Audio Engineering Society
Convention Paper
Presented at the 110th Convention
2001 May 12–15 Amsterdam, The Netherlands

NON-LINEAR CONVOLUTION: A NEW APPROACH FOR THE AURALIZATION OF DISTORTING SYSTEMS

Angelo Farina, Alberto Bellini and Enrico Armelloni
University of Parma, Via delle Scienze 181/A
Parma, 43100 ITALY

ABSTRACT

This work defines a new method for processing audio signals, with the aim to recreate an audible simulation (auralization) of the modification imposed on the original signal by a complex system. The new method is the extension of the classic auralization process based on the linear convolution of the "dry" original signal with the impulse response of the system.

The extension allows for the emulation of non-linear systems, characterized in terms of harmonic distortion at several orders. The work first presents the mathematical framework of the proposed implementation, then it is shown how a not linear system can be experimentally characterized by a new measurement method of multiple impulse responses at various harmonic orders, and finally it is shown how these impulse responses can be employed in a multiple convolution process: an experimental demonstration is given of the similarity of the numerically processed sound with the live recording coming from a highly distorting device.

INTRODUCTION

The traditional auralization process is in common use since some years [1]. The method is usually employed for adding to dry music or speech recordings a set of information related to an acoustic space (and optionally to the sound system installed in it) such as reverberation and frequency response. Usually the system is modeled as a linear, time invariant process, and thus it is completely characterized by its impulse response.

Several measurement techniques have been developed for measuring the impulse response of a system [2,3,4] or for predicting it in the case of large rooms [5,6,7], small rooms [8,9] and taking into account the directivity of sources and receivers [10]. Furthermore, efficient implementations are available for realizing the convolution process in frequency domain, for example with the traditional select-and-save algorithm [11], or with partitioned-block frequency filtering methods, pioneered by Soo and Pang [12] for the case of an impulse response subdivided in equally-sized blocks, and further refined by Gardner and McGrath [13,14] with the introduction of the partition of the impulse response in different-size blocks, which allows for very little Input/Output delay.

In all those cases, anyway, the two constraints imposed on the system (linearity and time invariance) must be closely respected. In fact, even minor deviations from linearity or time invariance can disrupt completely the measurement of the impulse response [15,16], and even if these problems are circumvented with more advanced measurement techniques [15], the auralization obtained by the linear convolution process cannot represent faithfully the non-linear effects, which instead are often present in real-world systems, and which happen to be subjectively well noticeable. This situation is responsible for the fact that the sounds obtained with the traditional auralization method usually are perceived as slightly unrealistic, artificial, or unnaturally “clean”, due to the complete absence of the “natural” artifacts that are usually present in the real systems.

These effects are particularly important when the auralization method is employed for subjective comparison of different electro acoustic reproduction systems, as it is common in the car-audio field of application. The non-linear effects are an important part of the evaluation of the perceived sound quality, and their complete removal causes a harmful bias of the subjective responses.

Recently the authors developed a novel measurement method [15], which allows for the complete characterization of the linear and not linear behavior of a complex system with a single measurement. The result of this measurement procedure is a set of impulse responses, the first being the traditional linear response, and the other the responses at several harmonic orders. From these measurement results, all the traditional metrics for describing the distortion of a reproduction system can be derived easily.

Here it is proposed to employ this set of impulse responses in a multiple convolution process, capable of reconstructing the complete modification that happens to a signal passing through the complex system.

The theory behind the new processing method is briefly recalled, then the results of actual measurement on distorting systems are shown, and finally it is demonstrated (also with audible examples) how the proposed method can accurately reproduce the distortion effects produced by non-linear reproduction systems

THEORY

The following picture describes the flow diagram of a system obtained by a distorting transducer (memory-less distortion) driving a subsequent linear system with memory:

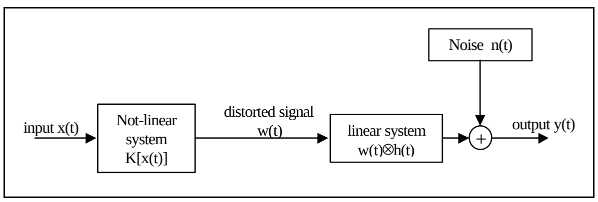

Figure 1. Flow diagram of the complex system

Neglecting the noise, the transfer function of this system can be described, in general, by means of a Volterra series expansion:

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

This general formulation also takes into account non-linear behaviour with memory (i.e., hystheresis), whilst in our case we are supposing that memory effects can be present only in the linear part of the two-block system illustrated in fig. 1. This means that, for orders higher than 1, the Volterra kernels h_2, h_3,… instead of being large multidimensional matrices reduce only to the terms on the diagonal, and thus can be represented by simple linear vectors having the same size as the first-order (linear) kernel. Under these hypotheses, eqn. 1 reduces to:

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

If these simplified (one-dimensional) Volterra kernels are known, eqn. 2 makes it possible to reconstruct the output signal y for any given input signal x.

MEASUREMENT TECHNIQUE

Employing a proper test signal, and sampling the output of the system, it is possible in general to compute not only the simplified, one-dimensional Volterra kernels, but even the complete n-dimensional ones. For example, Reed and Hawksford developed such a measurement technique based on varying-amplitude repetitions of an MLS (maximum length sequence) signal [16]. This technique revealed to be too complex and slow for practical applications, so the authors developed an alternative method [15].

The following pictures explain the measurement method. First, a test signal of length T, and covering the frequency range from w_1 to w_2 is generated, with this analytical expression:

$$x(t) = \sin[\omega_{var}] = \sin\left[\omega_1 \cdot T \Big/ \ln\left(\frac{\omega_2}{\omega_1}\right)\cdot\left(e^{\frac{t}{T}\cdot\ln\left(\frac{\omega_2}{\omega_1}\right)} - 1\right)\right] \qquad (3)$$

When the signal is introduced in the non-linear system, its output also contains harmonic distortion products, as shown here:

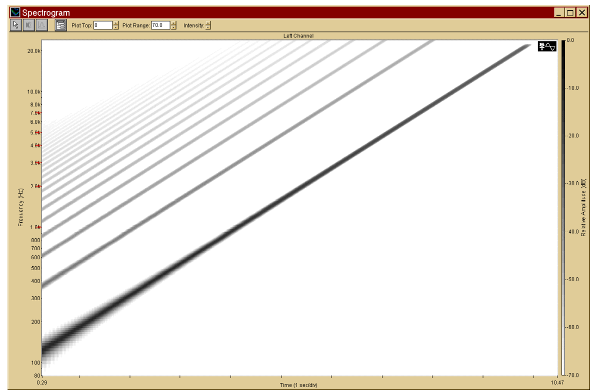

Figure 2. Spectrogram of the system’s response

It is possible to deconvolve the impulse response by applying to this response, by convolution, a proper inverse filter, which is simply the time-reversal of the excitation signal (3), equalized with a slope of 6dB/oct (time-reversal mirror plus whitening filter). This is the result:

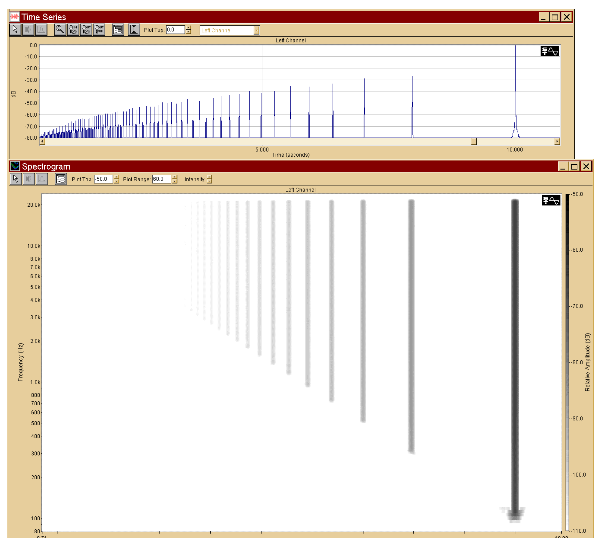

Figure 3. Spectrogram of deconvolved impulse responses

The rightmost impulse response is the linear one, which is preceded by the second-order harmonic response, and so on. The measured impulse responses are not directly the Volterra kernels, but these are easily computed by solving a linear equation system.

COMPUTATION OF THE VOLTERRA KERNELS

In practice the measurement procedure described in the previous chapter produces ordered impulse responses. The measured output signal can be represented as the sum of the linear convolution of the measured ordered impulse responses h'_i with the original input signal and the corresponding frequency-shifted version (obtained introducing respectively a factor 2, 3, etc. inside the sine function of eqn. 3):

$$y(t) = h'_1 \otimes \sin[\omega_{var}] + h'_2 \otimes \sin[2 \cdot \omega_{var}] + h'_3 \otimes \sin[3 \cdot \omega_{var}] + \ldots \quad (4)$$

The same result can also be espressed in terms of the diagonal Volterra kernels h_i defined in chapter 2:

$$y(t) = h_1 \otimes \sin[\omega_{var}] + h_2 \otimes \sin^2[\omega_{var}] + h_3 \otimes \sin^3[\omega_{var}] + \ldots \quad (5)$$

The powers up to 5th order of sine functions are:

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

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

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

$$\sin^5(\omega \cdot \tau) = \frac{5}{8} \cdot \sin(\omega \cdot \tau) - \frac{5}{16} \cdot \sin(3 \cdot \omega \cdot \tau) + \frac{1}{16} \cdot \sin(5 \cdot \omega \cdot \tau) \quad (6)$$

Taking the Fourier transform of eq. 4, and showing the result for a given value of the output frequency ω, we obtain:

$$Y(\omega) = \overline{H'}_1[\omega] \cdot X[\omega] + \overline{H'}_2[\omega] \cdot X[\omega/2] + \overline{H'}_3[\omega] \cdot X[\omega/3] + \ldots \quad (7)$$

This must be equal to the Fourier transform of eq. 5, where the expressions 6 have been substituted:

$$\begin{aligned} Y(\omega) = & \left[\overline{H}_1 + \frac{3}{4} \cdot \overline{H}_3 + \frac{5}{8} \cdot \overline{H}_5\right] \cdot X[\omega] + \left[-\frac{1}{2} \cdot \overline{H}_2 - \frac{1}{2} \cdot \overline{H}_4\right] \cdot j \cdot X[\omega/2] + \\ & + \left[-\frac{1}{4} \cdot \overline{H}_3 - \frac{5}{16} \cdot \overline{H}_5\right] \cdot X[\omega/3] + \frac{1}{8} \cdot \overline{H}_4 \cdot j \cdot X[\omega/4] + \\ & + \frac{1}{16} \cdot \overline{H}_5 \cdot X[\omega/5] + \ldots \end{aligned} \quad (8)$$

If the Fourier's theorem holds, this must be true for any input signal X, so the corresponding terms at second member of eqs. 7 and 8 must be the same. This allows for the following linear equation system to be set up for each frequency ω:

$$\begin{cases} \overline{H'}_1 = \overline{H}_1 + \frac{3}{4} \cdot \overline{H}_3 + \frac{5}{8} \cdot \overline{H}_5 \\ \overline{H'}_2 = -j \cdot \frac{1}{2} \cdot \left[\overline{H}_2 + \overline{H}_4\right] \\ \overline{H'}_3 = -\frac{1}{4} \cdot \overline{H}_3 - \frac{5}{16} \cdot \overline{H}_5 \\ \overline{H'}_4 = j \cdot \frac{1}{8} \cdot \overline{H}_4 \\ \overline{H'}_5 = \frac{1}{16} \cdot \overline{H}_5 \end{cases} \quad (9)$$

The solution of this system allows for the computation of the unknown diagonal Volterra kernels h_i starting from the measured ordered impulse responses h'_i.

This is the solution for the first 5 orders:

$$\begin{cases} \overline{H}_1 = \overline{H'}_1 + 3 \cdot \overline{H'}_3 + 5 \cdot \overline{H'}_5 \\ \overline{H}_2 = 2 \cdot j \cdot \overline{H'}_2 + 8 \cdot j \cdot \overline{H'}_4 \\ \overline{H}_3 = -4 \cdot \overline{H'}_3 - 20 \cdot \overline{H'}_5 \\ \overline{H}_4 = -8 \cdot j \cdot \overline{H'}_4 \\ \overline{H}_5 = 16 \cdot \overline{H'}_5 \end{cases} \quad (10)$$

After computing the values of the kernels from the measured multiple impulse response, the non-linear convolution can be efficiently implemented following eq. 2.

IMPLEMENTATION OF THE MULTIPLE CONVOLUTION AS A COOL EDIT PLUGIN

In principle, eqn (2) can be easily implemented as the sum (or mix) of many "usual" convolutions. The overall algorithm is in some way similar to the partitioned impulse response frequency domain convolution by Soo and Pang [12], but it differs because the input block of data is preliminarily subjected to a power elevation before being transformed and convolved with each impulse response block, and because the results of the convolution of the input block with the many IR blocks are not delayed, but summed one over the other. For making the task easier, a special *plugin* for CoolEditPro [17] was developed: it takes the current waveform, and processes it by convolution with the set of impulse responses (which have to be stored in the Windows clipboard before invoking the plug-in). In practice, the whole waveform resulting from the measurement is loaded into the clipboard (it is a sort of multiple IR, as described in [15]), and then the plugin automatically segments the multiple impulse response in the different orders, based on the information supplied by the user and related to the original input signal employed for the measurement.

The following picture shows the user interface of the Multiconvolve plugin.

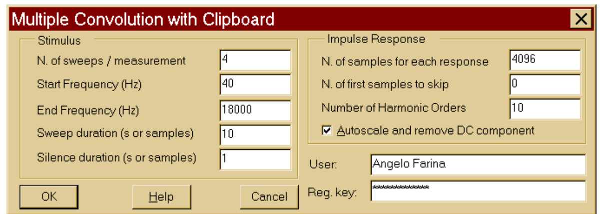

Figure 4. User's interface of the effect

Despite of the large number of multiply-add operations theoretically required for computing eqn. 2, the computation is very fast (real-time on modern computers), thanks to a frequency-domain implementation of each convolution operator, based on many FFTs and the well-known select-save algorithm [18]. This makes it possible to keep the process in real-time, even for values of M (impulse response length) of ten thousands coefficients and more.

Having all the FFT blocks the same size, this partitioned-frequency-domain block convolution is directly descending from the Soo and Pang algorithm, and thus it is outside of the applicability of the subsequent patent actually covering the not-uniform-size partitioned block algorithm [13,14].

SUBJECTIVE VERIFICATION OF THE NEW NON-LINEAR CONVOLUTION METHOD

First, a small, low quality loudspeaker was measured with the above-described technique, when driven with a signal of amplitude exceeding its linear limits. On the same loudspeaker, four music pieces were also played and the emitted sound was simultaneously recorded, imposing that the RMS amplitude of each music piece was the same as the test sound.

Then eqn. (2) was employed for mathematically reconstructing the distortion of the device, based on a digital copy of the music pieces and on the measured monodimensional Volterra kernels.

The experimental recordings and the artificially filtered samples were compared in a subjective listening test: 12 subjects had to fill-up a questionnaire for each pair-comparison, as shown in fig. 5.

The questionnaire was implemented thanks to a special software, which allows for instantaneous switching between the two samples (A/B comparison). A separate questionnaire is filled up for each of the 4 music pieces. The listening was made through an high quality reproduction system (24 bits sound board, high quality bi-amplified studio monitors, anechoic listening room). Each listener was unaware of what of the two samples in each pair was synthetic: the presentation order was randomly shuffled for the 12x4 questionnaires.

Each response is mapped on a numerical scale, ranging from 1 (leftmost mark) to 5 (rightmost mark).

Figure 5. Interface of the software for questionnaires

The following table shows the main statistical results coming from an ANOVA of the responses.

Table 1: statistical results

Question Number	Average score	Standard Deviation
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

Please note that for question 2..9, the responses were re-aligned in such way that the leftmost sample (A) is always associated to the original recordings, and the rightmost sample (B) is the synthetic signal.
The response of question 1 means that the two samples were considered very often identical, although a small percentage of listeners (17%) detected some difference.
It can be concluded that the perceived differences are below the significance level, and thus that the synthetic approximation of the non-linear device is a reasonable representation of the reality.

CONCLUSIONS

The proposed method for measurement and subsequent reproduction of the behaviour of a non-linear system revealed to be fast, robust and capable of producing realistic results. Although some strong simplifications were introduced for reducing the dimensionality of high-order Volterra kernels, the multiple convolution method revealed to fill almost completely the gap existing, till now, between linear auralization and real-world recordings made with non-linear devices.

The new technique will be employed in the next future for laboratory subjective comparisons of different sound reproduction systems, with the goal of better understating the relationship between human preference and objective electroacoustic parameters.

ACKNOWLEDGEMENTS

David Johnston, author of CoolEditPro [17], is acknowledged for his excellent software, which was kindly made available free for this research.
Many of the graphs presented here were obtained through post-processing made with the program SpectraLab by SoundTechnology [19], during the 30-days free license period.
This work was supported through a research convention between ASK Industries, Reggio Emilia, Italy and the University of Parma, co-funded by the Italian Ministry for University and Research (MURST) under the grant MURST-98 #9809323883-007.

DOWNLOADABLE AUDIBLE DEMONSTRATION

The presentation of this paper is accompanied by the reproduction of the four sound samples employed for the subjective experiments, each of them both in its real-world version and in the synthetic one. The same samples are downloadable in WAV format from HTTP://pcangelo.eng.unipr.it/Public/AES-110. The plugins for CoolEdit, referred into this paper and regarding impulse response measurements and convolution, can be freely downloaded from HTTP://www.ramsete.com/aurora.

References

[1] Kleiner M., Dalenback B., Svensson P., "Auralization - An Overview", *AES Journ., vol. 41 N.11, p.861-875, November 1993.*

[2] Rife, D.D., and Vanderkooy, J., "Transfer-function measurements with maximum length sequences", *J. Audio Eng. Soc., June, 1989, 37, pp. 419–444.*

[3] Vanderkooy, J., "Aspects of MLS measuring systems", *J. Audio Eng. Soc., April 1994, 42, pp. 219–231.*

[4] Farina A., Righini F., "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.

[5] Naylor G.M., "ODEON - Another hybrid room acoustical model" - *Applied Acoustics Vol. 38 no.s 2-4, pag. 131 (1993).*

[6] Dalenbäck B.I., "Room acoustic prediction and auralization based on a unified treatment of diffuse and specular reflection" – 15th Intern. Congress on Acoustics, Trondheim, Norway, pp. 425-428, June 1995.

[7] Farina A., "RAMSETE - a new Pyramid Tracer for medium and large scale acoustic problems", Proc. of EURO-NOISE 95 Conference, Lyon 21-23 march 1995.

[8] E. Granier, M. Kleiner et al., "Experimental auralization of car audio installations", *J.A.E.S., vol. 44, n. 10, pp.835-849, 1996 October.*

[9] Farina A., Ugolotti E., "Numerical model of the sound field inside cars for the creation of virtual audible reconstructions", DAFX-98 Conference, November 19-21, 1998 Barcelona, Spain.

[10] Tronchin L., Farina A., Pontillo M., Tarabusi V., "Binaural hearing and its numerical representation with an Eulerian approach", Seventh International Congress On Sound And Vibration ICSV7 - Garmisch-Partenkirchen, Germany July 4-7, 2000.

[11] A. Farina - "An example of adding spatial impression to recorded music: signal convolution with binaural impulse responses" - Proc. of International Conference "Acoustics and recovery of spaces for music", Ferrara 27-28 October 1993.

[12] J.S. Soo, K.K. Pang – "Multidelay block frequency domain adaptive filter", *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 39, no. 2, pp. 373-376, Feb. 1990.

[13] Really A., McGrath D.S., "Convolution processing for realistic reverberation", Pre-Prints of the 98th AES Convention, 1995 February 25-28, Paris

[14] Gardner W.G. – "Efficient convolution without input-output delay", *JAES vol. 43, n. 3, 1995 March, pp. 127-136.*

[15] Farina A., "Simultaneous measurement of impulse response and distortion with a swept-sine technique", 108th AES Convention, Paris 18-22 February 2000.

[16] Hawksford M.O.J., Reed M.J., "Identification of discrete Volterra series using maximum length sequences", *IEE Proc.-Circuits Devices Syst., Vol. 143, No. 5, pp. 241-248, October 1996.*

[17] Johnston, D., "Cool Edit Pro v. 1.2A", HTTP://www.syntrillium.com, 2000.

[18] Oppheneim A.V., Schafer R., *Digital Signal Processing*, Prentice Hall, Englewood Cliffs, NJ 1975, p. 242

[19] SpectraLab v. 4.32.16 - HTTP://www.soundtechnology.com, January 2001.