The Emulation of Nonlinear Time-Invariant Audio Systems with Memory by Means of Volterra Series

LAMBERTO TRONCHIN, *AES Member*
(lamberto.tronchin@unibo.it)

DIENCA-CIARM, University of Bologna, Bologna, Italy

The measurement and emulation of audio systems (devices, environments, and sound boxes) have been studied in recent years. The most-used methods to obtain information about an audio system are those based on measuring its impulse response (IR). Once the IR has been caught it is possible to recreate, by the use of linear convolution, the output signal that the audio system will generate when it is physically driven by any input signal. This method gives great results if the system is linear and time-invariant (environmentals behavior is very linear and, therefore, its reverberant effect can be faithfully recreated using IRs) but not satisfactory in other cases, such as the emulation of tube preamps (mainly nonlinear) and musical instruments. Since musical instruments cannot be considered completely linear, their musical performance might be analyzed properly considering also their nonlinear behavior. By using Volterra series it is possible to represent the input-output relationship of nonlinear systems. This mathematical theory uses a set of impulse responses to describe the system and not only one as before. By an enhanced impulse response measurement method it is possible to obtain this set of impulses and, then, with Volterra series it would be possible to have the output of the audio system driven by any input. A numerical tool has been developed to recreate the system behavior by using this method. Satisfactory results have been obtained in comparison with the traditional linear convolution based approach.

0 INTRODUCTION

For almost 100 years, it has been theoretically possible to model linear time-invariant (LTI) systems with memory [1]. The model is based on the convolution between the input signal and the impulse response (IR) of a system [2, 3]. This approach fails when we try to predict the behavior of a non-linear time-invariant system with memory.

The objective of this paper is to show how the emulation of a non-linear time-invariant system with memory based on a Volterra series is possible. There currently exists a measurement technique that is able to provide, under certain conditions, the data required to calculate kernels [4]. These conditions require a special phase equalization of the signal and the correct value of its amplitude.

1 NON-LINEAR SYSTEMS

A non-linear system is a system in which superposition and homogeneity properties are generally not valid. The superposition of a system indicates that $F[x_1(\tau) + x_2(\tau)] = F[x_1(\tau)] + F[x_2(\tau)]$, where $F[]$ is a generic function. Moreover, the homogeneity of a system indicates that $F[\alpha x(\tau)] = \alpha F[x(\tau)]$. These features, which distinguish a linear system, indicate that the output will be a linear combination of the input: no new frequencies will be created by the system itself. This is not true for non-linear systems that are able to create components at frequencies not present in the input signal [5], [6]; these new frequencies are defined as harmonic and intermodulation distortions, where the harmonic ones represent new frequencies at multiple values of those present in the input, and the latter ones are new frequencies that are linear combinations of those present in the input. The non-linear systems could be modeled in different way, from a more generic description to a more specific case. A non-linear time variant system with memory could be described using the following Volterra series [7, 8]

$$\begin{aligned} y(t) = & \int_{-\infty}^{+\infty} h_1(t, \tau_1)\, x(\tau_1)\, d\tau_1 \\ & + \int \int_{-\infty}^{+\infty} h_2(t, \tau_1, \tau_2)\, x(\tau_1)\, x(\tau_2)\, d\tau_1 d\tau_2 \\ & + \int \int \int_{-\infty}^{+\infty} h_3(t, \tau_1, \tau_2, \tau_3)\, x(\tau_1)\, x(\tau_2) \\ & \times x(\tau_3)\, d\tau_1 d\tau_2 d\tau_3 + \ldots\ldots \end{aligned} \tag{1}$$

where the $h_n(t, \tau_1, \tau_2, \ldots \tau_n)$ terms are the kernels. The term for $n = 1$ characterizes the linear part of the system; it is simply the familiar impulse response. The non-linear

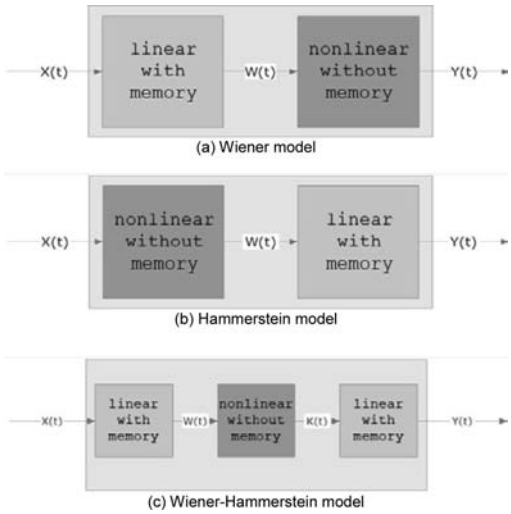


Fig. 1. Nonlinear system models.

part is characterized as starting from $h_2(t, \tau_1, \tau_2)$. These kernels can be viewed as higher-order impulse responses.

The time-invariant version of Eq. (1) is represented by the following:

$$\begin{aligned} y(t) &= h_0 + \sum_{n=1}^{+\infty} \frac{1}{n!} \int_{-\infty}^{+\infty} d\tau_1 \ldots \int_{-\infty}^{+\infty} d\tau_n h_n(\tau_1, \ldots, \tau_n) \\ &\times \prod_{r=1}^{n} x(t-\tau_r) = h_0 + \frac{1}{1!} \int_{-\infty}^{+\infty} h_1(\tau_1)\, x(t-\tau_1)\, d\tau_1 \\ &+ \frac{1}{2!} \int \int_{-\infty}^{+\infty} h_2(\tau_1, \tau_2)\, x(t-\tau_1)\, x(t-\tau_2)\, d\tau_1 d\tau_2 \\ &+ \frac{1}{3!} \int \int \int_{-\infty}^{+\infty} h_3(\tau_1, \tau_2, \tau_3)\, x(t-\tau_1)\, x(t-\tau_2) \\ &\times x(t-\tau_3)\, d\tau_1 d\tau_2 d\tau_3 + \ldots . \end{aligned} \tag{2}$$

where h_0 represents a DC offset value, which is often omitted, assuming the system is passive (i.e., no input and no output). Moreover, $h_n(\tau_1, \ldots, \tau_n)$ are the n-order kernels. It is possible to further simplify the model in Eq. (2) by introducing other constraints. Fig. 1 shows three simplified models based on the assumption that memory effects reside in the linear part of the system and that the non-linearity is purely algebraic (i.e., represented by a Taylor series expansion).

The Wiener model [8] represents the non-linear system as a series of linear time-invariant systems with memory and a purely algebraic non-linear time-invariant system, as shown in Eq. (3).

$$\begin{aligned} y(t) &= a_0 + \sum_{n=1}^{+\infty} a_n [w(t)]^n \\ &= a_0 + \sum_{n=1}^{+\infty} a_n \left[\int_{-\infty}^{+\infty} h(t-\tau)\, x(\tau)\, d\tau \right]^n \\ &= a_0 + a_1 h(t) * x(t) + a_2 [h(t) * x(t)]^2 \\ &+ a_3 [h(t) * x(t)]^3 + \ldots \end{aligned} \tag{3}$$

where $w(t)$ is the output of the linear part and is therefore substituted with $\int_{-\infty}^{+\infty} h(t-\tau)x(\tau)d\tau = h(t) * x(t)$, where $h(t)$ is the impulse response.

The Hammerstein model [9] represents the non-linear system as a series of a non-linear time-invariant systems without memory and an LTI system with memory; its output is as follows:

$$\begin{aligned} y(t) &= \int_{-\infty}^{+\infty} h(t-\tau)\, \omega(\tau)\, d\tau \\ &= \int_{-\infty}^{+\infty} h(t-\tau) \left\{ a_0 + \sum_{n=1}^{+\infty} a_n [x(t)]^n d\tau \right\} \\ &= \int_{-\infty}^{+\infty} h(t-\tau)\, a_0 d\tau + \sum_{n=1}^{+\infty} \int_{-\infty}^{+\infty} a_n h(t-\tau) [x(t)]^n\, d\tau \\ &= h_0 + h_1(t) * x(t) + h_2(t) * [x(t)]^2 \\ &+ h_3(t) * [x(t)]^3 + \ldots \end{aligned} \tag{4}$$

where $w(t)$ is the output of the non-linear purely algebraic part and is therefore substituted with $a_0 + \sum_{n=1}^{+\infty} a_n [x(t)]^n$, where $x(t)$ is the input of the whole system, and the $h(t-\tau)\, a_n$ terms have been substituted with $h_n(t-\tau)$.

The Wiener-Hammerstein model [9, 10] represents the system as the series of a linear time-invariant system with memory, a non-linear time-invariant system without memory and another linear time-invariant system with memory. Its output is represented by the following equation:

$$\begin{aligned} y(t) &= \int_{-\infty}^{+\infty} \beta(t-\tau)\, k(\tau)\, d\tau \\ &= \int_{-\infty}^{+\infty} \left\{ \beta(t-\tau) \left\{ a_0 + \sum_{n=1}^{+\infty} a_n [\omega(t)]^n \right\} \right\} d\tau \\ &= \int_{-\infty}^{+\infty} \left\{ \beta(t-\tau) \left\{ a_0 + \sum_{n=1}^{+\infty} a_n \right. \right. \\ &\left. \left. \times \left[\int_{-\infty}^{+\infty} \alpha h(t-\gamma)\, x(t)\, d\gamma \right]^n \right\} \right\} d\tau \end{aligned} \tag{5}$$

where $x(t)$ is the input of the whole system, $\alpha(t)$ is the IR of the first LTI system of the chain, $\alpha_1, \alpha_2, \ldots \alpha_i, \ldots \alpha_n$ are the terms of the Taylor series that describe the non-linear time-invariant without memory system and $\beta(t)$ is the impulse response of the downstream LTI system.

The Hammerstein model is a particular case of the general Volterra model. Considering its equation, we can see that it equals a Volterra series in which, for each kernel, only the values for $\tau_1 = \tau_2 = \ldots = \tau_n$ differ from zero, and therefore, even the kernels of order higher than two can be considered two-dimensional; this particular case of the Volterra series model is called the Diagonal Volterra model.

The Diagonal Volterra model is described by the following equation, where the $h_n(t)$ terms are the kernels of the

system and x(t) the input:

$$y(t) = h_0(t) + h_1(t) * x(t) + h_2(t) * [x(t)]^2 + h_3(t) * [x(t)]^3 + \ldots \quad (6)$$

The discrete version of (6), assuming that all of the h_n(t) are functions defined for $t \in (0, +\infty)$ and the input signal x(t) is zero prior to t = 0 becomes:

$$y[n] = h_0 + \sum_{i=0}^{M-1} h_1[i]\, x[n-i] + \sum_{i=0}^{M-1} h_2[i]\, x^2[n-i] + \sum_{i=0}^{M-1} h_3[i]\, x^3[n-i] + \ldots \quad (7)$$

Because a real non-linear system could have different characteristics of non-linearity, the method herein described could accurately emulate these systems, which physically behave very similarly to the Hammerstein model. This is possible by means of an existing measuring technique that is able to precisely produce the kernels [4], [11].

2 MEASUREMENT OF THE SYSTEM AND KERNELS CALCULUS

Many measurement methods have been developed to test systems and determine their features. They are based on the hypothesis that the system being tested is linear and time-invariant (LTI) and linear shift invariant (LSI) if the technique requires digital signal processing and the time domain then becomes discrete. The more the system differs from this assumption, for example, if it is non-linear, the more the impulse response obtained from the measurements will differ from the real one. Some methods are widely used to capture IRs. Among them are the pistol shot, TDS and MLS techniques, and sine sweeps.

A sine sweep (SS) is simply a sine with a frequency that varies from f_0 to f_1 in T seconds. Its equation is

$$SS(t) = A \sin[2\pi g(t)]\,; A \in \Re \quad (11)$$

where $g(t)$ is the function that controls the sweep. There are no constraints on the $g(t)$ function, and theoretically it could be chosen *ad libitum*. However, two kinds of functions are typically used: linear and exponential.

As the time derivative of $g(t)$ provides the equation for instantaneous values of frequency, it is easy to obtain the sine sweeps of interest.

2.1 Linear Sine Sweep (LSS)

We want to linearly sweep frequencies from f_0 to f_1 in T seconds; therefore, if we call f(t) the instantaneous frequency function, its equation will be

$$f(t) = f_0 + \frac{f_1 - f_0}{T} t \quad (12)$$

The instantaneous frequency function is the time derivative of $g(t)$. To obtain this, we must integrate.

$$g(t) = \int f(t)\, dt = \int \left(f_0 + \frac{f_1 - f_0}{T} t \right) dt = \frac{f_1 - f_0}{2T} t^2 + f_0 t + \vartheta \quad (13)$$

The linear sine sweep will be:

$$LSS(t) = A \sin\left[2\pi \left(\frac{f_1 - f_0}{2T} t^2 \right) + 2\pi f_0 t \right] \quad (14)$$

$$LSS(t) = A \sin\left[\left(\frac{\omega_1 - \omega_0}{2T} t^2 \right) + \omega_0 t \right] \quad (15)$$

where, without loss of generality, we have set $\vartheta = 0$.

2.2 Exponential Sine Sweep (ESS)

In an exponential sine sweep, f(t) will be of the form:

$$f(t) = e^{\gamma_0 + \frac{\gamma_1 - \gamma_0}{T} t} = e^{\gamma_0} e^{(\gamma_1 - \gamma_0)\frac{t}{T}} \quad (16)$$

If we set

$$f(0) = e^{\gamma_0} = f_0;$$
$$f(T) = e^{\gamma_0} e^{(\gamma_1 - \gamma_0)} = f_0 e^{(\gamma_1 - \gamma_0)} = f_1 \quad (17)$$

and therefore:

$$e^{(\gamma_1 - \gamma_0)} = \frac{f_1}{f_0} = e^{\ln\left(\frac{f_1}{f_0}\right)} \Rightarrow \gamma_1 - \gamma_0 = \ln\left(\frac{f_1}{f_0}\right) \quad (18)$$

we find $f(t)$ as the following equation:

$$f(t) = f_0 e^{\frac{t}{T} \ln\left(\frac{f_1}{f_0}\right)} \quad (19)$$

Eq. (19), once integrated, will produce $g(t)$ in the exponential case:

$$g(t) = \frac{f_0 T}{\ln\left(\frac{f_1}{f_0}\right)} e^{\frac{t}{T} \ln\left(\frac{f_1}{f_0}\right)} + \vartheta \quad (20)$$

As before, if we set

$$\vartheta = -\frac{f_0 T}{\ln\left(\frac{f_1}{f_0}\right)} \quad (21)$$

we finally obtain the following:

$$ESS(t) = A \sin \frac{\omega_0 T}{\ln\left(\frac{\omega_1}{\omega_0}\right)} \left[e^{\frac{t}{T} \ln\left(\frac{\omega_1}{\omega_0}\right)} - 1 \right] \quad (22)$$

with $\omega_0 = 2\pi f_0$ and $\omega_1 = 2\pi f_1$. One more evolution of the ESS technique is represented by the silence sweep [12], which is still based on the ESS but could excite all frequencies simultaneously.

Other measurement techniques that are able to calculate distortion in signals have been recently developed but cannot be used to predict output signals [13].

2.3 Sine Sweep Test

The two aforementioned sine sweeps could be used to calculate the impulse response of a system. Figs. 2a and 2b represent the input and the output of the system excited by a linear sine sweep. The non-linear behavior of the system

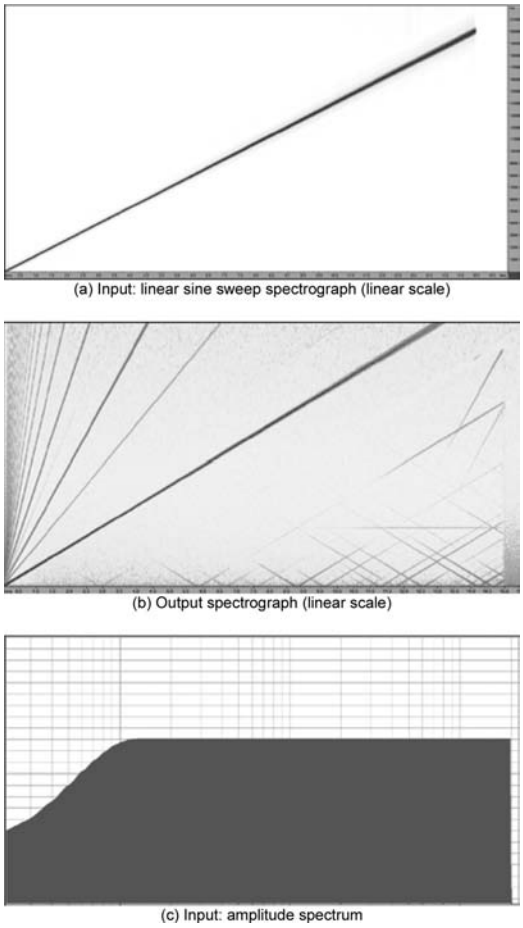


Fig. 2. Linear sine sweep test signal.

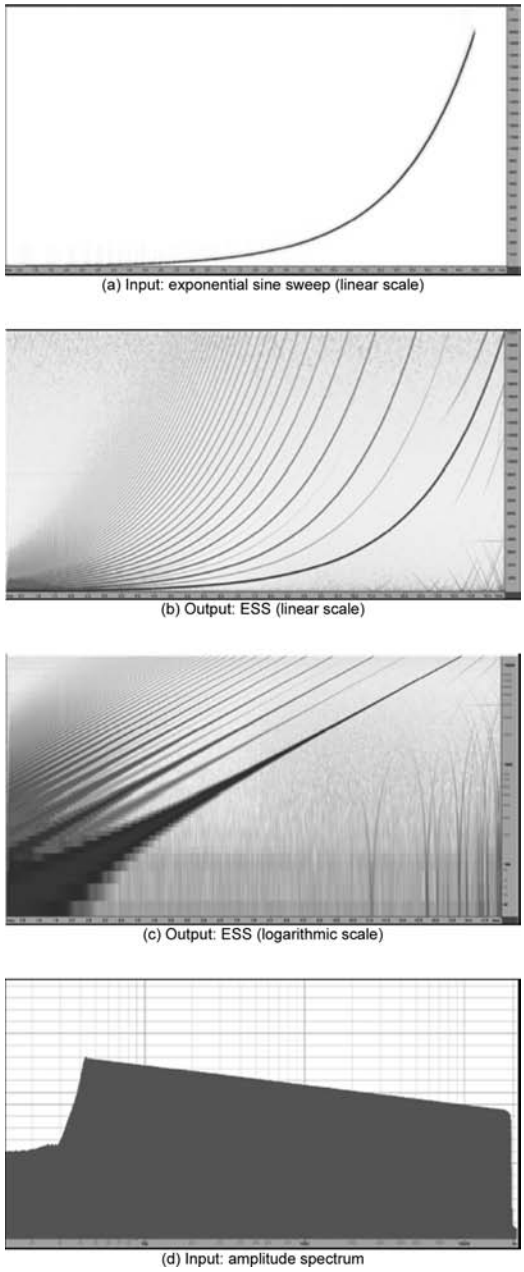


Fig. 3. Linear sine sweep test signal.

is underlined by all of the lines, which are non-parallel to the sine sweep in the output spectrograph; this spectrograph represents its distortion harmonics. In Fig. 2c, the amplitude spectrum of a linear sine sweep from 40 Hz to 20 kHz at 1/2 full scale is represented. On the other hand, Figs. 3b and 3c represent the output of the same non-linear system if the input were an exponential sine sweep (Fig. 3a) in a linear and logarithmic scale. It is of interest to notice that, visually, the logarithmic scale transforms the exponential curves in the linear scale output spectrograph to parallel lines.

2.4 Deconvolution of Sine Sweeps

Given a signal $x(t)$, we could define $\overline{x(t)}$ as the inverse of $x(t)$ if it is true that

$$x(t) * \overline{x(t)} = \delta(t - t_0), t_0 \geq 0 \tag{23}$$

A similar signal could allow for the derivation of the IR of an LTI system simply by convolving between it and the system's output. A powerful feature of the sine sweep, SS, is that its inverse approximately equals the time reversed signal:

$$SS(t) * SS(T - t) \cong \delta(t - t_0), 0 \leq t \leq T \tag{24}$$

where the approximation is only due to the fact that in practice the sine sweep covers only a limited range of frequencies, typically [40 Hz, 20 kHz]. The more the frequency range tends to [0 Hz, +μ Hz], the more the convolution tends to $\delta(t - t_0)$. It will not be considered erroneous to assume $SS(t) * SS(T - t) = \delta(t - t_0)$, if $f_0 = 40$ Hz and $f_1 = 20$ kHz, because this is the average range of perceptible frequencies for human beings.

By comparing the spectrographs in Fig. 4, a geometrical rule can be found to explain the deconvolution process. Every point of the first spectrograph is shifted to the right by convolution with its inverse. The amount of translation of every point equals the time at which the frequency of the "to be translated" point is reached in the inverse signal.

In this manner, all of the swept frequencies are compacted into a short interval, approximately a perfect impulse.

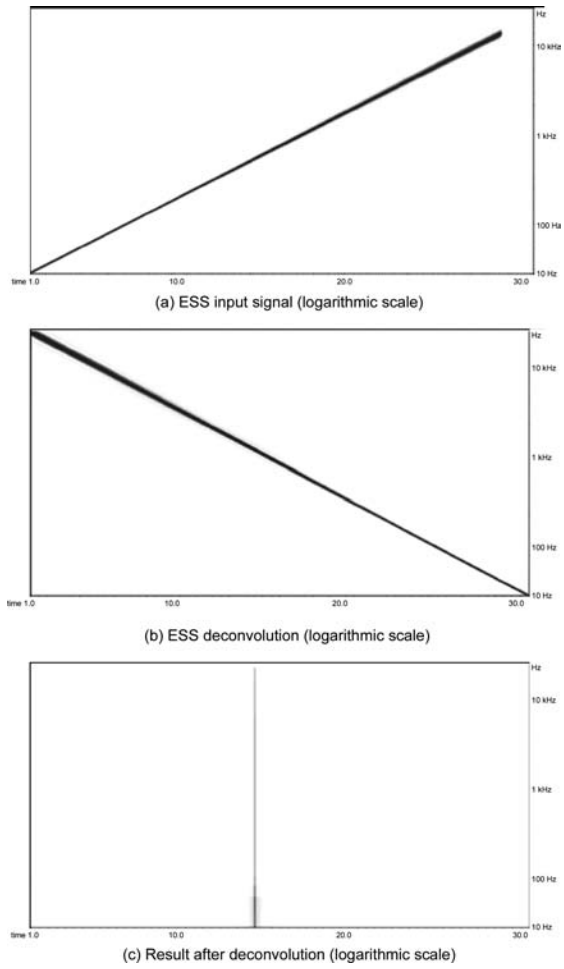


Fig. 4. Deconvolution effects.

2.4.1 Linear Sine Sweep Deconvolution

Figs. 5a and 5b show the spectrogram of the linear sine sweep LSS and of a non-linear system's output after passing the LSS. Figs. 5c and 5d show the results of the convolution between the output of a non-linear system stimulated with a linear sine sweep and its inverse. The new lines in Fig. 5b are due to harmonic distortions and represent the non-linearity of the system. These new lines after the deconvolution become non-vertical lines, as shown in Fig. 5c.

Because they all mix together after the deconvolution, they cannot be used to investigate non-linear properties, as depicted in Fig. 5d. Only an approximation of the linear part of a non-linear system can be extracted through the use of a linear sine sweep as a test signal.

2.4.2 Exponential Sine Sweep Deconvolution

By using an ESS as a test signal, the non-linear system adds new frequencies, which are parallel lines, and shown in Fig. 6a, whereas the same non-linear system adds new frequency to a LSS test signal that are not parallel. The spectrogram of the deconvolution of the ESS test signal shown in Fig. 6c, differs from the spectrogram of the deconvolution of the LSS shown in Fig. 5c.

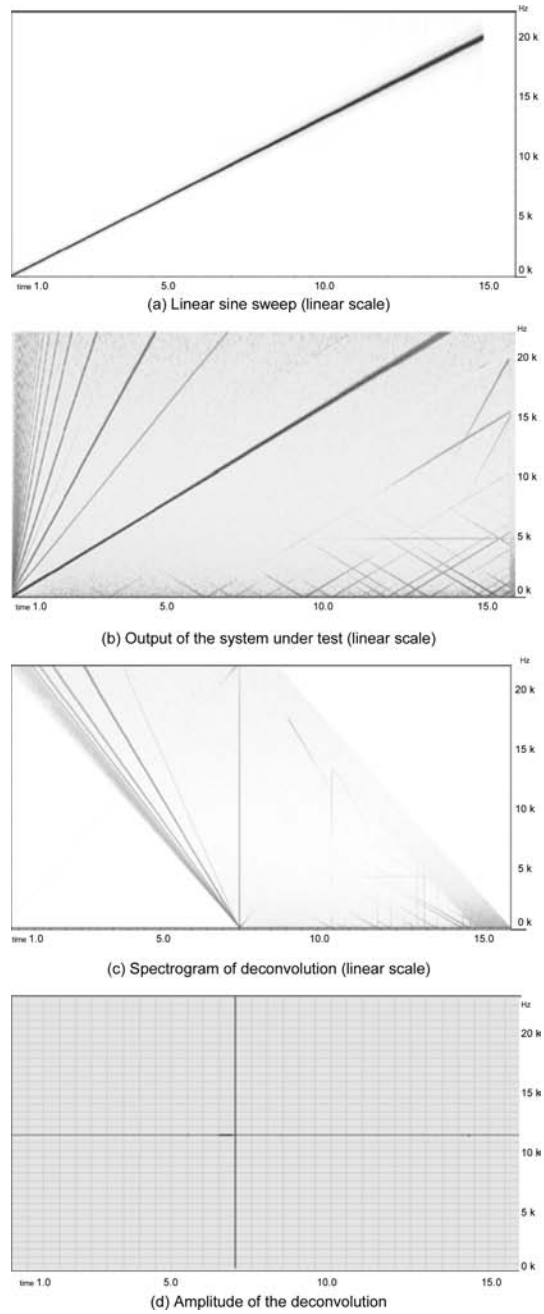


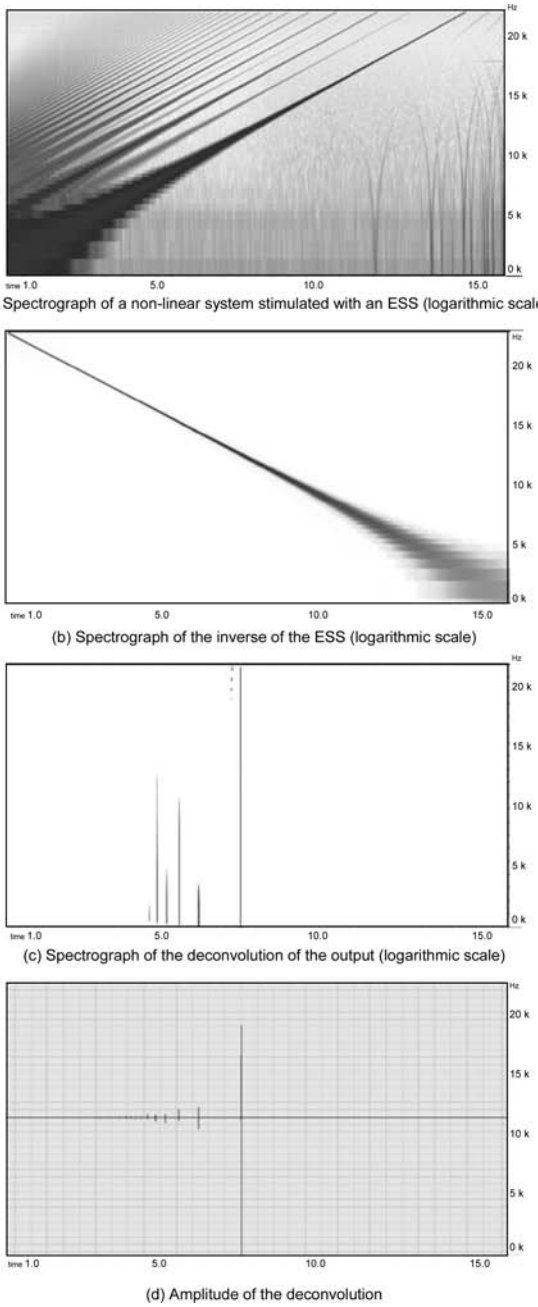
Fig. 5. Linear sine sweep test deconvolution.

In the ESS test signal, the lines are parallel, and in time domain they result in separated impulses. In the LSS test signal the lines are not parallel, and in time domain they cannot be separated.

2.5 Kernel Calculus

It is possible to determine the relationship between these vertical lines and the kernels of the diagonal Volterra model of the system being tested: first, it is necessary to isolate every line. Fig. 7 shows the theoretical extension for $t < 0$ of a spectrograph of the output of a non-linear system stimulated with an ESS.

The gaps $\Delta t_2, \Delta t_3, \ldots \Delta t_n$ between these lines remain unchanged after the deconvolution. Every Δt_n represents



(a) Spectrograph of a non-linear system stimulated with an ESS (logarithmic scale)

(b) Spectrograph of the inverse of the ESS (logarithmic scale)

(c) Spectrograph of the deconvolution of the output (logarithmic scale)

(d) Amplitude of the deconvolution

Fig. 6. Exponential sine sweep test deconvolution.

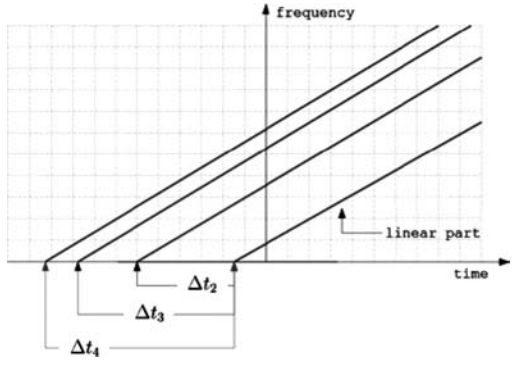


Fig. 7. Time delay gaps in ESS.

Table 1. Trigonometric identities

$sin^2(\omega_0 t)$	$\frac{1}{2} - \frac{1}{2} cos(2\omega_0 t)$
$sin^3(\omega_0 t)$	$\frac{3}{4} sin(\omega_0 t) - \frac{1}{4} sin(3\omega_0 t)$
$sin^4(\omega_0 t)$	$\frac{3}{8} - \frac{1}{2} cos(2\omega_0 t) + \frac{1}{8} cos(4\omega_0 t)$
$sin^5(\omega_0 t)$	$\frac{5}{8} sin(\omega_0 t) - \frac{5}{16} sin(3\omega_0 t) + \frac{1}{16} sin(5\omega_0 t)$
$sin(\omega_0 t)\, sin(\omega_1 t)$	$\frac{1}{2} cos[(\omega_0 - \omega_1) t] - \frac{1}{2} cos[(\omega_0 + \omega_1) t]$

the delay the sine sweep spends in duplicate, triplicate, etc., multiplied by n times its instantaneous frequency $f(t_*), \forall t_* > 0$.

Recalling that the instantaneous frequency function for ESS is (20) and considering the trigonometric identities in Table 1, the time gaps can be calculated as follows:

$$\begin{aligned} f(t_* + \Delta t_N) &= N f(t_*)\ f_0 e^{\frac{t_* + \Delta t_N}{T} \ln\left(\frac{f_1}{f_0}\right)} \\ &= N f_0 e^{\frac{t_*}{T} \ln\left(\frac{f_1}{f_0}\right)}\ e^{\frac{\Delta t_N}{T} \ln\left(\frac{f_1}{f_0}\right)} \\ &= N\ e^{\frac{\Delta t_N}{T} \ln\left(\frac{f_1}{f_0}\right)} \\ &= e^{\ln N\ \frac{\Delta t_N}{T} \ln\left(\frac{f_1}{f_0}\right) = \ln N} \end{aligned} \tag{25}$$

and finally

$$\Delta t_N = \ln N \frac{T}{\ln\left(\frac{f_1}{f_0}\right)}, \qquad with \quad N \geq 2 \tag{26}$$

For our purposes, it is useful to know the absolute starting time of every impulse. Knowing that the last impulse starts at $t = T$, it is possible to use (26) and obtain:

$$t_N = T - \Delta t_N = T \left[1 - \frac{\ln N}{\ln\left(\frac{f_1}{f_0}\right)} \right] \quad with \quad N \geq 1 \tag{27}$$

It is then possible to isolate the impulses and find a relationship between them and the Volterra kernels. Calling $ss[\omega(t)]$, the exponential sine sweep of length T seconds from f_0 to f_1, it is possible to rewrite (6), assuming that

$$x(t) = \alpha ss[\omega(t)], \alpha \in \Re \tag{28}$$

and thereby obtain

$$\begin{aligned} y(t) = h_0 + h_1(t) * \alpha\, ss[\omega(t)] + h_2(t) * \alpha^2\, ss^2[\omega(t)] \\ + \ldots + h_n(t) * \alpha^n\, ss^n[\omega(t)] \end{aligned} \tag{29}$$

Using the identities in Table 1, it is possible to write (29) as follows

$$\begin{aligned} y(t) \approx\ & h_0 + \alpha h_1(t) * ss[\omega(t)] \\ & + \alpha^2 h_2(t) * \left\{ \frac{1}{2} - \frac{1}{2} cs[2\omega(t)] \right\} \\ & + \alpha^3 h_3(t) * \left\{ \frac{3}{4} ss[\omega(t)] - \frac{1}{4} ss[3\omega(t)] \right\} \\ & + \alpha^4 h_4(t) * \left\{ \frac{3}{8} - \frac{1}{2} cs[2\omega(t)] + \frac{1}{8} cs[4\omega(t)] \right\} \\ & + \alpha^5 h_5(t) * \left\{ \frac{5}{8} ss[\omega(t)] - \frac{5}{16} ss[3\omega(t)] \right. \\ & \left. + \frac{1}{16} ss[5\omega(t)] \right\} \end{aligned} \tag{30}$$

where the series expansion is truncated at the fifth order. In (32), the expression $cs\,[\omega\,(t)]$ indicates the *cosine sweep*, a signal defined as a sine sweep with a phase delay of $\pi/2$. Equation (30) can be rewritten as follows:

$$\begin{aligned} y\,(t) \;\approx\; & \left[h_0 + \alpha^2 h_2\,(t) * \frac{1}{2} + \alpha^4 h_4\,(t) * \frac{3}{8}\right] \\ & + \left[\alpha h_1\,(t) + \frac{3}{4}\alpha^3 h_3\,(t) + \frac{5}{8}\alpha^5 h_5\,(t)\right] * ss\,[\omega\,(t)] \\ & + \left[-\frac{1}{2}\alpha^2 h_2\,(t) - \frac{1}{2}\alpha^4 h_4\,(t)\right] * cs\,[2\omega\,(t)] \\ & + \left[-\frac{1}{4}\alpha^3 h_3\,(t) - \frac{5}{16}\alpha^5 h_5\,(t)\right] * ss\,[3\omega\,(t)] \\ & + \left[\frac{1}{8}\alpha^4 h_4\,(t)\right] * cs\,[4\omega\,(t)] \\ & + \left[\frac{1}{16}\alpha^5 h_5\,(t)\right] * ss\,[5\omega\,(t)] \end{aligned} \tag{31}$$

Convolving with $\overline{ss\,[\omega\,(t)]}$, produces

$$\begin{aligned} y\,(t) * \overline{ss\,[\omega\,(t)]} \;\approx\; & A\,(t) * \overline{ss\,[\omega\,(t)]} \\ & + B\,(t) * ss\,[\omega\,(t)] * \overline{ss\,[\omega\,(t)]} \\ & + C\,(t) * cs\,[2\omega\,(t)] * \overline{ss\,[\omega\,(t)]} \\ & + D\,(t) * ss\,[3\omega\,(t)] * \overline{ss\,[\omega\,(t)]} \\ & + E\,(t) * cs\,[4\omega\,(t)] * \overline{ss\,[\omega\,(t)]} \\ & + F\,(t) * ss\,[5\omega\,(t)] * \overline{ss\,[\omega\,(t)]} \end{aligned} \tag{32}$$

where we have considered the following equivalences

$$\begin{aligned} A\,(t) &= h_0 + \alpha^2 h_2\,(t) * \frac{1}{2} + \alpha^4 h_4\,(t) * \frac{3}{8} \\ B\,(t) &= \alpha h_1\,(t) + \frac{3}{4}\alpha^3 h_3\,(t) + \frac{5}{8}\alpha^5 h_5\,(t) \\ C\,(t) &= -\frac{1}{2}\alpha^2 h_2\,(t) - \frac{1}{2}\alpha^4 h_4\,(t) \\ D\,(t) &= -\frac{1}{4}\alpha^3 h_3\,(t) - \frac{5}{16}\alpha^5 h_5\,(t) \\ E\,(t) &= \frac{1}{8}\alpha^4 h_4\,(t) \;\; F\,(t) = \frac{1}{16}\alpha^5 h_5\,(t) \end{aligned} \tag{33}$$

As it is necessary to Fourier transform the equation (32), it is also necessary to note that

$$\begin{aligned} A\,(t) &= constant\; h_0 = constant\; \alpha^2 h_2\,(t) * \frac{1}{2} \\ &= \int_{-\infty}^{+\infty} \alpha^2 h_2\,(t)\,\frac{1}{2}dt = \frac{1}{2}\alpha^2 \int_{-\infty}^{+\infty} h_2\,(t)\,dt \\ &= constant\; \alpha^4 h_4\,(t) * \frac{3}{8} = \int_{-\infty}^{+\infty} \alpha^4 h_4\,(t)\,\frac{3}{8}dt \\ &= \frac{3}{8}\alpha^4 \int_{-\infty}^{+\infty} h_4\,(t)\,dt = constant \end{aligned} \tag{34}$$

where:

- $A\,(t) * \overline{ss\,[\omega\,(t)]}$ is also constant and represents a DC offset that could be removed by applying a high-pass filter. This offset is not relevant to this calculus, so there is no need to consider it;
- $X\,(\omega) = \Im\,[ss\,(\omega\,(t))]$ and $\overline{X\,(\omega)}$ define the inverse ESS such that $\Im^{-1}\left[X\,(\omega)\,\overline{X\,(\omega)}\right] = \delta\,(t - t_0)$;
- If $\Im\,[g\,(\omega\,(t))] = X\,(\omega)$, then $\Im\,[g\,(a\omega\,(t))] = \frac{X\left(\frac{\omega}{a}\right)}{|a|}$;
- All of the signals are real; therefore, it is sufficient to consider the positive side of the spectrum and leave the Hermitian symmetry properties (i.e., even amplitude spectrum and odd phase spectrum) to manage the negative one. This makes it possible to affirm:

$$\Im\,[cs\,(\omega\,(t))] = j\Im\,[ss\,(\omega\,(t))] = jX\,(\omega) \tag{35}$$

The term $j = e^{j\frac{\pi}{2}}$ delays every component of the ESS by $\pi/2$ degrees. As ESS only has one frequency component instant by instant, the product between j and $X\,(\omega)$ transforms sine to cosine and the sine-sweep to a cosine-sweep. Fourier transforming (34) and deleting the DC offset produces the following:

$$\begin{aligned} \Im\left[y\,(t) * \overline{ss\,[\omega\,(t)]}\right] \;\approx\; & \overline{B}\,(\omega)\,X\,(\omega)\,\overline{X\,(\omega)} \\ & + \overline{C}\,(\omega)\,j\,\frac{X\left(\frac{\omega}{2}\right)}{|2|}\overline{X\,(\omega)} \\ & + \overline{D}\,(\omega)\,\frac{X\left(\frac{\omega}{3}\right)}{|3|}\overline{X\,(\omega)} \\ & + \overline{E}\,(\omega)\,j\,\frac{X\left(\frac{\omega}{4}\right)}{|4|}\overline{X\,(\omega)} \\ & + \overline{F}\,(\omega)\,\frac{X\left(\frac{\omega}{5}\right)}{|5|}\overline{X\,(\omega)} \end{aligned} \tag{36}$$

where $\overline{B}\,(\omega), \overline{C}\,(\omega), \ldots \overline{F}\,(\omega)$ represent the Fourier transform of $B\,(t), C\,(t), \ldots F\,(t)$. As

$$\begin{aligned} \Im^{-1}\left[X\,(\omega)\,\overline{X\,(\omega)}\right] &\cong \delta\,(t - \tau_1) \\ \Im^{-1}\left[\frac{X\left(\frac{\omega}{2}\right)}{|2|}\overline{X\,(\omega)}\right] &\cong -\delta\,(t - \tau_2) \\ \Im^{-1}\left[\frac{X\left(\frac{\omega}{3}\right)}{|3|}\overline{X\,(\omega)}\right] &\cong \delta\,(t - \tau_3) \\ \Im^{-1}\left[\frac{X\left(\frac{\omega}{4}\right)}{|4|}\overline{X\,(\omega)}\right] &\cong -\delta\,(t - \tau_4) \\ \Im^{-1}\left[\frac{X\left(\frac{\omega}{5}\right)}{|5|}\overline{X\,(\omega)}\right] &\cong \delta\,(t - \tau_5) \end{aligned} \tag{37}$$

the inverse-transform of (36) will be:

$$\begin{aligned} deconv\,(t) \;\cong\; & \Im^{-1}\left[\overline{B}\,(\omega)\right] * \delta\,(t - \tau_1) \\ & + \Im^{-1}\left[-j\overline{C}\,(\omega)\right] * \delta\,(t - \tau_2) \\ & + \Im^{-1}\left[\overline{D}\,(\omega)\right] * \delta\,(t - \tau_3) \\ & + \Im^{-1}\left[-j\overline{E}\,(\omega)\right] * \delta\,(t - \tau_4) \\ & + \Im^{-1}\left[\overline{F}\,(\omega)\right] * \delta\,(t - \tau_5) \end{aligned} \tag{38}$$

that can be written as follows

$$\begin{aligned} deconv\,(t) \;\cong\; & k_1\,(t - \tau_1) + k_2\,(t - \tau_2) + k_3\,(t - \tau_3) \\ & + k_4\,(t - \tau_4) + k_5\,(t - \tau_5) \end{aligned} \tag{39}$$

The addenda in (39) are, from right to left, the first five impulses in Fig. 6d; their starting point is described by (27). Once they have been isolated, it is necessary to anti-transform the $H_n(\omega)$ to obtain the Volterra kernels $h_n(t)$ using the following:

$$K_1(\omega) = \overline{B}(\omega) = \alpha H_1(\omega) + \frac{3}{4}\alpha^3 H_3(\omega) + \frac{5}{8}\alpha^5 H_5(\omega)$$
$$K_2(\omega) = -j\overline{C}(\omega) = -j\left[-\frac{1}{2}\alpha^2 H_2(\omega) - \frac{1}{2}\alpha^4 H_4(\omega)\right]$$
$$K_3(\omega) = \overline{D}(\omega) = -\frac{1}{4}\alpha^3 H_3(\omega) - \frac{5}{16}\alpha^5 H_5(\omega)$$
$$K_4(\omega) = -j\overline{E}(\omega) = -j\frac{1}{8}\alpha^4 H_4(\omega)$$
$$K_5(\omega) = \overline{F}(\omega) = \frac{1}{16}\alpha^5 H_5(\omega) \quad (40)$$

that can be rewritten as:

$$H_1(\omega) = \frac{K_1(\omega) + 3K_3(\omega) + 5K_5(\omega)}{\alpha}$$
$$H_2(\omega) = \frac{-2jK_2(\omega) - 8jK_4(\omega)}{\alpha^2}$$
$$H_3(\omega) = \frac{-4K_3(\omega) - 20K_5(\omega)}{\alpha^3}$$
$$H_4(\omega) = \frac{8jK_4(\omega)}{\alpha^4}$$
$$H_5(\omega) = \frac{16K_5(\omega)}{\alpha^5} \quad (41)$$

Once the kernels of the system are known, it is possible to calculate the system's output using (6).

It should be noted that the approximations in (37) become worse the higher the degree of non-linearity becomes. However, assuming that all of these terms are similar, both to each other and to $\delta(t - \tau_n)$, makes it possible to easily solve the operations, and in practice, even with these kind of known errors, the results will be satisfactory.

Figs. 8, 9, and 10 show how this approximation worsens when the system becomes more non-linear, and therefore the higher kernels, where the mismatch increases, cannot be neglected.

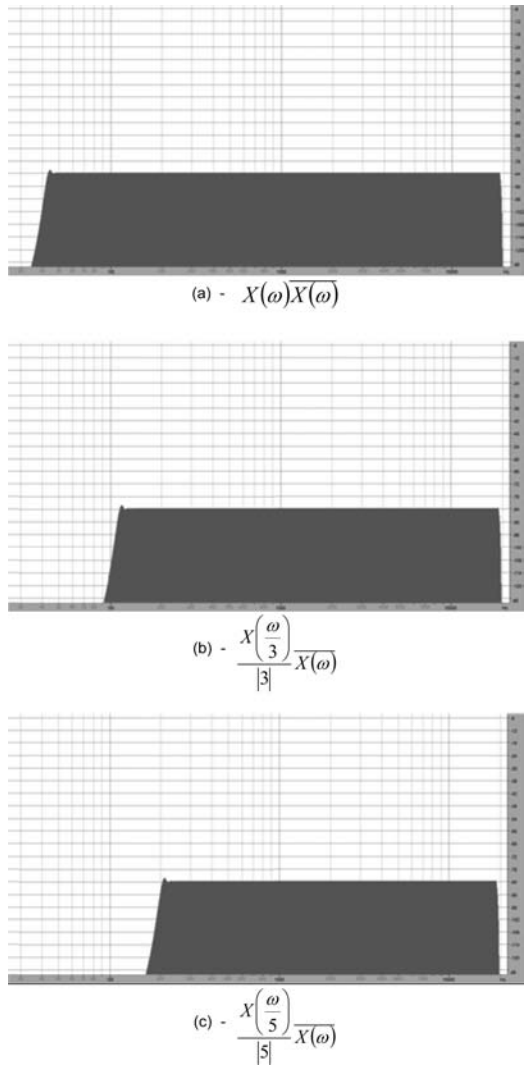


Fig. 8. Worsening of the assumption in (39) as the degree of nonlinearity rises.

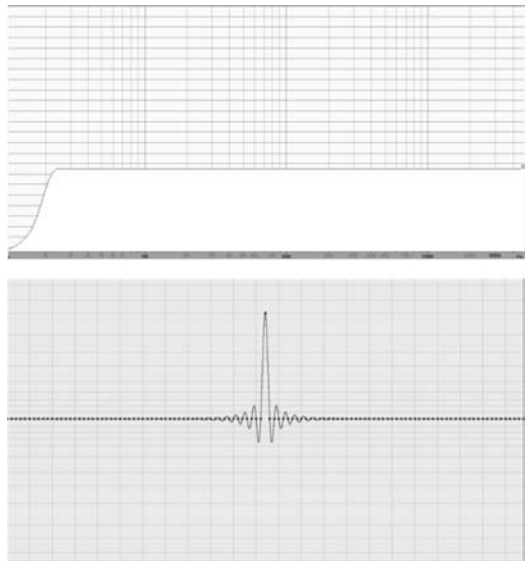


Fig. 9. Delta Dirac from the first equation of (37).

2.6 Phase Mismatch

The harmonic components of the sine sweeps have a phase distortion that provokes an incorrect emulation of the non-linear system. This could be avoided by modifying the ESS signal as proposed in [14]. Considering Eq. (37), the first equation states that convolving a sine sweep, e.g., of 15 s from 20 Hz to 48 kHz, with its inverse results in a waveform of the Dirac Delta type. This is always verified as shown in Fig. 9b. The amplitude diagram of the frequency response shows a flat spectrum, and the waveform also has a shape that can be compared to that of a Dirac Delta.

The second equation in Eq. (37) states convolving the aforementioned inverse sine sweep with a sine sweep of 15 seconds between 40 Hz and 96 kHz should equally obtain the Dirac Delta. Fig. 10b shows how this expectation failed to be met in this case. The problem is not caused by aliasing limitations but is due to a phase distortion of the harmonic components of the signal. To solve these problems, it is nec-

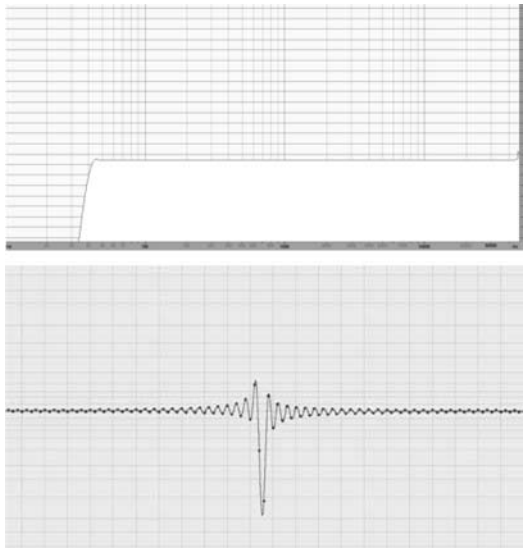


Fig. 10. Delta Dirac from the second equation of (37).

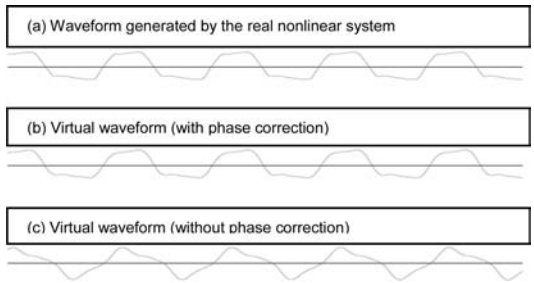


Fig. 11. Emulation of a nonlinear system:(a) real waveform, (b) virtual waveform without phase correction, (c) virtual waveform with phase correction.

essary to calculate four FIR filters, which, once they have been applied to the Dirac Deltas, are able to "re-align" the phase, returning the signal to the type of shapes shown in Fig. 9. The FIR filters may be easily calculated for example by means of the Nelson/Kirkeby frequency-domain regularization [15]. Once the four corrective filters have been calculated and applied to the corresponding harmonic responses in Eq. (16), it is finally possible to correctly emulate the non-linear system, as shown in Fig. 11.

3 EXPERIMENTAL RESULTS

To test the efficiency of the theoretical formulation given above, a non-linear hardware device was tested (namely, the Ibanez Tube Screamer guitar overdrive pedal, Fig. 12). This choice, made also by other authors [15, 16, 17], was due to the natural non-linearity of this device and its fame in music production.

Fig. 13a shows the input signal for the test. A set of five exponential sine-sweeps, different in amplitude, with the following features were used: $f_0 = 40$ Hz, $f_1 = 20$ kHz, duration 15 s, amplitude = 10%, 25%, 50%, 75%, 90% full scale. This signal stimulates the device in five different regions of operation, mapping its behavior roughly across its whole dynamic.

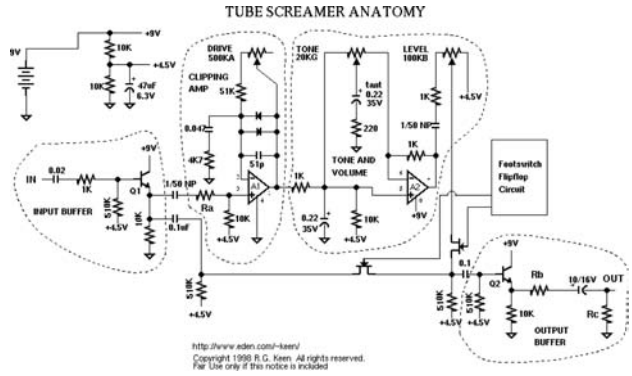


Fig. 12. Tube screamer anatomy.

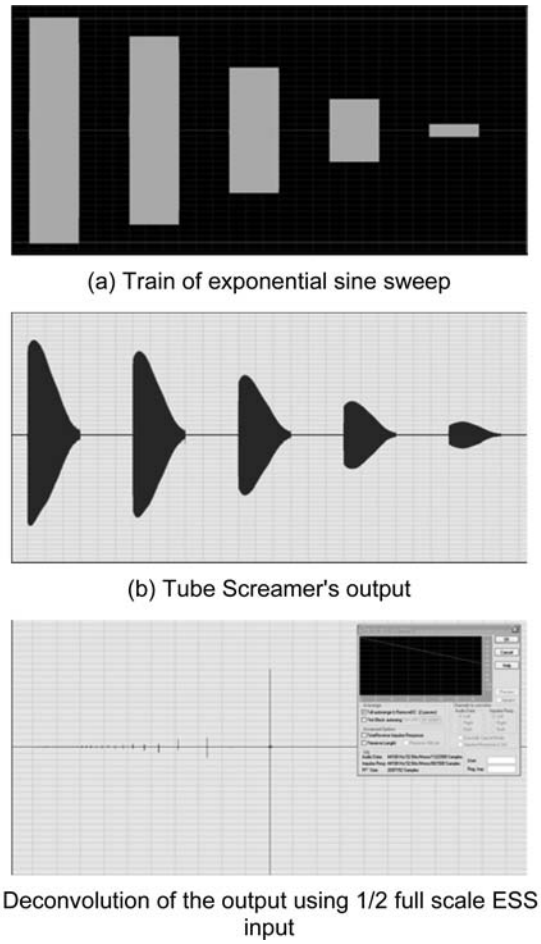


Fig. 13. Tube Screamer's test.

Fig. 13b shows the Tube Screamer's output, and Fig. 10c shows the deconvolution of the output signal with an ESS 50% full scale as input. The large number of impulses unmasks the non-linear nature of the device.

Table 2 shows the RMS values of the five exponential sine sweeps; it could be possible to calculate five sets of kernel (one set per amplitude value) and, according to the power of the input, choose which set must be used during emulation, depending on the power input of the overdrive.

Table 2. RMS value of the five ESS input signals

Amplitude	RMS
90%	0.96 dB
75%	2.54 dB
50%	6.06 dB
25%	12.08 dB
10%	26.06 dB

3.1 Numerical Simulation

A numerical code was developed to properly recreate the non-linear effect up to the tenth order. It could handle the output of the non-linear system, starting from the knowledge of the ESS utilized for the measurements, i.e., f_0, f_1, T, amplitude.

3.2 ESS Emulation

The first test consisted of emulating the output of the device if the input was the train of ESS used for the measurement, as depicted in Fig. 14.

The spectral analysis in Fig. 14 highlights the circled case in Fig. 14f as the best emulated, and Fig. 14e shows the worst one. The difference between the two cases is due to the different level of matching between the kernel set and the input file: the chosen deconvolution file and, consequently, the kernel set is relative to the 1/2 full scale of the ESS signal; therefore, the best emulation is clearly achieved with the 1/2 full scale ESS.

As the numerical code emulates a finite number of orders, it is expected that the real output has more harmonics. However, it is significant to note how well the amplitudes of the emulated harmonics match the original ones.

3.3 Music Emulation

The second test consists of the emulation of the device with a piece of music as input. The chosen song (namely "I Thank You Child," Book of Shadows, Zakk Wylde) contains a clean guitar arpeggio; this is the type of input with which the Tube Screamer has been designed to work. The mean RMS value of the input file is –11.73 dB; from Table 2, the better deconvolution file to be used is the 25% full scale ESS relative one. To estimate the results, a comparison between the real-out spectrum and the virtual-out spectrum could be misleading because too many frequencies are present. The comparison between the waveform and a simple listening test can prove the effects of the emulation.

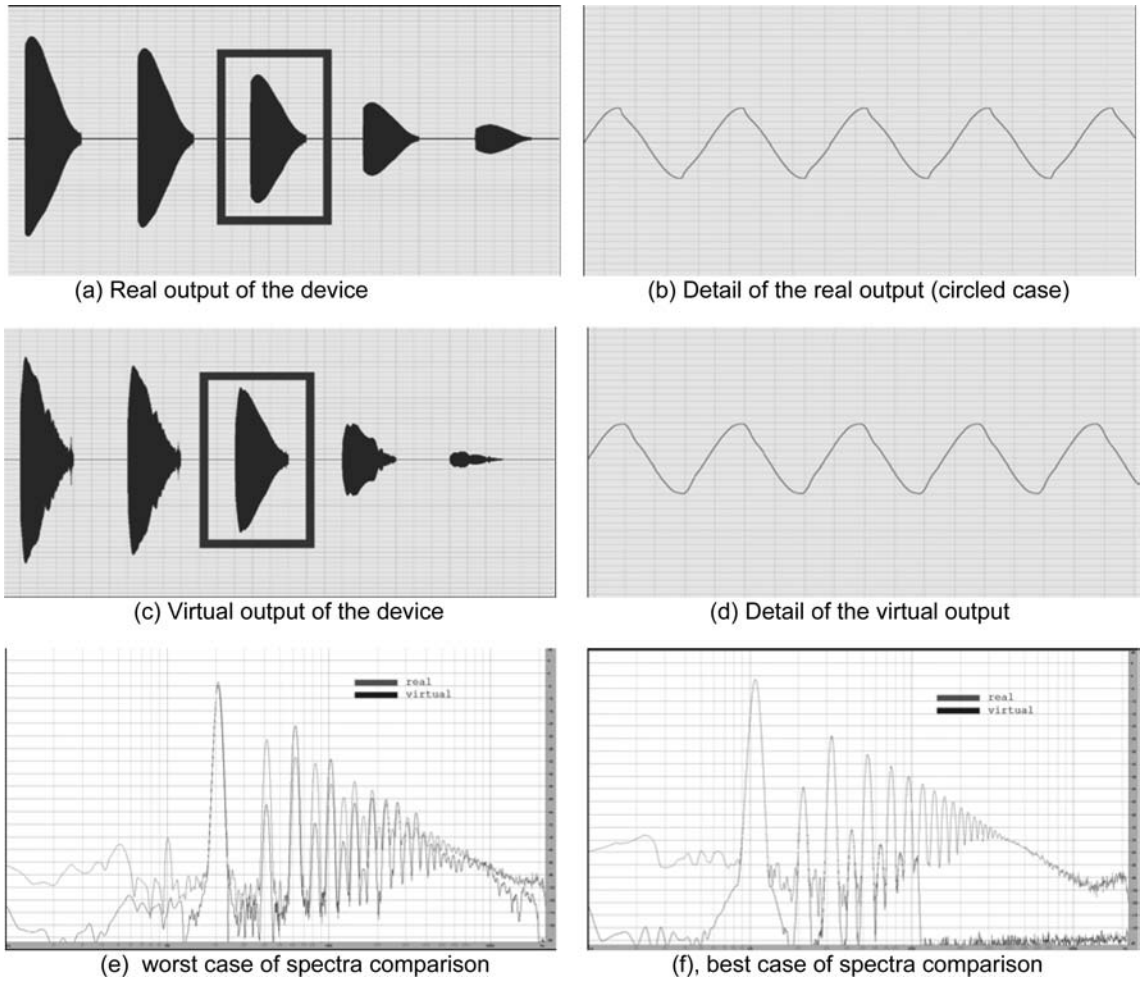
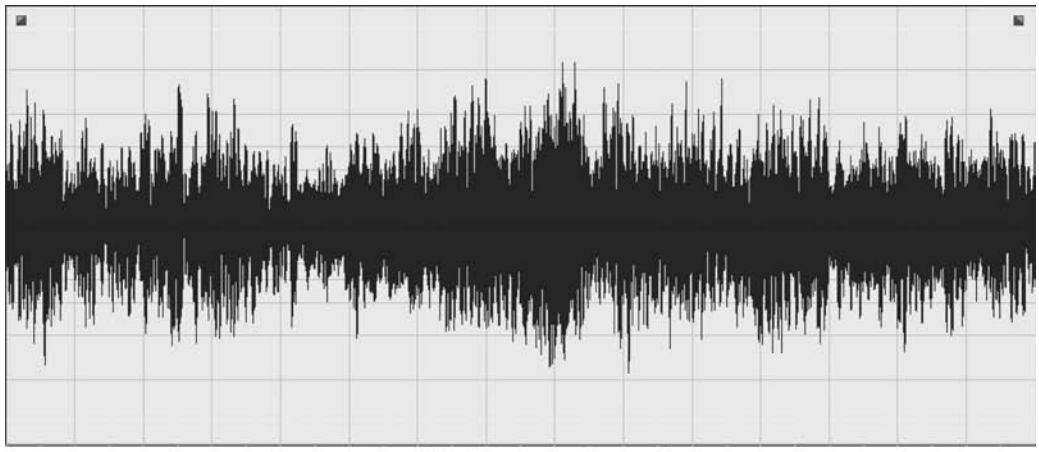
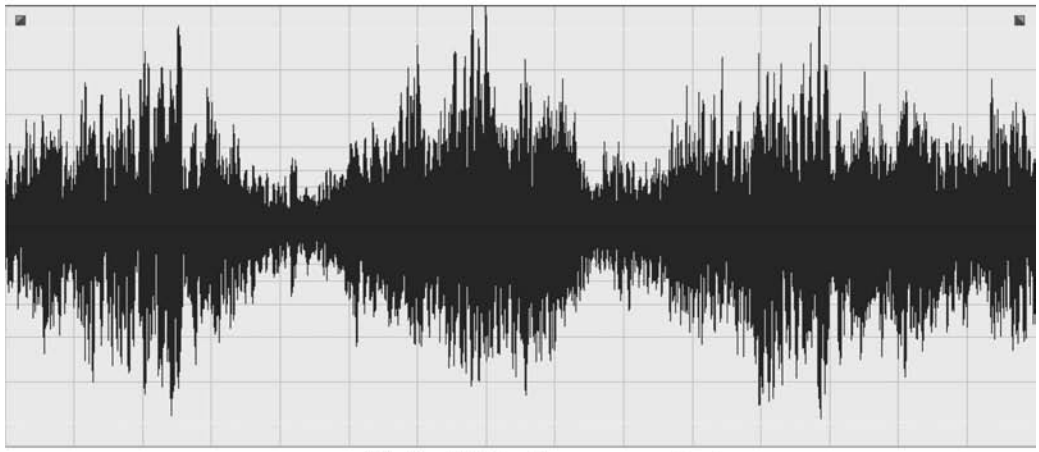


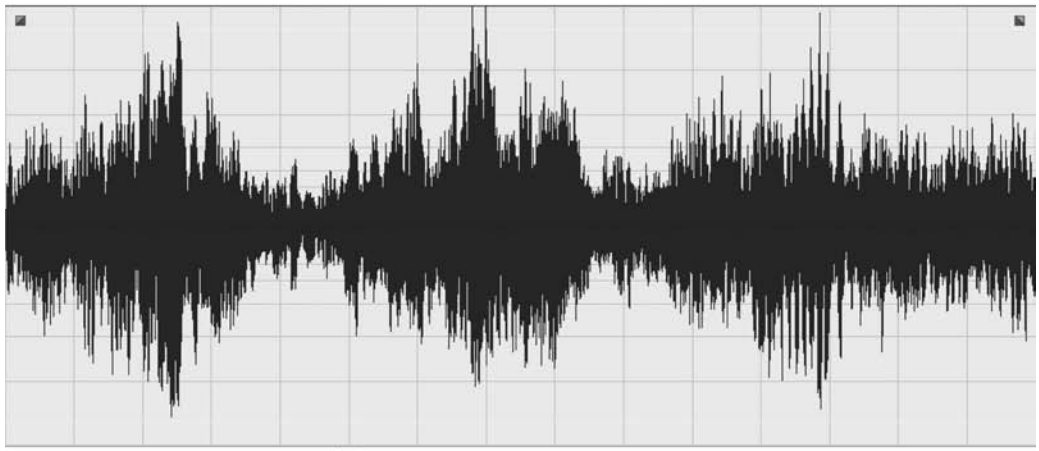
Fig. 14. Real and virtual output of the device.



(a) Clen guitar sound



(b) Real Tube Screamer output



(c) Emulated Tube Screamer output

Fig. 15. Clean guitar sound emulation.

4 CONCLUSIONS

The sine sweep measurement techniques developed by Farina and the mathematical framework developed in this paper provide a way to obtain an approximated version of the kernels to use in Eq. (7). It is known that by using a sine sweep (one frequency at time), only information about harmonic distortions at given amplitudes can be captured. Nothing is known about intermodulation distortions. They are generated by the Volterra diagonal model but are treated with the same kernels used to equalize the harmonic ones.

Even with these known limitations, sine sweeps and the Volterra diagonal model provide excellent results when adopted to emulate non-linear devices such as the Tube Screamer in which the richness of sound resides mainly in harmonic distortions.

5 ACKNOWLEDGMENTS

The author wishes to thank Andrea Venturi for his very precious help with the mathematical portion of this article. He also wishes to thank Angelo Farina for the useful discussion about the measurement methods.

6 REFERENCES

[1] P. Svensson, M. Kleiner, and B. Dalenback, "Auralization—An Overview," *J. Audio Eng. Soc.*, vol. 41, pp. 861–875 (1993 Nov.).

[2] A. Farina, A. Langhoff, and L. Tronchin "Acoustic Characterisation of 'Virtual' Musical Instruments: Using MLS Technique on Ancient Violins," *J. New Music Research*, vol. 27, pp. 359–379 (1998 Apr.).

[3] L. Farina and Tronchin, "On the 'Virtual' Reconstruction of Sound Quality of Trumpets," *Acustica United with Acta Acustica*, vol. 86, pp. 737–745 (2000 Apr.).

[4] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique," presented at the *108th Convention of the Audio Engineering Society* (2000 Feb.), convention paper 5093.

[5] T. Ogunfunmi, *Adaptive Nonlinear System Identification: The Volterra and Wiener Model Approaches* (Springer, 2007).

[6] V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations* (Dover Phoenix Editions, January 27, 2005).

[7] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems* (Krieger Publishing Co., Inc., Melbourne, FL, USA, 2006).

[8] J. Kemp and H. Primack, "Impulse Response Measurement of Nonlinear Systems: Properties of Existing Techniques and Wide Noise Sequences," *J. Audio Eng. Soc.,* vol. 59, pp. 953–963 (2011 Dec.).

[9] E. Bedrosian and S. O. Rice, "The Output Properties of Volterra Systems (Nonlinear Systems with Memory) Driven by Harmonic and Gaussian Inputs," *Proceedings of the IEEE*, vol. 59, pp. 1688–1707 (1971 Dec.).

[10] A. Y. Kibangou and G. Favier, "Wiener-Hammerstein Systems Modeling Using Diagonal Volterra Kernels Coefficients," *Signal Processing Letters, IEEE*, vol. 13, pp. 381–384 (2006 June).

[11] A. Farina, A. Bellini, and E. Armelloni, "Non-Linear Convolution: A New Approach for the Auralization of Distorting Systems," presented at the *110th Convention of the Audio Engineering Society* (2001 May), convention paper 5359.

[12] A. Farina, "Silence Sweep: A Novel Method for Measuring Electro-Acoustical Devices," presented at the *126th Convention* of the Audio Engineering Society (2009 May), convention paper 7700.

[13] S. Temme and P. Brunet, "A New Method for Measuring Distortion Using a Multitone Stimulus and Noncoherence," *J. Audio Eng. Soc.*, vol. 56, pp. 176–188 (2008 Mar.).

[14] A. Novák, L. Simon, F. Kadlec, andP. Lotton, "Nonlinear System Identification Using Exponential Swept-Sine Signal," *IEEE Instrumentation and Measurement*, vol. 59 (2010 Aug.).

[15] R. Cauduro Dias de Paiva, J. Pakarinen, and V. Välimäki, "Reduced-Complexity Modeling of High-Order Nonlinear Audio Systems Using Swept-Sine and Principal Component Analysis," *45th AES International Conference: Applications of Time-Frequency Processing in Audio* (2012 Mar.), paper number 5-4.

[16] R. Cauduro Dias De Paiva, H. Penttinen, and M. Cable, "Instrument Cables Affect the Frequency Response of Electric Guitars," *presented at the 131st Convention of the Audio Engineering Society,* (2011 Oct), convention paper 8466.

[17] L. Tronchin and A. Venturi, "The Use of Volterra Series for Simulating the Nonlinear Behavior of Musical Instruments," *Second Vienna Talks* (2010 Sept.).

THE AUTHOR

Lamberto Tronchin

Dr. Lamberto Tronchin is Associate Professor in Environmental Physics at the University of Bologna. A pianist himself, with a diploma in piano from the Conservatory of Reggio Emilia, Dr. Tronchin's principal area of research has been room acoustics, musical acoustics, 3-D auralization, and signal processing. He is the author of more than 190 papers and was Chair of the Musical Acoustics Group of the Italian Association of Acoustics from 2000 to 2008. Dr. Tronchin is a member of the Scientific Committee of the CIARM, the Inter-University Centre of Acoustics and Musical research, has chaired sessions of architectural and musical acoustics during several international symposiums, been a referee for a number of international journals. He was a visiting researcher at the University of Kobe in Japan, a visiting professor at the University of Graz in Austria, and special honored international guest at the International Workshop, "Analysis, Synthesis, and Perception of Music Signals," at Jadavpur University of Kolkata, India, in 2005. He has chaired the International Advanced Course on Musical Acoustics (IACMA), organized with the European Association of Acoustics, which was held in Bologna, in 2005. In 2008 and 2009 he gave plenary lectures at International Congresses on Acoustics in Vancouver, Prague, Bucharest, Santander, Kos, Malta, Cambridge and Paris. He designed theaters and other buildings as acoustic consultant in collaboration with several architects, among them Richard Meier and Paolo Portoghesi.