Audio Engineering Society

Convention Paper

Presented at the 121st Convention
2006 October 5–8 San Francisco, CA, USA

A New Method for Measuring Distortion using a Multitone Stimulus and Non-Coherence

Steve Temme[1] and Pascal Brunet[2]

[1] Listen, Inc., Boston, MA, 02118, USA
stemme@listeninc.com

[2] Listen, Inc., Boston, MA, 02118, USA
pbrunet@listeninc.com

ABSTRACT

A new approach for measuring distortion based on dual-channel analysis of non-coherence between a stimulus and response is presented. This method is easy to implement, provides a continuous distortion curve against frequency, and can be used with a multitone stimulus, noise, or even music.

Multitone is a desirable test signal for fast frequency response measurements and also for assessing system non-linearities. However, conventional single-channel multitone measurements are challenging because the number of intermodulation tones grows rapidly with the number of stimulus tones and makes it extremely difficult to separate harmonics from intermodulation products. By using dual-channel measurement techniques, only well-known, standard signal processing techniques are used, resulting in simplicity, accuracy and repeatability.

1. INTRODUCTION

Unlike linear measurements, non-linear distortion measurements depend heavily on the stimulus signal's spectral content and level. Traditional harmonic distortion measurements using a single test tone or sweeping tone are easy to calibrate and perform. However, they do not reveal intermodulation products and are not at all similar to music.

System theory shows that more than two tones are needed to fully characterize a non-linear system [3, 4] Multitone is a popular test signal for fast frequency response measurements. It is also a more rigorous test signal for assessing system non-linearities because it excites many frequencies simultaneously and produces both harmonic and intermodulation distortion products.

Statistically, it is closer to music than a sine wave, two tones, or noise. See Ref [7, 8] for a comprehensive review of multitone testing.

Multitone measurements, however, are challenging because the number of intermodulation tones grows extremely fast with the number of stimulus tones. Most measurement methods to date have focused on analyzing the single channel spectrum and power summing the distortion products between stimuli tones. These methods require careful selection of the stimulus frequencies so they match the individual FFT filters/bins to prevent frequency aliasing. The curve versus frequency is discontinuous, giving an incomplete picture. Furthermore, the individual distortion components and orders are lost in the process.

The new method presented here for measuring distortion based on dual-channel analysis of non-coherence between a stimulus and response is easy to implement, provides a continuous distortion curve against frequency, and can be used with many simultaneous tones, noise or even music.

2. NON-LINEAR SYSTEMS THEORY

2.1. Volterra-Wiener Approach

Mathematician, Vito Volterra, in the 1900s introduced an extension of Taylor Series expansion to functionals. A functional associates a real value to a function [1]. An example is the function that associates a real output value to a continuum of past input values. The use of Volterra series was introduced in non-linear system theory by Norbert Wiener in the 1940s [3].

Since then the Volterra/Wiener approach has grown in popularity because it gives a consistent explanation of nonlinear system behavior. It has proven handy because it yields a time domain characterization of a system as well as a frequency domain. It explains a system's response to common test signals such as multitones. Furthermore, using a multidimensional extension of the Laplace transform and a set of combination rules, it predicts the global nonlinear response of systems made of nonlinear blocks interconnected with addition, multiplication, cascade and feedback. Differential equations can also be used as a starting point with the Volterra/Wiener approach and this has previously been used to model the non-linear behavior of loudspeakers [5, 10]

In general the behavior of nonlinear systems is well described by that approach for regular and moderate nonlinearities. However, it does not describe well irregular distortions such as loudspeaker Rub & Buzz or loose particles and extreme nonlinearities such as amplifier clipping. It is generally accepted that the theory is valid for regular distortions below 10%.

The Volterra theory is an extension of linear system theory. For linear systems, the output at t is a weighted sum of the input values.

$$\mathrm{y}(t) = \int \mathrm{h}(\tau)\, x(t-\tau)\, d\tau \qquad (1)$$

The weighting function h(t) is related to the frequency response H(ω) of the system by a Fourier Transform. In Volterra theory h(t) is the linear kernel.

In a similar fashion, the output of a nonlinear system at a certain time is the result of nonlinear interaction of the input values. Here is the characteristic equation for a pure n^{th} order system.

$$y_n(t) = \int h_n(\tau_1, \ldots \tau_n)\, x(t-\tau_1) \ldots x(t-\tau_n)\, d\tau_1 \ldots d\tau_n \qquad (2)$$

For convenience of notation, the multiple integral is reduced to one. The single integral represents here a multiple summation on all variables of integration $\tau_1 \ldots \tau_n$. The n-dimensional time function $h_n(\tau_1, \ldots \tau_n)$ is called the order-n Volterra kernel.

A physical world nonlinear system will be represented by a Volterra series that is an infinite sum of subsystems of increasing order.

$$y(t) = \int h_1(\tau)\, x(t-\tau)\ d\tau + \sum_{n \geq 2} \int h_n(\tau_1, \ldots \tau_n)\, x(t-\tau_1) \ldots x(t-\tau_n)\, d\tau_1 \ldots d\tau_n \qquad (3)$$

The equivalent block diagram is shown hereafter.

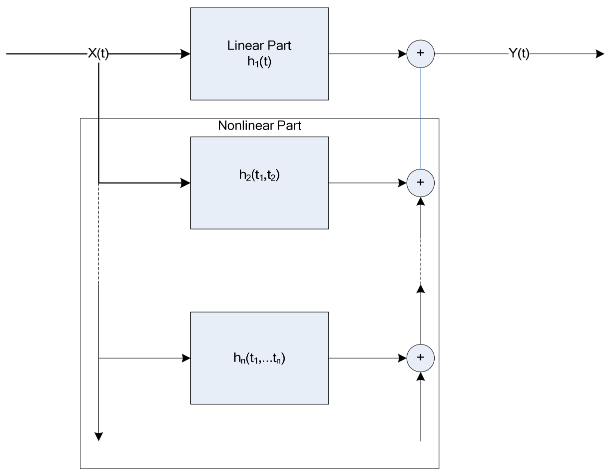

Figure 1: Equivalent Block Diagram of a Volterra Series

It should be noted that:

1) The output of a nonlinear system is a sum of a linear output and several nonlinear outputs. The nonlinear distortion adds to the linear output. That means that the nonlinear distortion could be removed if the different kernels are identified. Much work is being done in that direction with increasing success [6].

2) The Volterra kernel itself is level independent. The nonlinearity arises from the "multi-convolution" that occurs in each subsystem (2). For an nth order system, when the input is multiplied by A the output is multiplied by A^n. E.g. if the input gains 10 dB the output of a 5th order system will gain 50 dB.

3) As Martin Schetzen puts it [3], "the Volterra series is a power series with memory". That memory yields a finite bandpass. A linear system without memory has an infinite bandpass:

$$h(t) = A\delta_0(t) \Leftrightarrow y(t) = A\,x(t) \tag{4}$$

This system cannot exist in the physical world. In a equivalent manner a Volterra subsystem has a memory: at time t, the output for the nth degree, $y_n(t)$ is a weighted sum of a product of n past values of the input (2). Consequently, the bandpass is limited. On the contrary, a memory less (static) nonlinearity such as defined by:

$$h_n(t_1,..t_n) = A\prod_{i=1}^{n}\delta_0(t) \Leftrightarrow y(t) = A\,x^n(t) \tag{5}$$

has an infinite bandpass. As before, that system cannot exist in the physical world. By extension, it follows that a static polynomial model of a nonlinear system is not realistic. The Volterra theory by introducing a memory mechanism, explains the fact that distortion depends on frequency.

A complete and rigorous description of the Volterra/Wiener theory can be found in [3, 4].

2.2. Response of Volterra Systems to Different Signals

2.2.1. Response to an Impulse

A Dirac impulse fed to the system described in Figure 1 is determined by its characteristic equation (3) and yields:

$$\delta_0(t) \rightarrow h(t) + h_2(t,t) + \ldots + h_n(t,\ldots t) + \ldots \tag{6}$$

The impulse response is the sum of the linear impulse response plus all the non-linear impulse responses. Each nonlinear term is the diagonal slice of its nth degree kernel.

2.2.2. Response to a Single Tone

First of all, the frequency response associated with an nth degree kernel is defined as n-dimensional Fourier transform:

$$H_n(\omega_1,\ldots\omega_n) = \int h_n(t_1,..t_n)\,e^{-j\omega_1 t_1 \ldots - j\omega_n t_n}\,dt_1 \ldots dt_n \tag{7}$$

The multi-dimensional frequency response shapes the response of a Volterra kernel for different sets of n frequencies (ω_1,…. ω_n) It is of course a complex function that conveys amplitude and phase. It is usually considered symmetric: that is the order of frequencies doesn't change its value.

Then feeding a single phasor to the Volterra series in figure 1, using (3) is:

$$e^{j\omega t} \rightarrow H_1(\omega)\, e^{j\omega t} + \sum_{n \geq 2} H_n(\omega,...\omega)\, e^{jn\omega t} \qquad (8)$$

That is a single phasor input gives the expected linear output plus a series of harmonic phasors, each of them weighted with the diagonal slice of the nth degree frequency response defined in (7).

The nonlinear behavior really starts to unfold with a sine input. Let's consider, to minimize the complexity, a cosine input:

$$x(t) = 2|A|\cos(\omega t + \angle A) = Ae^{j\omega t} + \overline{A}\, e^{-j\omega t} \qquad (9)$$

and evaluate the result for one nth degree kernel only. By substitution into equation (2), the following response can be found [4]:

$$2|A|\cos(\omega t + \angle A) \rightarrow \sum_{k=0}^{n} \binom{n}{k} \overline{A}^k A^{n-k} H_n(\{-\omega\}^{(k)}, \{\omega\}^{(n-k)})\ e^{j(n-2k)\omega t} \qquad (10)$$

$\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$ is the binomial coefficient.

$\{\omega\}^{(p)}$ denotes the set of p identical frequencies ω.

The phasor frequencies range from nω down to –nω.

For n even, we have: nω, (n-2)ω,…,2ω,0,-2ω,…,-nω.

For n odd, we have: nω, (n-2)ω,…,ω,-ω,…,-nω.

Then using conjugate symmetry of the Volterra frequency response and the symmetry of the binomial coefficient, the phasors can be associated by pairs to yield real sine components. Details of this process can be found in [4]. The end result is the well known fact that an even distortion order yields even order harmonics plus a DC component and odd order distortion yields odd order harmonics plus a component at the fundamental frequency. The harmonics arise from intermodulation between negative and positive frequencies. The amplitude and phases are given by the subsystem frequency response, and are a function of the fundamental frequency. It is worth noting that the amplitude of each order k harmonic is proportional to the amplitude of the single tone raised to the power k.

So far we have seen the response to only one nth order subsystem. In a real system made of several subsystems (see Figure 1), the different harmonics overlap and add up with one another. Each harmonic has an amplitude which is a polynomial expression of the input tone amplitude.

2.2.3. Response to a Multitone

In the case of two tones:

$$x(t) = A_1 e^{j\omega_1 t} + \overline{A}_1 e^{-j\omega_1 t} + A_2 e^{j\omega_2 t} + \overline{A}_2 e^{-j\omega_2 t} \qquad (11)$$

The result is expressed by:

$$x(t) \rightarrow \sum_{\substack{p_1,...p_4 \geq 0 \\ p_1+...+p4=n}} A_1^{p_1+p_2} A_2^{p_3+p_4} \left(\frac{n!}{p_1!...p_4!} \right). H(\{\omega_1\}^{(p_1)};...;\{\omega_n\}^{(p_4)})\, e^{j(\{p_1-p_2\}\omega_1+\{p_3-p_4\}\omega_2)t} \qquad (12)$$

For example, 2 tones applied to a 3rd order subsystem will generate 8 tones with the following frequencies: ω_1, ω_2, $|2\omega_1 - \omega_2|$, $|2\omega_2 - \omega_1|$, $2\omega_1 + \omega_2$, $2\omega_2 + \omega_1$, $3\omega_1$, $3\omega_2$.

Let's now define a multitone by a sum of m phasors (input tones), with complex amplitudes:

$$x(t) = \sum_{k=1}^{m} A_k\ e^{i\omega_k t} \qquad (13)$$

In the case of a real signal, half of the components will be the conjugate of the other half (except for a DC component which stays single).

By substituting a multitone in (2) for a single subsystem, we get the formidable expression:

$$x(t) \rightarrow \sum_{k_1=1}^{m} ... \sum_{k_n=1}^{m} A_{k_1} ... A_{k_n} H(\omega_{k_1},...\omega_{k_n})\, e^{j\omega_{k_1}t+...j\omega_{k_n}t}$$
$$= \sum_{\substack{p_1,...p_m \geq 0 \\ p_1+...+p_m=n}} A_1^{p_1} ... A_m^{p_m} \left(\frac{n!}{p_1!...p_m!} \right) H(\{\omega_1\}^{(p_1)};...;\{\omega_n\}^{(p_n)})\, e^{jp_1\omega_1 t+...jp_n\omega_n t} \qquad (14)$$

The summation is made on all sets of m positive integers $p_1, \ldots p_m$ that sum up to n. So m phasors input to an n^{th} order subsystem, yields all distinct linear combinations of m frequencies (half positive, half negative) such that the sum of their weightings adds up to n. That number of combinations grows very fast with the number of frequencies m, even for a low order n, and the growth is on the side of the IM products.

For example, in [7], it is shown that twenty tones through a 5^{th} order system yields 20 5^{th} order harmonics plus 437,560 intermodulation products.

The total number of tones produced by an n^{th} order system is roughly given by:

$$N = 0.4 \binom{2m + n - 1}{n} \tag{15 }$$

where m is the number of input tones.

The amplitude of each tone in the response is proportional to the cross-product of the amplitudes of each primary tone raised to a power equal to its contribution in the final frequency. That amplitude is weighted by the Kernel Frequency response for that specific set of frequencies.

The phase of each component is constructed in a similar manner: it is the weighted sum of all input tone phases plus the phase given by the frequency response.

On top of that, much overlap occurs between the tones created by the different kernels of a Volterra series (e.g. harmonics of same parity kernels).

Note that nonlinear amplitude behavior is not due to the Volterra Frequency response but to the power cross-products.

Finally it is worth noting that a multitone probes a grid of points in each kernel transfer function (see ref. [5]), so is potentially able to excite some high gain locations.

2.2.4. Response to Noise

The Volterra series representation is not helpful for noise and stochastic signals. The combination of cross-correlations that arises with the application of each kernel is inextricable.

To overcome that problem Wiener came up with Wiener Operators [2] which are derived from the Volterra Kernels by polynomial combination. These new kernels exhibit statistical orthogonality properties that simplify the outcome for a noise input. Basically the output of a nonlinear system in that case becomes a sum of uncorrelated noises, each of them produced by one of the Wiener operators.

3. NON-COHERENCE & DISTORTION

3.1. Definition of Non-Coherence

The model used to define non-coherence is the following:

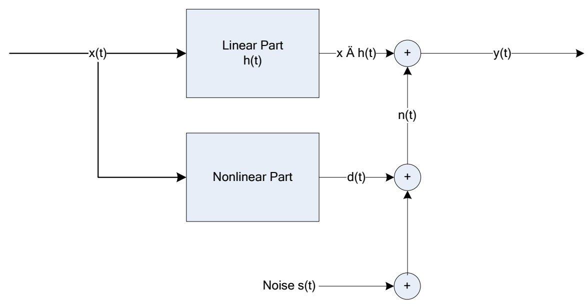

Figure 2: Block Diagram for Non-Coherence Analysis

It assumes that n(t), the sum of the non-linear distortion and noise, is uncorrelated with x(t) and $x \otimes h(t)$. That is obvious with noise and can be assumed with nonlinearity considering the fact that the distortion signal d(t) contains new frequencies that are different from those found in x(t). For weak distortion we can neglect the feedback on the fundamental.

Let's define the following:

Auto-spectrum:

$$G_{XX}(\omega) = E\left[\overline{X}(\omega) X(\omega)\right] \tag{16 }$$

With X(ω) being the Fourier spectrum of x(t) and E. being the mathematical expectation. G_{XX} is a real positive function.

Input-output Cross-spectrum:

$$G_{XY}(\omega) = E\left[\overline{X}(\omega)\,Y(\omega)\right] \tag{17}$$

With X and Y being respectively the input and output Fourier spectrums, G_{XY} is a complex function.

Then the coherence is defined by:

$$\gamma^2_{(\omega)} = \frac{|G_{XY}|^2}{G_{XX}\cdot G_{YY}}(\omega) \tag{18}$$

And the Non-Coherence is simply: $1-\gamma_{(\omega)}^{\;2}$

By construction, the coherence is a positive number smaller than 1. It can be seen as the ratio of Output Coherent Power to the Total Output Power as shown below.

If H(ω) is the linear frequency response of the system, then:

$$G_{XY}(\omega) = H(\omega)\cdot G_{XX}(\omega) \tag{19}$$

And, because the noise is uncorrelated with the input signal:

$$G_{YY}(\omega) = |H|^2(\omega)\cdot G_{XX}(\omega) + G_{NN}(\omega) \tag{20}$$

Then combining (19) and (20) into (18), one obtains:

$$\gamma^2_{(\omega)} = \frac{|H|^2 G_{XX}}{G_{YY}}(\omega) \tag{21}$$

From (21) the important relationship is derived:

$$(1-\gamma^2_{(\omega)})\cdot G_{YY}(\omega) = G_{NN}(\omega) \tag{22}$$

The non-coherence applied to the output auto-spectrum yields the uncorrelated noise spectrum present at the output.

All of the above are derived from [11, 12].

We can then introduce a Non-Coherent Distortion measure:

$$\eta^2(\omega) = \frac{(1-\gamma^2_{(\omega)})\cdot G_{YY}(\omega)}{\sum_{\omega} G_{YY}(\omega)} = \frac{G_{NN}(\omega)}{\sum_{\omega} G_{YY}(\omega)} \tag{23}$$

The Non-Coherent Distortion (NCD) is the Non-Coherent Power Spectrum G_{NN} normalized against the Total Output Power. As for the coherence we have: $0 \le \eta^2 \le 1$. And if we sum against frequency, we end up with the ratio of the Total Non-Coherent Power to the Total Output Power.

$$\lambda = \sqrt{\sum_{\omega} \eta^2(\omega)} = \sqrt{\frac{\sum_{\omega} G_{NN}(\omega)}{\sum_{\omega} G_{YY}(\omega)}} \tag{24}$$

The Total Non-Coherent Distortion (TNCD) λ is an extension of the THDN for a multitone or a broadband signal. It can be expressed in % as well.

3.2. Non-Coherence Measurement

With an FFT analyzer, a collection of successive time limited spectra is measured for both input and output. Each individual spectrum is then a function of time and frequency that can described by:

$$X(t,\omega) = \sum_{u=-N/2}^{N/2} w(u).x(t+u).e^{-j\omega u} \tag{25}$$

With the following notations:

- t,u discrete time positions,
- ω discrete angular
- w(u) windowing function (e.g. Hanning) of support [-N/2, N/2].

Then the auto and cross-spectrum are obtained by time averaging.

$$G_{XX}(\omega) = \frac{1}{M}\sum_{t=0}^{M-1}\left|X(t,\omega)\right|^2 \quad (26)$$

$$G_{XY}(\omega) = \frac{1}{M}\sum_{t=0}^{M-1}\overline{X}.Y(t,\omega) \quad (27)$$

Now, by looking more closely to Eq (24), we can see that the spectrum X(t,ω) is convolution product between x(t) and a complex kernel, limited to the support [-N/2, N/2]:

$$g(t,\omega) = w(-t)\cdot e^{j\omega t} \quad (28)$$

For a given frequency ω, this function g is the finite impulse response of a linear phase complex filter. With a usual windowing function like Hanning or Blackman-Harris, the filter is a bandpass centered on the frequency ω_0. Considering the windowing function we can even express the frequency response of the equivalent filter:

$$g(t,\omega_0) \xrightarrow{\Im} W(\omega_0 - \omega) = W(\omega - \omega_0) \quad (29)$$

W(ω) being the Fourier transform of w(t).

The FFT analyzer is a filter bank with linearly spaced bandpass filters that spans the range from 0 to the sampling frequency. Because of the symmetry property of spectra, only the frequencies up to Nyquist are usually shown.

Now let's see how coherence works in a simple case.

If the input signal is a single phasor of frequency ω_0, we get the output signal y(t):

$$x(t) = e^{j\omega_0 t} \rightarrow y(t) = H(\omega_0)\cdot e^{j\omega_0 t} + n(t) \quad (30)$$

Then the output of the FFT bin centered on ω_0 is the phasor itself for x(t) (we assume that W(0)=1 for the sake of simplicity):

$$X(t,\omega_0) = e^{j\omega_0 t} \quad (31)$$

And at the output the same FFT bin for y(t):

$$Y(t,\omega_0) = H(\omega_0).e^{j\omega_0 t} + n(t)\otimes g(t,\omega_0) \quad (32)$$

⊗ denotes the convolution product.

The instantaneous cross-spectrum is:

$$\overline{X}\cdot Y(t,\omega_0) = H(\omega_0) + n(t)\otimes g(t,\omega_0).e^{-j\omega_0 t} \quad (33)$$

The first term in (33) is constant with time. The second term is the noise + distortion filtered around ω_0 and shifted back around DC. Averaging the cross-spectra will leave the constant term $H(\omega_0)$ and cancel out the filtered noise term. Therefore:

$$E\left[\overline{X}\cdot Y(t,\omega_0)\right] = H(\omega_0) \quad (34)$$

And we also have:

$$E\left[\overline{X}\cdot X(t,\omega_0)\right] = 1 \quad (35)$$

$$E\left[\overline{Y}\cdot Y(t,\omega_0)\right] = \left|H(\omega_0)\right|^2 + B\sigma_n^2 \quad (36)$$

B is the energy noise bandwidth of the FFT filter and σ_n^2 is the mean spectral density of the noise on the bin.

When the number of averages is high enough, G_{XX}, G_{YY} and G_{XY} meet their expectations given above.

Then we get the value of the coherence measured for the FFT bin at frequency ω_0:

$$\gamma^2(\omega_0) = \frac{\left|H(\omega_0)\right|^2}{\left|H(\omega_0)\right|^2 + B\sigma_n^2} \quad (37)$$

The coherence is then the ratio of the coherent power to total power, both collected by the FFT filter. The cross-spectrum cancels the rotation of the linearly correlated phasor present in the output signal. That cancellation is performed by the multiplication with the conjugate input phasor that rotates in the opposite direction. This works because the two phasors are proportional: $Y(\omega_0) = H(\omega_0).X(\omega_0)$. The noise and distortion cannot be stopped because the frequencies are different and there is no linear ratio with the input.

In the general case where:

$$x(t) = \sum_{\omega} X(\omega) \cdot e^{j\omega t} \qquad (38)$$

And:

$$y(t) = \sum_{\omega} H(\omega) X(\omega) \cdot e^{j\omega t} + n(t) \qquad (39)$$

It can be shown [9] that:

$$G_{XX}(\omega_0) = \sum_{\omega} | W(\omega - \omega_0) |^2 \cdot | X(\omega) |^2 \qquad (40)$$

$$G_{YY}(\omega_0) = \sum_{\omega} | W(\omega - \omega_0) |^2 \cdot | H(\omega)\, X(\omega) |^2 + B\sigma_n^2 \qquad (41)$$

$$G_{XY}(\omega_0) = \sum_{\omega} | W(\omega - \omega_0) |^2 \cdot H(\omega) \cdot | X(\omega) |^2 \qquad (42)$$

The measured auto (respectively cross)-spectrum is equal to the auto (resp. cross)-spectrum convolved with the window auto-spectrum. With the usual window (e.g. Hanning), it is equivalent to smoothing. It is worth emphasizing the fact that, for the estimated cross-spectrum, the convolution is applied to complex values (42) and if the phase of H(ω) rotates too much within the width of the window, complete cancellation and other dramatic effects can occur. Because the estimated coherence is proportional to the cross-spectrum square modulus, the type and width of the window used can have a big effect on it.

3.3. Non-Coherence versus Nonlinear Distortion

If the excitation signal is periodic, distortion will add new frequencies different from the input frequencies. These frequencies will not be halted in the instantaneous cross-spectrum.

If the excitation signal is stochastic, the Volterra/Wiener theory shows that the distortion noise is uncorrelated with the input noise.

In both cases, the cross-spectrum will average out the distortion signal and keep only the linear power. Then the coherence will be the ratio of linear power over the total power and the non-coherence will yield the nonlinear power present at the output.

3.4. Non-Coherence Applied to Test Signals

In the case of a single tone or a combination of tones (2 tones, multitone), as stated before the distortion will add new tones. Some of the new tones will overlap with the fundamentals, but the vast majority of them will be at new frequencies. The non-coherence calculus is able to distinguish these new tones, and yield the distortion power.

Some caution must used for the calculus though: the input auto-spectrum of a set of tones will be about null in many bins. Therefore some regularization must be applied in order to avoid the coherence ratio becoming infinite.

For noise signals, the subject is well covered in the technical literature [18]. It is the traditional input signal for coherence analysis. The main information to keep in mind is that the precision of the coherence estimate is proportional to the number of averages.

The non-coherent distortion analysis can also be applied on music or speech. Music may contain long stable periodic parts (even very long for some minimalist pieces). Provided that a suitable regularization is applied on the input autospectrum the coherence can be measured.

4. EXPERIMENTS

In order to better understand the significance and practicality of Non-Coherent Distortion, many distortion measurements were made on various loudspeakers. All the loudspeakers tested showed similar results when comparing the different distortion measurement techniques. Consequently, for this paper, only one full range 6 by 9 inch car loudspeaker was selected to demonstrate the comparison. Figure 3 shows the frequency response of the car loudspeaker when excited with pink noise from 20 to 20 kHz at 1 Vrms.

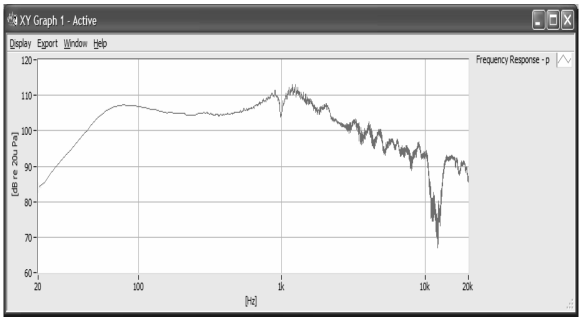

Figure 3: Loudspeaker Frequency Response

4.1. Stimulus level

It is very difficult to compare the excitation level of a single test tone at one frequency to a broadband test signal like pink noise.

Set up of RMS level is meaningful when the DUT exhibits soft clipping like a loudspeaker because the distortion can then be related to the output power.

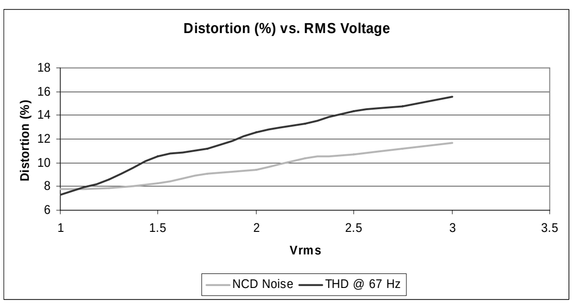

Figure 4: Loudspeaker Distortion vs. level

On the other hand, for a DUT with hard clipping (e.g. an amplifier), it is more interesting to set the stimulus in peak level to check its saturation limit.

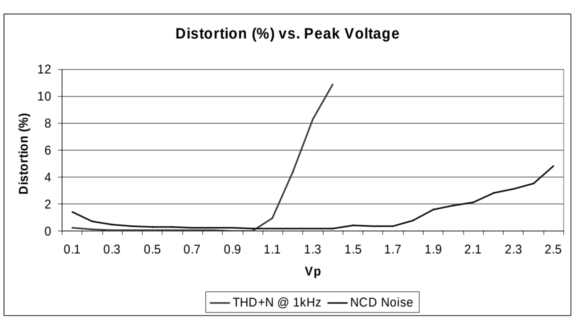

Figure 5: Amplifier Distortion vs. level

Anyway it is important to know that stimuli with the same peak value but different amplitude distributions will not produce the same distortion. The Volterra series is sensitive to the instantaneous value of the signal by design and the higher the order the more sensitive it is. Roughly speaking, distortion depends on the crest-factor. For example, the amplitude distribution for a sine is a "bath-tub" curve. For multitone and noise it is a bell curve (Figure 6). Sine signal spends most of the time near the extreme values. For multitone and noise, the signal spends most of its time around zero value. This explains why noise produces less distortion than a sine when it exceeds the clipping limit (Figure 5).

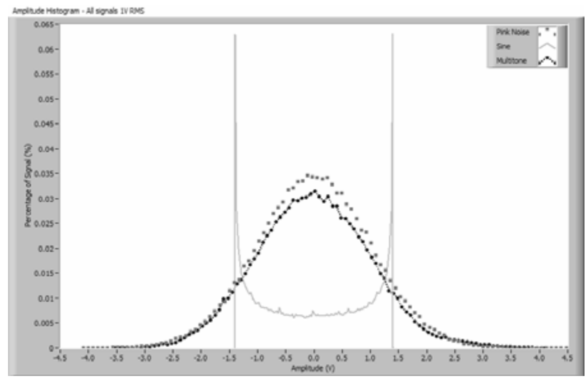

Figure 6: Amplitude Distributions Comparison

In addition, the stimulus spectral content changes the distortion spectrum in term of frequencies and levels because the IM products are the cross-products of the signal components (14). Then the overall distortion depends on the stimulus spectrum. It is obvious for a sine versus a multitone, but even a two tone stimulus with different tone spacing won't generate the same distortion.

4.2. Distortion Measurements using Different Test Signals

Five different test signals were used for measuring distortion on this loudspeaker;

- Harmonic Distortion (THD): 1/12th octave stepped sine sweep from 20 to 20 kHz @ 1 Vrms
- Difference Frequency Distortion (Total DFD): 1/12th octave two tone stepped sine separated by 83 Hz swept from 100 to 20 kHz both tones @ 707m Vrms for a total power of 1 Vrms

- Multitone Non-Coherent Distortion (Multitone NCD): 1/12th octave spaced multitones with random phases (crest factor approximately 4) from 20 to 20 kHz @ 1 Vrms

- Pink Noise Non-Coherent Distortion (Pink Noise NCD): 5 seconds of pink noise band limited to 20 to 20 kHz @ 1 Vrms

- Techno Music NCD: 10 seconds of techno music @ 1 Vrms

For starters, it would be interesting to compare traditional total harmonic distortion, difference frequency distortion, and the new non-coherent distortion to see if there are any similarities. Since loudspeaker non-linear distortion is very level dependent, it was decided to set the excitation level for the various test signals to the same total RMS level.

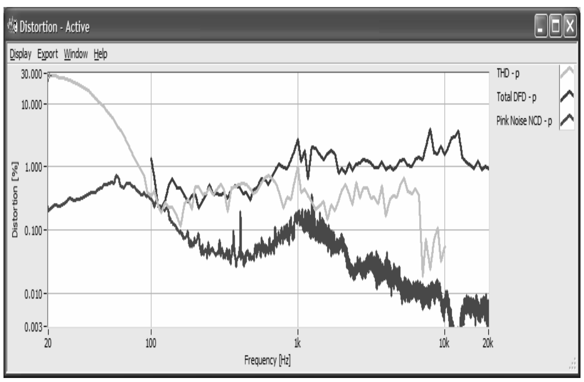

Figure 7: Loudspeaker THD, Total DFD, and Pink Noise NCD

At first glance, it looks like these three different distortion metrics have nothing in common. THD is almost 30% at low frequencies and Pink Noise NCD is under 1%. It is impractical to measure DFD below 100 Hz due to the 83 Hz difference frequency [19].

But it is not really fair to compare these three different distortion metrics without first removing the influence of the frequency response on the distortion products. A distortion product that falls on the same frequency as a peak, dip, or roll-off in the frequency response will be amplified or attenuated when compared to the stimulus frequency in a flat region of the frequency response. This is explained in detail in an AES preprint describing amplitude normalized distortion in reference [20].

In Figure 8, the influence of the frequency response on the distortion is removed by amplitude normalizing the distortion products to the fundamental at their measured frequency before plotting them at their excitation frequency.

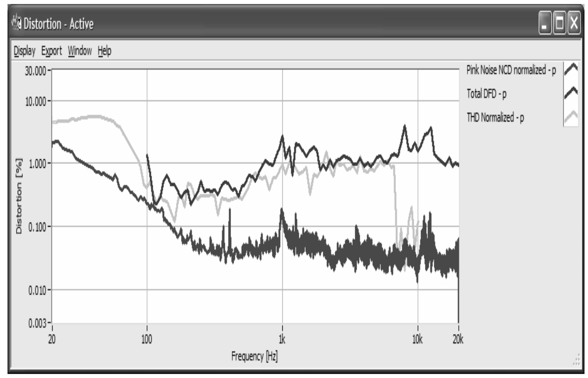

Figure 8: Amplitude Normalized THD, Total DFD, and Pink Noise NCD

By amplitude normalizing, the three different distortion curves start to look more similar especially in the middle of the frequency range (e.g. all 3 show a bump at approximately 1 & 9 kHz). Only the pink noise NCD covers the entire frequency range from 20 to 20 kHz. THD is difficult to measure above the cutoff frequency of the analysis (e.g. the 3rd harmonic of 10 kHz is 30 kHz which is above the cutoff frequency of 20 kHz). Also, there seems to be an overall level shift between the Pink Noise NCD and the one and two tone distortion measurements.

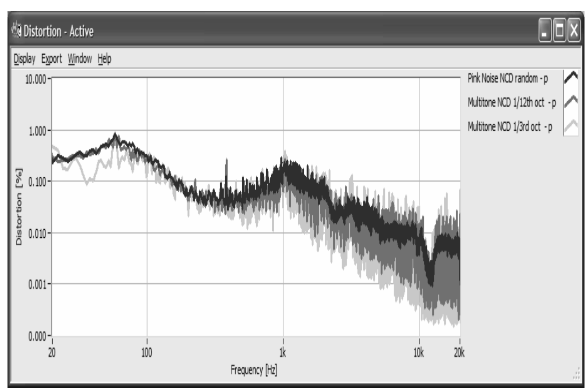

Figure 9: Multitone NCD vs. Pink Noise NCD

We decided to see if increasing the number of tones made much difference to the NCD curve. What we found is that the Non-Coherent Distortion just spreads out with increasing number of tones (see figure 9). Since noise is effectively an infinite number of tones, we decided to see what affect using pink noise had on the NCD curve. As suspected, the pink noise NCD curve looked like a smoother version of the multitone NCD curve. Also, we decided to power sum the NCD energy into one overall level and compare. Not surprisingly, the overall levels were extremely close (3% for multitone and 3.5% for pink noise).

4.3. Distortion measurements using Music as the Test Signal

With dual channel analysis, non-coherent distortion can be measured using any stimulus signal. We thought it might be interesting to look at a real life signal like music. We choose 10 seconds of some popular techno music and analyzed its spectrum to make sure there was enough energy over the entire frequency range. It was hard to find a piece of techno music with a lot of energy below 60 Hz and above 10 kHz but there seemed to be enough in this piece to get good coherence with a lot of spectrum averaging over a 10 second music sample.

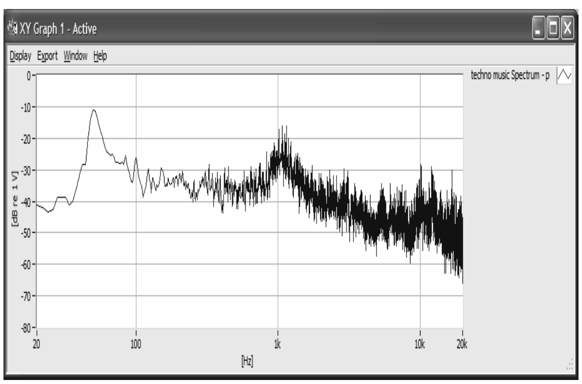

Figure 10: Spectrum of Techno Music 10 sec. sample

In Figure 11, the Pink Noise Non-Coherent Distortion for 1 Vrms input is compared to the techno music NCD at 1 Vrms. Again, there are striking similarities around 60 Hz (near the speaker's resonance frequency) and 1 kHz where the speaker starts to break up with cone resonances. Interestingly, after power summing the two curves, the Total NCD for pink noise is 3.5% and techno music is 3.7%

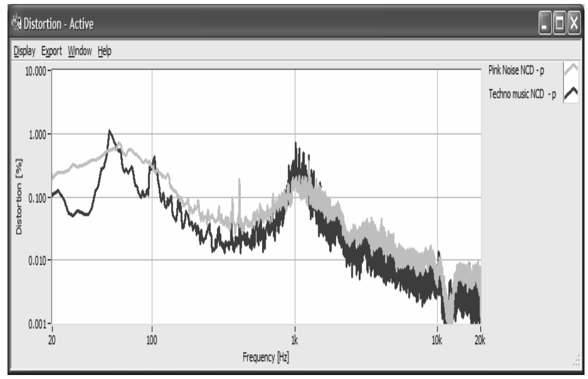

Figure 11: Pink Noise NCD vs. Techno Music NCD

We also tried to see if non-coherent distortion could reveal defects such as rub & buzz and loose particles. When comparing non-coherent distortion curves for both good and bad loudspeakers, there was no discernable difference in the NCD curves or overall level. This makes sense because rub & buzz and loose particle defects are much lower in energy compared to typical 2nd and 3rd order distortion components which tend to dominate the other higher distortion orders in level. So, it probably still makes sense to look at these defects with traditional high order harmonic distortion and filtered time envelope analysis [21].

4.4. Distortion versus Level

As mentioned earlier, it is interesting to look at distortion versus level to see how well the loudspeaker behaves to changes in test level. Ideally, the distortion should increase gradually and not hard limit. In figure 12, the loudspeaker distortion is gradually increasing with test level.

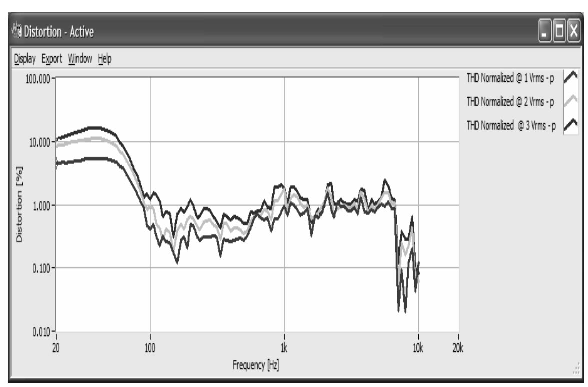

Figure 12: THD Normalized at 1, 2, & 3 Vrms

We tested the same loudspeaker with pink noise at the same RMS voltage levels. Even though the pink noise NCD distortion level looks much lower than the THD for the same levels (note the change in scales), the relative increase in distortion between levels is about the same.

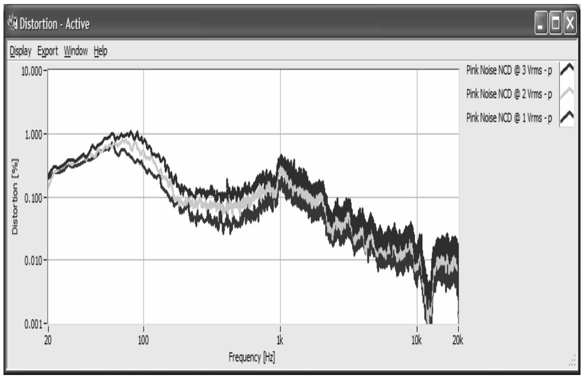

Figure 13: Pink Noise NCD at 1, 2, & 3 Vrms

Because the energy is spread out over the entire frequency range for pink noise NCD and not just one frequency as for THD, it might be interesting to compare the total distortion level for pink noise NCD to the maximum or average THD. In figure 13, the Total Pink Noise NCD was 3.5% at 1 Vrms, 5.0% at 2 Vrms, and 6.9% at 3 Vrms. These levels seem more comparable to the THD maximum or average values.

5. ADVANTAGES & LIMITATIONS

5.1. THD vs. Non-Coherent Distortion

THD is a ratio of the total RMS level of the harmonics to the total RMS level of the signal. This measurement is done for a one tone stimulus. The result is dependent on the stimulus frequency. One can see the THD as a normalized level of distortion. On the other hand, the NCD (23) is the normalized distortion spectrum. Therefore THD (or THD+N) cannot be compared to the NCD, because one is homogeneous to a global level and the other to a spectrum. The only legitimate comparison that can be done is between the TNCD, which is the power sum of NCD, and the THD+N for some frequency or its average.

5.2. Proper Analysis Parameters

There are two approaches to studying non-coherent distortion: an analytical approach that searches to identify specific individual IM components and a global assessment approach that searches to qualify the DUT distortion.

To analyze distinct IM components, one will use a sparse multitone (e.g. five tones) or a multitone with rounded frequencies (e.g. to the nearest 20 Hz). The Non-coherence function applied to the multitone will then act as a mask. That mask applied on G_{YY} will reject all the tones but pass everything else (22). The quality of the measurement will depend on the selectivity of that mask. Ideally we should have 100% between the tones and 0% on the tones. To get valid results the closest tones must be separated enough in the FFT. A bin alignment of the FFT lines on the tone frequencies will help because it will reduce spectral leakage. To be able to discriminate individual distortion components, the best results are obtained with a Hanning or Blackman-Harris window, bin alignment and 5 bins separation between the tones. Overlap improves the estimation of the non-coherent distortion and should be as high as possible (typically 90% or higher). This analytic approach will give similar results to the currently used multitone measurement methods [13-15].

On the other hand when applied to high-density multitone or noise, the coherence analysis will demonstrate its ability to sort out the distortion inside each FFT filter. The Non-Coherence will act then as a weighting function applied to G_{YY} (22) that sorts out the Non-Coherent Power for each filter. A window with low leakage is preferred (Hanning or better yet, Blackman-Harris) to limit its influence on the cross-spectrum (40-42). The frequency resolution and window applied must be broad enough to minimize the variance of the curves but sharp enough not to corrupt the coherence of the frequency response. Some ad-hoc tuning is therefore necessary.

5.3. Limitations

The first obvious limitation is that non-coherence analysis includes background noise in the analysis of distortion, as shown in section 3. They cannot be distinguished from one another but with careful attention to the measurement environment, e.g. using a quiet lab, the signal to background noise can be sufficiently large as to discount the affect of background noise on the loudspeaker distortion.

The other limitation or drawback is that the level of the distortion spectrum depends on the FFT resolution when the distortion products are merged together. That can be overcome by using Spectral Density instead of RMS as a level unit. But then we end up with a distortion density in %/√Hz.

6. FURTHER DEVELOPMENTS

To further distinguish distortion from noise, we could use a multitone with rounded frequencies and measure the noise between the distortion products [14]. That is possible because with frequencies multiple of a common value, all the produced tones align to a harmonic pattern.

We could also identify and highlight the individual harmonics and IM components produced by a specific order of distortion but it becomes unwieldy if there are too many tones.

We could try to weight the non-coherent distortion curve with a psychometric filter such as an A, B, or C weighting curve or masking curves to better predict what levels of distortion are audible.

The multitone phase distribution is another important subject. The phase between tones determines its crest factor and affects the global distortion level. The random phase used for this paper yields a noise-like signal but can lead to variable distortion results. Some deterministic phase rule could be applied instead [17].

The multitone frequency distribution could be another area of work to minimize the distortion products overlap for fine resolution analysis purposes [16].

Finally the next step should be the direct measurement of Volterra-Wiener kernels using multitone or noise. That is the only way to fully characterize a non-linear system.

7. CONCLUSION

It is almost impossible to compare the different distortion measurement techniques (namely one tone, two tones, multitone, and noise) because the distortion obtained is very dependent on the crest-factor and the spectral content of the stimulus. However, there are some similarities in the shape of the distortion curves between the different techniques after removing the effect of the linear response on the non-linear response. Also, the distortion levels between the different techniques are not so different if the non-coherent distortion curves are power summed when comparing to the average or maximum level of the one tone, THD or two tone, DFD curves. However, non-coherent distortion does not appear to detect very low level but quite audible non-regular distortions such as Rub & Buzz, loose particles and hard limiting (e.g. clipping).

When the number of tones increases, the distortion products gradually merge into a continuum and approach the result obtained with noise. Beyond a certain density of tones the distortion spectrum of multitone and noise are almost identical. Noise is the ultimate multitone because it contains all frequencies and excites all parts of the Volterra kernels. Pink noise and white noise are de-facto standard test signals that anybody can agree on. This is not the case with multitone.

The total NCD yields a single distortion ratio. It is then easy to use for loudspeaker and probably any DUT qualification and technical specification. It is more meaningful than say the usual THD+N @ 1 kHz because it will better predict the distortion obtained with real signal material such as music.

With this technique we can use any complex test signal such as music or speech at realistic levels and measure the non-coherent distortion. Ultimately, is this not what everyone wants to do?!

8. ACKNOWLEDGEMENTS

We thank Douglas Preis, Professor of Signal Processing at Tufts University, for his previous work and useful insights on this subject.

9. REFERENCES

[1] V.Volterra, “Theory Of Functionals and of Integral and Integro-Differential Equations”, Dover, New-York, 1958

[2] N.Wiener, "Non Linear Problems In Random Theory", MIT Press, Cambridge, MA, 1958

[3] M.Shetzen, "The Volterra & Wiener Theories of Nonlinear Systems", John Wiley, New York, 1980

[4] W.J.Rugh, "Nonlinear System Theory - The Volterra/Wiener Approach", John Hopkins University Press, 1981

[5] A.M.Kaizer, "Modeling of The Nonlinear Response of an Electrodynamic Loudspeaker by a Volterra Series Expansion", JAES, Vol 35, pp 421-433, 1987 June

[6] Khosrow Lashkari, "A Novel Volterra-Wiener Model for Equalization of Loudspeaker Distortions", 1-4244-0469-X/06, ICASSP 2006

[7] E.Czerwinsky, A.Voishvillo & Al., "Multitone Testing of Sound System Components- Some Results and Conclusions- Part 1: History and Theory", JAES, Vol 49, No 11, 2001 November

[8] E.Czerwinsky, A.Voishvillo & Al., "Multitone Testing of Sound System Components- Some Results and Conclusions- Part 2: Modeling and Application", JAES, Vol 49, No 12, 2001 December

[9] Alan V. Oppenheim & Ronald W.Shafer, "Digital Signal Processing", Prentice-Hall, 1975

[10] Martin J.Reed and Malcolm O.J. Hawksford, "Comparison of Audio System Nonlinear Performance in Volterra Space", AES 103rd Convention, September 1997, Preprint 4606

[11] U. Totzek, D. Preis, "How to Measure and Interpret Coherence Loss in Magnetic Recording", JAES, Vol 35,No 11,1987 November

[12] D.Preis, R.Gregg, "Coherence-Based, Wide-Band, Signal-to-Distortion Ratio Versus Total Harmonic Distortion of Slightly Non-Linear Audio Systems", AES 109th, 2000 September, Preprint 5262

[13] J.Vanderkooy, Scott G.Norcross, "Multitone Testing of Audio Systems", AES 101st Convention, 1996, Preprint 4378

[14] Richard C.Cabot, "Fundamentals of Modern Audio Measurement", JAES, Vol 47, No 9, 1999 September

[15] E.Czerwinsky, A.Voishvillo & Al., "Graphing, Interpretation, and Comparison of Results of Loudspeaker Nonlinear Distortion Measurements", JAES, Vol 52, No 4, 2004 April

[16] Jon M.Risch, "A New Class of In-Band Multitone Test Signals", AES 105th, 1998 September, Preprint 4803

[17] Alexander Potchinkov, "Low-Crest Factor Multitone Test Signals for Audio Testing", JAES, Vol 50, No 9, 2002 September

[18] R.B.Randall, "Frequency Analysis", Brüel & Kjær, 1987 September

[19] S.F. Temme, "Why and How to Measure Distortion in Electroacoustic Transducers", presented at the AES 11th International Conference on Audio Test and Measurement, Portland, Oregon (29-31 May, 1992).

[20] S.Temme, "How to Graph Distortion Measurements" AES 94th Convention, March 1993, Preprint 3602

[21] P.Brunet, S.Temme, "Loose Particle Detection in Loudspeakers" AES 115th Convention, October 2003