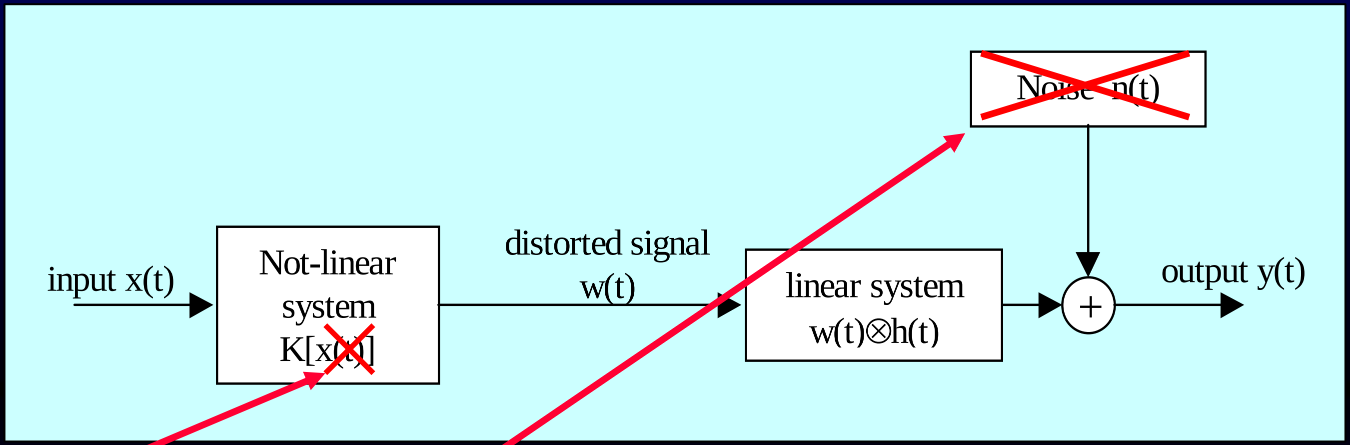
- The general Volterra series expansion is defined as:

$$y(n) = \sum_{i_1=0}^{M-1} h_1(i_1)\cdot x(n-i_1) + \sum_{i_1=0}^{M-1}\sum_{i_2=0}^{M-1} h_2(i_1,i_2)\cdot x(n-i_1)\cdot x(n-i_2) +$$

$$+ \sum_{i_1=0}^{M-1}\sum_{i_2=0}^{M-1}\sum_{i_3=0}^{M-1} h_3(i_1,i_2,i_3)\cdot x(n-i_1)\cdot x(n-i_2)\cdot x(n-i_3) + \ldots..$$

This explains also nonlinear effect with memory, as the system output contains also products of previous sample values with different delays

Memoryless distortion followed by a linear system with memory



- The first nonlinear system is assumed to be memory-less, so only the diagonal terms of the Volterra kernels need to be taken into account.
- Furthermore, we neglect the noise, which is efficiently rejected by the sine sweep measurement method.

Volterra kernels from the measurement results

The measured multiple IRs h' can be defined as:

$$y(t) = h'_1 \otimes \sin[\omega_{var}] + h'_2 \otimes \sin[2 \cdot \omega_{var}] + h'_3 \otimes \sin[3 \cdot \omega_{var}] + \ldots$$

We need to relate them to the simplified Volterra kernels h:

$$y(t) = h_1 \otimes \sin[\omega_{var}] + h_2 \otimes \sin^2[\omega_{var}] + h_3 \otimes \sin^3[\omega_{var}] + \ldots$$

Trigonometry can be used to expand the powers of the sinusoidal terms:

$$\sin^2(\omega \cdot \tau) = \frac{1}{2} - \frac{1}{2} \cdot \cos(2 \cdot \omega \cdot \tau)$$

$$\sin^3(\omega \cdot \tau) = \frac{3}{4} \cdot \sin(\omega \cdot \tau) - \frac{1}{4} \cdot \sin(3 \cdot \omega \cdot \tau)$$

$$\sin^4(\omega \cdot \tau) = \frac{3}{8} - \frac{1}{2} \cdot \cos(2 \cdot \omega \cdot \tau) + \frac{1}{8} \cdot \cos(4 \cdot \omega \cdot \tau)$$

$$\sin^5(\omega \cdot \tau) = \frac{5}{8} \cdot \sin(\omega \cdot \tau) - \frac{5}{16} \cdot \sin(3 \cdot \omega \cdot \tau) + \frac{1}{16} \cdot \sin(5 \cdot \omega \cdot \tau)$$

Finding the connection point

Going to frequency domain by taking the FFT, the first equation becomes:

$$Y(\omega) = \overline{H'_1}[\omega] \cdot X[\omega] + \overline{H'_2}[\omega] \cdot X[\omega/2] + \overline{H'_3}[\omega] \cdot X[\omega/3] + \ldots$$

Doing the same in the second equation, and substituting the trigonometric expressions for power of sines, we get:

$$Y(\omega) = \left[\overline{H_1} + \frac{3}{4}\cdot\overline{H_3} + \frac{5}{8}\cdot\overline{H_5}\right]\cdot X[\omega] + \left[-\frac{1}{2}\cdot\overline{H_2} - \frac{1}{2}\cdot\overline{H_4}\right]\cdot j \cdot X[\omega/2] +$$
$$+\left[-\frac{1}{4}\cdot\overline{H_3} - \frac{5}{16}\cdot\overline{H_5}\right]\cdot X[\omega/3] + \frac{1}{8}\cdot\overline{H_4}\cdot j\cdot X[\omega/4] + \frac{1}{16}\cdot\overline{H_5}\cdot X[\omega/5] + \ldots$$

The terms in square brackets have to be equal to the corresponding measured transfer functions H’ of the first equation

Solution

- Thus we obtain a linear equation system:

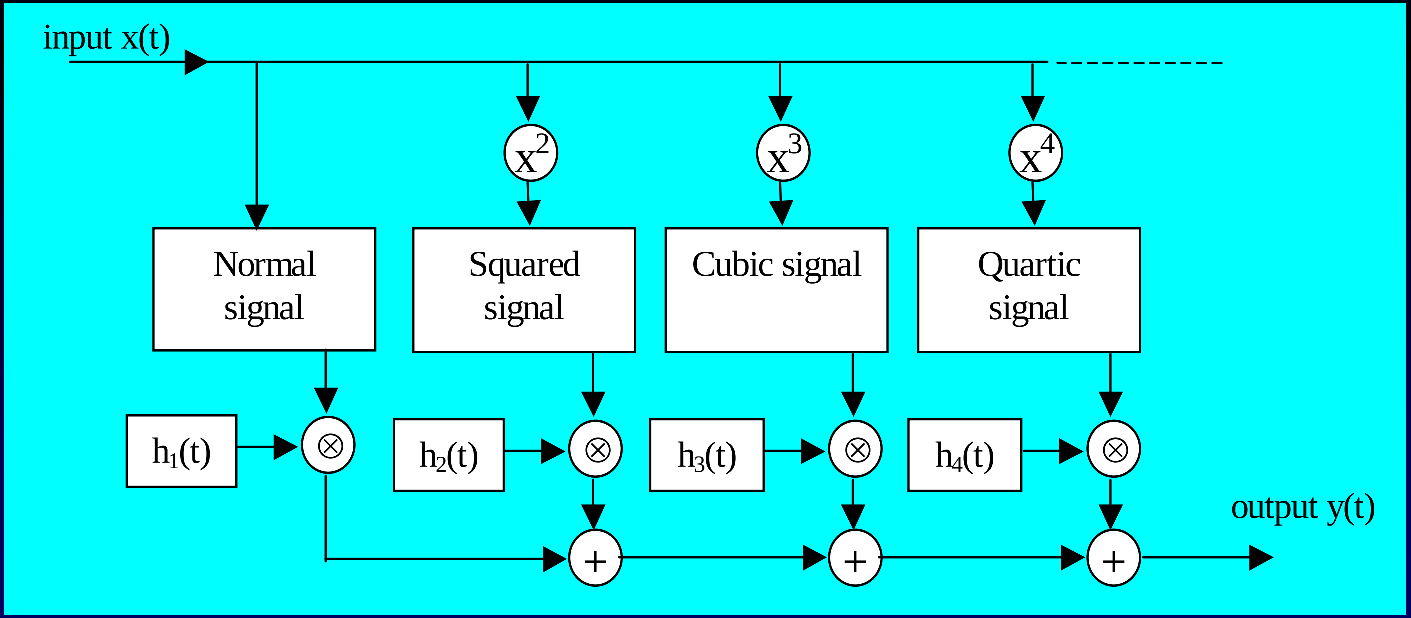
$$\begin{cases} \overline{H'_1} = \overline{H_1} + \frac{3}{4} \cdot \overline{H_3} + \frac{5}{8} \cdot \overline{H_5} \\ \overline{H'_2} = -j \cdot \frac{1}{2} \cdot \left[\overline{H_2} + \overline{H_4} \right] \\ \overline{H'_3} = -\frac{1}{4} \cdot \overline{H_3} - \frac{5}{16} \cdot \overline{H_5} \\ \overline{H'_4} = j \cdot \frac{1}{8} \cdot \overline{H_4} \\ \overline{H'_5} = \frac{1}{16} \cdot \overline{H_5} \end{cases}$$

We can easily solve it, obtaining the required Volterra kernels as a function of the measured multiple-order IRs:

$$\begin{cases} \overline{H_1} = \overline{H'_1} + 3 \cdot \overline{H'_3} + 5 \cdot \overline{H'_5} \\ \overline{H_2} = 2 \cdot j \cdot \overline{H'_2} + 8 \cdot j \cdot \overline{H'_4} \\ \overline{H_3} = -4 \cdot \overline{H'_3} - 20 \cdot \overline{H'_5} \\ \overline{H_4} = -8 \cdot j \cdot \overline{H'_4} \\ \overline{H_5} = 16 \cdot \overline{H'_5} \end{cases}$$

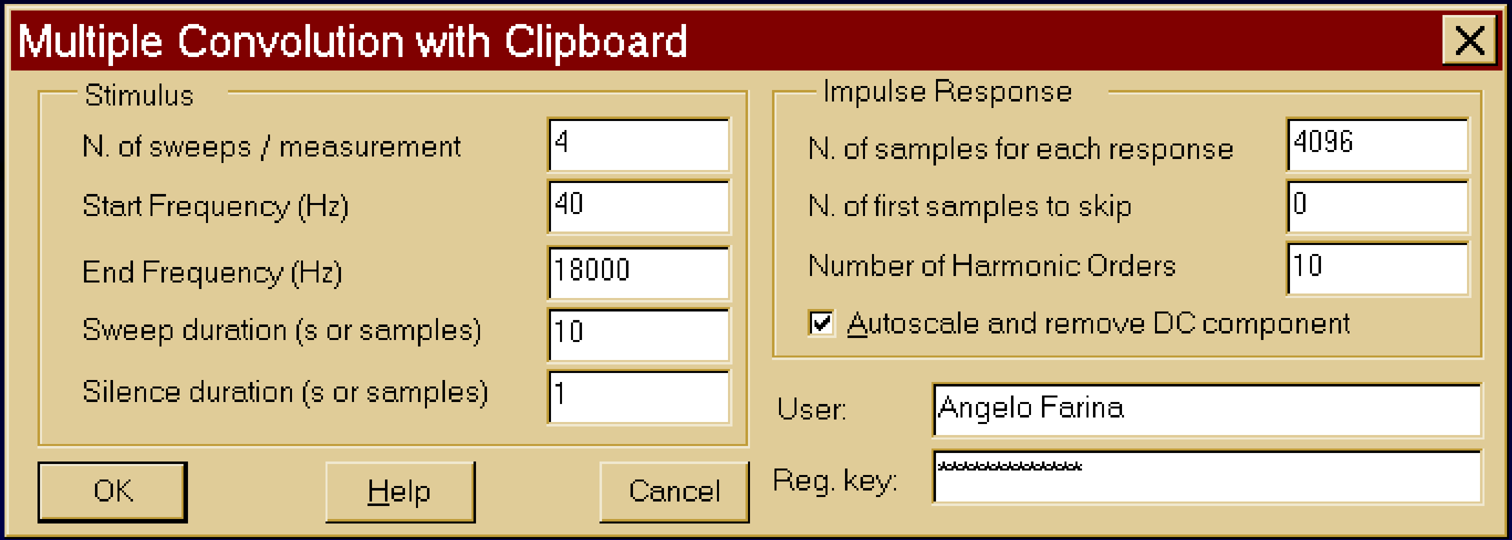
Non-linear convolution

As we have got the Volterra kernels already in frequency domain, we can efficiently use them in a multiple convolution algorithm implemented by overlap-and-save of the partitioned input signal:



Software implementation

Although today the algorithm is working off-line (as a mix of manual operations performed with Adobe Audition), a more efficient implementation as a plugin is being worked out:



This will allow for real-time operation even with a very large number of filter coefficients

Audible evaluation of the performance

Original signal

Linear convolution

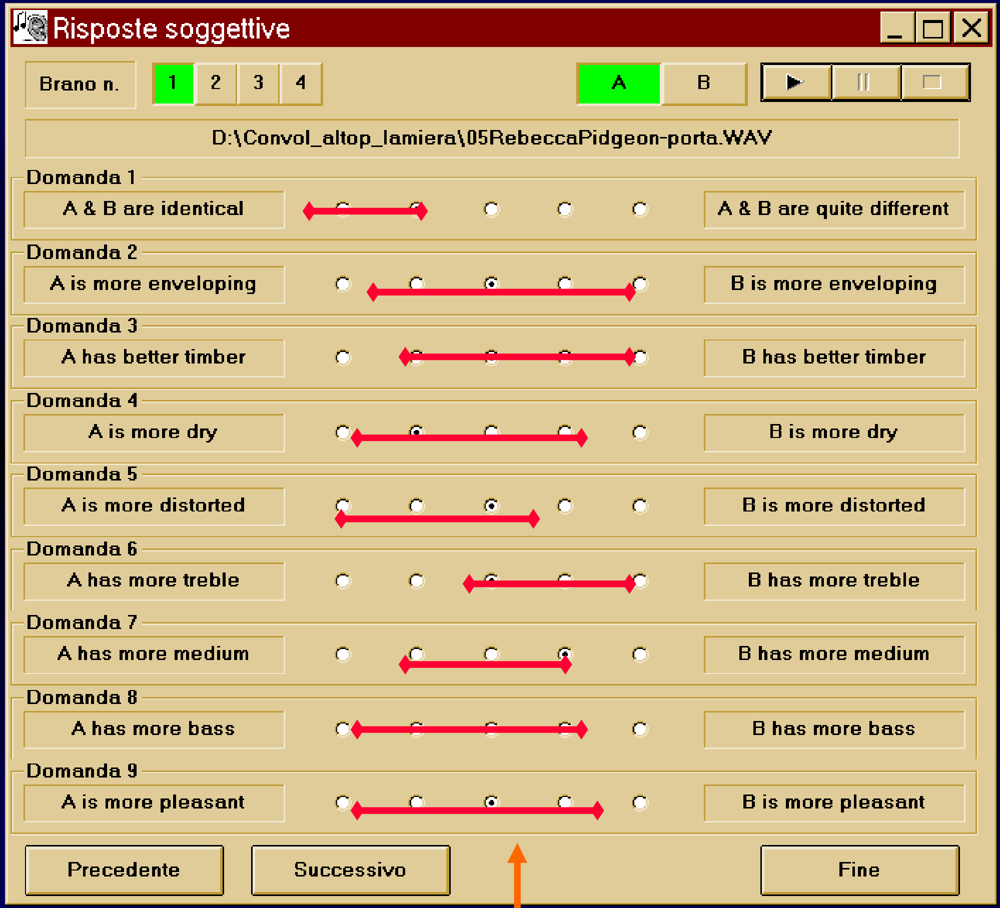
These last two were compared in a formalized blind listening test

Live recording

Non-linear multi convolution

Subjective listening test

- A/B comparison
- Live recording & non-linear auralization
- 12 selected subjects
- 4 music samples
- 9 questions
- 5-dots horizontal scale
- Simple statistical analysis of the results
- A was the live recording, B was the auralization, but the listener did not know this



95% confidence intervals of the responses

Conclusion

Statistical parameters – more advanced statistical methods would be advisable for getting more significant results

Question Number	Average score	2.67 * Std. Dev.
1 (identical-different)	1.25	0.76
3 (better timber)	3.45	1.96
5 (more distorted)	2.05	1.34
9 (more pleasant)	3.30	2.16

Remarks

- The Audition plugins shown here are freely downloadable from HTTP://www.aurora-plugins.com

- The sound samples employed for the subjective test are available for download at HTTP://pcangelo.eng.unipr.it/public/AES110

Future developments

- In the “IR switching” techniqque it is posssible to obtain some “memory effect” employing a fast block convolution algorithm, instead of processing “sample by sample”.
- The choice of the lenght of the processing block has to correspond to the latency to level variations of the not-time-invariant device

Future developments

- In the “diagonal volterra kernels” method, some meory effect can be obtained adding a variable gain control droven by a time averaging block
- Also in this case, the choice of the time constant of the averaging block needs to be aligned with the latency to level variations of the not-time-invariant device