Simultaneous Measurement of Impulse Response and Distortion With a Swept-Sine Technique

Article · November 2000
Source: CiteSeer

1 author:

Angelo Farina
Università di Parma
175 PUBLICATIONS **2,680** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Reverberance View project

Project   Virtual Microphone Arrays View project

Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique.

5093 (D - 4)

Angelo Farina,
Dipartimento di Ingegneria Industrial, Universita' di Parma, Italy.

Presented at
the 108th Convention
2000 February 19-22
Paris, France

AN AUDIO ENGINEERING SOCIETY PREPRINT

Simultaneous measurement of impulse response and distortion with a swept-sine technique

Angelo Farina

Dipartimento di Ingegneria Industriale, Università di Parma,
Via delle Scienze - 43100 PARMA - tel. +39 0521 905854 - fax +39 0521 905705
E-MAIL: farina@pcfarina.eng.unipr.it - HTTP://pcfarina.eng.unipr.it

Abstract

A novel measurement technique of the transfer function of weakly not-linear, approximately time-invariant systems is presented. The method is implemented with low-cost instrumentation; it is based on an exponentially-swept sine signal. It is applicable to loudspeakers and other audio components, but also to room acoustics measurements.

The paper presents theoretical description of the method and experimental verification in comparison with MLS.

1. Introduction

The actual state-of-the art of audio measurements is represented by two different kinds of measurements: characterisation of the linear transfer function of a system, through measurement of its impulse response, and analysis of the nonlinearities through measurement of the harmonic distortion at various orders. These two measurements are actually well separated: for the impulse response measurement the most employed technique are MLS (Maximum Length Sequence) and TDS (Time-Delay Spectrometry). Both these methods are based on the assumption of perfect linearity and time-invariance of the system, and give problems when these assumptions are not met. In particular MLS is quite delicate, it does not tolerate very well nonlinearity or time-variance, and requires that the excitation signal is tightly synchronised with the digital sampler employed for recording the system's response. The novel technique employed here was developed while attempting to overcome to the MLS limitations through TDS measurements. It was discovered that employing a sine signal with exponentially varied frequency, it is possible to deconvolve simultaneously the linear impulse response of the system, and separate impulse responses for each harmonic distortion order. In practice, after the deconvolution of the sampled response, a sequence of impulse responses appears, clearly separated along the time axis. By FFT analysing each of them, the linear frequency response and the corresponding spectra of the distortion orders can be displayed. This means that the system is characterised completely with a single, fast and simple measurement, which proved to compare very well with traditional techniques for measuring the linear impulse response and the harmonic distortion. Furthermore, the system revealed to be very robust to minor time-variance of the system under test, and to mismatch between the sampling clock of the signal generation and recording. The paper presents the theoretical background of the measurement method, and attempts to explain physically what happens and how the results are obtained. Then some experimental results are reported, which demonstrate the capabilities of the new technique in comparison with established measurement methods.

2. Theory

We start taking into account a single-input, single-output system (a black box), in which an input signal x(t) is introduced, causing an output signal y(t) to come out. Common assumptions for the system are to be linear and time-invariant, but we will able to release these constraints in the following. Inside the system, some noise could be generated, and added to the "deterministic" part of the output signal. Usually this noise is assumed to be white gaussian noise, completely uncorrelated with the input signal. Fig. 1 shows the flow diagram of such a system.

In practice, the output signal can be written as the sum of the generated noise and a deterministic function of the input signal:

$$y(t) = n(t) + F\big[x(t)\big]$$

If the system is linear and time-invariant, the function F assumes the form of the convolution between the input signal and the system's impulse response h(t):

$$y(t) = n(t) + x(t) \otimes h(t)$$

If now we release the constraint for the system to be linear, we have a much complex case, which cannot be studied easily. But often the nonlinearities of the system happen to be at its very beginning, and are substantially memoryless. After this initial distortion, the signal passes through a linear subsequent system, characterized by evident temporal effects (memory). This scenario is typical, for example, of a reverberant space excited through a loudspeaker: the distortion occurs in the electro-mechanical transducer, but as the sound is radiated into air, it passes through a subsequent linear propagation process, including multiple reflections, echoes and reverberation.

Fig. 2 shows such a composite system. In practice, we can assume that the input signal first passes through a memoryless not linear device, characterized by a N-th order Volterra kernel $k_N(t)$, and the result of such a distortion process (called w(t)) is subsequently reverberated through the linear filter h'(t).

A memory-less harmonic distortion process can be represented by the following equation:

$$w(t) = x(t) \otimes k_1(t) + x^2(t) \otimes k_2(t) + x^3(t) \otimes k_3(t) + ... + x^N(t) \otimes k_N(t)$$

As the convolution of w(t) with the following linear process h'(t) possesses the distributive property, we can represent the measured output signal as:

$$y(t) = n(t) + x(t) \otimes k_1(t) \otimes h'(t) + x^2(t) \otimes k_2(t) \otimes h'(t) + ... + x^N(t) \otimes k_N(t) \otimes h'(t)$$

In practice, it is difficult to separate the linear reverberation from the not-linear distortion, and we can assume that the deterministic part of the transfer function is described by a set of impulse responses, each of them being convolved with a different power of the input signal:

$$y(t) = n(t) + x(t) \otimes h_1(t) + x^2(t) \otimes h_2(t) + x^3(t) \otimes h_3(t) + ... + x^N(t) \otimes h_N(t)$$

Other considerations are needed for describing not-time-invariant systems. In such systems, the impulse responses $h_N(t)$ do not remain always the same, but change slowly in time. The variation is usually slow enough for avoiding audible effects such as tremolo or other form of modulation, and in most cases there are not significant differences in the objective acoustical parameters or in the subjective effects connected with different "instantaneous" values of the changing transfer function. Simply, this continuous variation poses serious problems during the measurements, as it impedes to use the averaging

technique for removing the unwanted extraneous noise n(t): increasing the number of averages, in fact, not only the contaminating noise n(t), but also the variable part of the transfer function is rejected.

Now, let we go back to the most common assumptions of linear, time invariant system characterised by a single transfer function h(t). A common practice for measuring the unknown transfer function is to apply a known signal to the input x(t), and to measure the system's response y(t). For this task, the most commonly used excitation signals are wide-band, deterministic and periodic: these include

- MLS (Maximum-Length-Sequence) pseudo-random white noise
- Sine sweeps and chirps

The Signal-To-Noise ratio (S/N) is improved by taking multiple synchronous averages of the output signal, usually directly in time domain, prior to attempt the deconvolution of the system's impulse response. Let we call $\hat{y}(t)$ the averaged output signal. As both the input and output signal are periodic, a circular convolution process relates the input and the output. If we suppose that the noise n(t) has been reasonably averaged out thanks to the large number of averages, we can employ FFTs and IFFTs transforms for deconvolving h(t):

$$h(t) = IFFT\left[\frac{FFT(\hat{y}(t))}{FFT(x(t))}\right]$$

Another common approach is to perform the averages directly in the frequency domain (through the so-called auto-spectrum and cross-spectrum), computing the frequency response function known as H_2, and then taking the IFFT of the result:

$$h(t) = IFFT(H_2) = IFFT\left[\frac{\overline{G_{AB}}}{\overline{G_{AA}}}\right]$$

In both the above approaches, due to the continuous repetition of the test signal and the fact that a circular deconvolution is performed, there is the risk of the time aliasing error. This happens if the period of the repeated input signal is shorter than the duration of the system's impulse response h(t). This means that, with MLS, the order of the shift register employed for the generation of the sequence must be high enough, depending on the reverberation time of the system: modern MLS measurement equipment can produce very high-order MLS signals [1], but previous systems occurred easily in the time-aliasing problem, which causes the late part of the reverberant tail to fold-back at the beginning of the time window containing the deconvolved h(t).

With sine sweeps or chirps, it is common to add a segment of silence after each signal, for avoiding the time aliasing problem: if the data analysis window is still coinstrained to be of the same length as the sweep, the late part of the tail can be lost, but it will not come back at the beginning of the deconvolved h(t) (appearing as noise before the arrival of the direct wave). This is a first advantage of the traditional sine-sweep method over MLS.

What is not widely known is that also not-linear behavior of the system (i.e., harmonic distortion) can cause time aliasing artifacts, also if the length of the input signal is properly chosen. In practice, at various positions of the deconvolved impulse response strange peaks do appear: looking at these "distortion products" in details, reveals that they resemble scaled-down copies of the principal impulse response. This is clearly evident when making anechoic measurements of a loudspeaker, and applying to it too much voltage: the unwanted, spurious peaks appear after the anechoic linear response, both employing MLS and sine sweep.

A mathematical explanation of the appearance of the spurious peaks in the MLS case was given in [2]. Fig. 3 shows a typical MLS measurement affected by untolerable distortion, which produces evident spurious peaks.

Making use of sine sweeps in which the instantaneous frequency is made to vary linearly with time, the appearance of spurious peaks is not very evident: the distortion products simply cause a sort of noise to appear everywhere in the deconvolved h(t). This "noise" is actually correlated with the signal input, so it does not disappear by averaging. It usually sounds as a decreasing-frequency low-level multitone.

Instead, if the sine sweep was generated with instantaneous frequency varying exponentially with time (the so-called "logarithmic sweep"), the spurious distortion peaks clearly appear again, with their typical impulsive sound.

This was the starting point of the work presented here: a method was searched for "pushing out" the unwanted distortion products from the results of the deconvolution process. The most straightforward approach was to substitute the circular deconvolution with a linear deconvolution, directly implemented in the time domain. This is very easy, if a proper inverse filter f(t) can be generated, capable of packing the input signal x(t) into a delayed Dirac's delta function δ(t):

$$x(t) \otimes f(t) \Rightarrow \delta(t)$$

The deconvolution of the system's impulse response can then be obtained simply convolving the measured output signal y(t) with the inverse filter f(t):

$$h(t) = y(t) \otimes f(t)$$

Both fast convolution and inverse filter generation are nowadays easy and cheap tasks, due to recently developed software [1,3]. With this approach, any distortion products caused by harmonics produce output signals at frequencies higher than the instantaneous input frequency: figg. 4 and 5 show a not-linear system response with a linear and logarithmic sweep excitation respectively, in the form of a sonograph.

The convolution of the inverse filters causes these sonographs to deform (or to "stretch") counter-clockwise, so that the linear response becomes a straight vertical line (followed by some sort of tail, if the system is reverberant). The distortion products are pushed to the left of the linear response: in the case of linearly swept sine they spread along the time axis, whilst in the case of exponentially-swept sine they pack in "distortion peaks" at very precise anticipatory times before the linear response. Figgs. 6 and 7 show the inverse filter and the results of the deconvolution process, again in the form of sonographs, for the linear sweep case;. figgs. 8 and 9 show the inverse filter and the results of the deconvolution process for the log sweep case.

This different behavior can be explained by looking at the structure of the inverse filters (figs 6 and 8). First of all, in both cases the inverse filter is basically the input signal itself, reversed along the time axis (so that the instantaneous frequency diminishes with time). In the case of exponentially-swept sine, an amplitude modulation is added, for compensating the different energy generated at low and high frequencies.

It can be observed that the inverse filter has the effect to delay the signal which is convolved with it of an amount of time which varies with frequency: this causes the deformation of the sonographs, as it was clearly demonstrated by M. Poletti [4] for linearly-swept sine signal. This delay is linearly proportional to frequency for linear sweeps, and instead is proportional to the logarithm of frequency for the logarithmic sweep. This means that the delay is increasing, for example, of 1s each octave.

In practice, if the frequency axis of the sonograph is made linear when displaying measurements made with a linear sweep, and is made logarithmic when displaying measurements made with a log sweep, the excitation signal, the inverse filters and the system response always appears as straight lines on the sonographs (this was done in figgs. 4-9). Furthermore, also the harmonic distortions appear as straight lines: but these are parallel to the linear response in the case of the log sweep, whilst they are of increasing slope in the case of linear sweep (look at figures 4 and 5). Both inverse filters stretch the sonographs with a constant slope, corresponding to the inverse slope of the linear response: this packs the linear response onto a vertical line (at a precise time delay, which equals the inverse filter length). Obviously, also the harmonic distortion orders packs at very precise times in the case of the log sweep, as all the lines had the same slope (for examples 1 octave/s); instead, the harmonic distortion present in a response produced by a linear sweep tends to stretch over the time axis, producing a sort of sweeping-down multi-tone signal which precedes the linear impulse response (fig. 6).

It is clear at this point that the use of the linear deconvolution, instead of the circular one, pushes all the distortion artefacts well in advance than the linear response, and thus enables the measurement of the system's linear impulse response also if the loudspeaker is working in a not-linear region. This holds both for linear and log sweep, meaning that, if the goal of the measurement was simply to estimate the linear response, the log sweep has the only advantage over the linear sweep of producing a better S/N ratio at low frequencies.

In conclusion, the complete removal of distortion-induced artefacts is already a very important result compared with the traditional circular deconvolution approach.

But in the case of the log sweep another very important result can be obtained: if the sweep is slow enough, so that each harmonic distortion packs into a separate impulse response, without overlap with the preceding one, it is possible to window out each of them: and each of these impulse responses corresponds exactly to the rows of the Volterra kernel, convolved with the subsequent linear reverberation (if any), and thus to the terms previously named $h_1(t)$, $h_2(t)$ and so on.

For designing properly the excitation signal, and for retrieving each harmonic order response, what is needed at this point is a theoretical derivation of the starting time of each order's distortion.

A varying-frequency sine sweep can be mathematically described as:

$$x(t) = \sin\big(f(t)\big)$$

It must be noted that, following the general signal processing theory, the instantaneous frequency is given by the time derivative of the argument of the sine function. Thus, of course, if $f(t) = \omega t$, where ω is constant, the instantaneous frequency is also constant and equal to ω (in rad/s). But if, for example, we assume a linearly varying frequency, starting from ω_1 and ending to ω_2 in the total time T, we obtain:

$$\frac{d\big(f(t)\big)}{dt} = \omega_1 + \frac{\omega_2 - \omega_1}{T} \cdot t$$

which is satisfied if we pose:

$$f(t) = \omega_1 \cdot t + \frac{\omega_2 - \omega_1}{T} \cdot \frac{t^2}{2}$$

Following the same approach, we can find the rule for generating a log sweep, having a starting frequency ω_1, an ending frequency ω_2, and a total duration of T seconds; we start writing a generic exponential sweep in the form:

$$x(t) = \sin\left[K \cdot \left(e^{t/L} - 1\right)\right]$$

For obtaining the values of the two unknown K and L, we pose:

$$\left.\frac{d\left[K \cdot \left(e^{t/L} - 1\right)\right]}{dt}\right|_{t=0} = \omega_1 \qquad \left.\frac{d\left[K \cdot \left(e^{t/L} - 1\right)\right]}{dt}\right|_{t=T} = \omega_2$$

Which, after some passages, yields to:

$$K = \frac{T \cdot \omega_1}{\ln\left(\dfrac{\omega_2}{\omega_1}\right)} \qquad L = \frac{T}{\ln\left(\dfrac{\omega_2}{\omega_1}\right)}$$

So that the required equation for the log sweep is:

$$x(t) = \sin\left[\frac{\omega_1 \cdot T}{\ln\left(\dfrac{\omega_2}{\omega_1}\right)} \cdot \left(e^{\frac{t}{T} \cdot \ln\left(\frac{\omega_2}{\omega_1}\right)} - 1\right)\right]$$

Now we want to find for which time delay Δt the above function has an instantaneous frequency equal to N times the actual one: this represent the delay between the Nth order distortion and the linear response. So we impose that:

$$N \cdot \frac{d}{dt}\left[\frac{\omega_1 \cdot T}{\ln\left(\dfrac{\omega_2}{\omega_1}\right)} \cdot \left(e^{\frac{t}{T} \cdot \ln\left(\frac{\omega_2}{\omega_1}\right)} - 1\right)\right] = \frac{d}{dt}\left[\frac{\omega_1 \cdot T}{\ln\left(\dfrac{\omega_2}{\omega_1}\right)} \cdot \left(e^{\frac{t+\Delta t}{T} \cdot \ln\left(\frac{\omega_2}{\omega_1}\right)} - 1\right)\right]$$

And we obtain:

$$\Delta t = T \cdot \frac{\ln(N)}{\ln\left(\dfrac{\omega_2}{\omega_1}\right)}$$

It must be noted that the value of Δt is constant, and this ensures that each harmonic order will pack always at a very precise time lag before the linear response. Furthermore, Δt increases with the logarithm of N, and this means that the delay between each harmonic response and the previous one is not constant, but the higher orders are less spaced. The above equation correspond perfectly with the experimental results shown in fig. 5.

As a last theoretical consideration, we must notice that any kind of problems related with slightly time-variant systems are solved if we avoid to use the technique of multiple averages. The preferred technique is to employ a single, very long, logarithmic sine sweep: this produce a distortion-free linear response, well separated harmonic distortion responses up to very high orders, and the estimated response is not affected by the time variation, as a single measure was taken. The signal-to-noise ratio is indeed very good, as a lot of energy was diluted over a

long time, and then packed back to a short response, obtaining usually a S/N improvement of 60 dB or more in comparison with the generation of a single impulse having the same maximum amplitude.

3. Hardware Implementation

The novel measurement system has been implemented on a low-cost, PC-based hardware, avoiding the use of dedicated DSP boards or expensive audio analysers. Standard sound boards for high-level applications are on the market: these units are cheap (typically less than US $ 1000), have many input and output channels (typically 8 ins and 8 outs, plus digital interfaces such as SPDIF, TDIF or ADAT), and are equipped with top level A/D and D/A converters (with at least 20 bit effective resolution). The software drivers of these sound boards allow for the multichannel operation with 24-bit data depth and synchronous playback and record.

Obviously a proportionate computer is needed; for this work three hardware platforms were tested, as in the following table:

Configuration #1	Configuration #2	Configuration #3
PC Pentium-II 400 MHz	PC Pentium-II 350 MHz	PC Pentium-II 350 MHz
128 Mbytes RAM	128 Mbytes RAM	256 Mbytes RAM
HD SCSI (U2W) 9 Gbytes	HD EIDE (U-33) 6.4 Gbytes	HD SCSI (UW) 9 Gbytes
Echo Layla sound board (8in, 10 out, 20 bit converters)	GadgetLabs Wave8/24 sound board (8in, 8 out, 24 bit converters)	MOTU sound board (8in, 8 out, 20 bit converters)

It can be observed that these machines are nowadays substantially entry-level. Furthermore, it can be noted how it was considered more important to allocate resources for large memory and fast hard disk than for the processor itself.

In terms of hardware performance and practical results, all the three tested configurations worked with similar performances: no significant difference was found between the 20-bit converters and the 24-bit ones, although it was verified that reducing the data depth to 16 bit introduces a significant amount of discretisation noise and reduces the usable dynamic range. This means that actually there is no point in moving from 20 to 24 bits, as the analog electronic equipment which is part of the measurement chain introduces noise, which makes useless the 4 LS bits of 24 bits converters. Instead, the use of 20 bit converters (with 24-bits drivers) significantly enhances the performances, and set these high-level sound boards in a different class than 16-bit, multimedia sound boards.

It must be recalled that, in a previous comparative investigation among various measurement techniques [5], it was found that with the MLS technique there was no improvement in increasing the number of bits above 16, and in most cases the best results were obtained with the old MLSSA board, which is equipped with a single A/D converter with only 12 bits resolution.

It can be concluded that the new exponential sweep technique exploits the performances of modern sound boards, allowing for a much wider dynamic range than the one possible with MLS.

4. Software Implementation

The basis of the software implementation is the CoolEdit program by David Johnston [6]. It is a sound editor, already equipped with a lot of useful tools for filtering and manipulating

the digitised sound. It comes in two versions: Cool96 (shareware), which manages only a single stereo device, and CoolEditPro, which is a multi-track recorder, particularly useful when making measurement with a multichannel sound board and employing more than 2 channels.

Although CoolEditPro was employed for the experiments described here, all the software developed for implementing the new measurement technique also runs without any modification under Cool96.

It must be noted that CoolEditPro v. 1.2 already includes some tools which could make it possible to implement directly the new measurement without the addition of external software. In fact, the new Sine Sweep generator also includes the log sweep, and the program already incorporates a fast convolver. The generation of the inverse filter is simply matter of time-reversing the excitation signal, and then applying to it an amplitude envelope to reduce the level by 6 dB/octave, starting from 0 dB and ending to $-6 \cdot \log_2\left(\dfrac{\omega_2}{\omega_1}\right)$. Following these guidelines, probably also other programs could be used for the measurements, as long as they are capable of the generation of log sweeps and convolution.

In our case, anyway, a set of dedicated plug-ins was developed for CoolEdit: these make it easier to generate multiple repetitions of the log sweep, to produce automatically the inverse filter for the deconvolution, and to operate, if required, a synchronous average of the result for reducing the effect of the background noise in perfectly time-invariant systems. Furthermore, the convolution module does not suffer of the limitations about the length of the filter to be convolved, as it happens for the CoolPro convolver.

Fig. 10 shows the user's interface of the plug-in for the generation of sine sweeps. It can be seen that it is possible to set the start and end frequency, the sweep duration, the duration of silence between subsequent sweeps and the number of repetitions.

When a stereo waveform is generated, there are two possible options. In its basic mode, the plug-in generates first a sequence of sweeps on the left channel, followed by the same sweeps on the right channel, as it is shown in fig. 11. This makes it easy to measure automatically the transfer function matrix of a stereo system, for example the 2x2 matrix of a StereoDipole configuration [7].

If, instead, the flag marked "Generate control pulses on right channel" is set, the sine sweeps are generated only on the left channel, and on the right one, just after the end of each sweep, a short pulse is generated. This allows for the control of a motorised rotating board, which is commonly employed for the measurement of polar responses of loudspeakers, microphones and diffusing panels. Fig. 12 shows the signals obtained in this case, having set the number of sweeps to 4.

The generation of the inverse filter is automatically performed during the generation of the test signals. In fact, the Generate Sine Sweep plug-in loads into the Windows clipboard the proper inverse filter, obtained by the time reversal of a single sweep, properly amplitude-shaped in the case of the logarithmic sweep.

After the generation of the test signal is finished, CoolEditPro is placed in its multi-track mode, selecting the sequence of sweeps as the first waveform, set for play, and recording the response coming from microphones on the other waveforms. A typical case is the generation over a stereo loudspeaker pair and the recording of the response through a binaural microphone. Fig. 13 shows this case, during the playback/recording.

After the recording is complete, the deconvolution of impulse responses is easily accomplished. The Convolver plug-in is called, and the currently recorded signal is simply

convolved with the Windows clipboard, which contained the inverse filter. Fig. 14 shows the user's interface of the Convolver plug-in.

After the convolution process is terminated, a sequence of impulse responses appears in place of the recorded signals: the separation between each IR and the subsequent is equal to the length of the sine sweep (10s in the case shown).

If the system was perfectly time invariant, and we are interested only in the linear response, we can average together the IRs produced by subsequent repetitions of the same signal (4 repetitions in the example shown here), for improving the S/N ratio. Furthermore, all the unneeded data present before and after the significant responses can be stripped away, and only a significant number of data points can be extracted. These tasks are accomplished thanks to a dedicated plug-in, which performs such a synchronous averaging and data extraction process; its user's interface is shown in fig. 15. After the averaging is done, the results are stored onto the Windows clipboard, from where they can be retrieved: fig. 16 shows the results obtained from the above-described measurement procedure.

5. Comparison with other Impulse Response measurements

The first comparative tests between the novel measurement method and some traditional ones were performed during the AES Workshop on room acoustics measurements, which was organized by the Italian AES section in the Bergamo's Cathedral, in days 27/28 April 1999. A detailed report on the workshop and some of the experimental results can be found in [8].

The workshop was the occasion to test the new release 3.0 of the Aurora software suite, which incorporates the new log-sweep measurement technique [9].

In this case, the hardware system #1 was employed, as this unit is packaged in a flying-case together with a power amplifier (QSC 1202 PLX), the remote control unit of a rotating board (Outline R1), and the preamplifier of a Soundfield MKV microphone unit. Furthermore, in the chassis-mounted computer also a MLSSA A2D160 board was fitted for comparison.

Fig. 17 shows a scheme of the complete measuring system employed for the measurements: all the 8 signal inputs were employed, recording the 4 B-format signals from the Soundfield microphone, its stereo outputs in M-S (180°) configuration and the binaural signals coming from an Ambassador dummy head and torso. The sound was generated by means of an omnidirectional (dodechaedron) loudspeaker (Look Line mod. D1).

Also other researchers employed their measurement systems, so it was possible to compare the results. In particular, the following table reports the systems employed:

Researcher	Measuring system/method	Loudspeaker	Microphone
Angelo Farina	Aurora (synchronous measurement on PC+Layla) – MLS	Dodechaedron	Soundfield + binaural (Ambassador)
Angelo Farina	Aurora (synchronous measurement on PC+Layla) – log sweep	Dodechaedron	Soundfield + binaural (Ambassador)
Angelo Farina	MLSSA board - MLS	Dodechaedron	Soundfield channel W
A. Ricciardi	MLSSA board - MLS	Directional	Stage Accompany omnidirectional
Walter Conti	Techron TEF 20 – MLS & TDS	Directional	B&K Omnidirectional
Nicola Prodi	Aurora (asynchronous playback & record through a Tascam DA38 recorder) – log sweep	Dodechaedron	Soundfield + binaural (Neumann KU-100)

It is beyond the scope of this work to present here all the measurement results, and to compare the performances of different systems as regards the use of various loudspeakers and microphones.

So in the following only the results obtained by the author with his own equipment are presented. In particular, the comparison regards 3 measurements, made with the same loudspeaker, the same microphone (taking simply the omnidirectional channel of the Soundfield microphone) and the three possible measuring techniques: Aurora/MLS, Aurora/sweep and MLSSA/MLS. The first two are implemented with CoolEditPro, dedicated plug-ins and the Layla sound board, whilst the third one is implemented with the original MLSSA software (v. 10W2) and the MLSSA sound board.

As the church was quite reverberant (T60 = 4.5 s), it was necessary to employ a low sampling frequency with the MLSSA board (16 kHz) for reducing the time aliasing problems, whilst with Aurora the standard CD sampling frequency of 44.1 kHz was employed, as in this case there is no limitation regarding the order of the MLS sequence or the length of the sine sweep. An MLS of order 18 was employed, repeated 32 times, and the sine sweep duration was 15 seconds, repeated three times, but without averaging (the second sweep only was analyzed).

Figgs 18, 19 and 20 show the measured wide-band impulse responses with logarithmic amplitude scale. From fig. 18 it is clear how the Aurora/MLS method is severely affected by distortion products, which introduce evident spurious peaks in the late part of the impulse response (although at a level so low that the effect on the estimate of acoustical parameters is substantially negligible). Instead, the new logarithmic-sweep method (Fig. 19, also implemented within the CoolEdit/Aurora environment) appears perfectly free of any artifact, with a remarkable dynamic range of more than 80 dB. Fig. 20 shows the result of the measurement made with the old MLSSA board, which also appears free of evident artifacts, although in this case the dynamic range is less than 60 dB. It must be noted that with MLSSA the useful frequency range is reduced to less than 6 kHz, as the sampling frequency was set very low for avoiding time aliasing problems.

The fact that distortion products were evident in the Aurora/MLS measurement and not in the MLSSA measurement can be explained in two ways: first, the MLSSA measurement is shorter and with lower dynamic range, and the distortion artifacts visible in the Aurora/MLS measurement occur at low level, in the late part of the response. Second, it can easily be that the distortion occurred in analog components of the Layla sound board (both in the output and input sections), so that these causes of nonlinearity are completely removed by employing the MLSSA board. Of course, these distortion problems completely disappear with the new Aurora/sweep technique.

In conclusion, it resulted that the novel technique produces substantially robust estimates of the system's impulse response, without any artifact due to nonlinearities, and with a dynamic range which is approximately 20 dB better than with previously employed instrumentation.

6. Comparison with other distortion measurements

The novel measurement technique is also useful when a quantification of the harmonic distortion of a not-linear system is required. In this case, the traditional measurement technique was to apply a stable, high purity sine signal to the input of the system, and to measure the spectrum at the output through FFT analysis. In the case of very little distortion, and when the A/D converter employed for sampling the system's response has a too little dynamic range, it is common to apply a notch filter before the sampling, for reducing the amount of the linear response at the excitation frequency.

Nowadays, thanks to the incredibly wide dynamic range of modern A/D converters, and when components such as loudspeakers are measured (which often produce a substantial amount of harmonic distortion), there is no need for a notch filter, and the system response is directly sampled.

In this case, a comparison is made between a traditional measurement of the distortion of a headphone set and an application of the new log sine sweep.

In the first case, an high purity sine test signal at 1 kHz is generated with the proper tool of CoolEditPro. The test signal is continuously reproduced over the headphone, with an amplitude of 1V RMS, and its response is measured through the microphone incorporated in one ear of a B&K type 4100 dummy head, over which the headphone was mounted. It is obvious that an input signal of 1 V is quite high for the small headphone, inducing significant distortion.

The signal coming from the microphone is digitized through the Echo Layla sound board, and it is FFT analyzed with a 4096-points FFT and Hanning windowing, averaging 100 times.

As it is obvious, the measured spectrum exhibits a strong peak at 1 kHz, followed by a series of minor peaks at multiple frequencies (2, 3, 4 kHz and so on). The amplitude of these harmonic peaks, related to the amplitude of the main peak at 1 kHz, indicate the amount of harmonic distortion at various orders.

Then a second measurement was made, generating a log sine sweep ranging from 100 Hz to 5 kHz, and deconvolving the complete response of the system. Before the linear response peak, 3 very evident anticipatory peaks appear, which are the impulse responses of the 2nd, 3rd and 4th order distortions respectively.

The linear response and the three harmonic distortion responses were separately saved in 4 WAV files, for subsequent analysis. Then these 4 files were FFT analyzed, employing the same software already employed for the real-time measurement of the harmonic peaks.

The original FFT spectrum obtained with the 1kHz sine excitation was finally superposed to the four spectra obtained from the analysis of the 4 impulse responses measured with log sweep excitation. Fig. 21 shows this comparison.

It is easy to verify that the four peaks obtained with 1kHz excitation fall exactly over the corresponding continuous spectra coming from the analysis of the 4 IRs. The following table reports in more detail the exact values obtained at these 4 frequencies with the two measurement techniques:

Freq. (Hz)	1 kHz	2 kHz	3 kHz	4 kHz
1 kHz test tone	-62.2	-98.91	-88.39	-107.27
Log sweep	-61.96	-99.70	-88.75	-107.03

In practice, the minor deviations shown are probably due to measurement instability, because with both techniques, repeating the measurement, fluctuations of the same magnitude are found. This means that the differences are statistically not significant, and both the traditional single frequency method and the novel log sweep method produce substantially the same results. But the new technique has the advantage of producing directly the response for every excitation frequency, and thus a complete characterization of the not linear response as function of the excitation frequency is obtained with much less effort than with the traditional method.

7. Conclusions

A new measurement system for the complete characterization of complex sound systems has been developed. The new measurement technique works reliably also if the system includes parts which exhibit a not-linear behavior, and in these cases the measurement results include also the quantification of the harmonic distortion at various orders.

The measurements taken in comparison with widely diffused instruments have shown that the new method is at least as reliable and accurate as the others, and gives great benefits in terms of ease of use, signal-to-noise ratio and immunity from time variations of the system under test. It was also verified that there is no need to maintain tight synchronization between the sampling clock of the signal generator and of the digitizing unit employed for capturing the system response: this means that the measurement can be easily conducted also starting with a pre-recorded excitation signal, stored for example on an audio CD, and there is no need of synchronizing the digital clocks.

The measurement technique was implemented in a set of plug-ins for the CoolEdit program, making it possible to conduct the measurements with minimum effort and with a very cheap setup. This approach also enables the automatic measurement with multi-channel configurations.

In conclusion, the novel method of generating log sweeps, and deconvolving the system's response through a linear convolution with a proper inverse filter, revealed to possess only advantages over the already known, competing techniques such as MLS, TDS and Stretched Pulse. What's lacking, simply, is a short, appealing name for denoting the new technique: suggestions are welcome….

9. References

[1] A. Farina, F. Righini, "Software implementation of an MLS analyzer, with tools for convolution, auralization and inverse filtering", Pre-prints of the 103rd AES Convention, New York, 26-29 September 1997.

[2] J. Vanderkooy, "Aspects of MLS measuring systems", JAES vol. 42, n. 4, 1994 April, pp. 219-231.

[3] Ole Kirkeby, Per Rubak, Angelo Farina - "Analysis of ill-conditioning of multi-channel deconvolution problems" - 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics - Mohonk Mountain House New Paltz, New York October 17-20, 1999

[4] M. Poletti – "Linearly swept frequency measurements, time-delay spectrometry, and the Wigner distribution" – JAES vol. 36, n. 6, 1988 June, pp. 457-468.

[5] P. Fausti, A. Farina, R. Pompoli - "Measurements in opera houses: comparison between different techniques and equipment" - Proc. of ICA98 - International Conference on Acoustics, Seattle (WA), 26-30 june 1998.

[6] D. Johnston – "Cool Edit Pro v. 1.2" – HTTP://www.syntrillium.com, 1999.

[7] O. Kirkeby, P. A. Nelson, H. Hamada – "The "Stereo Dipole"-A Virtual Source Imaging System Using Two Closely Spaced Loudspeakers" – JAES vol. 46, n. 5, 1998 May, pp. 387-395.

[8] A. Farina, "Report on the Italian AES Workshop on room acoustics measurements", Bergamo (I), 27/28 April 1999 – HTTP://aurora.ramsete.com/AES-BG

[9] A. Farina – AURORA software suite – HTTP://aurora.ramsete.com – Acoustec ltd Publisher, London, 1999.

[10] SpectraLab v. 4.32.14 - HTTP://www.soundtechnology.com, 1999.
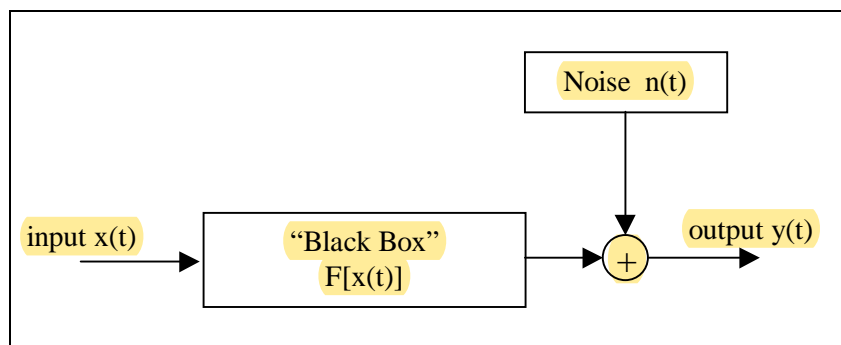
Fig. 1 – A basic input/output system
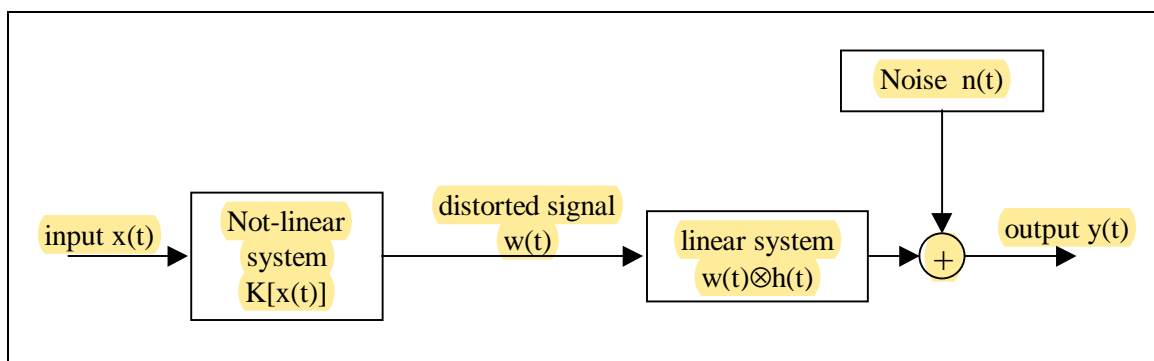
Fig. 2 – A more complex system, in which a not-linear, memoryless device drives a subsequent linear, reverberating system
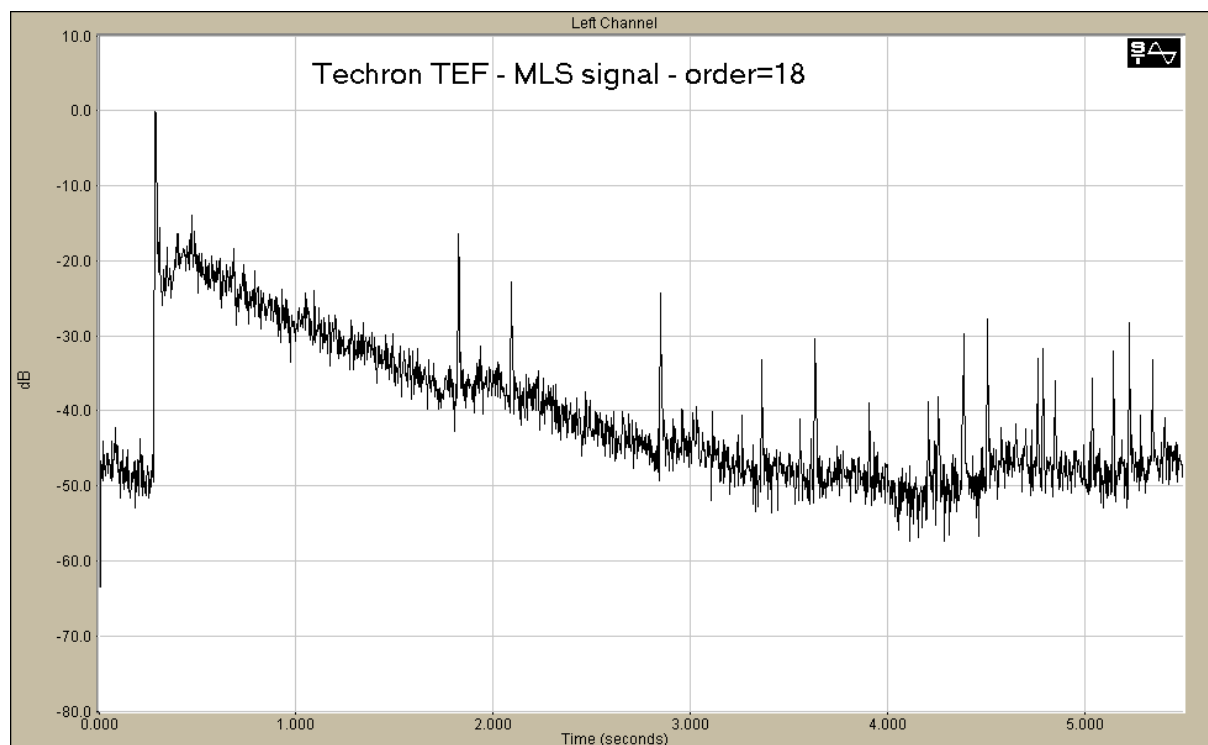
Fig. 3 – a MLS measurement made in presence of a strongly not-linear system
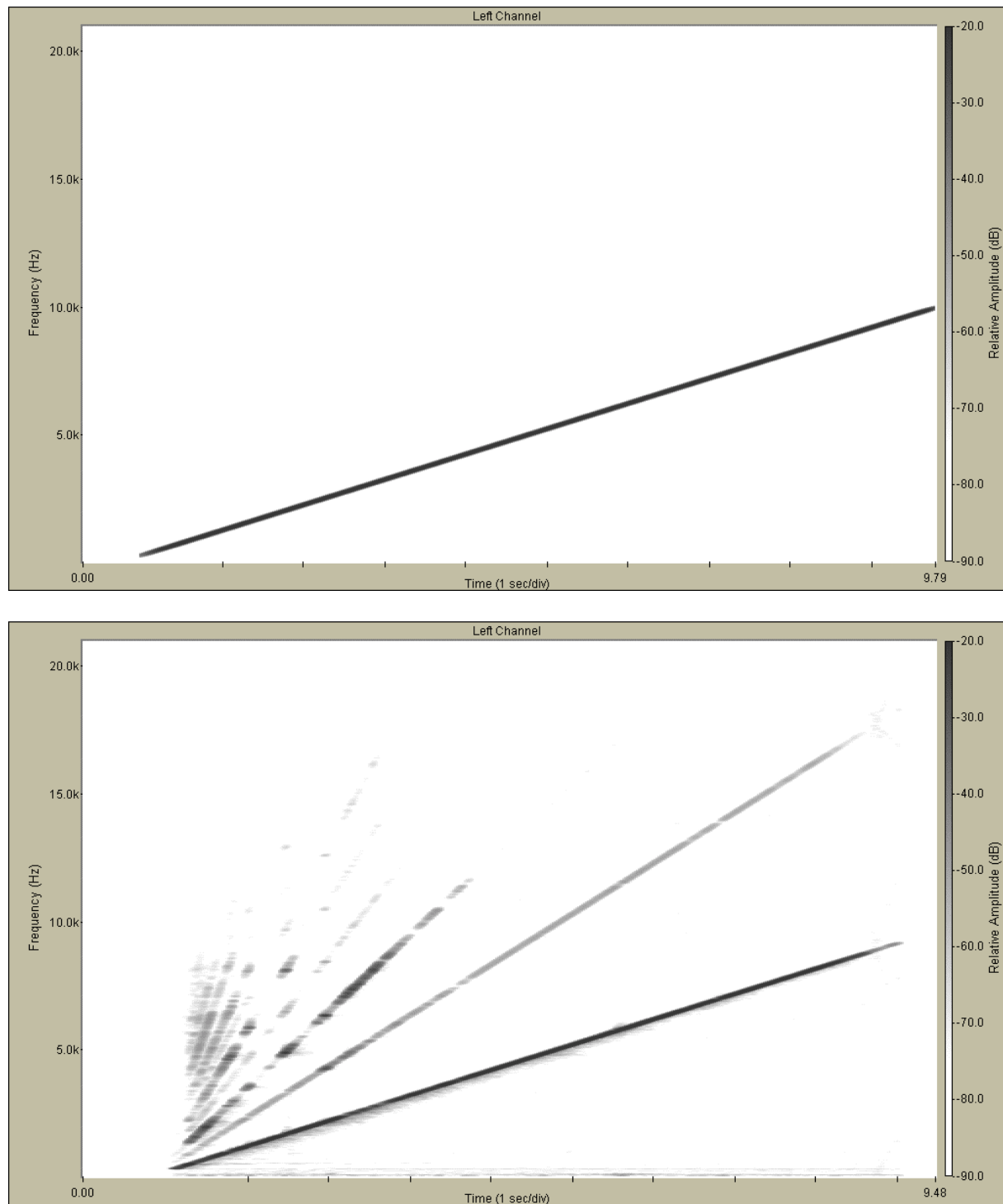
Fig. 4 – linear sine sweep: excitation signal (above) and system response (below) in the case of a weakly notlinear system exhibiting evident harmonic distortion.
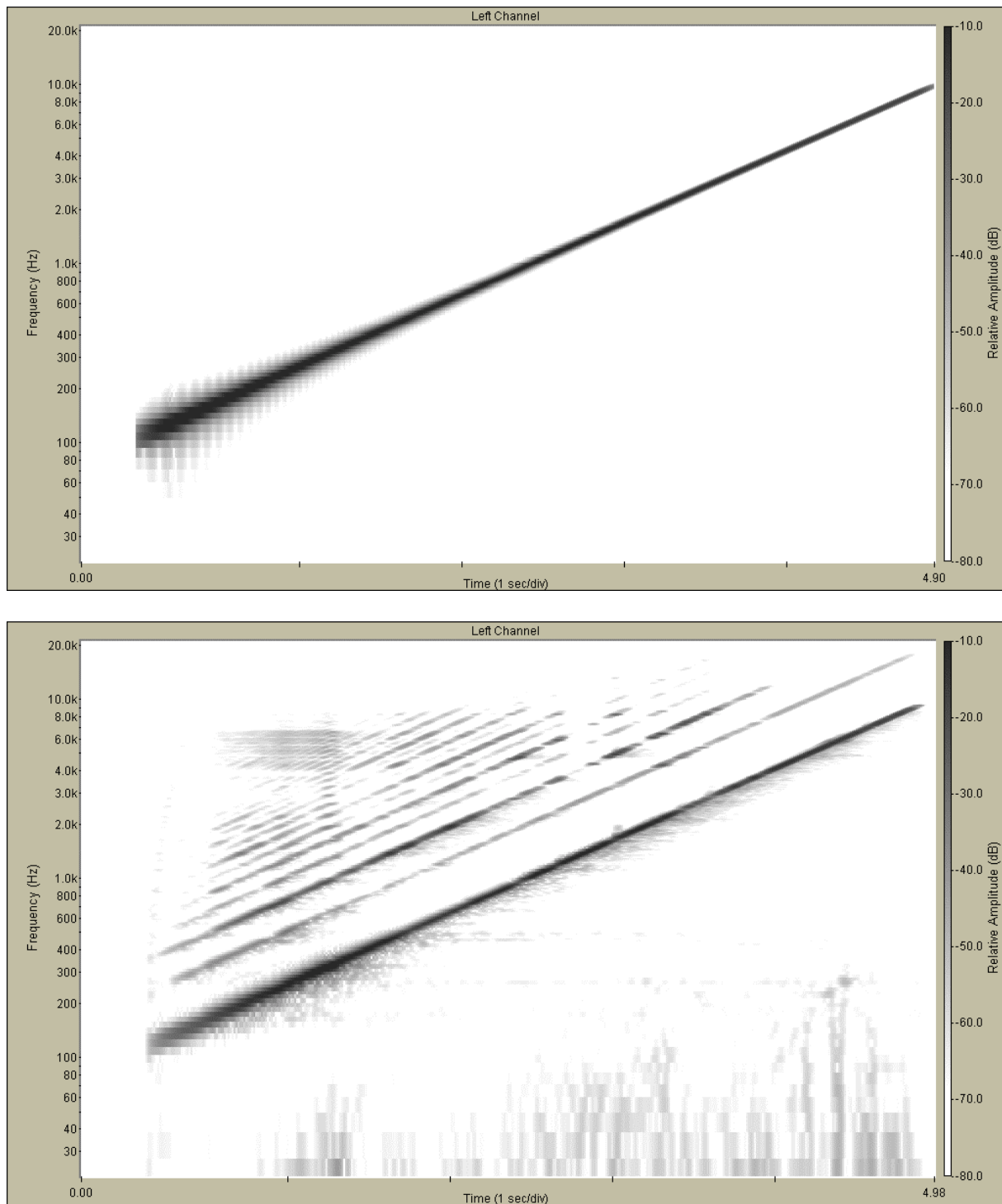
Fig. 5 – logarithmic sine sweep: excitation signal (above) and system response (below) in the case of a weakly notlinear system exhibiting evident harmonic distortion.
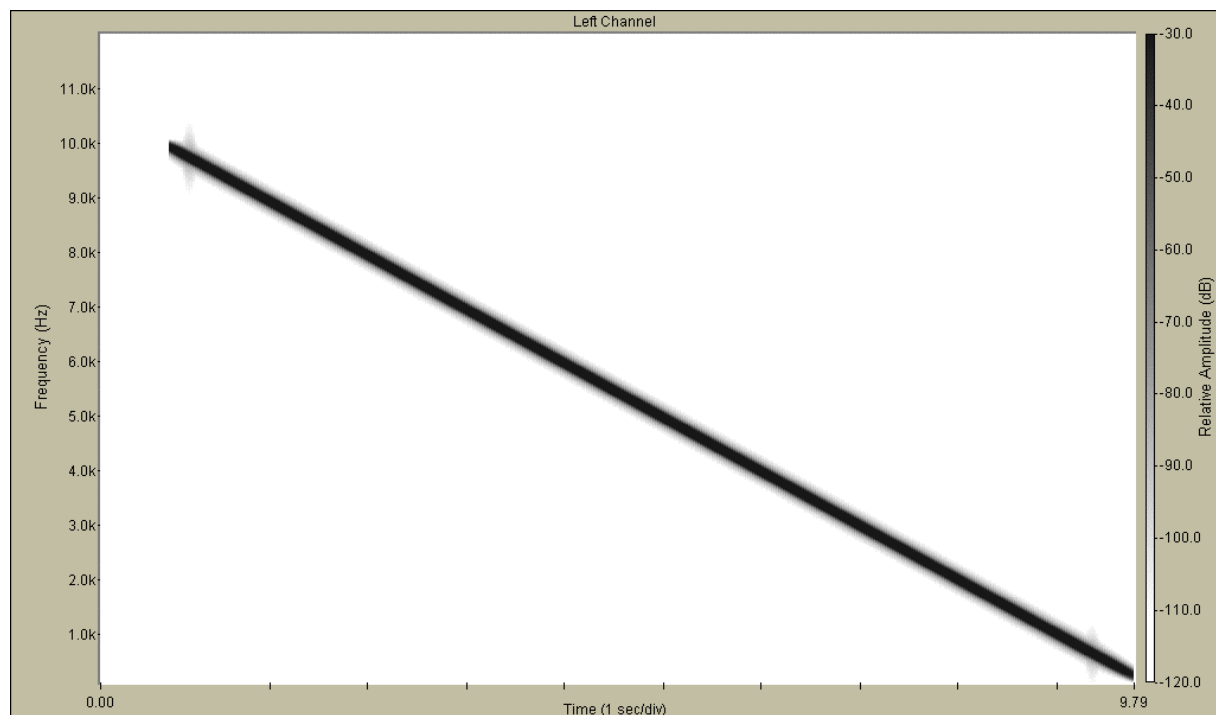
Fig. 6 – sonograph of the inverse filter – linear sweep
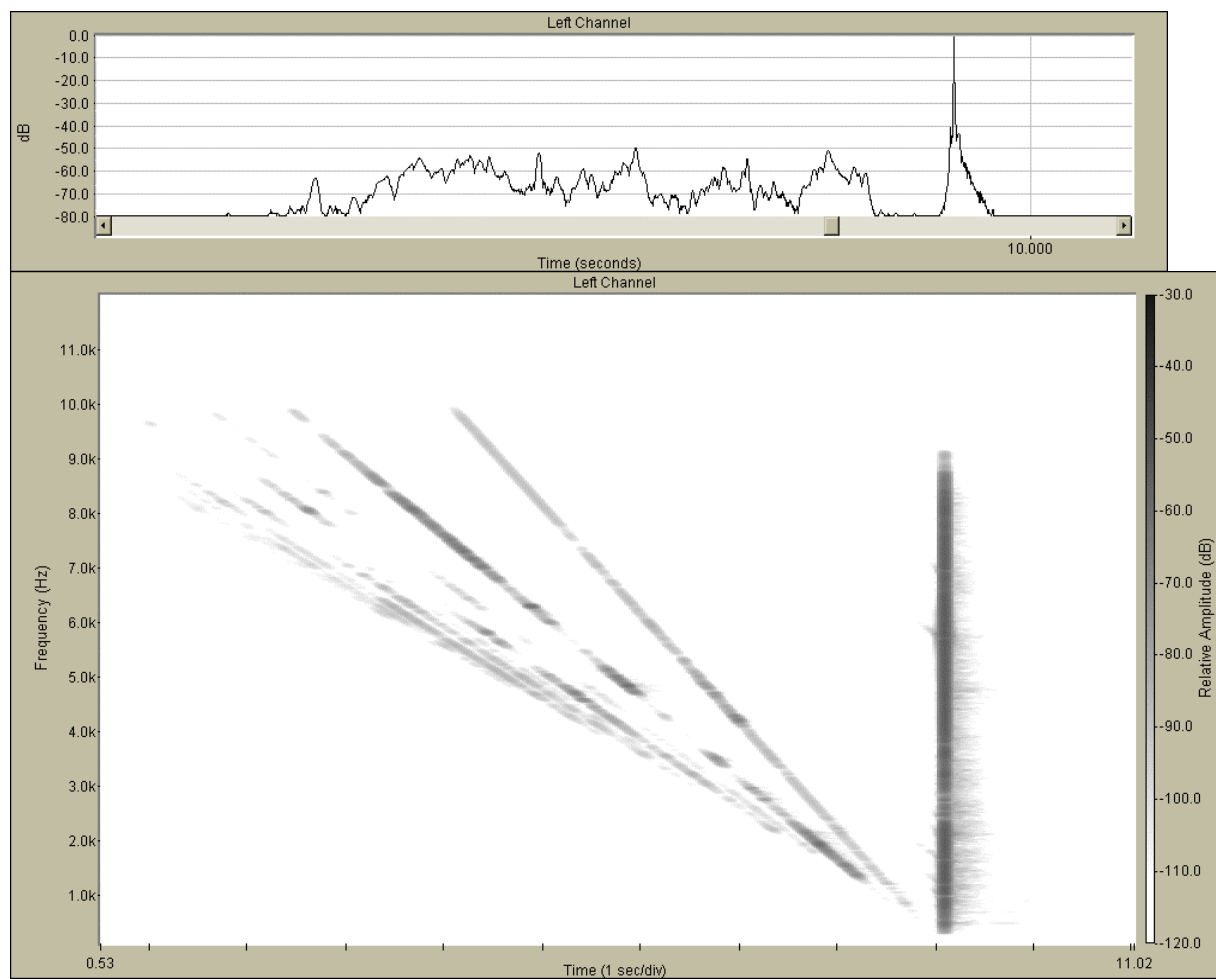
Fig. 7 – deconvolution of the system's impulse response after a linear sweep excitation
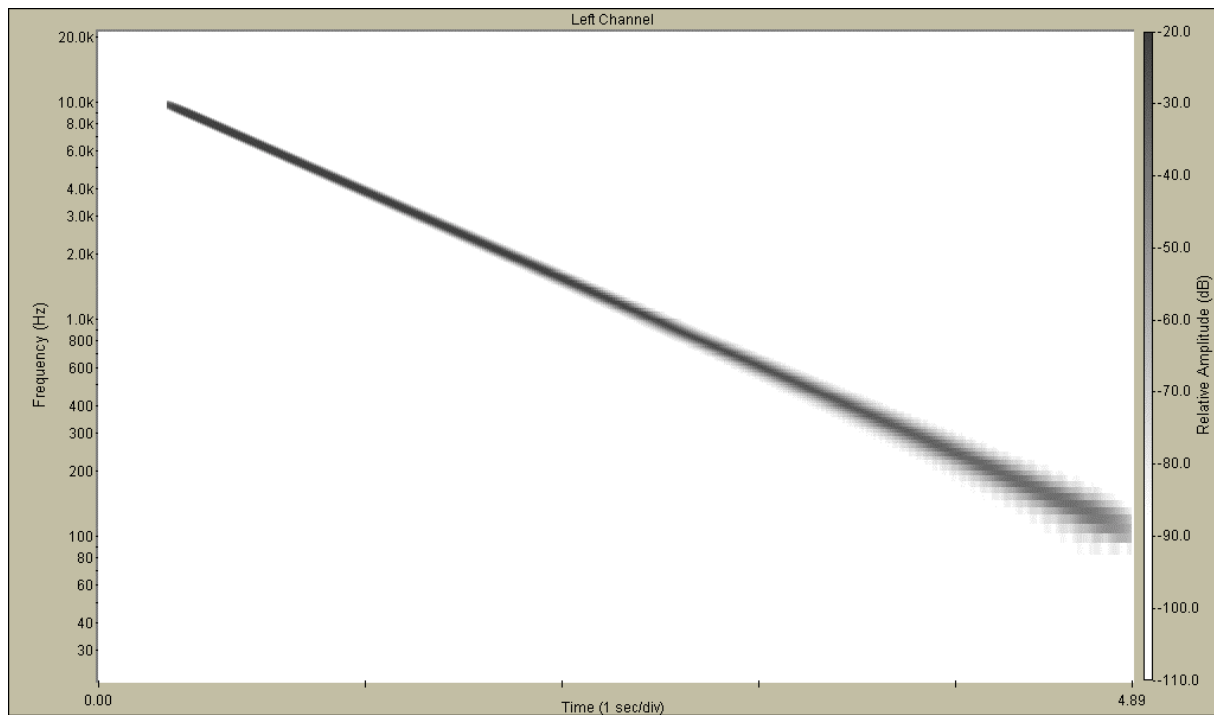
Fig. 8 – sonograph of the inverse filter – log sweep
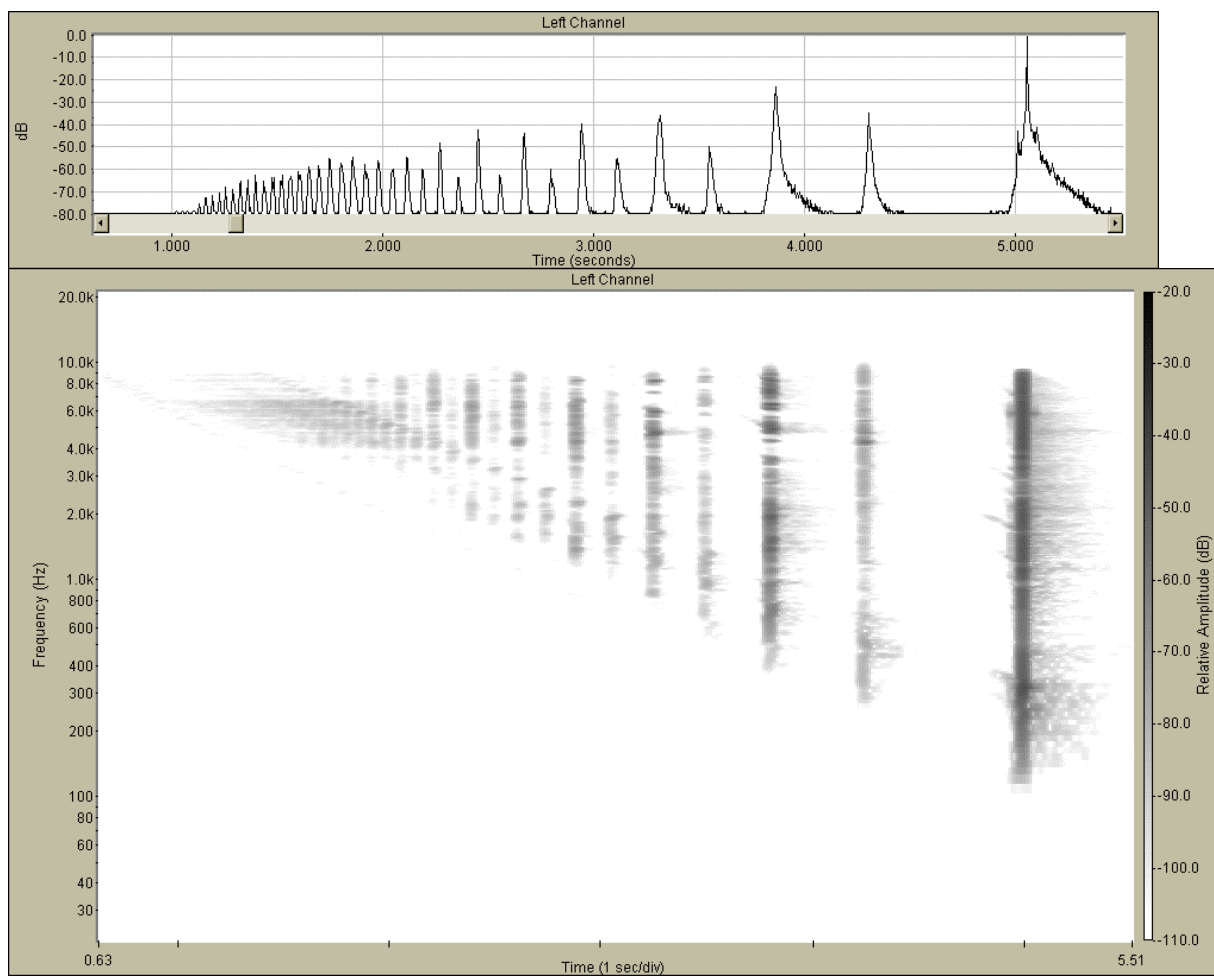
Fig. 9 – deconvolution of the system's impulse response after a log sweep excitation

Fig. 10 – user's interface of the plug-in for generating the sine sweeps
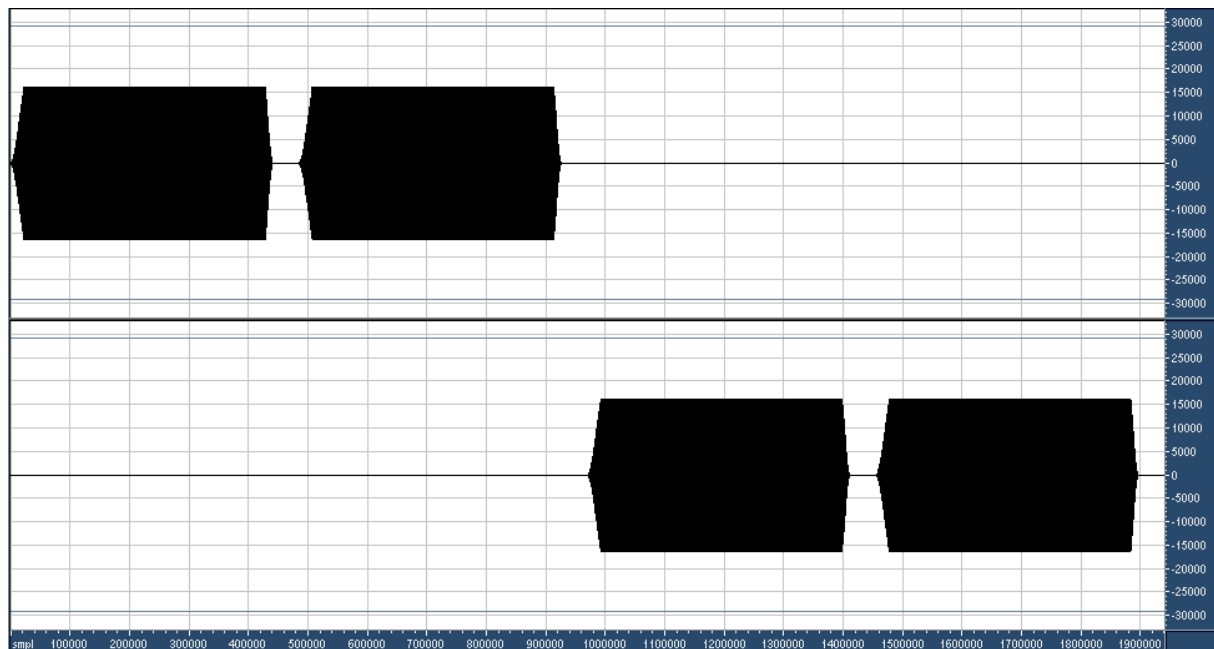
Fig. 11 – generation of a stereo sweep sequence (left first, then right)
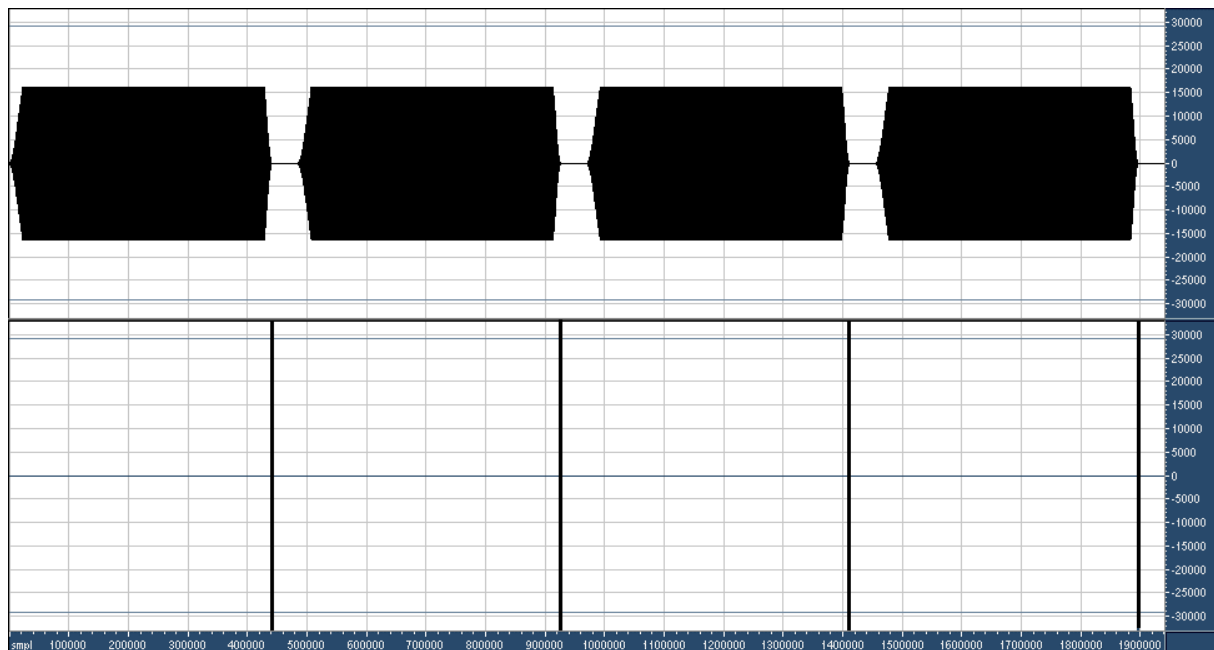
Fig. 12 – generation of multiple sweeps on the left channels, and control pulses on the right channel for stimulating the advancement of a motorized rotating board
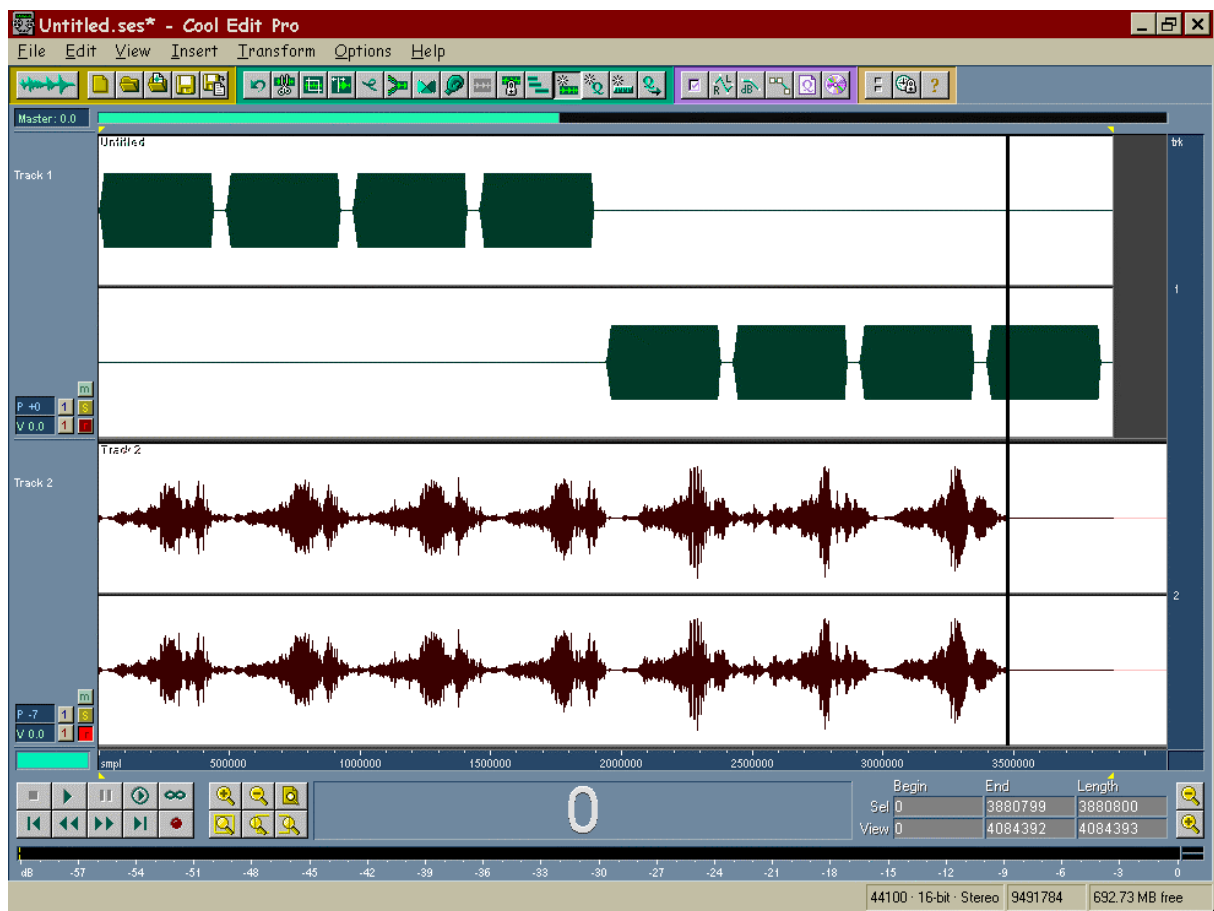
Fig. 13 – CoolEditPro during a multitrack session: sine sweeps are generated over a pair of loudspeakers (upper waveform), whilst the system's response is recorded through a pair of microphones (lower waveform)
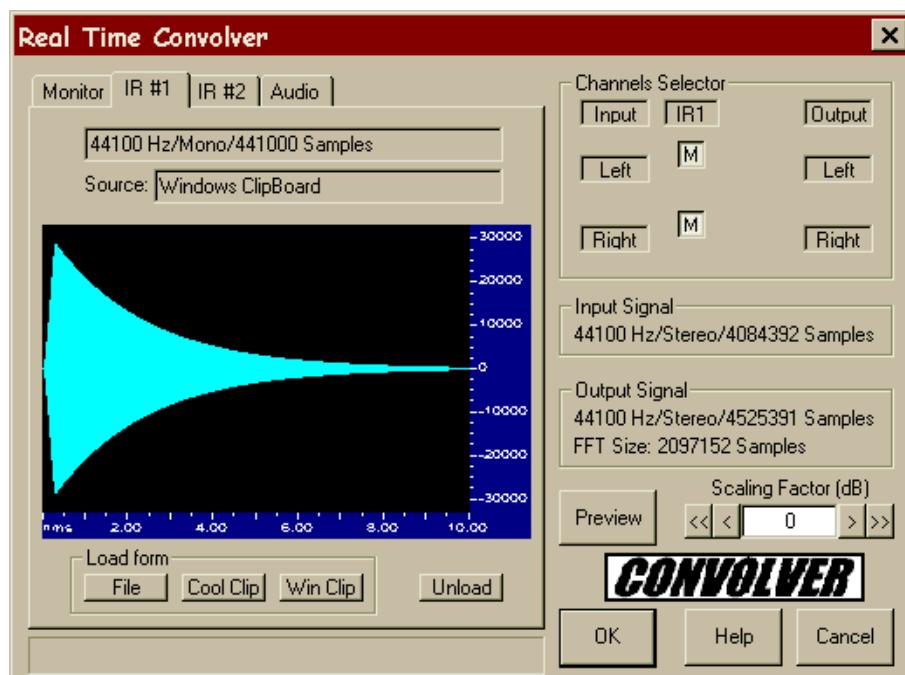
Fig. 14 – User's interface of the plug-in which performs the convolution of the measured data with the inverse filter stored in the Windows Clipboard.
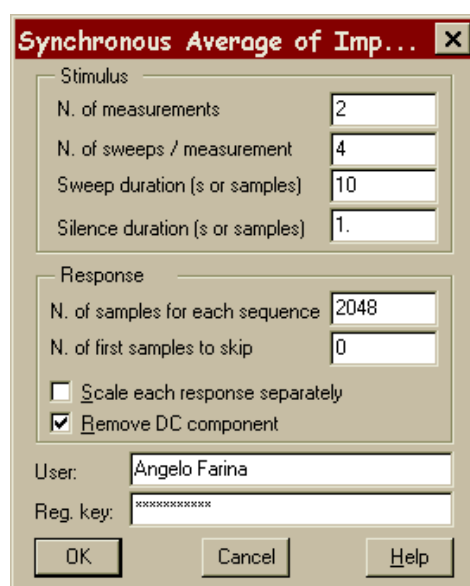
Fig. 15 – User's interface of the Synchronous-Average plug-in.
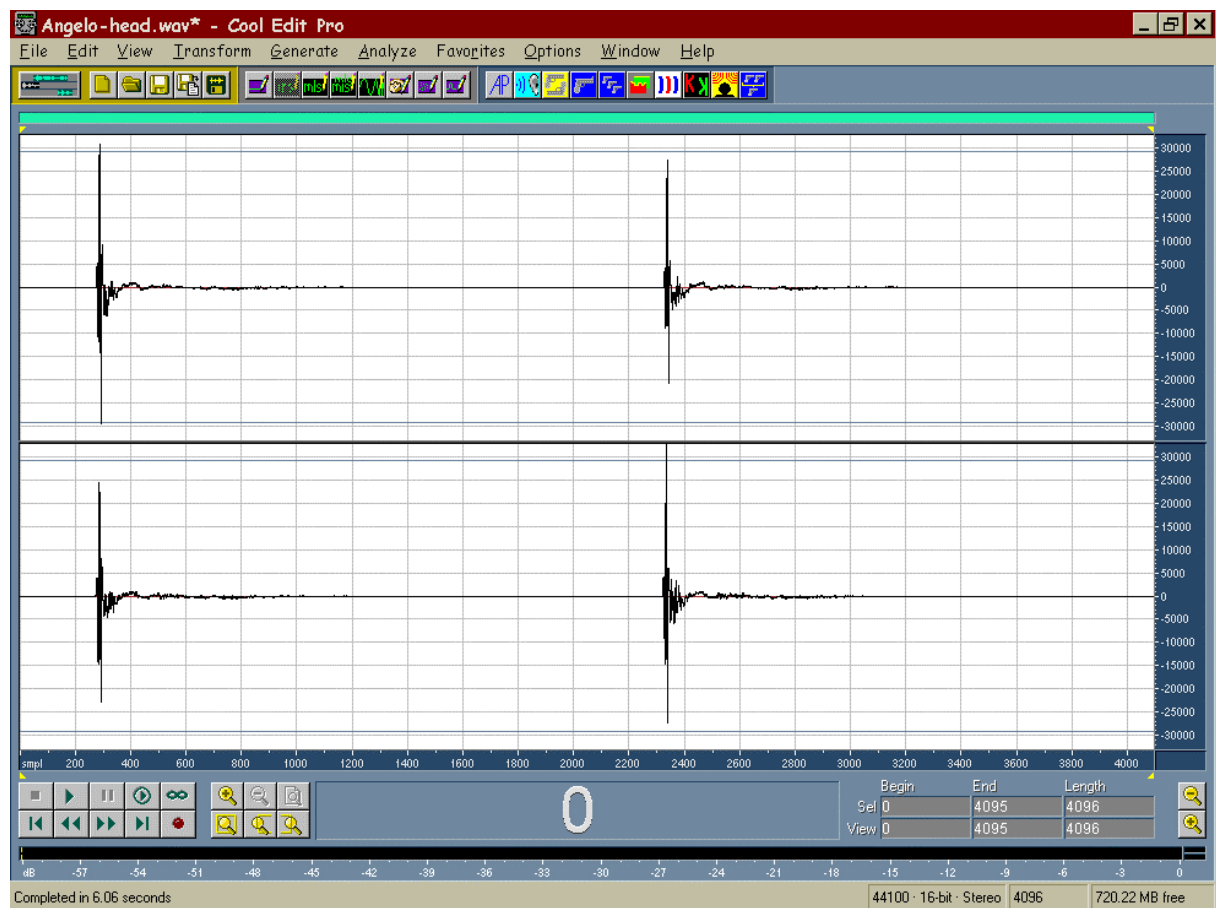
Fig. 16 – a set of 2x2 impulse responses obtained by a binaural measurement in front of a stereo-dipole loudspeaker pair, inside an anechoic chamber
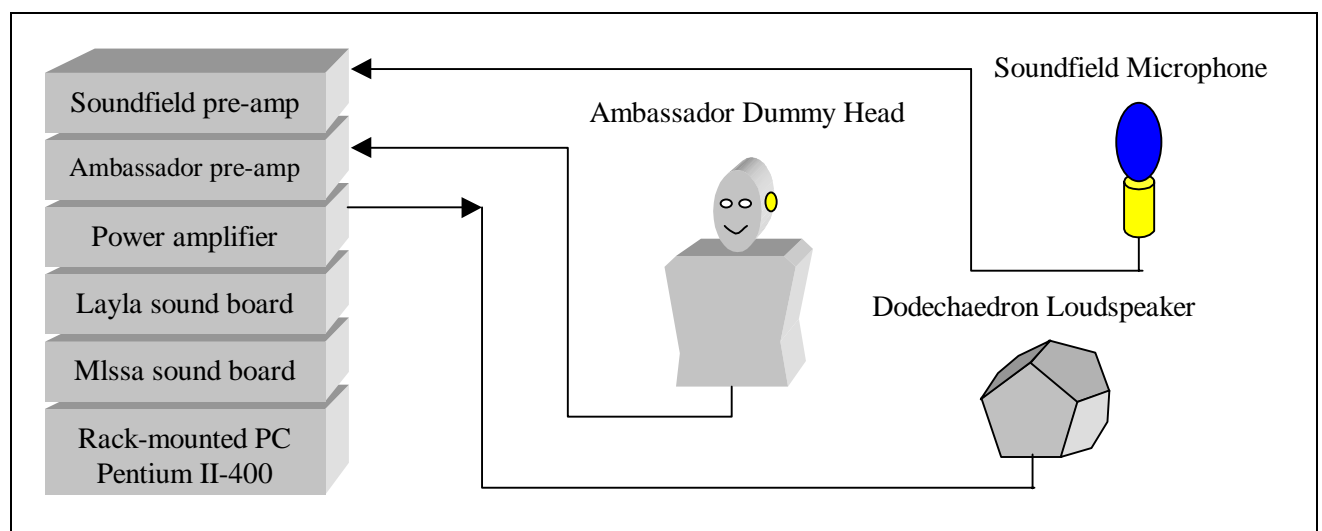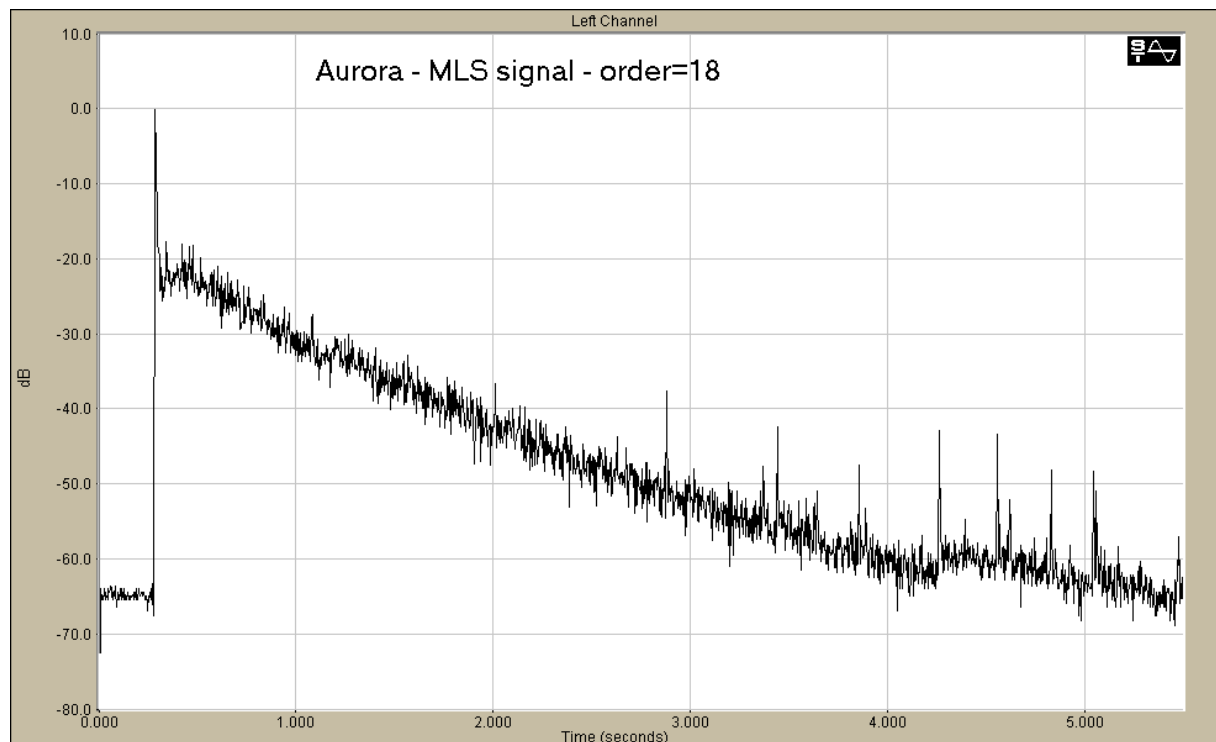

Fig. 17 – flow diagram of the measurement setup

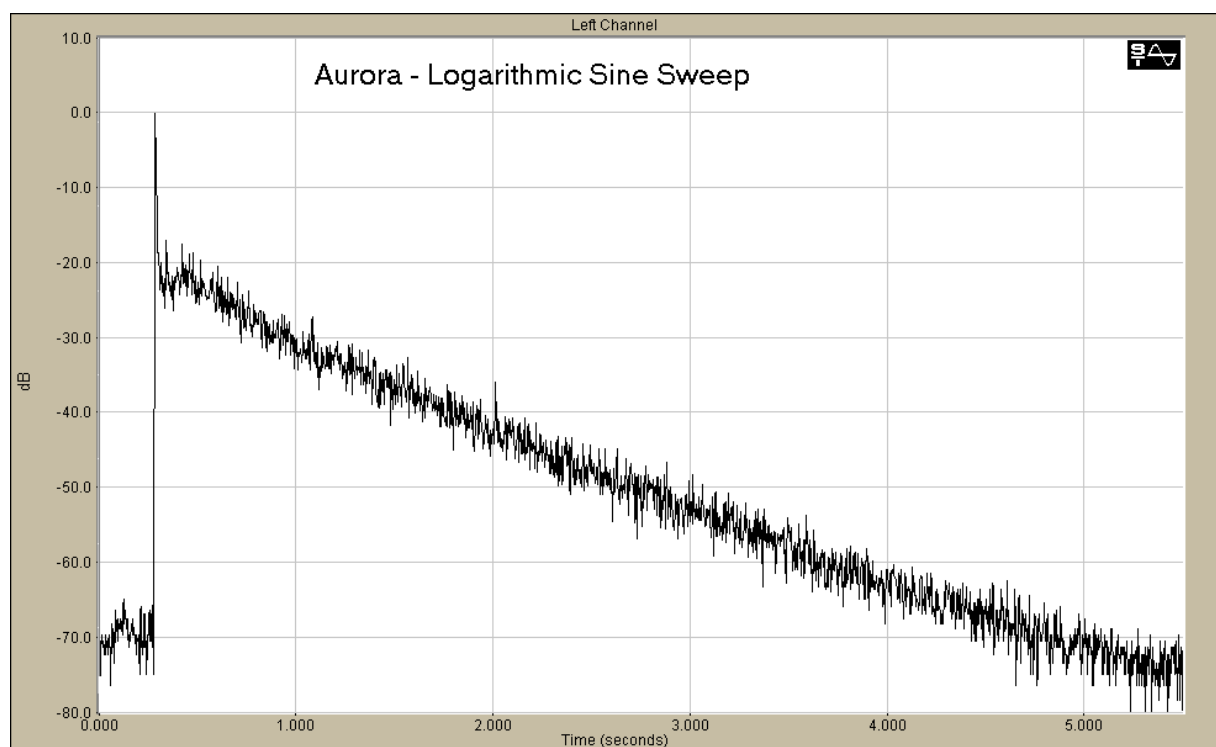Fig. 18 – Impulse response measurement with Aurora / MLS signal

Fig. 19 – Impulse response measurement with the new Aurora / log sine sweep method
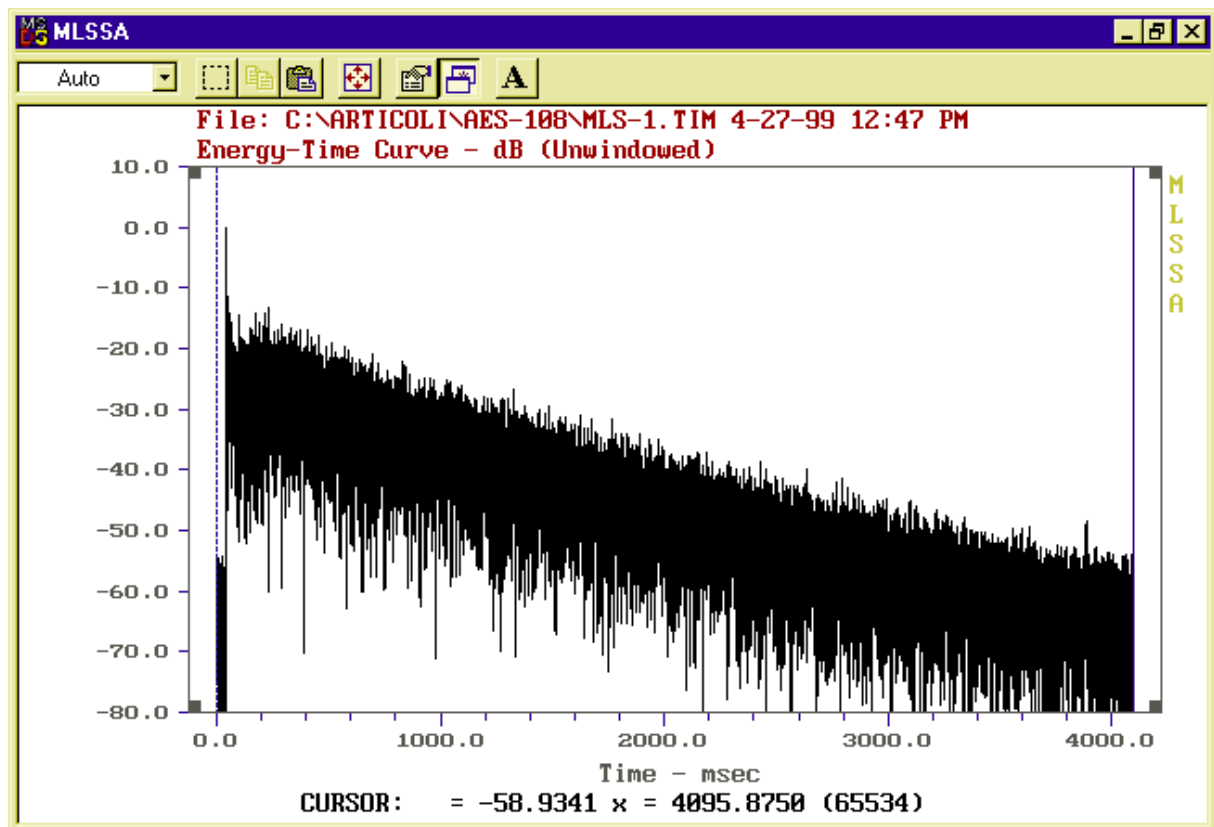
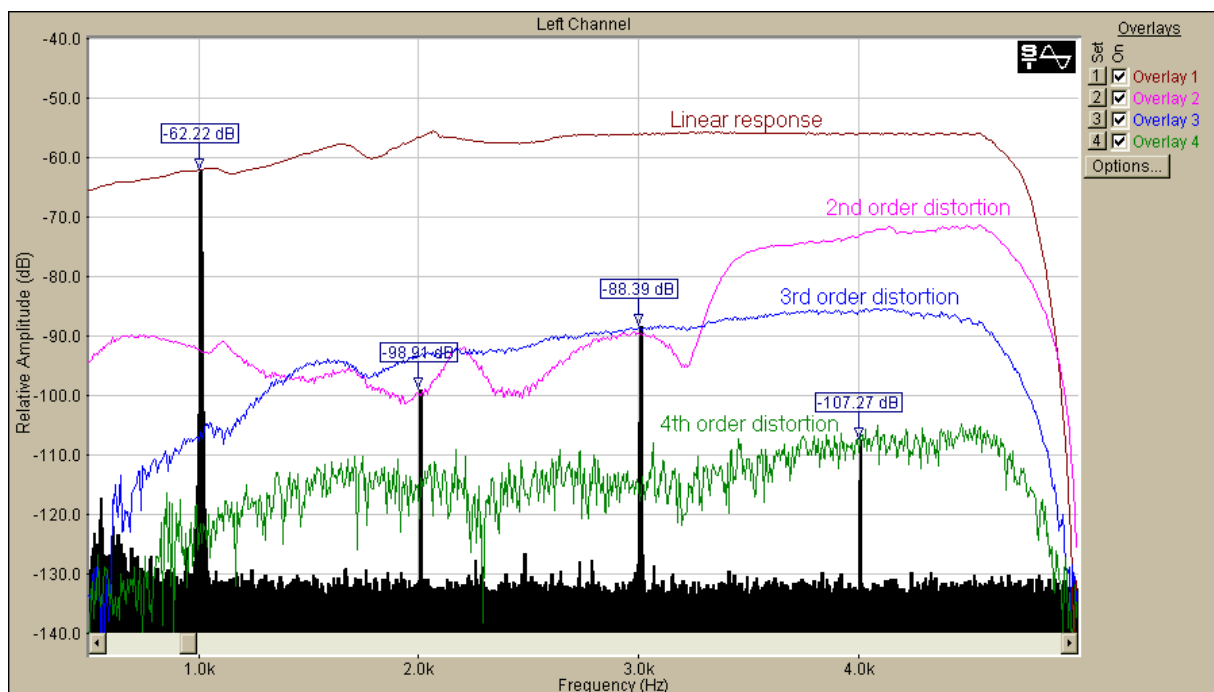Fig. 20 – Impulse response measurement with the MLSSA board.

Fig. 21 – comparison between traditional distortion measurement with fixed-frequency sine
(the black histogram) and the new log swept sine (the 4 narrow lines)