

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](http://pcfarina.eng.unipr.it)



AUDIO ENGINEERING SOCIETY
Italian Section



UNIVERSITA' 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



Convoluting 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

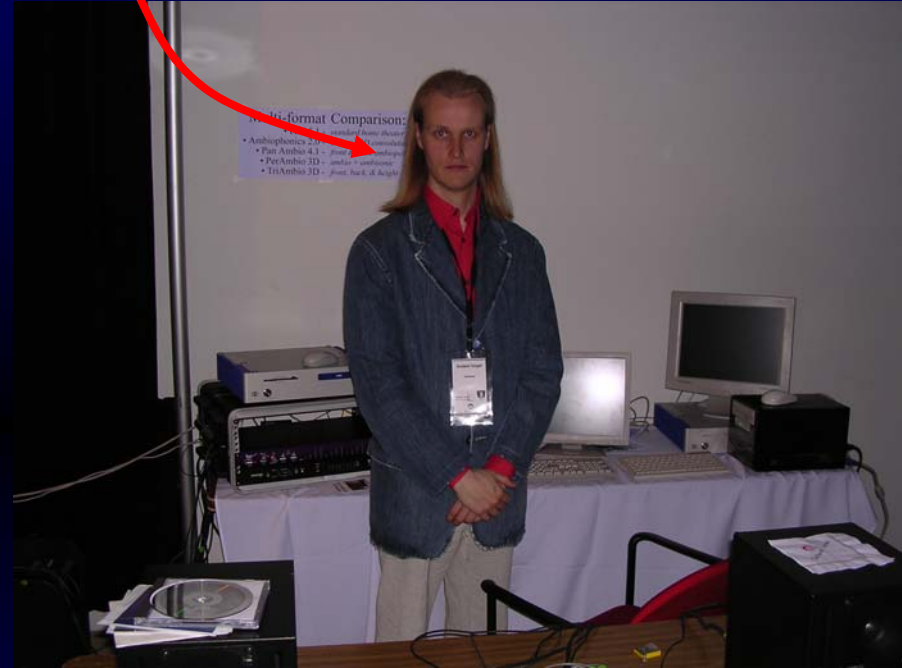


Photos taken on
16 december 2002



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

Convolve with clipboard

Graph showing amplitude vs. time (0 to 1.8 seconds). The amplitude starts at 0 and decays towards -100 dB.

Buttons: OK, Cancel, Help, Preview, Bypass

Autorange

- Full autorange & RemoveDC (2 passes)
- First Block autorangeGain (dB): 0

Advanced Options

- TimeReverse Impulse Response
- Preserve Length
- Preserve 16th bit

Channels to convolve

Audio Data: Left Right Both

Impulse Resp.: Left Right Both

Crosstalk Cancel Mode

Impulse Response is 2x2

Info

Audio Data: 44100 Hz/16 Bits/Stereo/422731 Samples
Impulse Resp. 96000 Hz/32 Bits/Stereo/192000 Samples
FFT Size: 524288 Samples

User: Angelo Farina
Reg. key: *****

SIR 1010.1 Impulse Response Processor
v1.010 written by Christian Knufinke www.knufinke.de/sir
This software is FREeware

File: C:\Users\Farina\Articol\AES-24\IR-Waves\IR_RM1200_ORTF_P1L.wav

reverse 2.50 sec
44100 Hz

Parameters:

- Predelay: 0 ms (0, 50, 100, 150, 200, 250, 300)
- Attack: 100% (0, 50, 100) 0 ms (0, 50, 100, 150, 300)
- Envelope: 0% (0, 17, 33, 50, 67, 83, 100)
- Length: 100% (0, 17, 33, 50, 67, 83, 100)
- Stretch: 100% (50, 67, 83, 100) (117, 133, 150)
- Stereo In/IR: 100% (0, 50, 100) 100% (0, 50)

Gain: Dry 0, Wet -3.1, -3.1, -7.2, -7.2, -12.6, -12.6, -18.6, -18.6, -23.1, -23.1, -42.9, -42.9, -66.8, -66.8, -∞ dB, -∞ dB

EQ: ON, ON, GAIN +5.6 dB, +12 dB

Open File:

- 1st-IR_Genelec - up ORTF_L.wav
- 1st-IR_RegioParma_Montarbo_C_P
- 1st-IR_Valli-SCenter-R11-Dodechaed
- DIRAC.wav
- IR_PR-Aud_ORTF_P1L.wav
- IR_RM1200_ORTF_P1L.wav
- IR_RM2800_ORTF_P1L.wav
- IR_RM700_ORTF_P1L.wav
- IR_Siracusa_ORTF-autorange.wav
- IR_Taormina_ORTF_P1L-autorange
- IR_Valli-SCenter_off_1m_R1-Genel
- IR-Concert Hall_Genelec_0deg_C_
- IR-Opera Theatre_Genelec_0deg_L
- IR-StateTheater-P11.wav
- IR-The Studio_Genelec_0deg_C_F
- Kir_Deconv_NORTF_P1L.wav
- Noh_Deconv_NORTF_P1L.wav
- Uha_Deconv_NORTF_P1L.wav

WavesPristineSpace_1

Presets: A | B | Copy | Reset | Pristine Space: BCH Convolution Processor v1.5

Title: IR-Lavillette.WAV ENGINEER: Copyright:

Status: Loaded (Stereo, 44100 Hz, 32-bit), 2.500 s

Offset Length: 0.000 s 2.500 s

Delay Gain: 0.000 s 0.0 dB

Graph showing amplitude vs. time (0 to 2.250 seconds). The amplitude starts at 0 and decays towards -18 dB.

Buttons: REVERS, A-GAIN

Buttons: Dry, Wet, Flow, Quality, Link To

Buttons: Show Mode, Show File, Slot 1 to Slot 8

Buttons: File...

Buttons: Slot/Chn, Aud In, Aud Out

Buttons: Dry, Wet, Flow, Quality, Link To

Buttons: Max S M, Default Mute W

WaveShell VST 5.0.1

IR-1 Setup A

Buttons: Full CPU, Reverse, Bypass, Gain Envelope, Clear

Buttons: R 1.00, F 600, R 1.00, F 2500

Name: Hall - 1
Type: Concert Hall
Date: 24 Mar 2004
SR: 96000Hz -> 44100Hz
Emitter: Genelec S30D

Convolution: 1.85s 1.85s
RT60: 1.4s 1.4s
Channels: 4 4
Size: 11267 11267
Distance: 13m 13m

Buttons: Zoom +, 0.500Sec, 1.000Sec, Reset

Buttons: Reverb Time, Size, Density, Reso, Decorr, Latency, Dry/Wet, Direct, Pre-delay, Output

Buttons: Cnv. Start, Cnv. Length, Full, RT60, Ratio

Buttons: ER Buildup, ER/ITR-X

Buttons: Damping, Equalizer (16, 62, 250, 1k, 4k, 16k), G, F, Q

Buttons: ER, Tail

What's missing in linear convolution ?

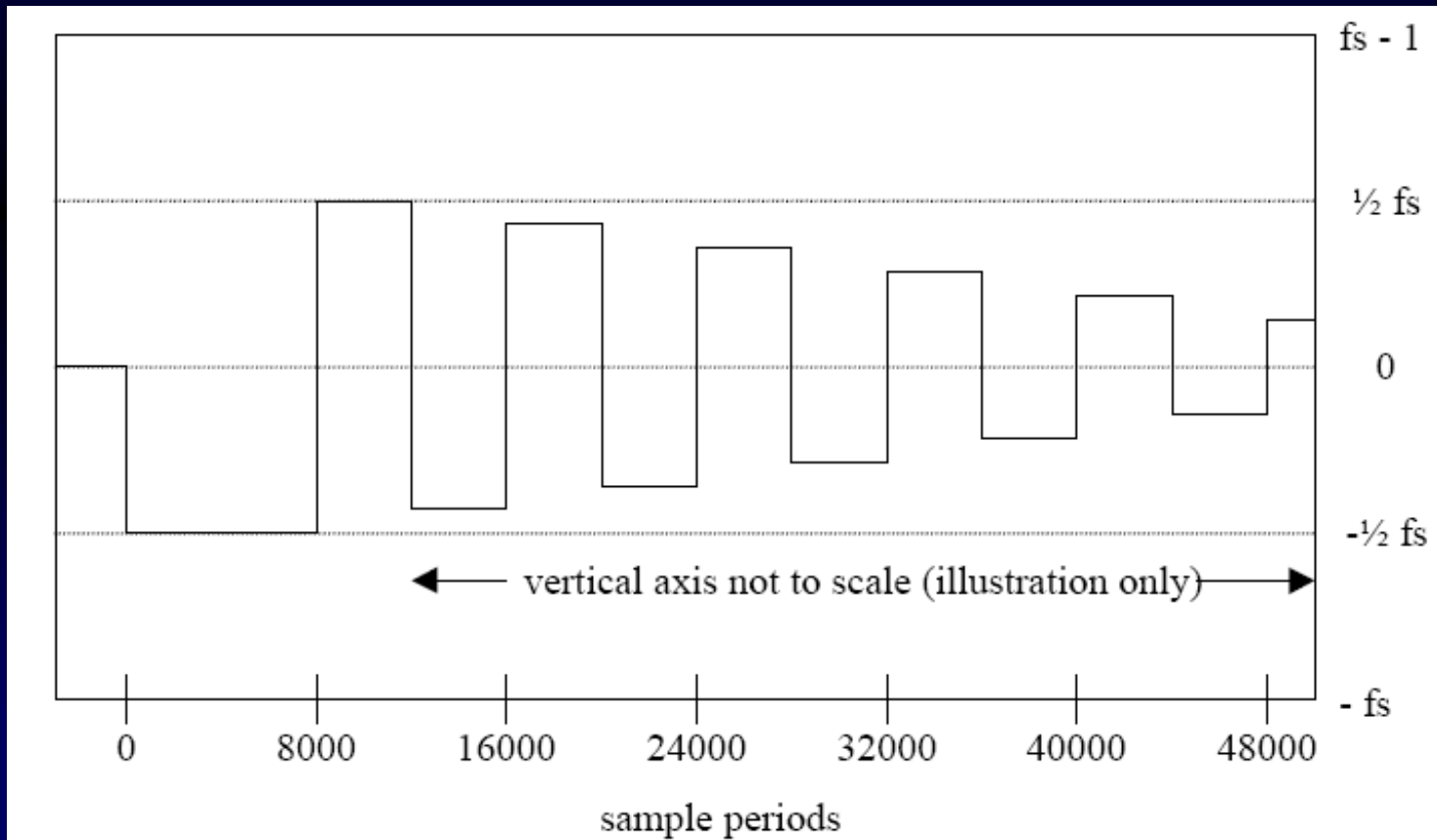
- No harmonic distortion, nor other nonlinear effects are being reproduced.
- From a perceptual point of view, the sound is judged “cold” and “innatural”
- A comparative test between a strongly nonlinear device and an almost linear one does not reveal any audible difference, because the nonlinear behavior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

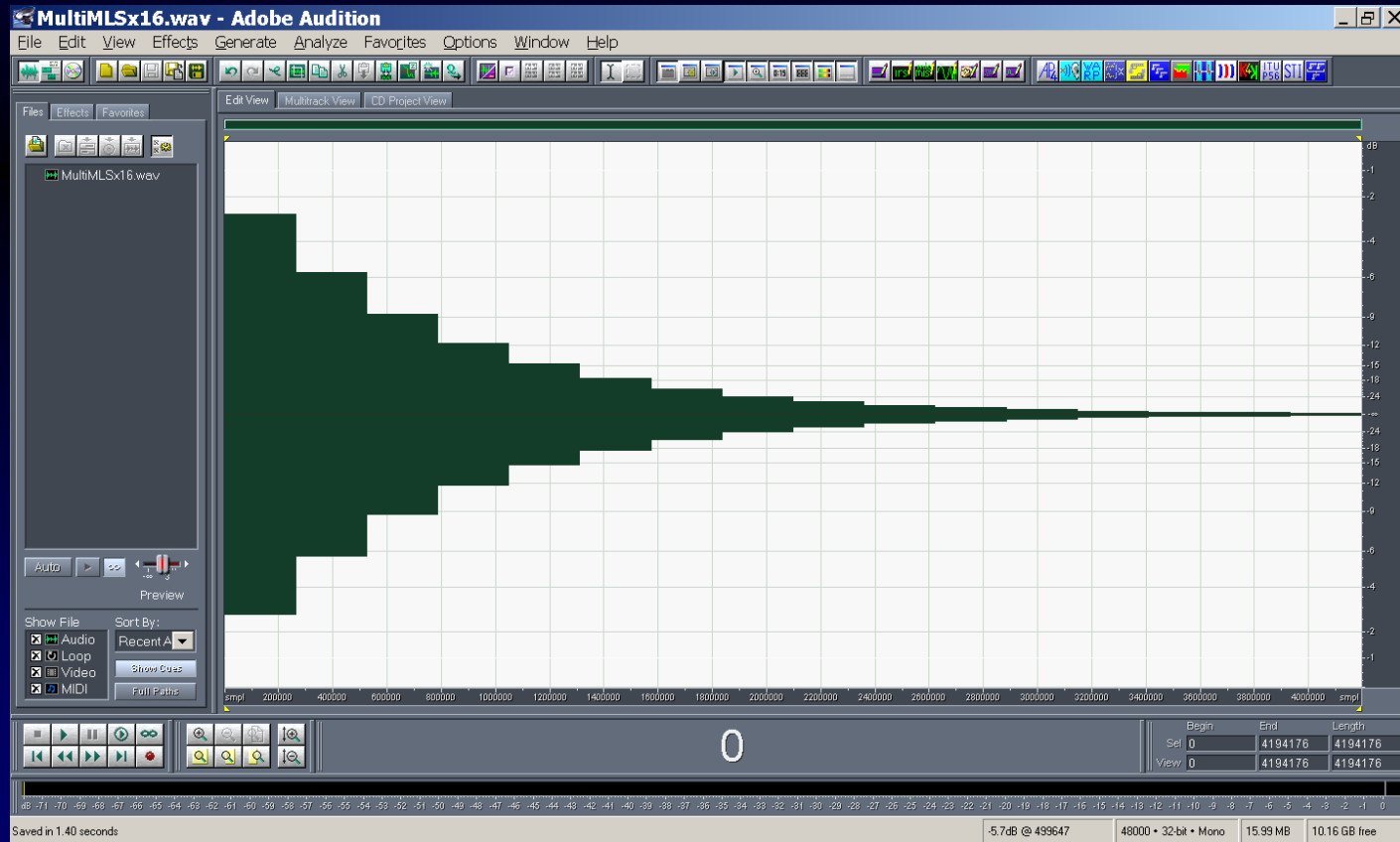
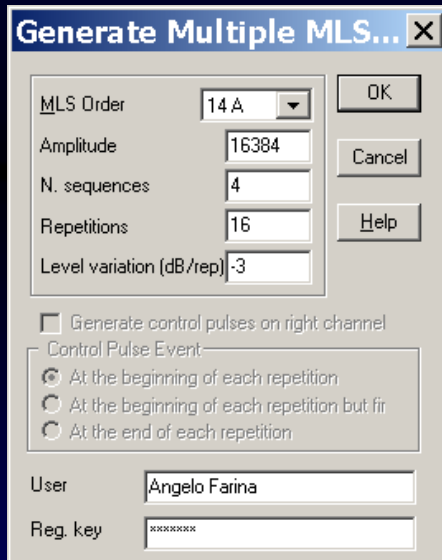
Measurement of multiple IRs

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



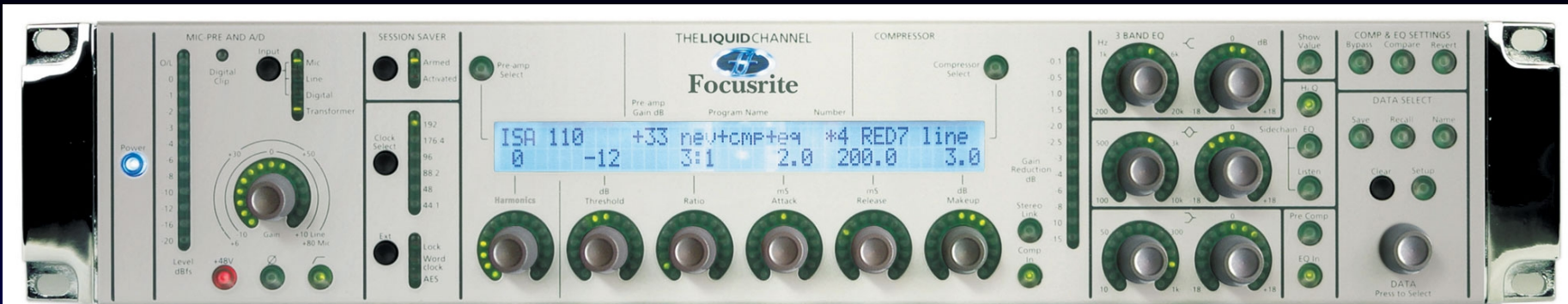
Misura delle IR multiple

- 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



z/rounds

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

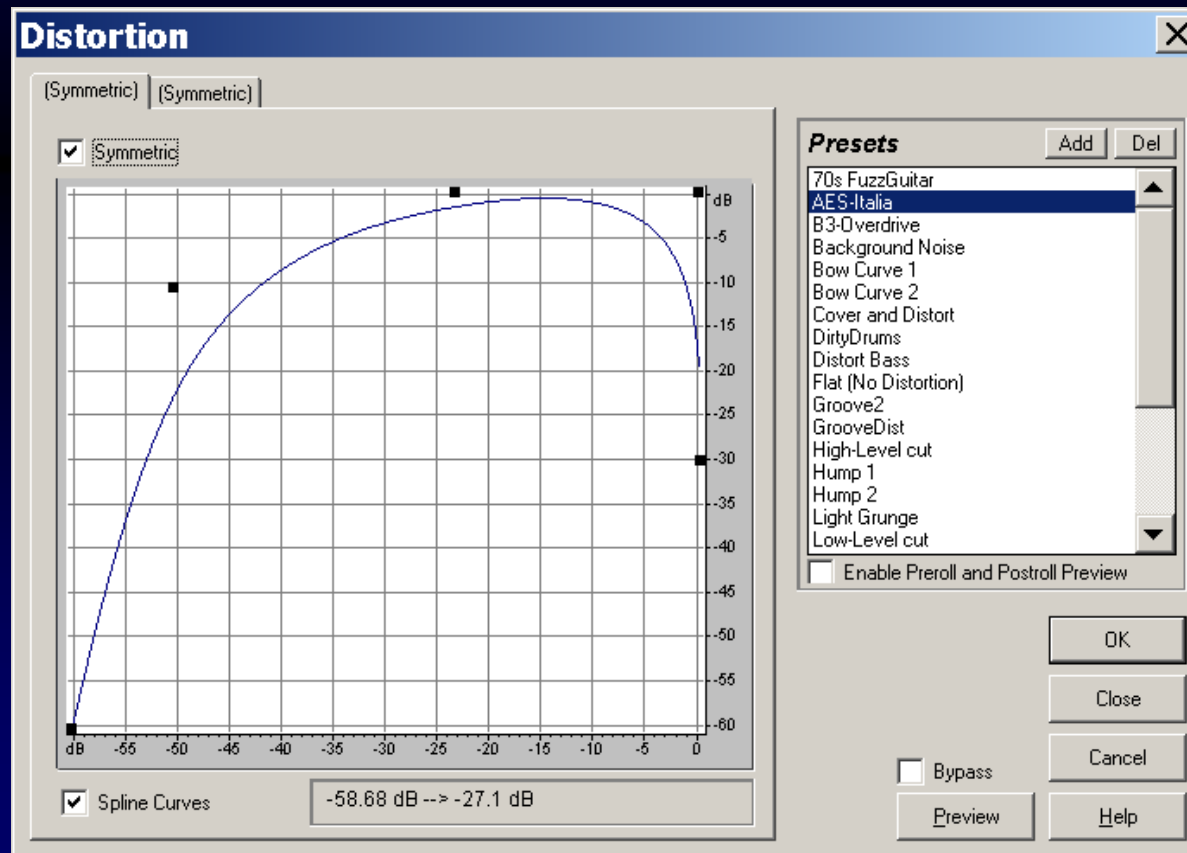
Implementation (Farina/Bellini)

- 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



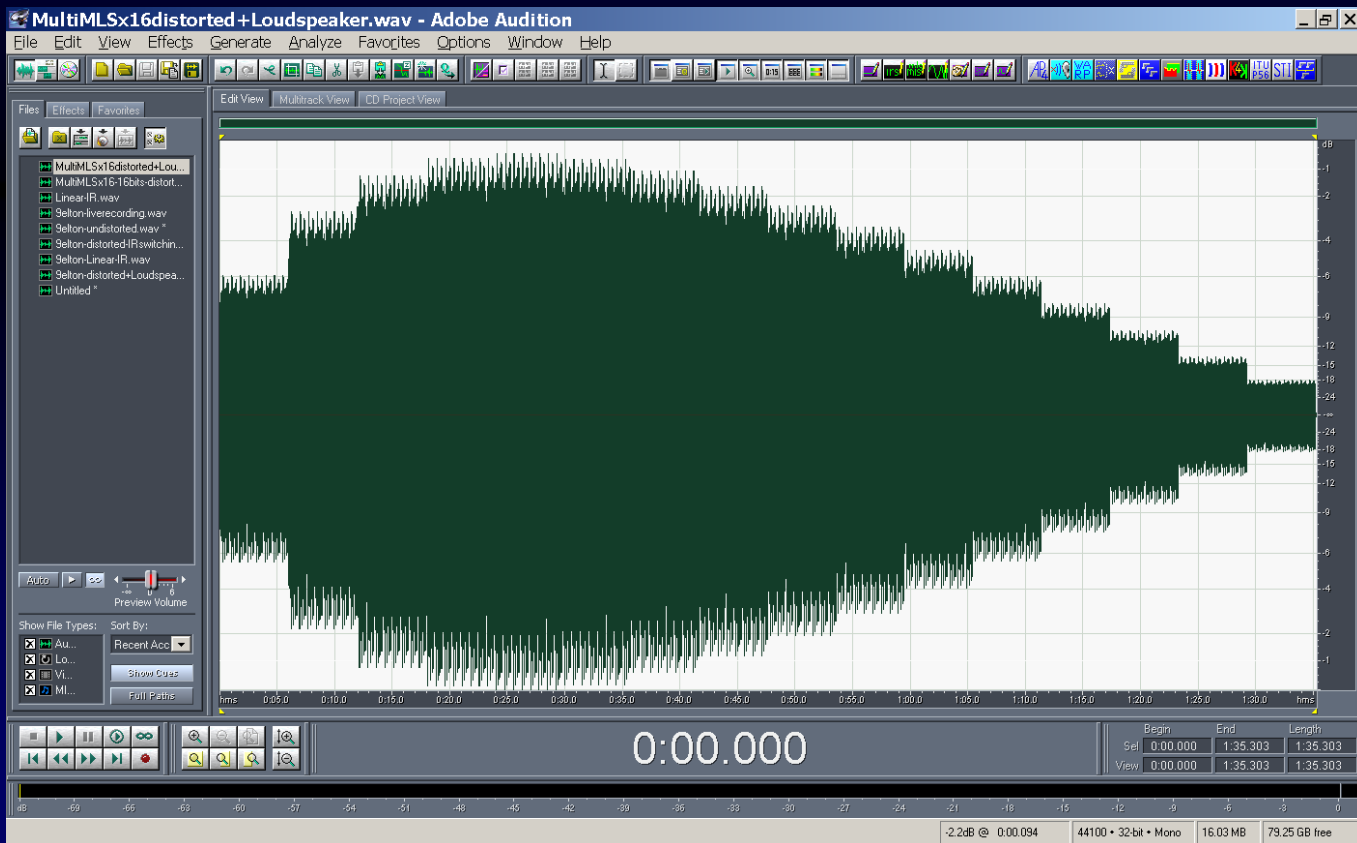
Example

- 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



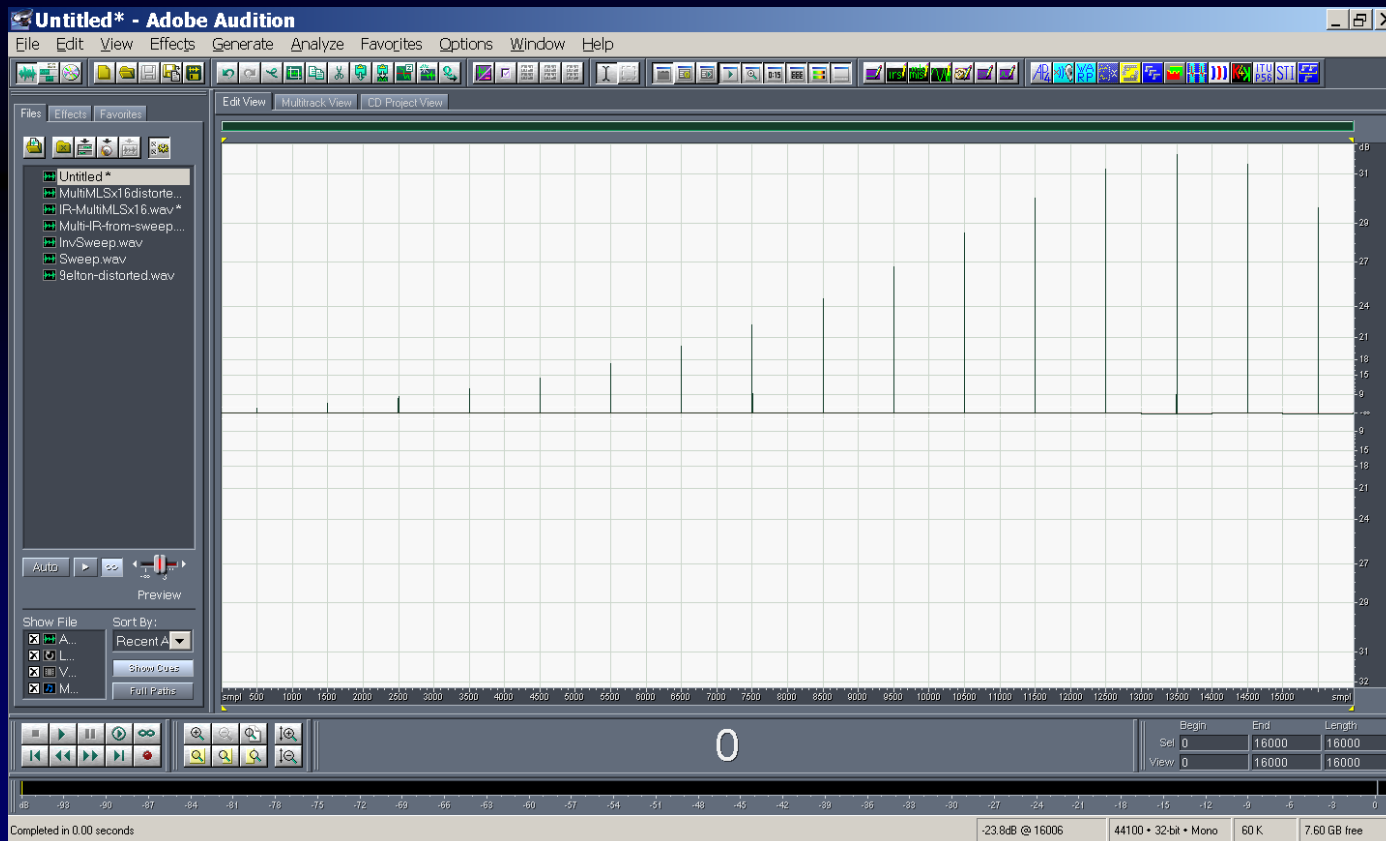
Example

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



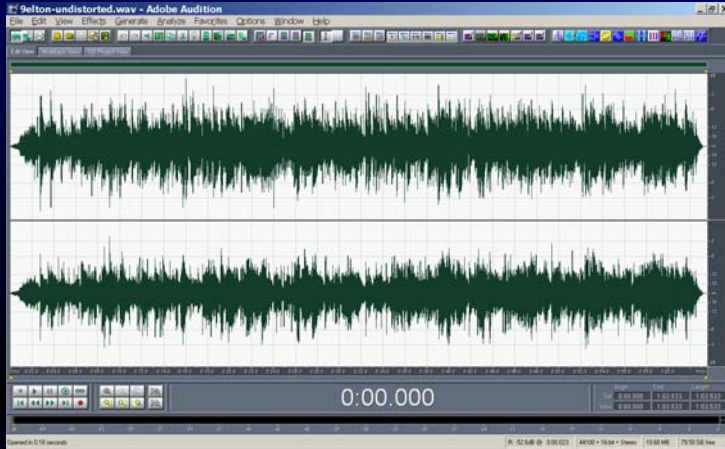
Example

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

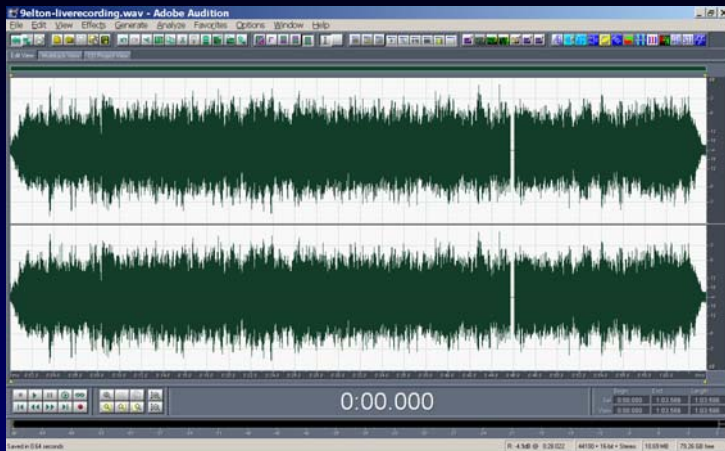
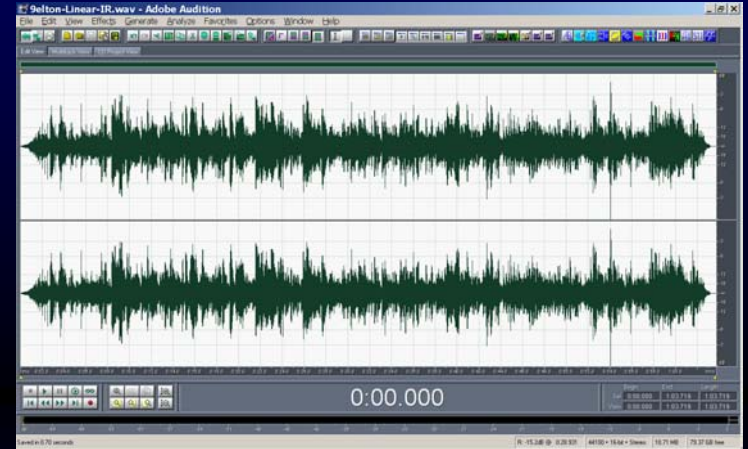


Audible evaluation of the performance

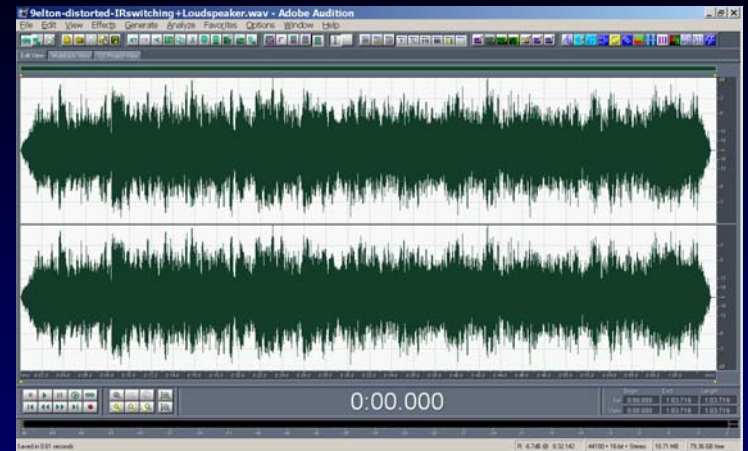
Original signal



Linear convolution



Live recording



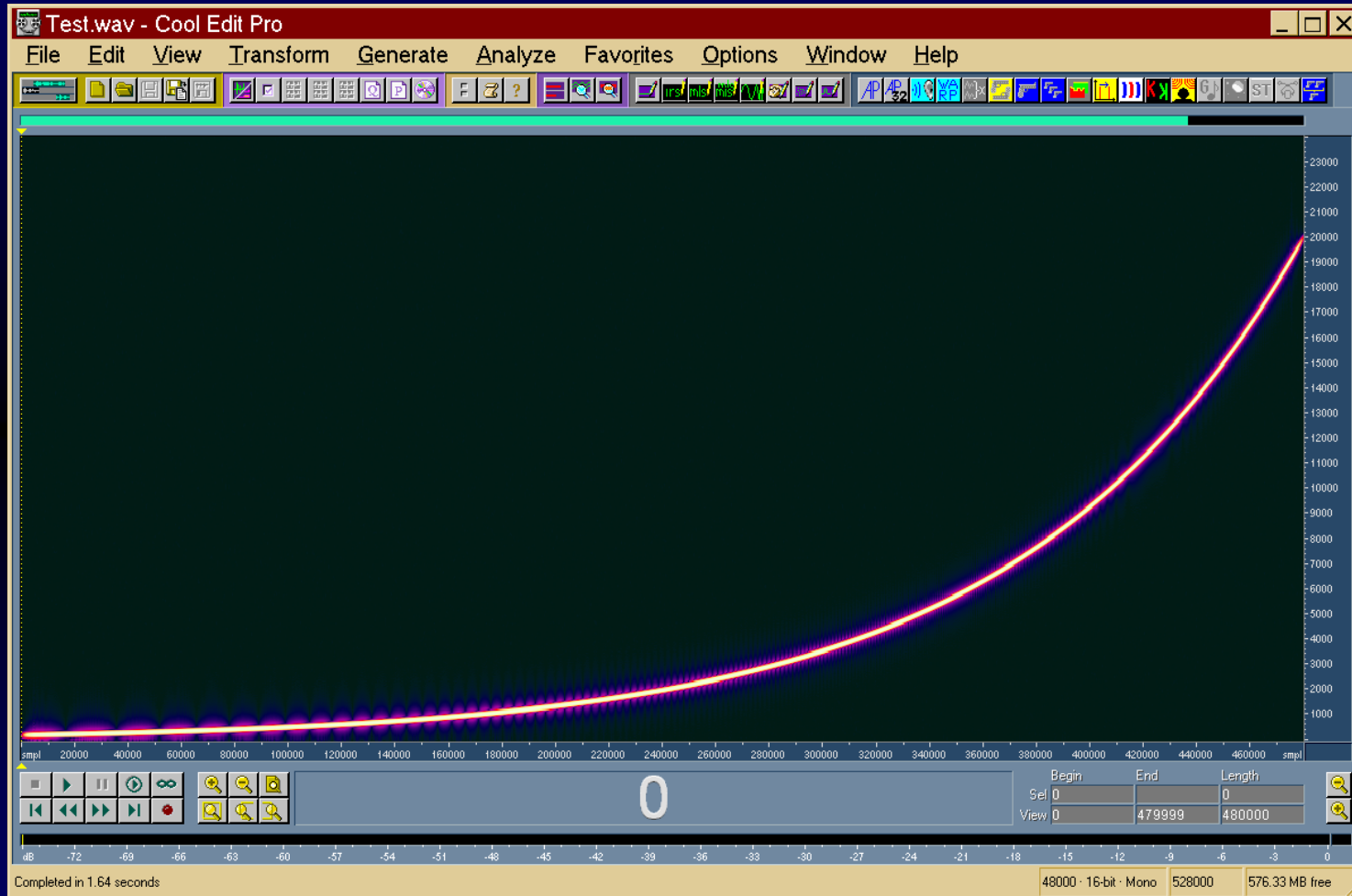
Non-linear (IR switching)



Method 2 – Diagonal Volterra Kernels

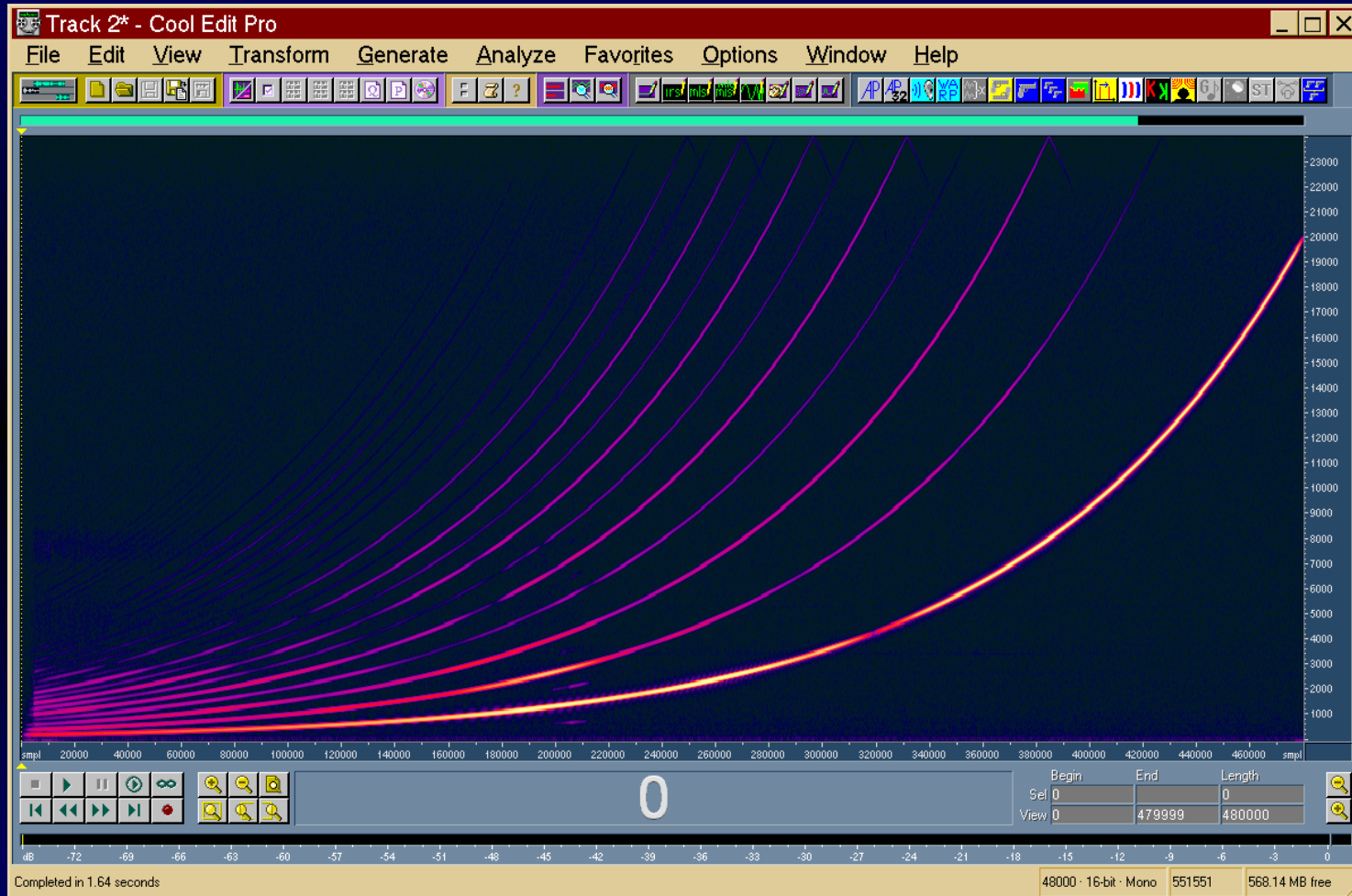
- We start from a measurement of the system based on exponential sine sweep (Farina, 108th AES, Paris 2000)
- Diagonal Volterra kernels are obtained by post-processing the measurement results
- These kernels are employed as FIR filters in a multiple-order convolution process (original signal, its square,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

Exponential sweep measurement



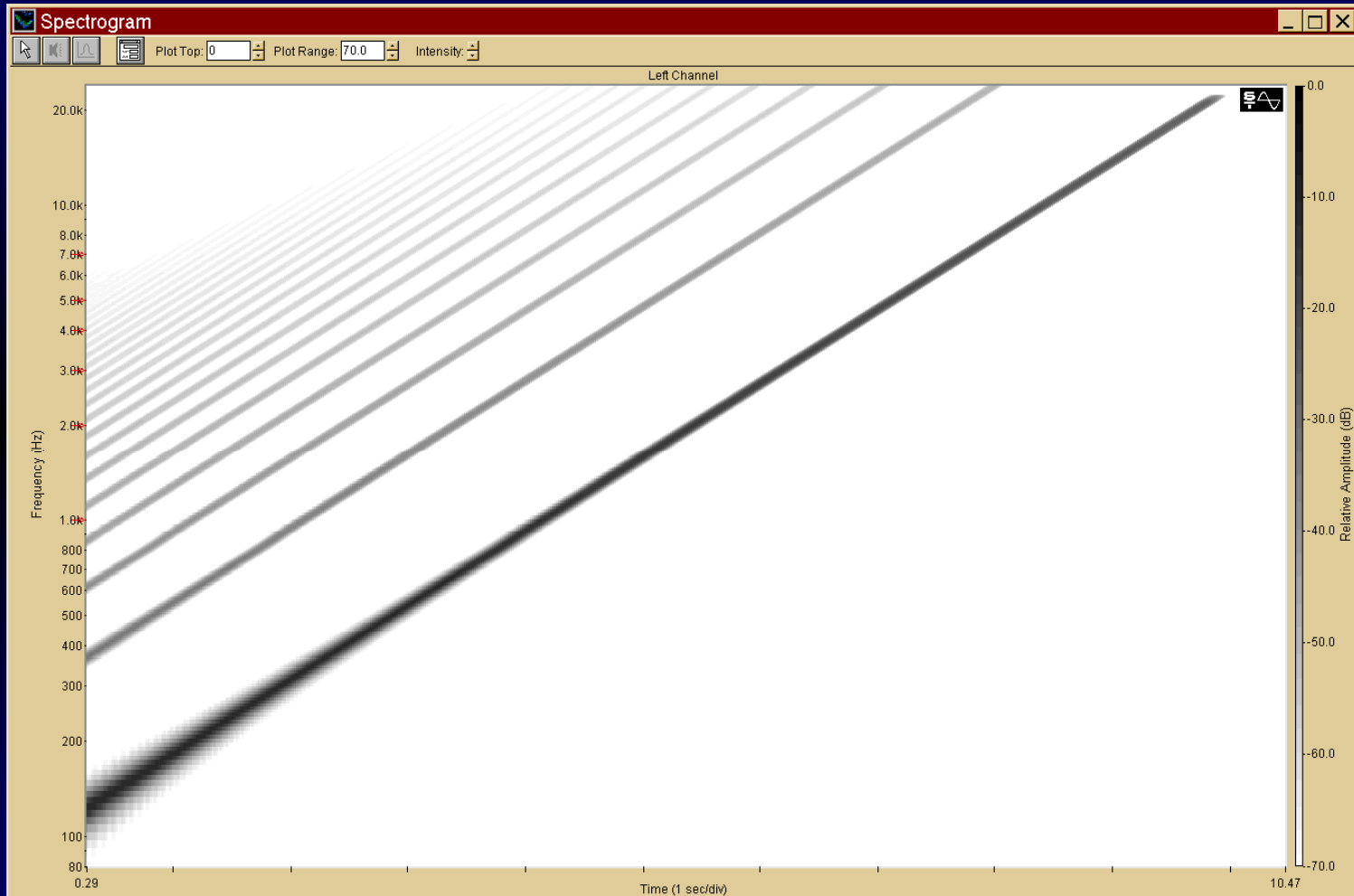
- The test signal is a sine sweep with constant amplitude and exponentially-increasing frequency

Raw response of the system



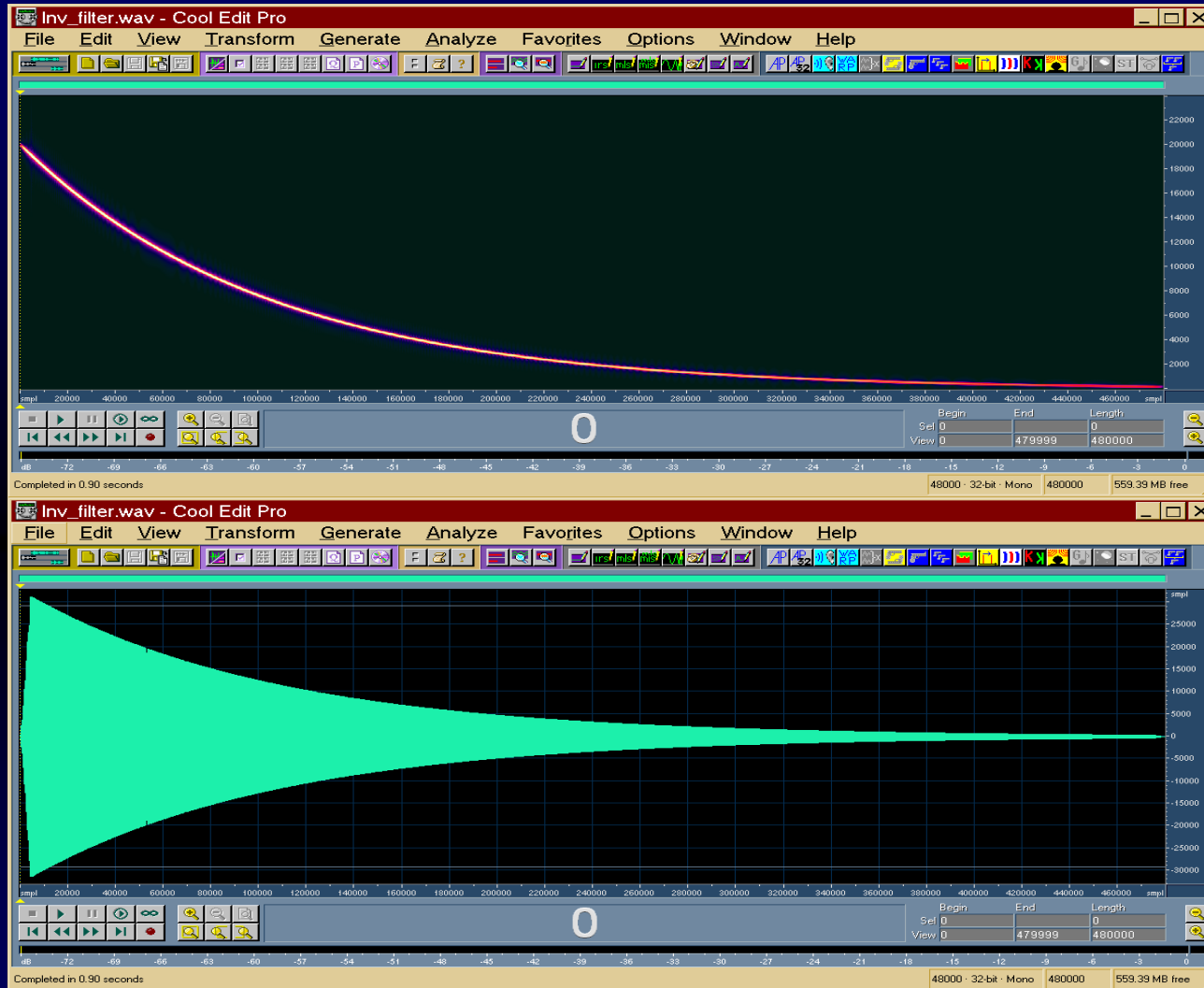
Many harmonic orders do appear as colour stripes

Raw response of the system



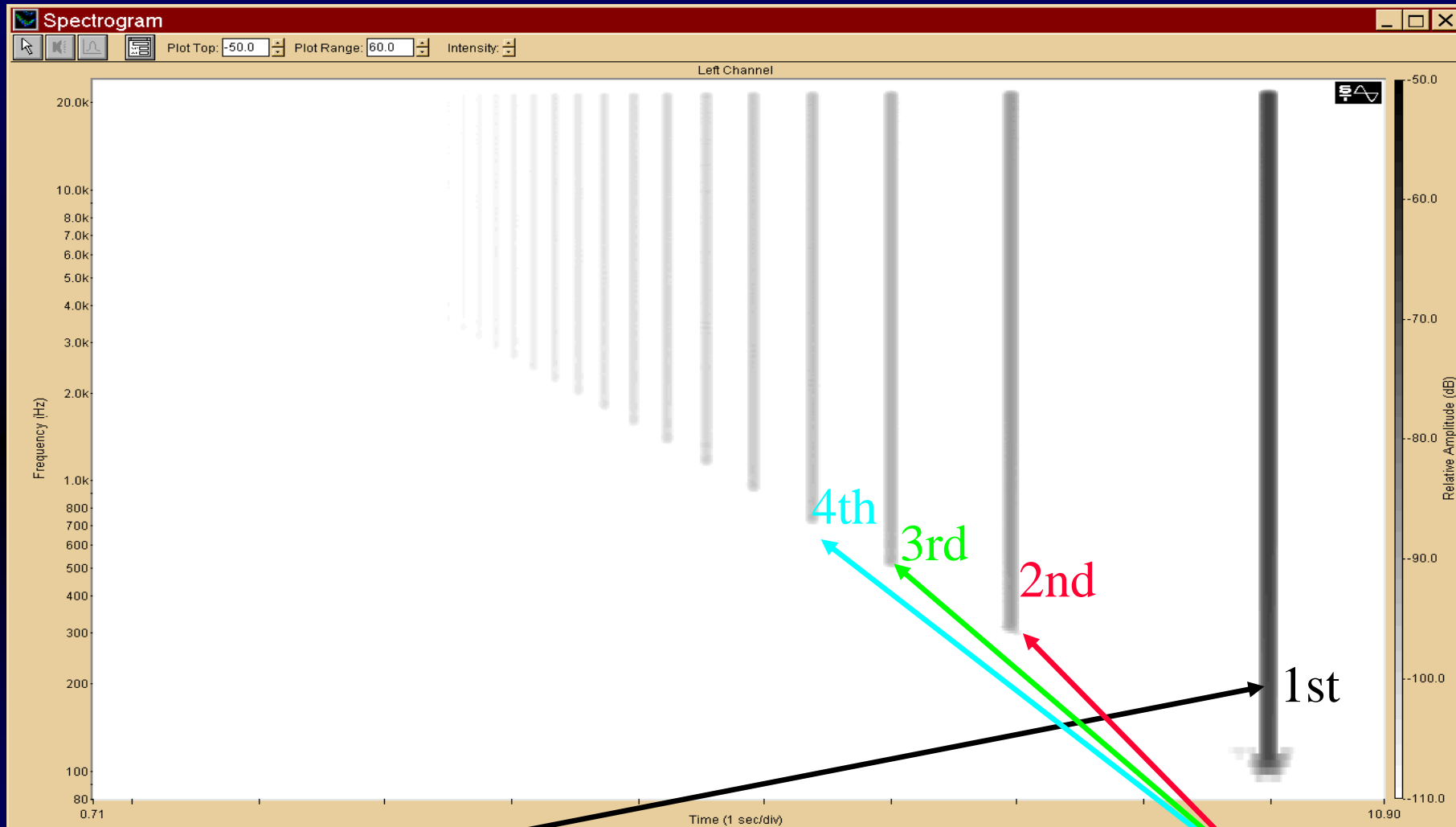
Many harmonic orders do appear as colour stripes

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

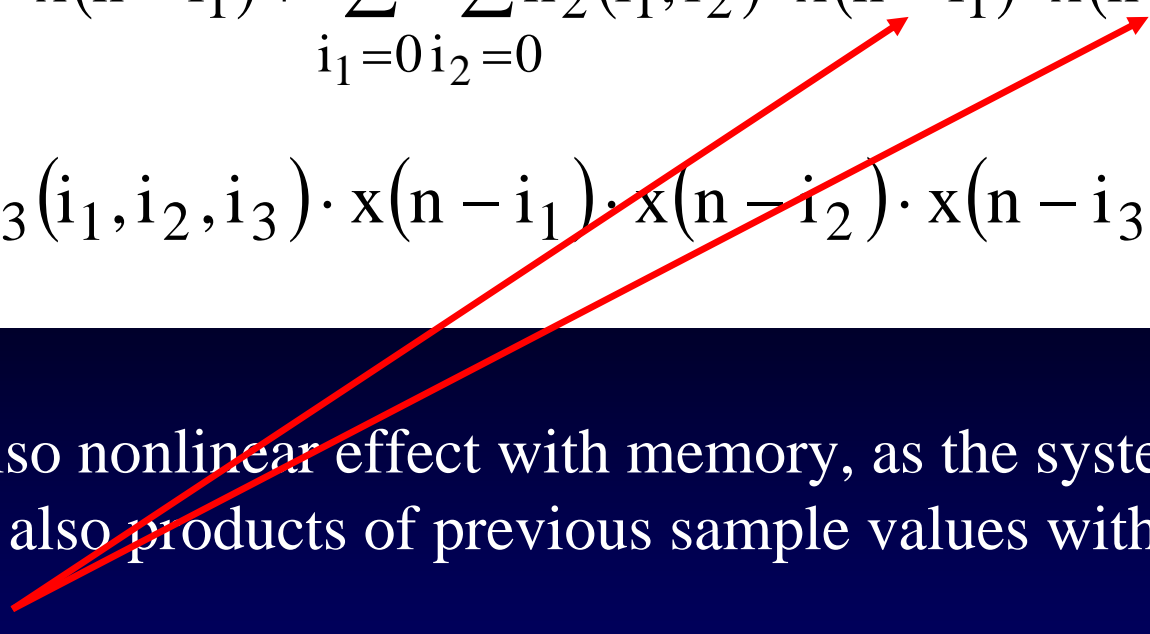
- The basic approach is to convolve separately, and then add the result, the linear IR, the second order IR, the third order IR, and so on.
- Each order IR is convolved with the input signal raised at the corresponding power:

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

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

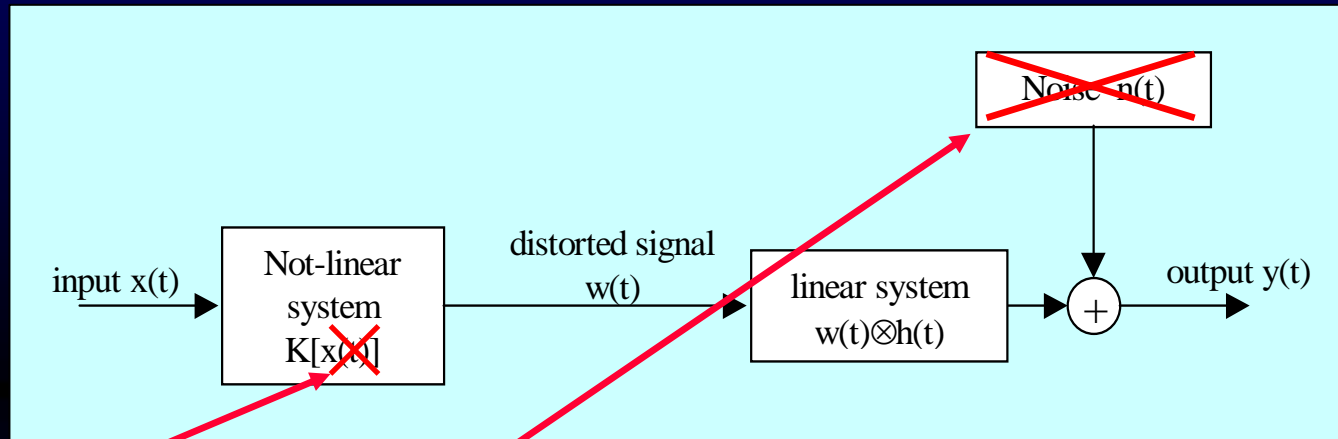
Volterra kernels and simplification

- The general Volterra series expansion is defined as:

$$y(n) = \sum_{i_1=0}^{M-1} h_1(i_1) \cdot x(n-i_1) + \sum_{i_1=0}^{M-1} \sum_{i_2=0}^{M-1} h_2(i_1, i_2) \cdot x(n-i_1) \cdot x(n-i_2) +$$
$$+ \sum_{i_1=0}^{M-1} \sum_{i_2=0}^{M-1} \sum_{i_3=0}^{M-1} h_3(i_1, i_2, i_3) \cdot x(n-i_1) \cdot x(n-i_2) \cdot x(n-i_3) + \dots$$


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:

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

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

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

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:

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

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

$$Y(\omega) = \left[\bar{H}_1 + \frac{3}{4} \cdot \bar{H}_3 + \frac{5}{8} \cdot \bar{H}_5 \right] \cdot X[\omega] + \left[-\frac{1}{2} \cdot \bar{H}_2 - \frac{1}{2} \cdot \bar{H}_4 \right] \cdot j \cdot X[\omega/2] + \left[-\frac{1}{4} \cdot \bar{H}_3 - \frac{5}{16} \cdot \bar{H}_5 \right] \cdot X[\omega/3] + \frac{1}{8} \cdot \bar{H}_4 \cdot j \cdot X[\omega/4] + \frac{1}{16} \cdot \bar{H}_5 \cdot X[\omega/5] + \dots$$

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:

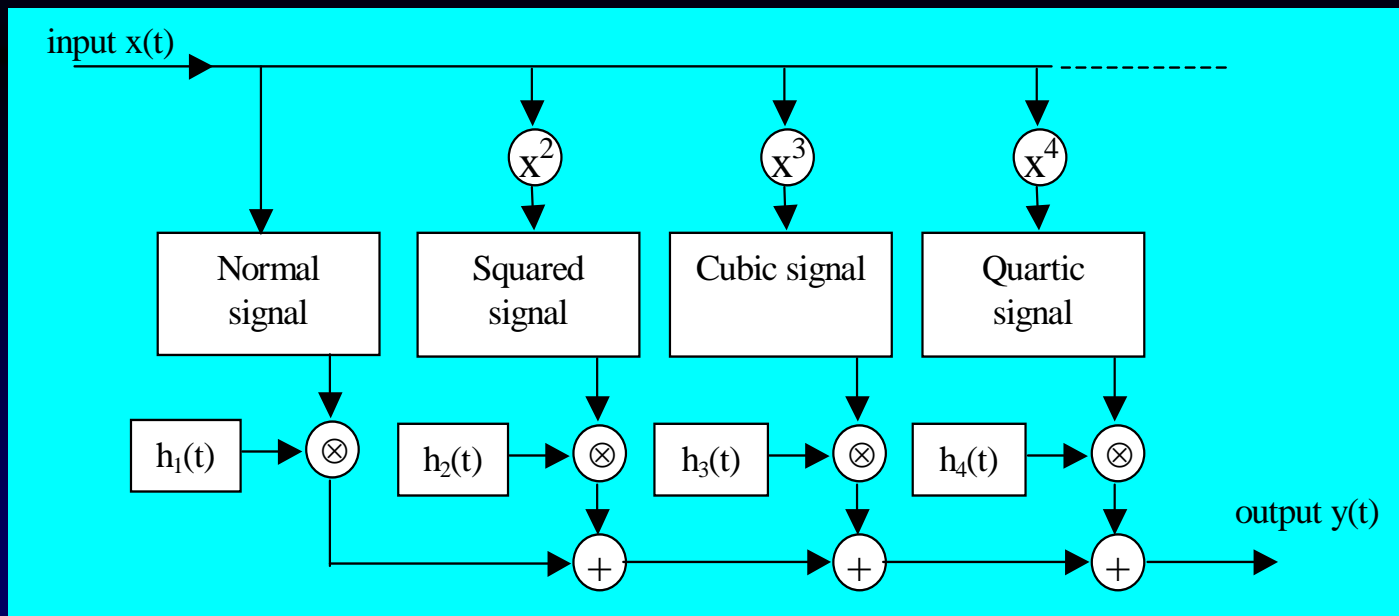
$$\begin{cases} \overline{H}'_1 = \overline{H}_1 + \frac{3}{4} \cdot \overline{H}_3 + \frac{5}{8} \cdot \overline{H}_5 \\ \overline{H}'_2 = -j \cdot \frac{1}{2} \cdot [\overline{H}_2 + \overline{H}_4] \\ \overline{H}'_3 = -\frac{1}{4} \cdot \overline{H}_3 - \frac{5}{16} \cdot \overline{H}_5 \\ \overline{H}'_4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

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

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

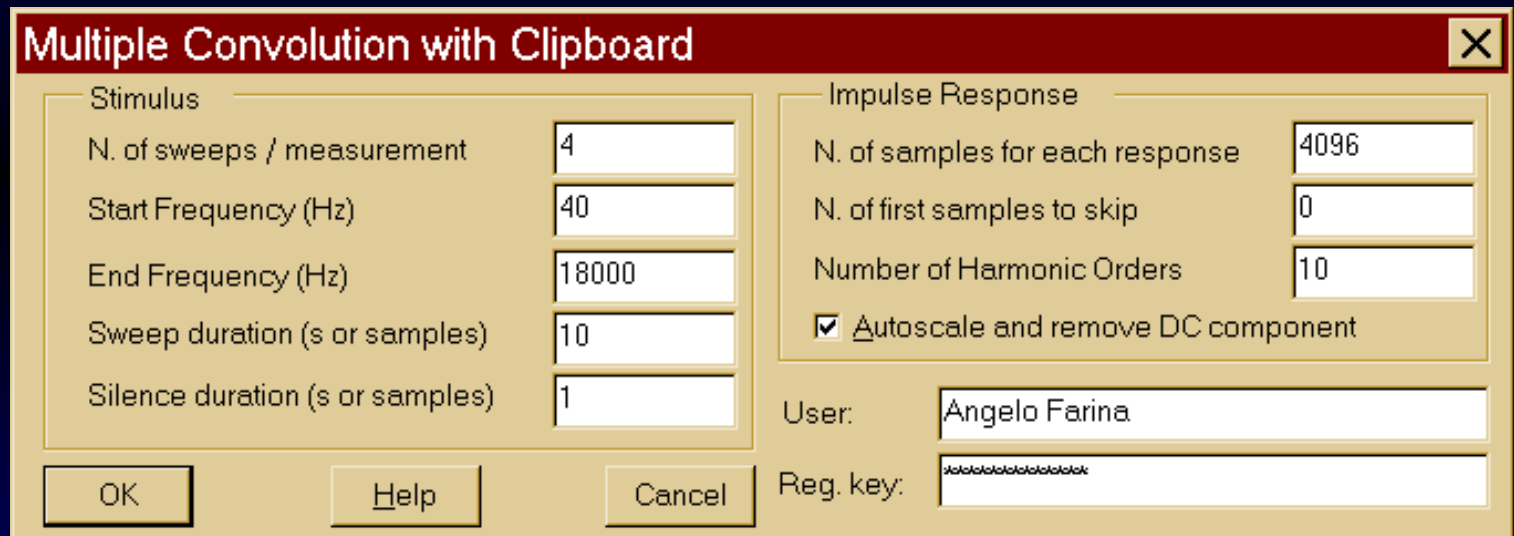
Non-linear convolution

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

The screenshot shows a software window titled "Risposte soggettive" with a red title bar. At the top, there are controls for "Brano n." (1, 2, 3, 4) and "A" (highlighted in green) and "B". Below this is a file path: "D:\Convol_altop_lamiera\05RebeccaPidgeon-porta.WAV". The main area contains nine questions, each with a 5-dot horizontal scale and two labels. Red arrows indicate the response for each question. At the bottom, there are buttons for "Precedente", "Successivo", and "Fine".

Domanda	Scale (1-5)	Response
Domanda 1: A & B are identical vs A & B are quite different	1 2 3 4 5	2
Domanda 2: A is more enveloping vs B is more enveloping	1 2 3 4 5	4
Domanda 3: A has better timber vs B has better timber	1 2 3 4 5	3
Domanda 4: A is more dry vs B is more dry	1 2 3 4 5	2
Domanda 5: A is more distorted vs B is more distorted	1 2 3 4 5	2
Domanda 6: A has more treble vs B has more treble	1 2 3 4 5	3
Domanda 7: A has more medium vs B has more medium	1 2 3 4 5	3
Domanda 8: A has more bass vs B has more bass	1 2 3 4 5	3
Domanda 9: A is more pleasant vs B is more pleasant	1 2 3 4 5	3

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](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](http://pcangelo.eng.unipr.it/public/AES110)

Future developments

- In the “IR switching” technique it is possible to obtain some “memory effect” employing a fast block convolution algorithm, instead of processing “sample by sample”.
- The choice of the length 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 memory effect can be obtained adding a variable gain control driven 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