Realtime auralization employing a not-linear, not-time-invariant convolver

Angelo Farina [1], Adriano Farina [2]

1) Industrial Engineering Dept., University of Parma, Via delle Scienze 181/A Parma, 43100 ITALY – **HTTP://www.angelofarina.it**
2) Liceo Classico G.D. Romagnosi, via Marialuigia 1, Parma, 43100 ITALY **HTTP://www.adrianofarina.it**

123rd AES Convention
2007 October 5–8
New York, NY, USA

UNIVERSITÀ DEGLI STUDI DI PARMA

Il passato del futuro

LICEO GINNASIO STATALE "G. D. ROMAGNOSI"

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 replicate 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

What’s linear convolution ?

- Also called “FIR filtering”
- Convolution is the mathematical operation performed when filtering a waveform x employing as filter coefficients the samples of a second waveform, usually denoted as h and called “impulse response”
- The samples of the input waveform are multiplied by the samples of the impulse response h, and the results accumulated (summed):

$$y(j) = \sum_{i=0}^{N-1} x(j-i) \cdot h(i)$$

Auralization by linear convolution

$\otimes$

$=$

Convolving linearly a suitable sound sample with the Imp.Resp., 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

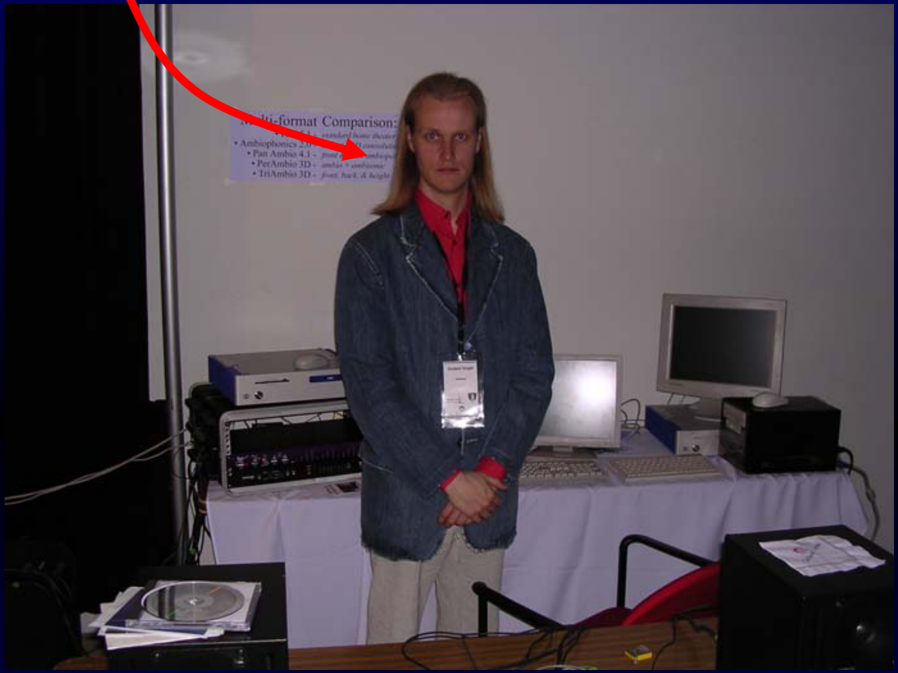
Sony DRES-777

The AMBIOPHONICS Institute: the home of convolution

Photos taken on
16 december 2002

Software Convolution: BruteFIR and AlmusVCU

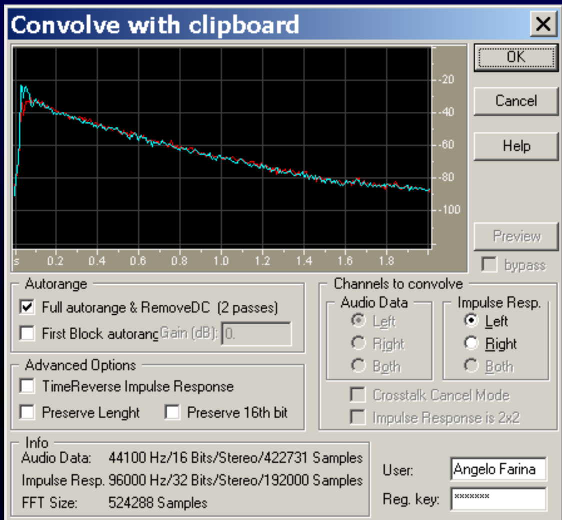
Open-source software for Linux by Anders Torger – AES 24° Conference

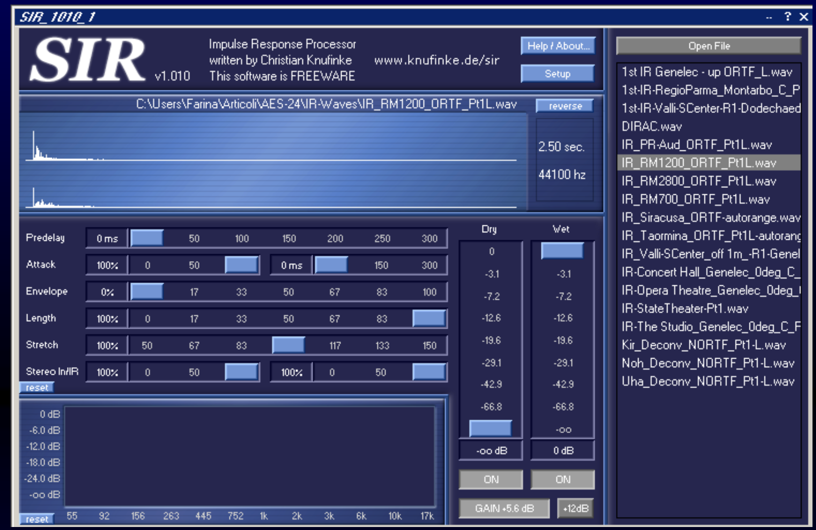


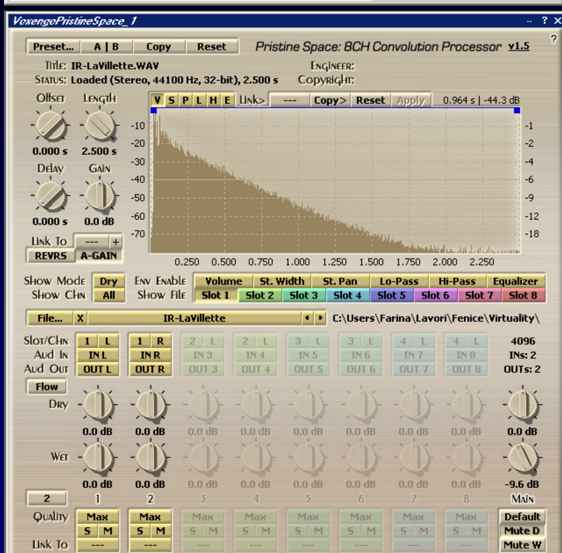
Performance: a fanless (silent) P-IV running at 2.5 GHz was capable of real time convolution of 2 inputs at 44.1 kHz, 24 bits, with 48 impulse responses, each 5s long, driving 24 Genelec loudspeakers (20 satellites + 4 subwoofers), employing 75% of the CPU time

Auralization Software

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







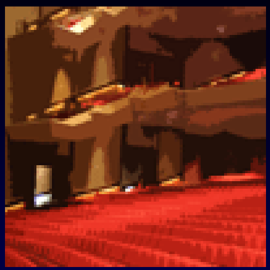
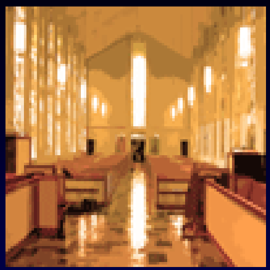
Linear Impulse Responses

Huge collections of impulse responses have been measured in famous theatres and concert halls all around the world, as well for renowned audio processing gear.

Recent advancements in the measuring technique, making use of the Exponential Sine Sweep signal, and the usage of multiple sound sources and microphones, make it possible to capture with minimal noise a detailed “acoustical photo” of existing rooms.

See, for example, the results of the Waves project on:

www.acoustics.net



Furthermore, all room acoustic modelling software is nowadays equipped with a “rendering” tool, which exports impulse responses suitable to be employed for auralization by linear convolution.

These synthetic IRs are even cleaner and with greater dynamic range than any measured IR

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 the simulation of a strongly nonlinear device and an almost linear one does not reveal any audible difference, because the nonlinear behavior is removed for both

How can we perform not-linear convolution?

There are actually two competing approaches available in the audio industry:

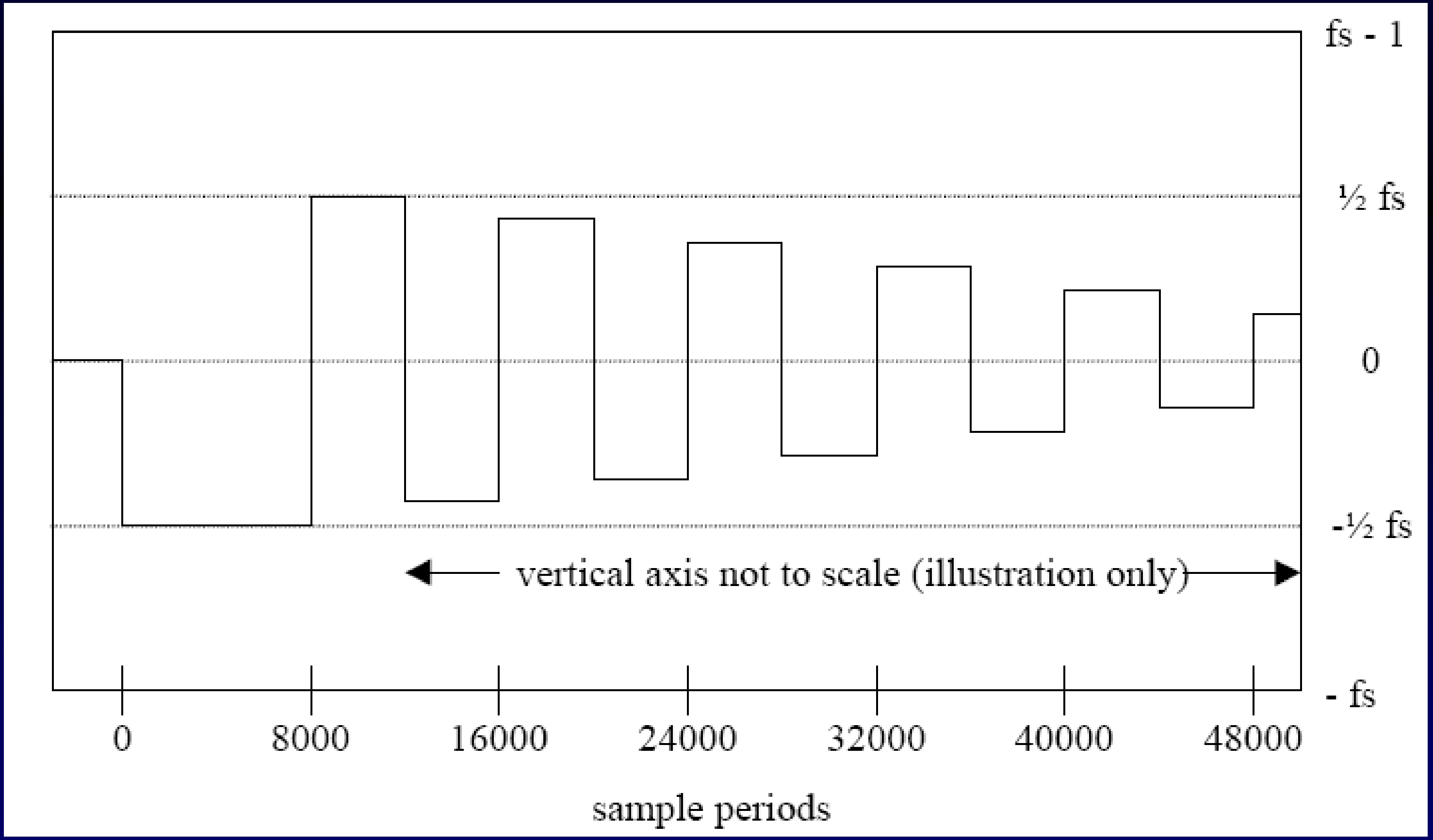
- IR-switching technique
- Diagonal Volterra Kernel

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 those of Bellini and Farina (1998) 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.

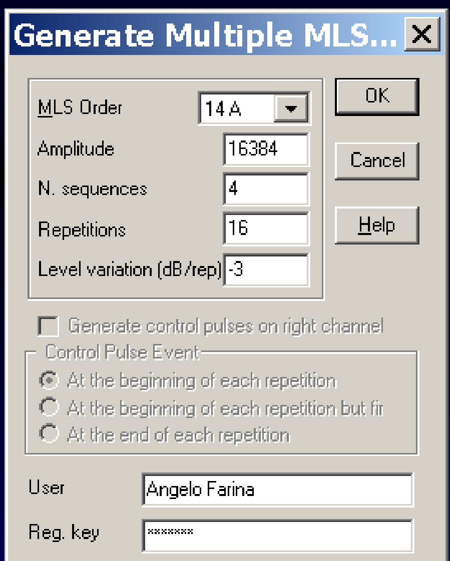
Measurement of multiple IRs

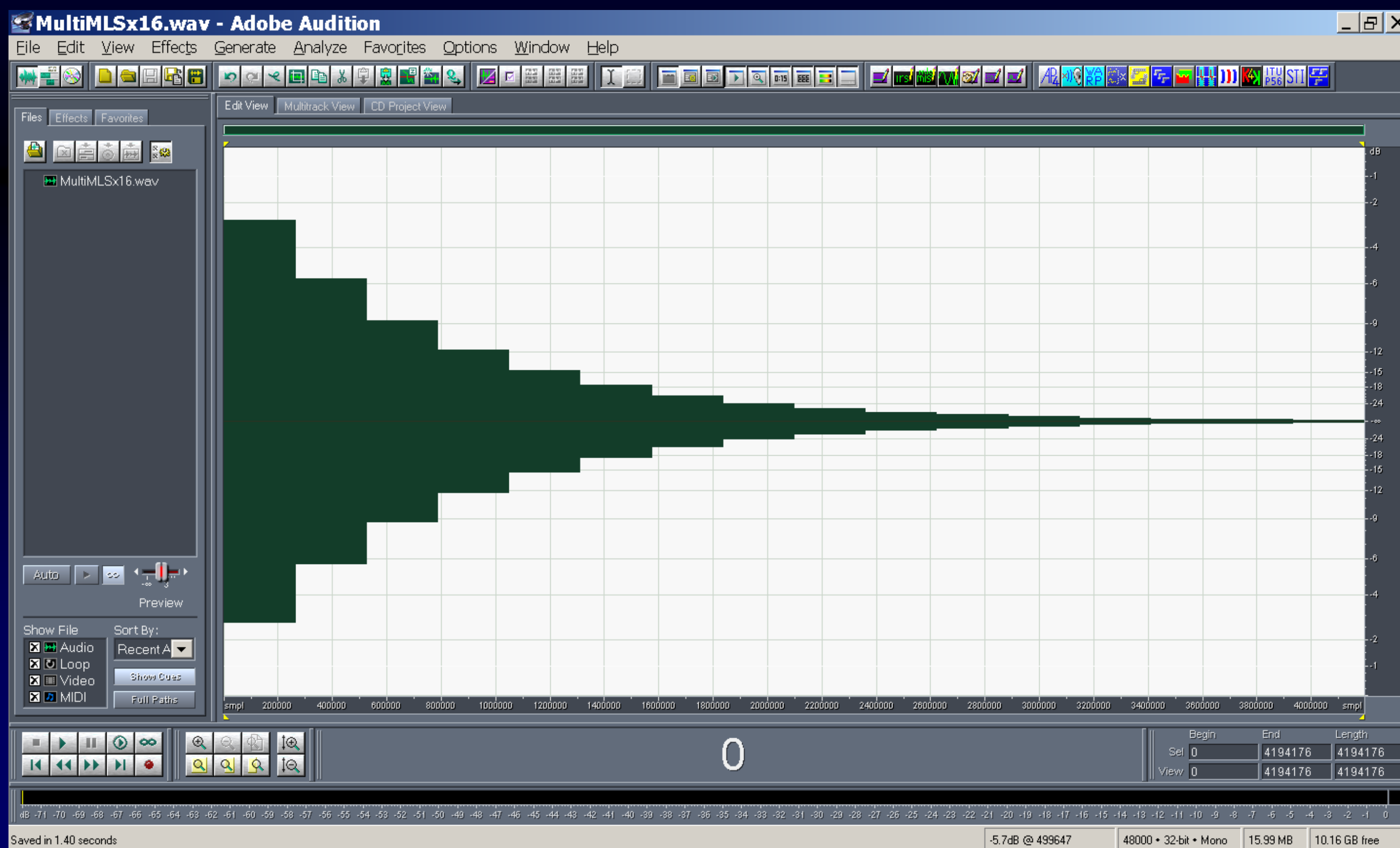
- Michael Kemp employed a step function, with several steps of decreasing amplitude



Measurement of multiple IRs

- Farina e Bellini did employ a sequence of MLS repetitions, each with decreasing amplitude, providing better S/N ratio in real-world measurements





Implementation (Michael J. Kemp)

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

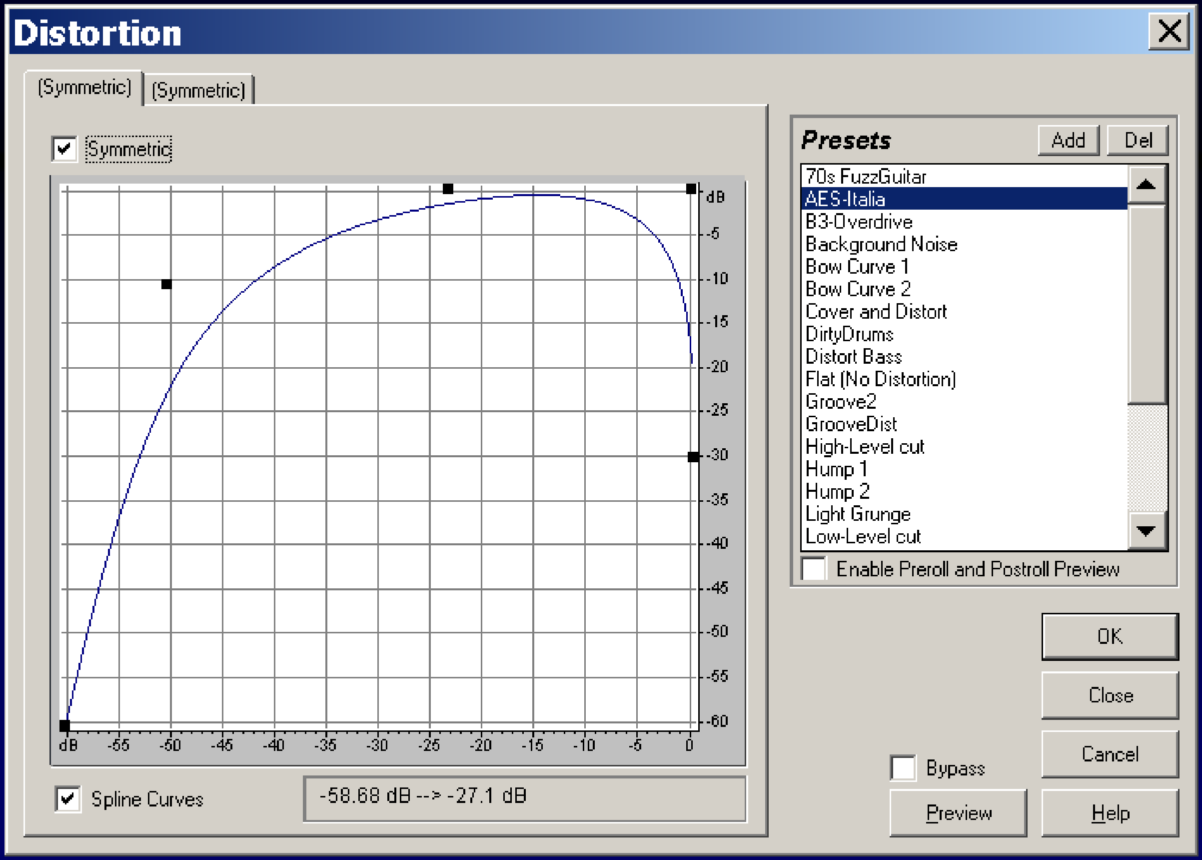
"The **Liquid Channel** is a revolutionary professional **channel** strip that can precisely replicate any classic mic-pre and compressor"

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

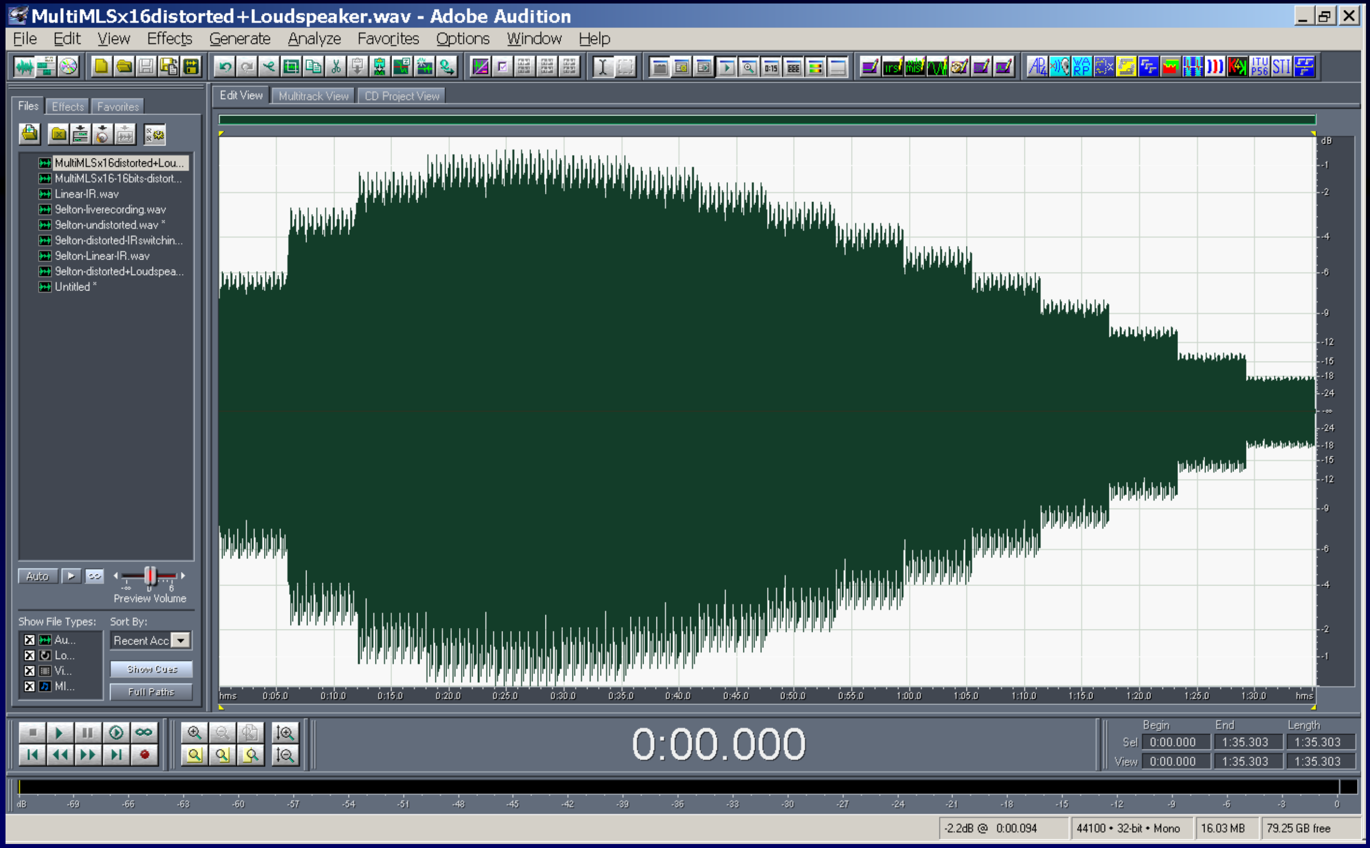
Example

- 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



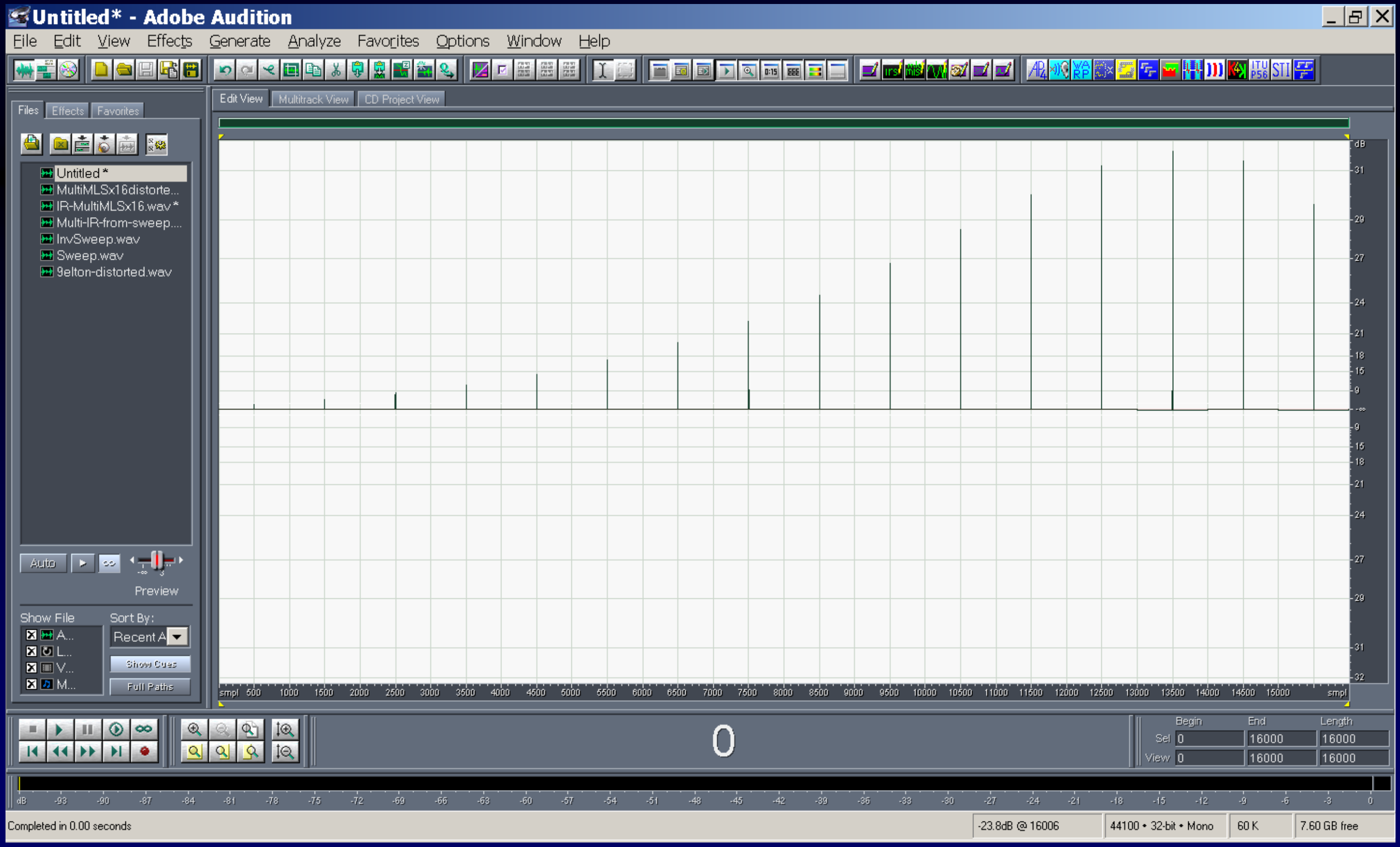
Example

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



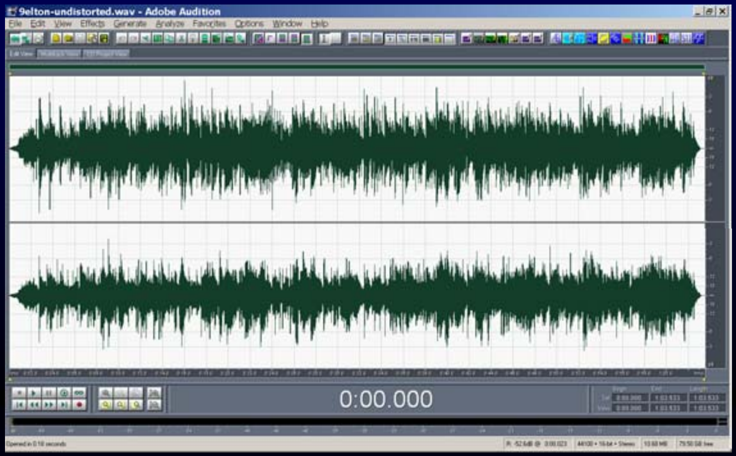
Example

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

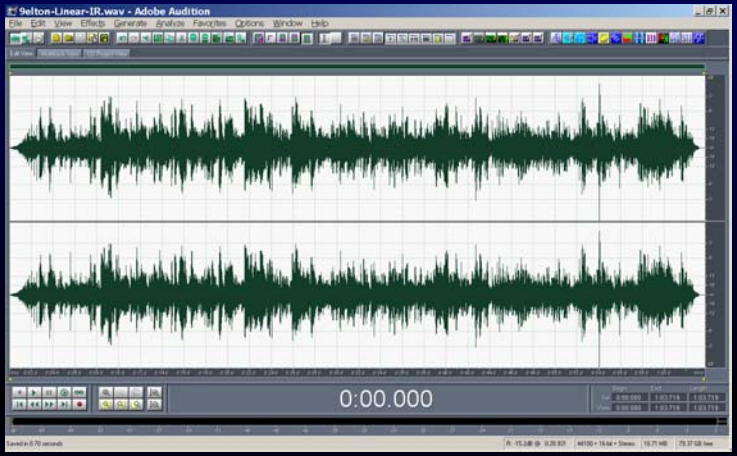


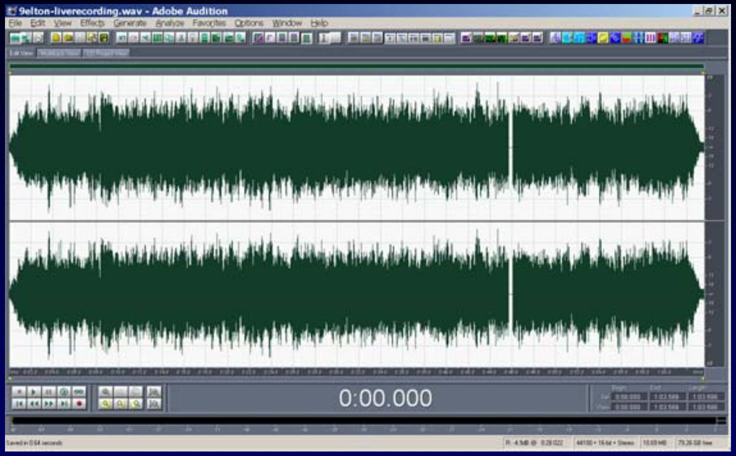
Audible evaluation of the performance

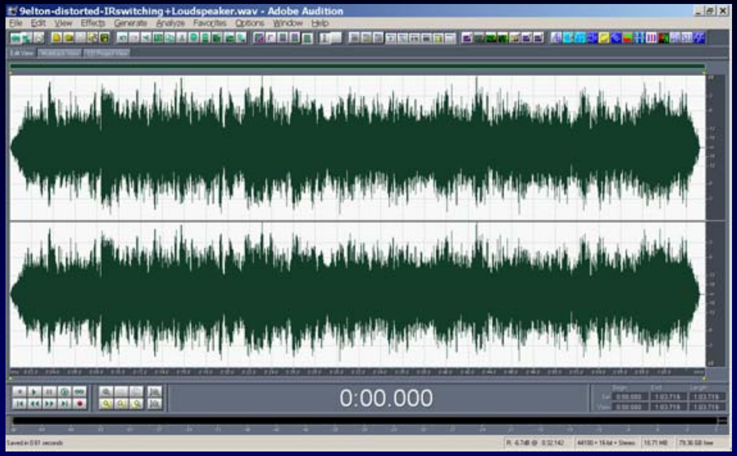
Original signal



Linear convolution







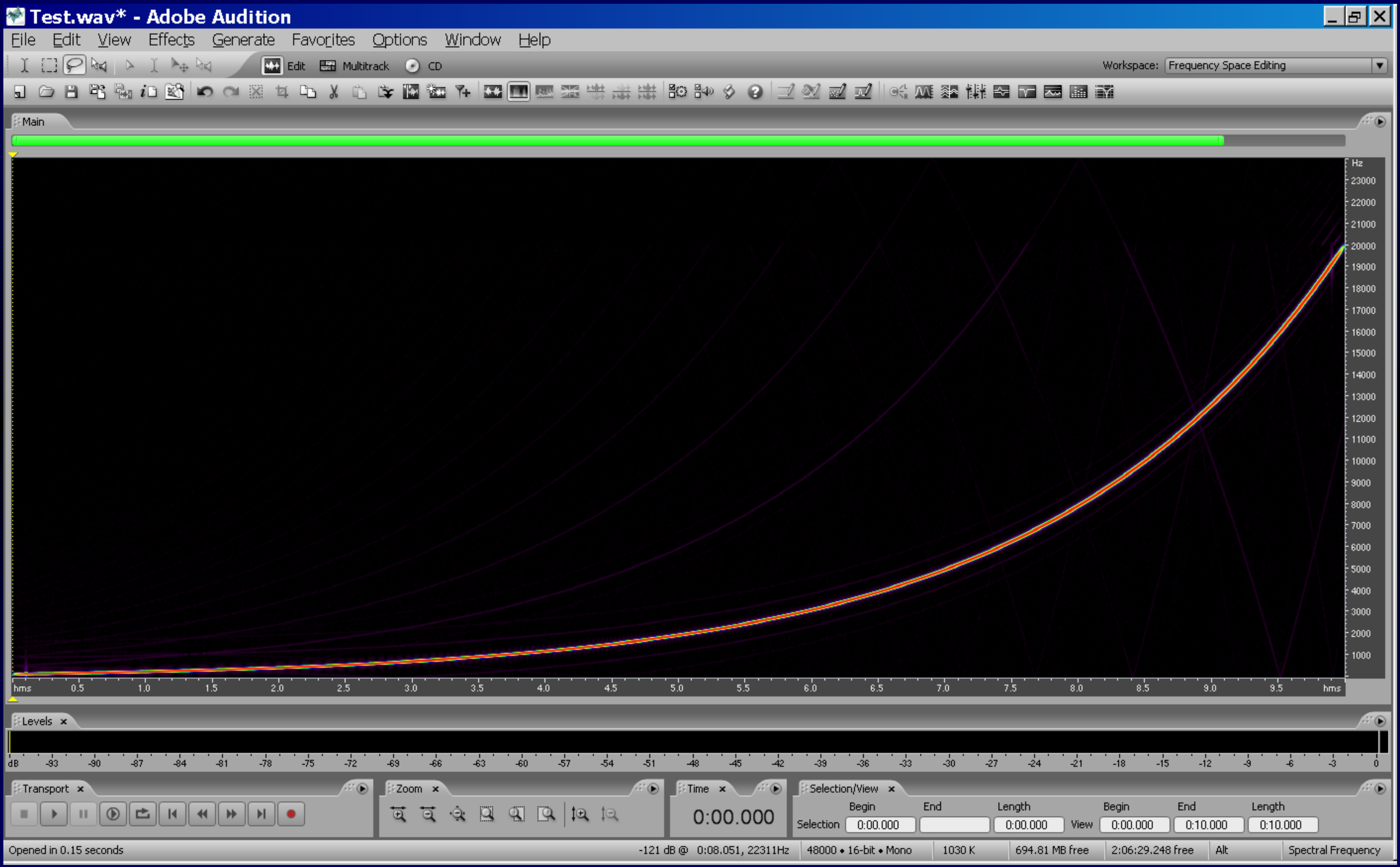
Live recording

Non-linear (IR switching)

Method 2 – Diagonal Volterra Kernels

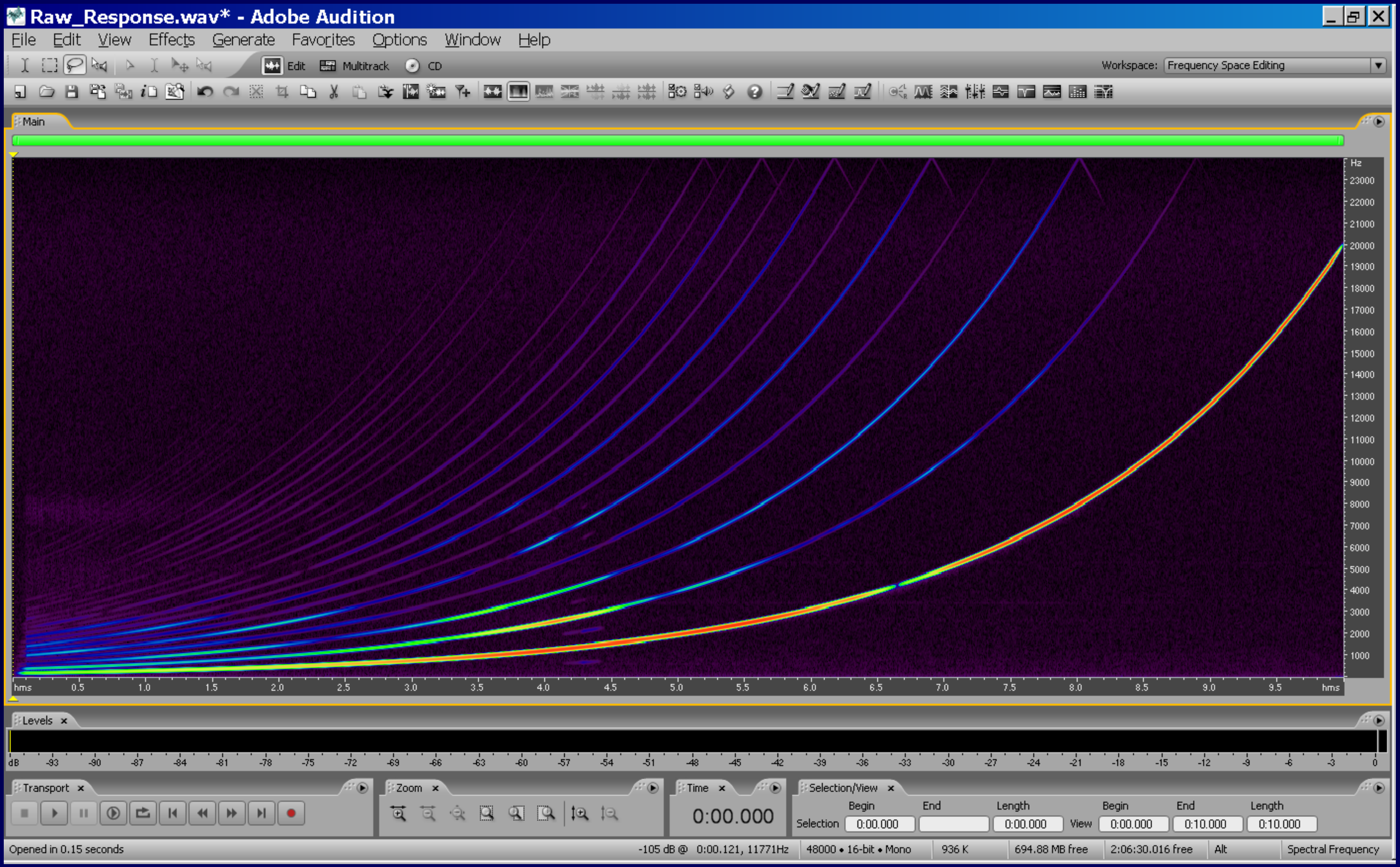
- We start from a measurement of the system based on exponential sine sweep (Farina, 108th AES, Paris 2000)
- Diagonal Volterra kernels are obtained by post-processing the measurement results
- These kernels are employed as FIR filters in a multiple-order convolution process (original signal, its square,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

Exponential sweep measurement



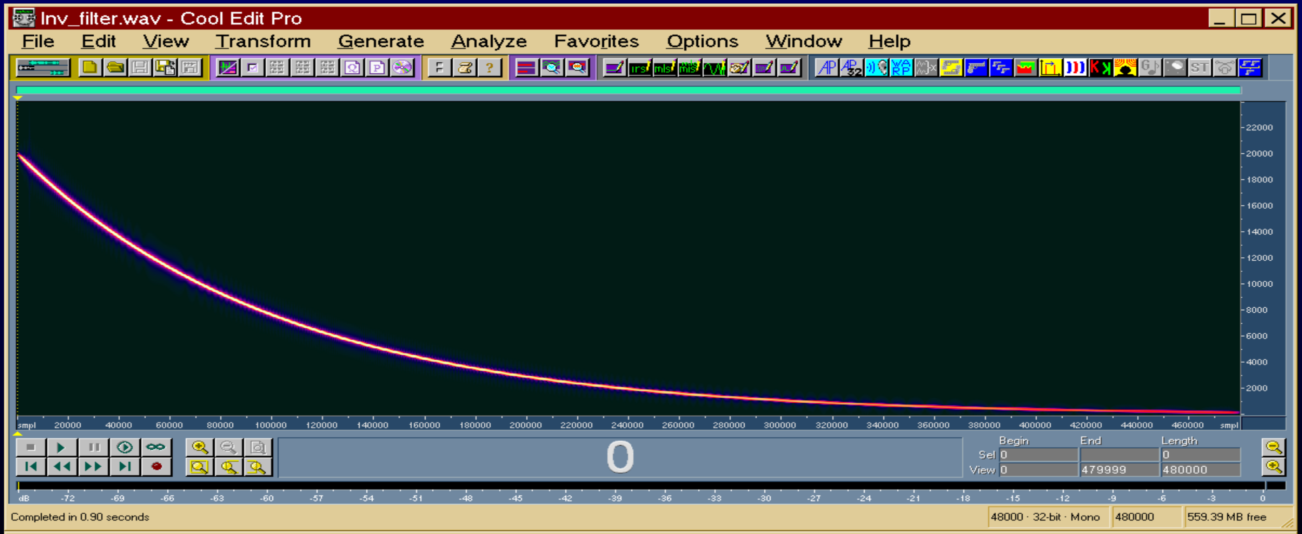
- The test signal is a sine sweep with constant amplitude and exponentially-increasing frequency

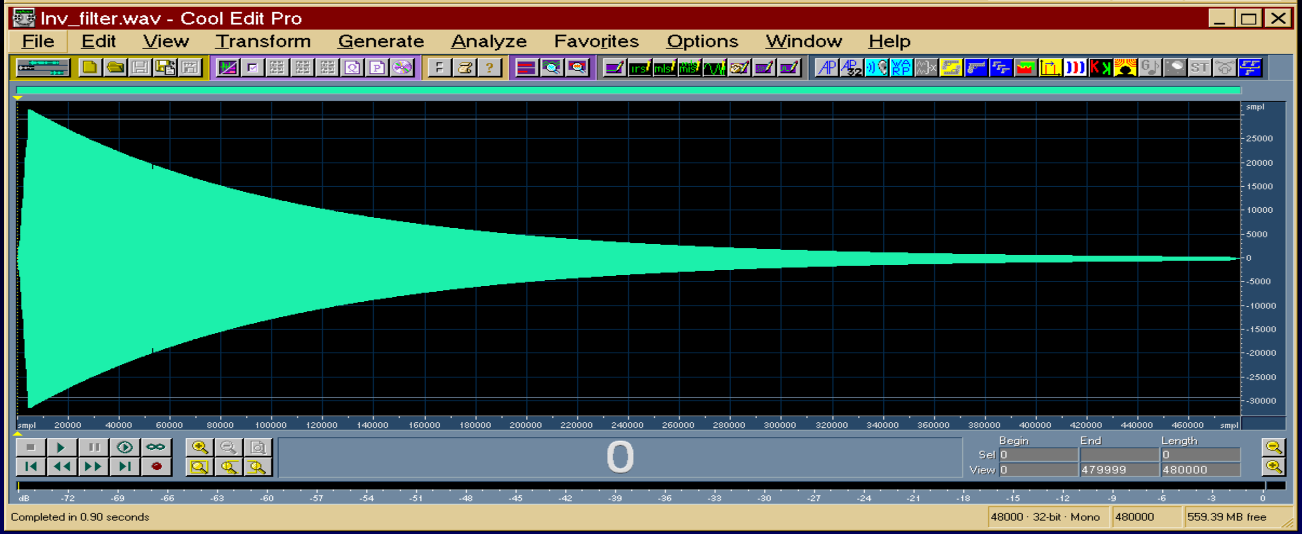
Raw response of the system



Many harmonic orders do appear as colour stripes

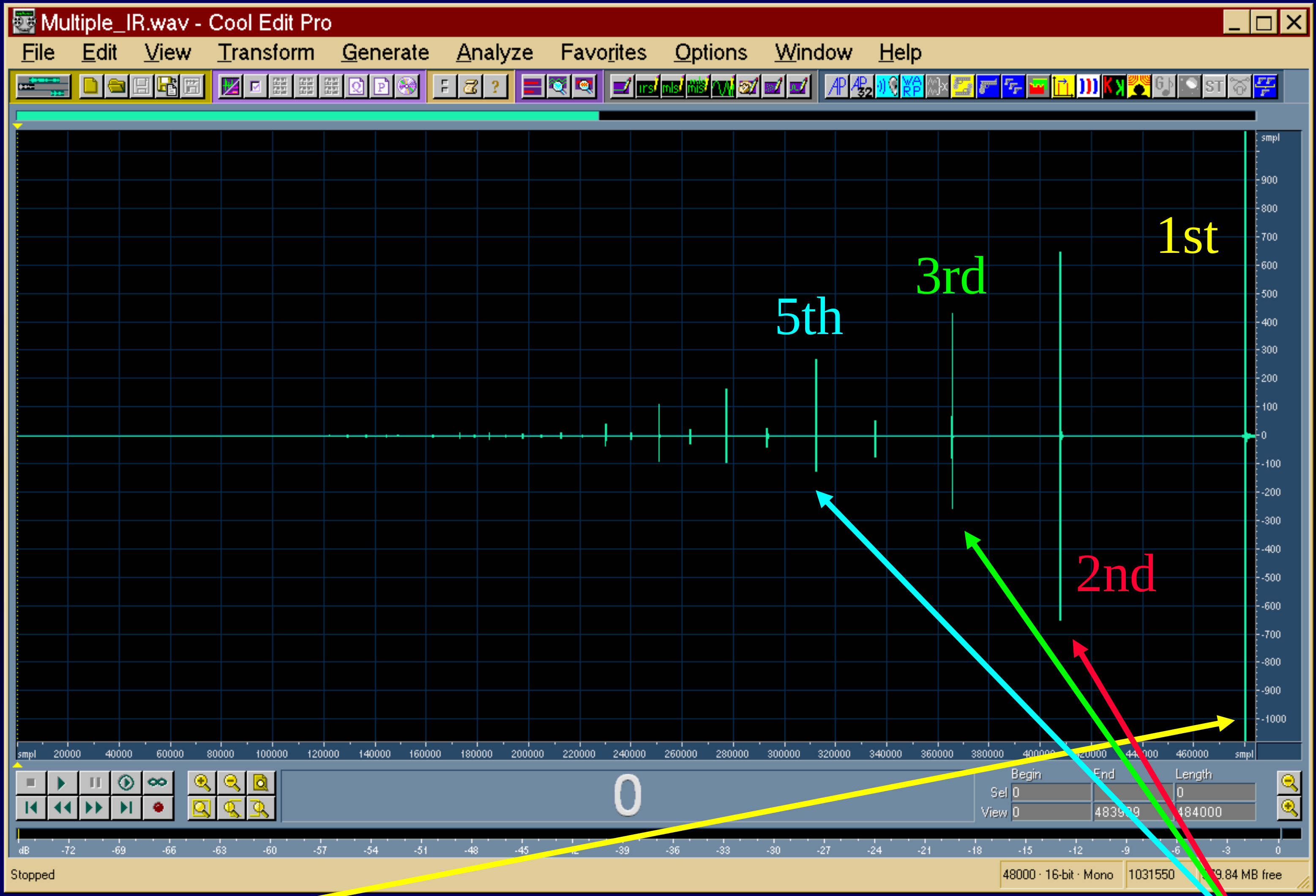
Deconvolution of system's impulse response





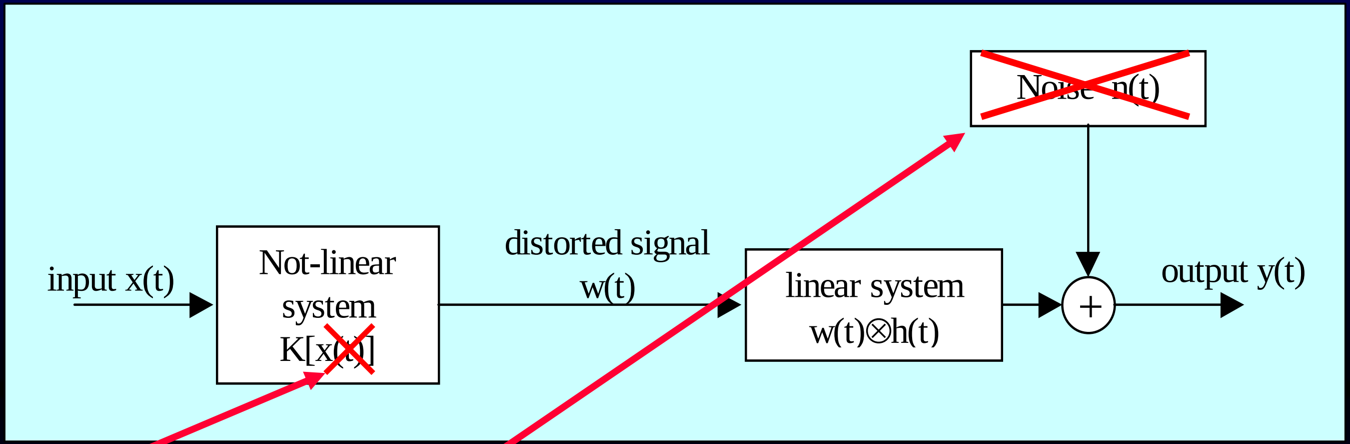
The deconvolution is obtained by convolving the raw response with a suitable inverse filter

Multiple impulse response obtained



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

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.

Theory of nonlinear convolution

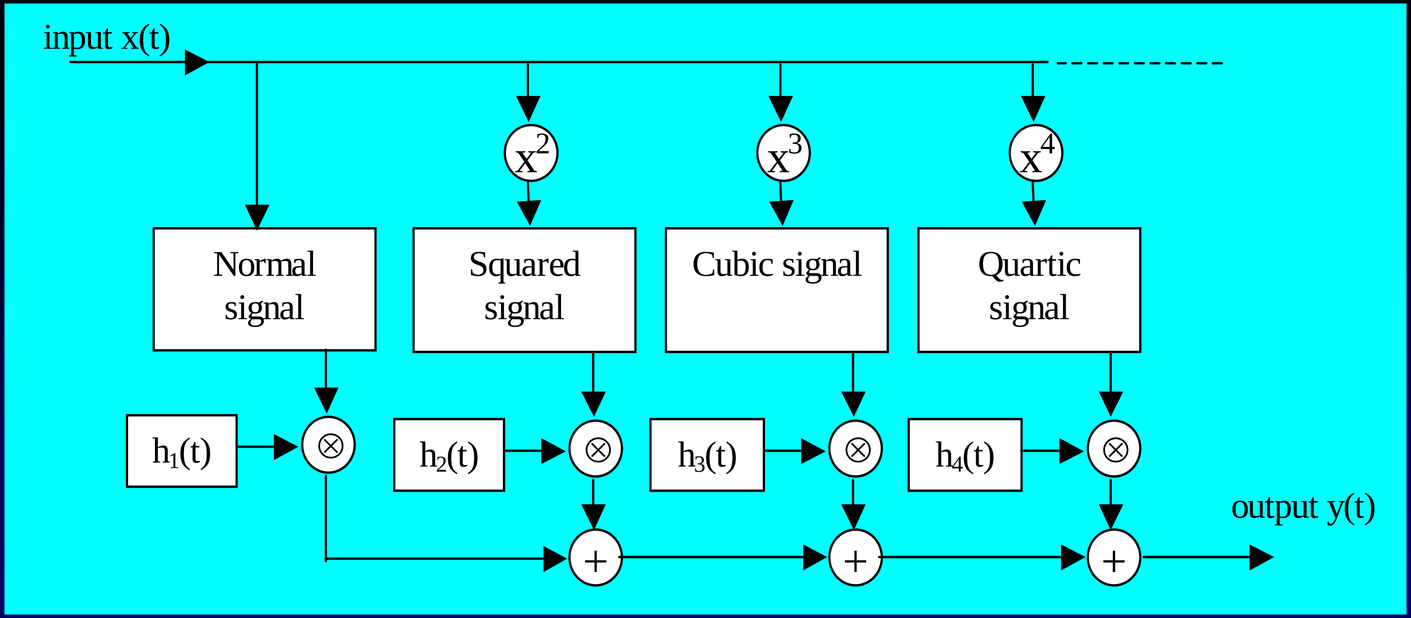
- The basic approach is to convolve separately, and then add the result, the linear IR, the second order IR, the third order IR, and so on.
- Each order IR is convolved with the input signal raised at the corresponding power:

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

The problem is that the required multiple IRs **are not** the results of the measurements: they are instead the diagonal terms of Volterra kernels, which can be derived from the set of measured Impulse Responses of several distortion orders

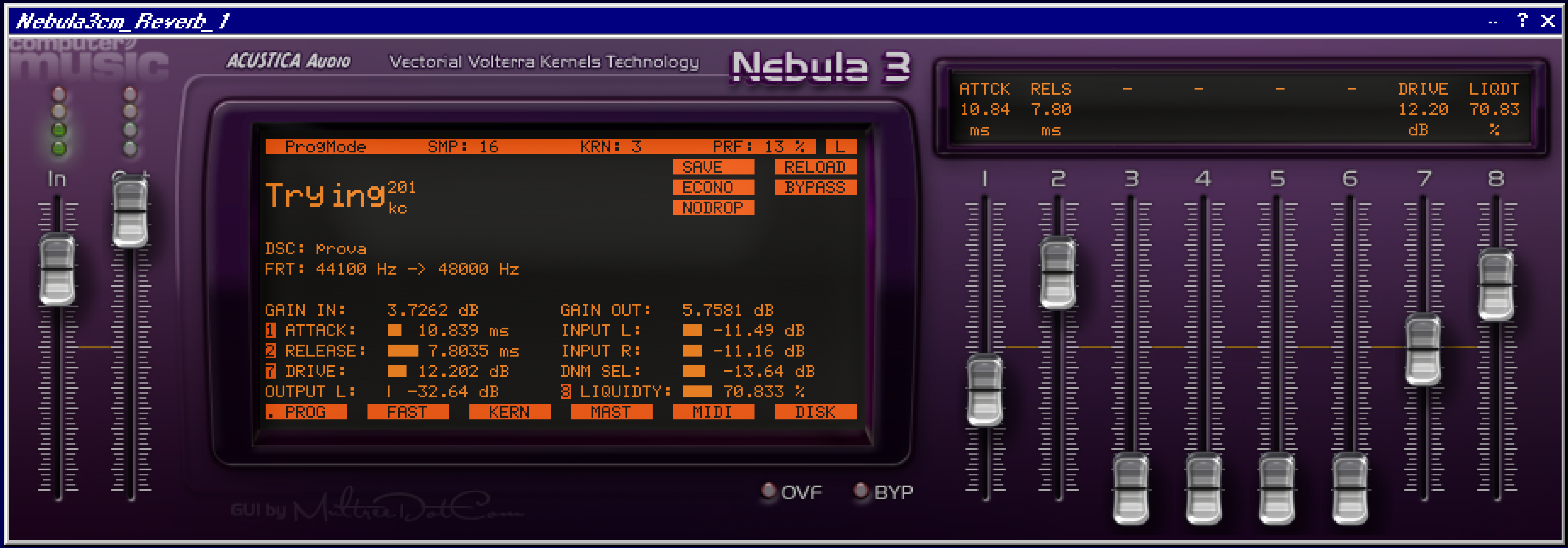
Efficient non-linear convolution

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

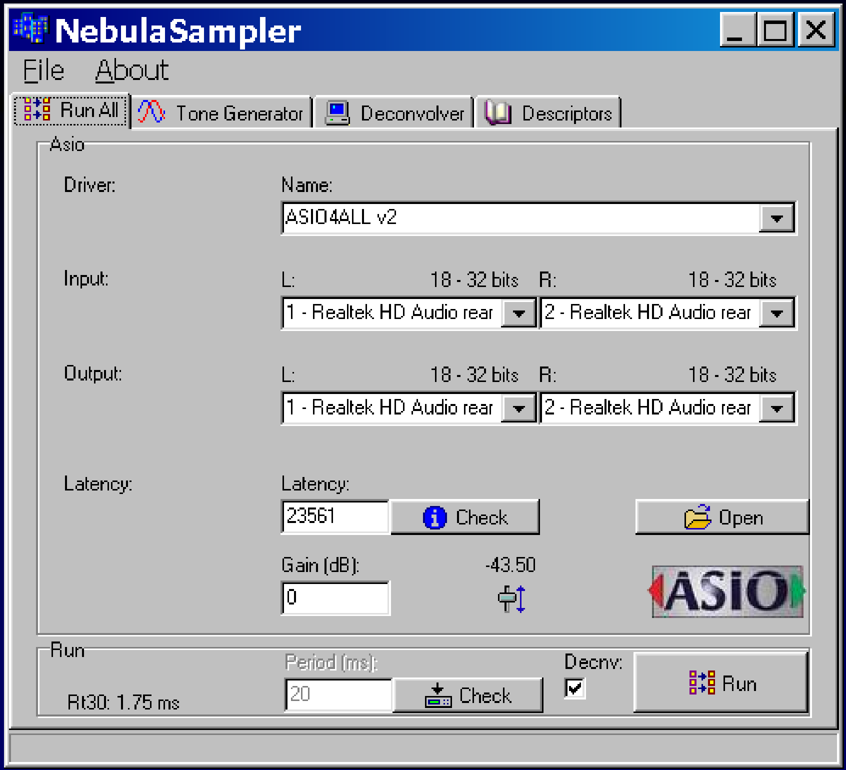
A small Italian startup company, Acustica Audio, developed a VST plugin based on the Diagonal Volterra Kernel method, named Nebula

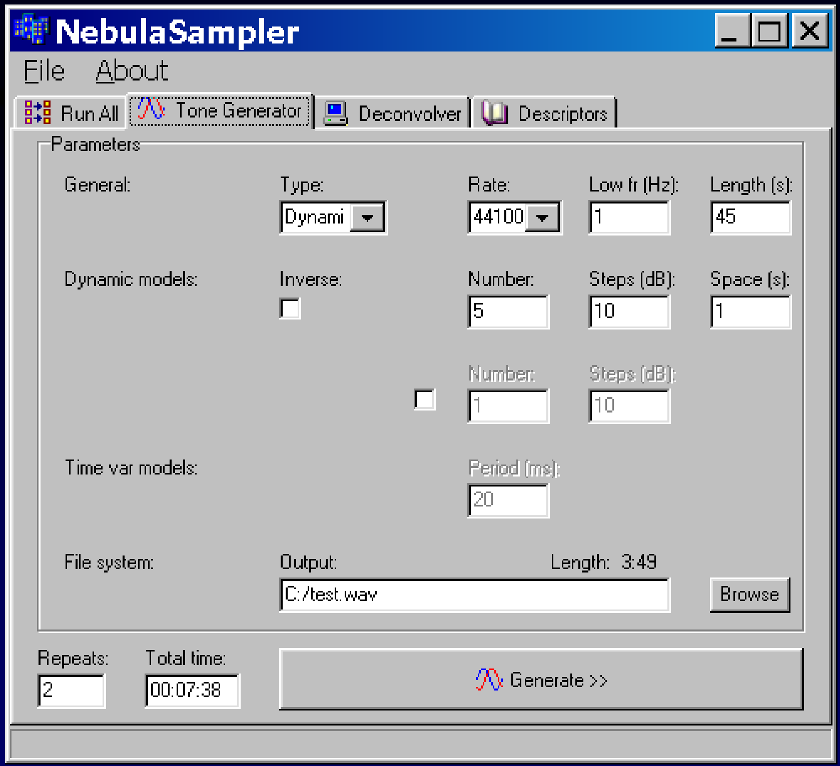


This is capable of real-time operation even with a very large number of filter coefficients

Software implementation

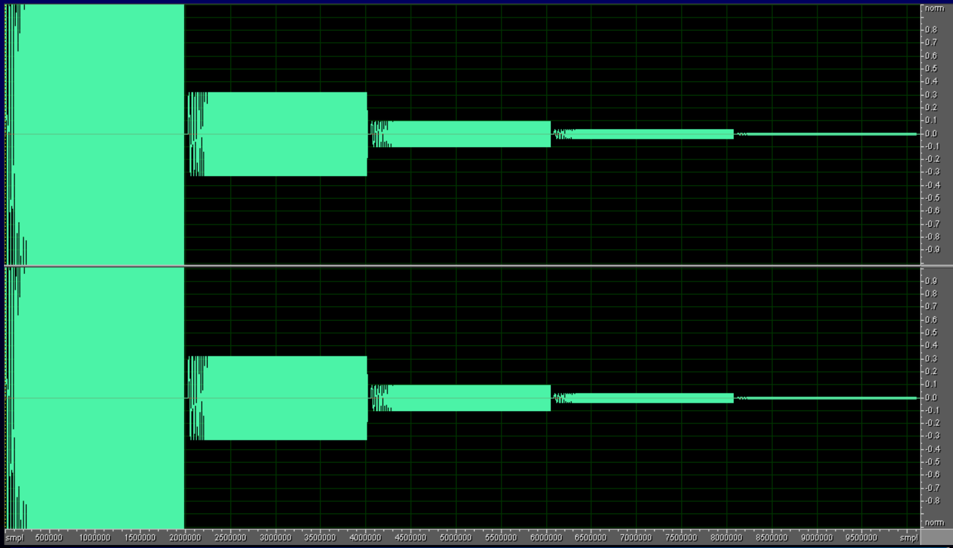
Nebula is also equipped with a companion application, Nebula Sampler, designed for automatizing the measurement of a not linear system with the Exponential Sine Sweep method:

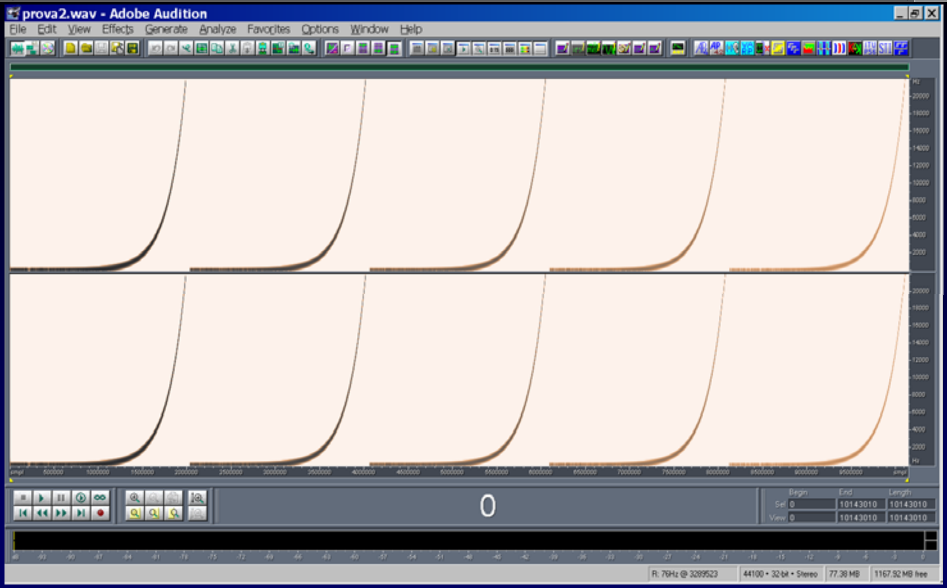




Time-variant systems

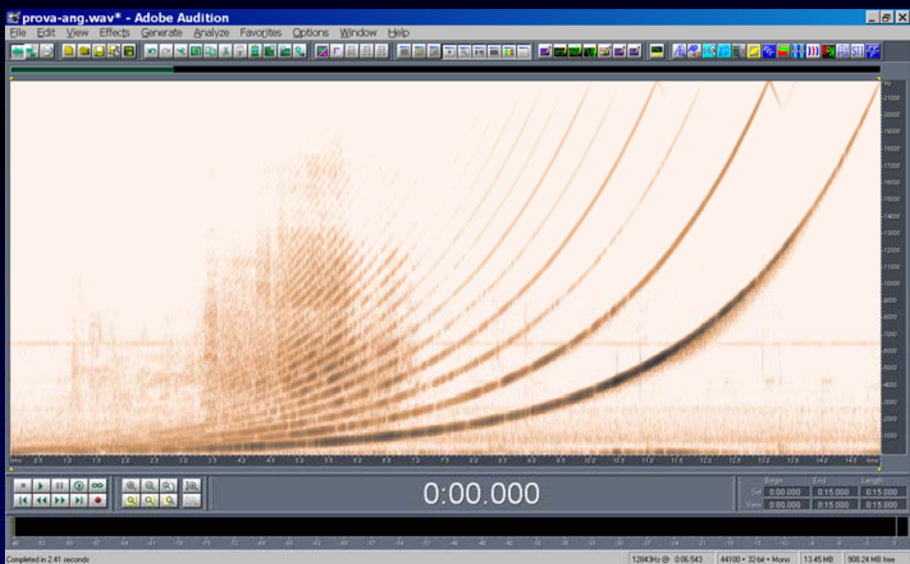
Nebula can sample also time-variant systems, such as flangers or compressors, by repeating the sine sweep measurement several times, along a repetition cycle or changing the signal amplitude

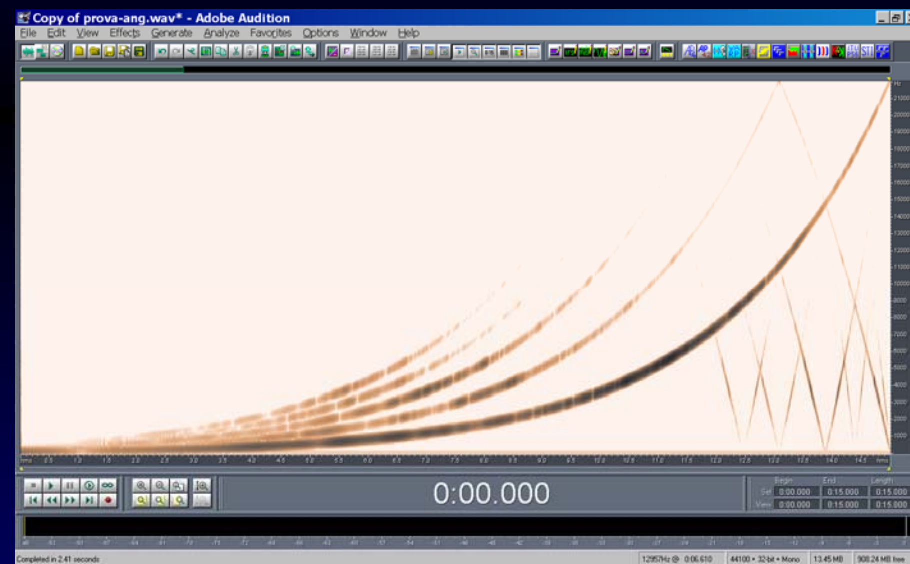




Reconstruction accuracy

Nebula is actually limited to Volterra kernels up to 5th order, and consequently does not emulate high-frequency harmonics:





Audible evaluation of the performance

Original signal

Linear convolution

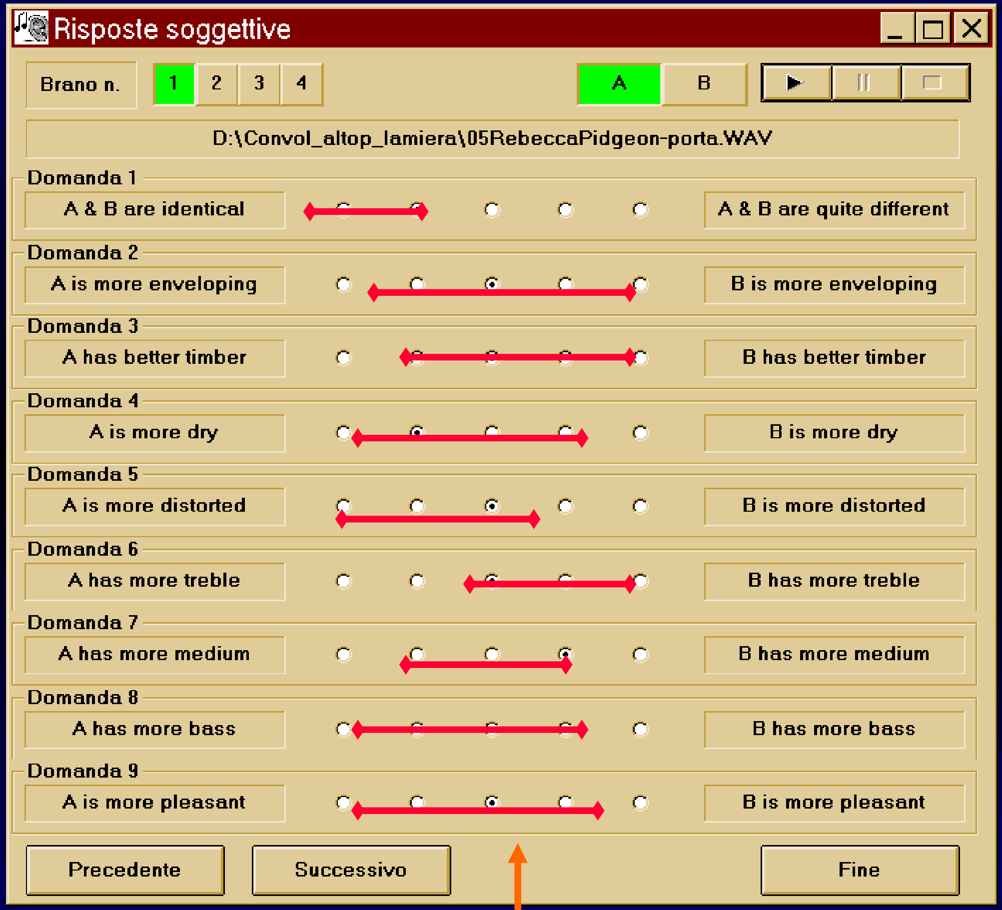
These last two were compared in a formalized blind listening test

Live recording

Non-linear
Diagonal Volterra Kernel

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 answers

Results

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

Comments

- Most listeneres judged the two samples identical
- However, sample B, on average, has slightly “better timber” (less distortion at high frequency), whilst sample A is “more distorted”.
- Despite of the slight reduction in perceived distortion, the not-linear emulation was slightly preferred to the real-world recording.

Another example

Original signal

Linear convolution

Live recording

Non-linear
Diagonal Volterra Kernel

Conclusions

- Traditional Linear Convolution is a powerful technique, but its reconstruction of the real world suffers for the limitations due to Linearity and Time Invariance
- Not-linear convolution is possible with two competing techniques:
  IR switching and Diagonal Volterra Kernels.
- The latter provides much more efficiency in computational load, and allows for easy emulation also of time-variant systems
- The Nebula software makes this technology available to everyone

Remarks

- The plugins for Adobe Audition shown here are freely downloadable from HTTP://www.aurora-plugins.com
- The Nebula VST plugin can be downloaded from HTTP://www.acusticaudio.com
- The sound samples employed for the subjective tests are available for download at HTTP://pcfarina.eng.unipr.it/public/AES123