Audio Engineering Society

Convention e-Brief 559

Presented at the 147th Convention
2019 October 16–19, New York, USA

Coherence as an Indicator of Distortion for Wide-Band Audio Signals such as M-Noise and Music

Merlijn van Veen[1], Roger Schwenke[1]

[1] *Meyer Sound Laboratories, 2832 San Pablo Ave, Berkeley, CA 94702, USA*

Correspondence should be addressed to Merlijn van Veen (merlijnv@meyersound.com)

ABSTRACT

M-Noise is a new scientifically derived test signal whose crest factor as a function of frequency is modeled after real music. M-Noise should be used with a complementary procedure for determining a loudspeaker's maximum linear SPL. The M-Noise Procedure contains criteria for the maximum allowable change in coherence as well as frequency response. When the loudspeaker and microphone are positioned as prescribed by the procedure, reductions in coherence are expected to be caused by distortion. Although higher precision methods for measuring distortion exist, coherence has the advantage that it can be calculated for wide-band signals such as M-Noise as well as music. Examples will demonstrate the perceived audio quality associated with different amounts of distortion-induced coherence loss.

1 Introduction

Coherence has a maximum value of 1 or 100% when a system's signal output power at a given frequency is caused exclusively by the system's signal input power at that frequency within the same analysis time window.

Coherence is affected by uncorrelated noise, distortion, misalignment (bias), loudspeaker-to-loudspeaker interaction and room-interaction such as reflections and reverberation arriving outside the analysis time window.

When a measurement microphone is placed close enough to a sole loudspeaker, the reduction in coherence due to uncorrelated noise and room-interaction is negligible. Under these circumstances, any further reduction in coherence is primarily caused by distortion.

While other metrics for distortion exist, none of them can be used with wide-band signals that contain energy at all frequencies simultaneously such as music. For this reason, coherence has been chosen as indicator of distortion using a specific procedure and accompanying test signal called M-Noise [1].

1.1 Coherence Definitions

Coherence is a statistical measure of the correlation between a system's input $\boldsymbol{X}$ and output $\boldsymbol{Y}$, and can be determined from quantities already produced in the course of calculating a transfer function.

What is colloquially referred to as simply "coherence" actually comes in two slightly different flavours [2].

$$\gamma^2(f) = \frac{\left|P_{XY}(f)\right|^2}{P_{XX}(f)P_{YY}(f)} \qquad (1)$$

Where $P_{XY}(f)$ is the cross-power spectral density of X and Y, and $P_{XX}(f)$ and $P_{YY}(f)$ the auto-power spectral densities of X and Y respectively. Equation (1) is the magnitude-squared coherence function (MSC)

$$\gamma(f) = \frac{P_{XY}(f)}{\sqrt{P_{XX}(f)P_{YY}(f)}} \qquad (2)$$

whereas Equation (2) is the coherence function. To avoid confusion, the authors will refer to the latter as "magnitude"-coherence (MC).

The difference between these two equations effectively constitutes a mere scaling issue (Fig. 1) which matters nonetheless in the interest of what follows.

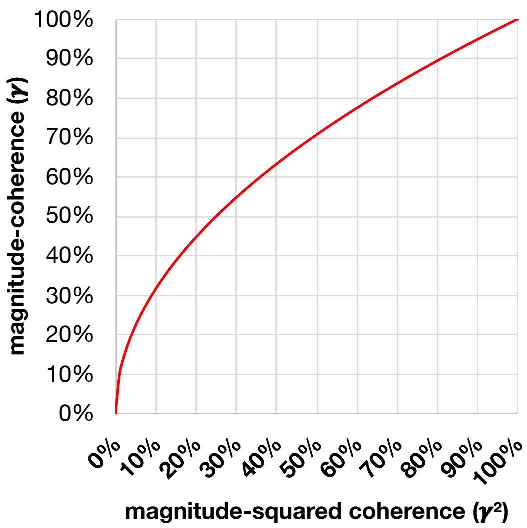

Figure 1. MSC (Eq. 1) vs. MC (Eq. 2)

The authors are aware of at least two commercially available transfer function analyzers, frequently used in the live sound reinforcement industry, where each of them uses one of these definitions and calls it "coherence".

1.2 Signal-to-Distortion Ratio

MSC (Eq. 1) or $\gamma^2(f)$ represents the fraction of a system's output signal power that is linearly dependent on its input where $0 \leq \gamma^2(f) \leq 1$. Conversely, the non-coherent part is simply the remainder $1 - \gamma^2(f)$.

The non-coherent "lump" part is a potpourri of uncorrelated noise or distortion. However, in the absence of noise, the signal-to-distortion ratio (SDR) can be defined as

$$SDR(f) = \frac{\gamma^2(f)}{1 - \gamma^2(f)} \qquad (3)$$

Equipped with this metric, we can compare identical values for MSC (Eq. 1) and MC (Eq. 2) and determine how they translate into SDR expressed in decibels.

MSC γ^2	SDR (dB)
99%	20.0
95%	12.8
90%	9.5
85%	7.5
80%	6.0
75%	4.8
70%	3.7

MC γ	SDR (dB)
99%	16.9
95%	9.7
90%	6.3
85%	4.2
80%	2.5
75%	1.1
70%	-0.2

Table 1. Comparison of MSC (Eq. 1) and MC (Eq. 2) values vs. SDR.

Notice that for similar percentages, SDR values can differ substantially (Table 1). Therefore, any assumptions about distortion based on coherence loss without prior knowledge of the analyzer's "coherence" regime, i.e., MSC (Eq. 1) or MC (Eq. 2), should be treated with scrutiny.

1.3 M-Noise Procedure

The M-Noise Procedure requires two microphone positions: a distant position which is used to measure the sound pressure level, and a close position which is used to calculate the transfer function (including coherence).

The distant microphone is placed at a desired point of interest, e.g., at a prescribed listener location when validating an installed loudspeaker system or at a distance of 1 meter when performing a measurement for a data sheet.

The second microphone should be sufficiently close to provide SNR and D/R values of at least 15 dB or more at a playback level where the loudspeaker is still comfortably running within its linear operational range, ensuring ample SNR and D/R values so that any loss in coherence can be attributed to distortion with reasonable certainty.

On an analyzer that displays MSC (Eq. 1) this yields coherence values of 97% or more whereas on an analyzer that displays MC (Eq. 2) this equals 98% or more.

1.3.1 M-Noise Excitation Signal

Knowing a signal's peak-to-average ratio or crest factor that a loudspeaker needs to reproduce, is important for determining the maximum level the loudspeaker can attain with such a signal.

Multiple test signals exist which feature an RMS level as a function of frequency meant to represent music but, until now, none of them feature crest factor as a function of frequency.

To measure the onset of non-linearity for loudspeakers, due to compression or distortion, a test signal must be used which exhibits the frequency dependent crest factor found in music as well as its average spectrum.

A large variety of music samples have been analyzed and it has been found that the crest factor increases with frequency without exception.

For this reason, a new test signal has been created [3] called M-Noise, whose crest factor increases with frequency as well.

Figure 2 shows the average spectra of M-Noise as well as other test signals.

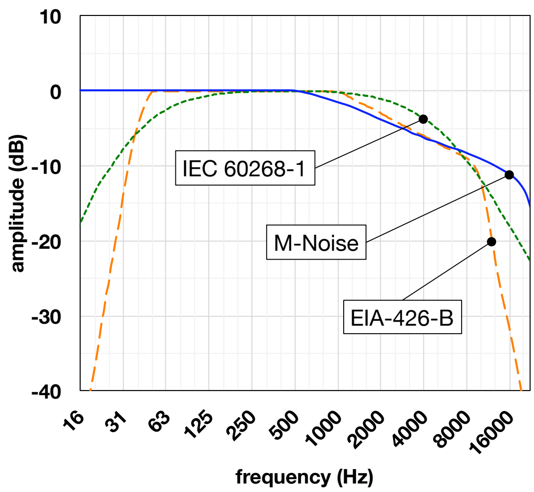

Figure 2. Various test signal spectra.

Figure 3 shows the crest factor of M-Noise as a function of frequency. Notice how the crest factor increases with frequency.

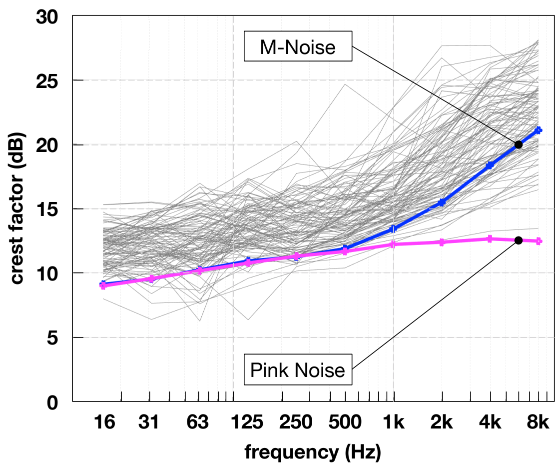

Figure 3. Crest factor per octave for pink noise, music and M-Noise.

Last but not least, music is sparse, i.e., not all notes are played at the same time. In contrast, M-Noise has energy at all frequencies, so that it is always possible to calculate the transfer function at each frequency.

For these reasons, M-Noise will be used for all subsequent measurements. Audio examples with clearly defined amounts of added harmonic distortion are presented during the e-Brief presentation and will be available from the authors subsequently.

2 Method

To determine distortion-induced coherence loss by non-linear loudspeaker behavior qualitatively, a proprietary analog distortion generator was used as "proxy" system under test (SUT).

2.1 System under Test

The distortion generator performs like a unity gain device which can be configured to introduce THD ranging from 0% to 100% independent of the drive level. Figure 4 shows a distortion graph of the distortion generator set to 10%, using a stepped sine wave excitation signal and a spectrum analyzer running at a sample rate of 96 kHz. Notice that the distortion generator adds even-order harmonics.

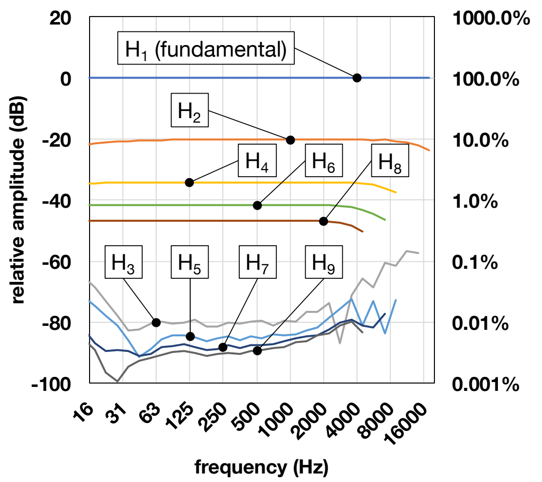

Figure 4. Distortion graph of distortion generator set to 10% THD.

2.2 Measurement Setup

Transfer functions, including coherence (MSC), of the distortion generator were obtained using a Dual-Channel FFT Transfer Function Analyzer featuring 48 points-per-octave resolution. All excitation signals were played back from a Digital Audio Workstation (DAW) which at the same time was also used to make high-quality recordings of the audio coming back from the distortion generator. Signal routing and conversion were done using a Loudspeaker Management System (LMS).

Figure 5 shows a diagram of the entire measurement setup where the entire digital side ran at a bit depth of at least 24 bits and a sample rate of 96 kHz except for the analyzer which runs at 48 kHz.

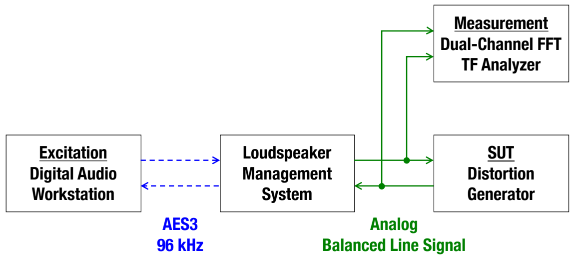

Figure 5. Diagram measurement setup.

3 Results

Figure 6 on the next page shows steady-state MSC values over frequency, accumulated over long time periods, for various amounts of THD using M-Noise as the source.

These trends are in very good agreement with results obtained using music as the source which are notoriously hard to capture in a single image due to the relatively sparse content and volatile behavior of music.

The coherence losses are proportional to frequency because distortion components accumulate at high frequencies while the RMS level of M-Noise decreases with frequency.

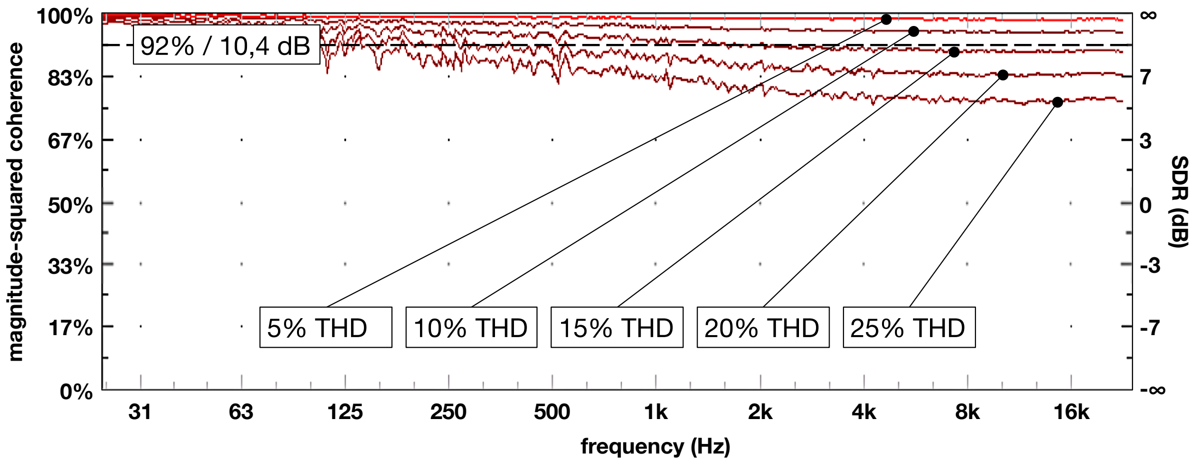

Figure 6. Transfer functions only showing MSC (Eq. 1) for various amounts of THD.

More importantly, this demonstrates that a loss of coherence that may appear to be trivial should be taken very seriously when the only known cause is distortion. Especially on analyzers that display MC (Eq. 2) instead of MSC (Eq. 1). Minor coherence losses indicate the onset of non-linear behavior in terms of distortion and justify further in-depth investigation using conventional means such as THD and THD+N.

4 Discussion

The M-Noise Procedure contains a criterion which states that coherence shouldn't drop below 91% (10 dB SDR) on an analyzer which displays MSC (Eq. 1) and 95% on an analyzer which displays MC (Eq. 2). Unlike THD or THD+N there currently is no good performance-metric for condensing coherence values, determined using wide-band signals, into a single figure. However, the authors would like to remind the readership of the work done by *Temme & Brunet* [4] who proposed Total Non-Coherent Distortion (TNCD)

$$TNCD = \sqrt{\frac{\sum_f \left\{\left[1-\gamma^2(f)\right] \cdot P_{YY}(f)\right\}}{\sum_f P_{YY}(f)}} \quad (4)$$

TNCD is an extension of THD+N for a multitone and could possibly augment established methods for determining distortion.

References

[1] M-Noise Procedure
Meyer Sound Part No.: 02.916.050.01 B2
https://m-noise.org/

[2] G. Carter et al., "Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing", *IEEE Transactions on Audio and Electroacoustics*, vol. 21, no. 4, pp. 337—344 (1973).

[3] R. Schwenke, "A New Signal for Measuring Loudspeaker Maximum Linear SPL", SMPTE 2019 Annual Technical Conference, Los Angeles, October 2019 (submitted).

[4] S. Temme, P. Brunet "A New Method for Measuring Distortion using a Multitone Stimulus and Non-Coherence", AES 121st Convention, October 2006, Paper 6877.