Audio Engineering Society

Convention Paper

Presented at the 123rd Convention
2007 October 5–8 New York, NY

Time-Frequency Characterization of Loudspeaker Responses Using Wavelet Analysis

Daniele Ponteggia[1], Mario Di Cola[2]

[1] *Audiomatica, Firenze, 50136, Italy*

[2] *Audio Labs Systems, Milano, 20068, Italy*

Correspondence should be addressed to Daniele Ponteggia (dp@audiomatica.com)

ABSTRACT

An electro-acoustic transducer can be characterized by measuring its Impulse Response (IR). Usually the collected IR is then transformed by means of the Fourier Transform to get the complex frequency response. IR and complex frequency response form a pair of equivalent views of the same phenomena. An alternative joint time-frequency view of the system response can be achieved using Wavelet Transform and a color-map display. This work illustrates the implementation of the Wavelet Transform into a commercial measurement software and presents some practical results on different kinds of electro-acoustic systems.

1. INTRODUCTION

This work stems from a previous work on linear phase crossover filters [1], where the authors wanted to visualize in a simple way the time smearing imposed by different crossover strategies. During that research we realized that the available tools to investigate the time behavior of loudspeaker systems are not easy to manage and if not used in a proper manner they can also be misleading.

The phase-time relationship of electro-acoustic systems is well documented in the technical literature. Meanwhile, probably due to the fact that this relationship is not easy to manage, the phase response is still often dismissed as less important than the magnitude response. It is our direct experience, confirmed by recent studies [2], that the distortion produced by the time smearing of the response of a loudspeaker system is correlated with the quality of the reproduced sound.

The use of the Wavelet theory in this field is not new: previous works in the application of this theory to the analysis of loudspeaker impulse responses must be credited to Keele [3], Gunness [4] and Loutridis [5]. The last of the above listed articles features a very detailed treatment of the subject and a good bibliography, even though in this article the visualization of the analysis results was partially missing. The use of waterfall displays is, in the opinion of the authors, not that clear to understand. Representations as color maps are generally preferable. In this respect, the work carried out by David Gunness was certainly a source of inspiration. This is not only due to the presentation of the analysis results in a good graphical form but also for his approach to the calculation of the time-frequency distribution through complex smoothing and, ultimately, for the applications of the new tool to the analysis and improvement of loudspeaker systems response.

The Wavelet theory is a quite recent topic, it was initially developed for seismic and geological applications, but its use has spread out in different fields. There is currently an enormous amount of scientific articles and books on this topic: a web search of the *wavelet* word can return thousands of hits. Such amount of resources may create confusion in the beginner, with the result of a very steep learning curve. We found particularly useful the article by Rioul and Vetterli [6] as a comprehensive and concise introduction on the topic.

2. LOUDSPEAKER CHARACTERIZATION

A loudspeaker, at least in its linear part, can be described by means of its Impulse Response IR. The IR is usually collected using computer based measuring instruments, placing the microphone at some distance in front of the loudspeaker system that is being tested (DUT)(Figure 1).

Due to the fact that the IR is stored into a computer, it is possible to post-process the data easily using Fast Fourier Transform FFT algorithms, passing back and forth between the IR and the complex frequency response. It is also possible to perform more complex post-processing as the calculation of the envelope of the analytic impulse (Energy Time Curve ETC) [7], joint time-frequency analysis as the Cumulative Spectral Decay CSD and the Short Time Fourier Transform STFT [8].

The IR and the complex frequency response are a pair of views of the same phenomena, linked by means of the Fourier Transform FT. Due to the nature of the FT, which is based on a kernel of infinite energy signals, it is difficult to get time information from the complex frequency response and vice-versa it is difficult to obtain frequency information from the IR.

In this work we will use the Fourier transform pair in its radial form:

$$h(t) \overset{\mathcal{F}}{\Longleftrightarrow} H(\omega) \tag{1}$$

$$H(\omega) = \int_{-\infty}^{\infty} h(t) e^{-j\omega t} dt \tag{2}$$

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega) e^{j\omega t} d\omega \tag{3}$$

2.1. Impulse Response

The Impulse Response $h(t)$ of a loudspeaker system describe completely the linear behavior of the system. By means of the convolution integral it is possible to predict the output $y(t)$ of the system given any stimulus $x(t)$. Looking at the IR the time behavior of the system can be easily understood. By the way, detailed information regarding the frequency domain cannot be easily read from it. By inspecting an impulse response, an experienced technician can just gain some rough indications on the frequency domain.

As an example, the IR of a two-way professional loudspeaker box is reported here (Figure 2). It can be clearly observed that the highs and the lows are not well packed in time, but at the same time, just looking at the IR graph it could not be that easy to state how many millisecond they are offset, or which frequency bands are involved.

In order to gain more information in the time domain there are some post-processing of the IR that can be used, we will show here the Step Response and the Energy Time Curve ETC.

The STEP response is, as implied by its name, the response of the system to a step function; it is a qualitative measure of the time coherence of a loudspeaker system but gives little information on the

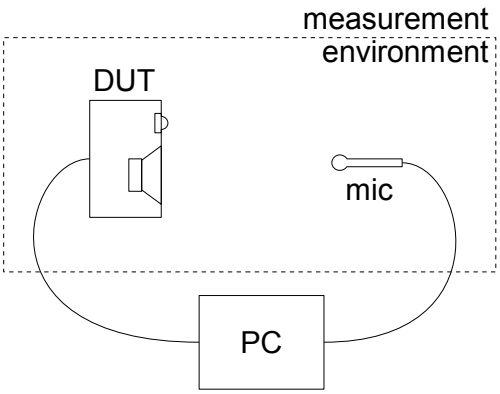

Fig. 1: Typical PC based measurement setup

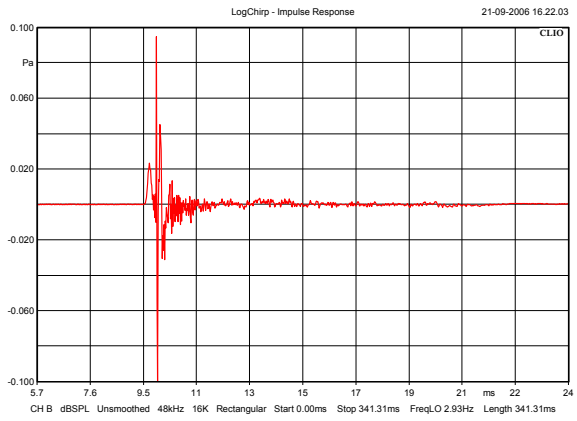

Fig. 2: Impulse response of a two-way loudspeaker box

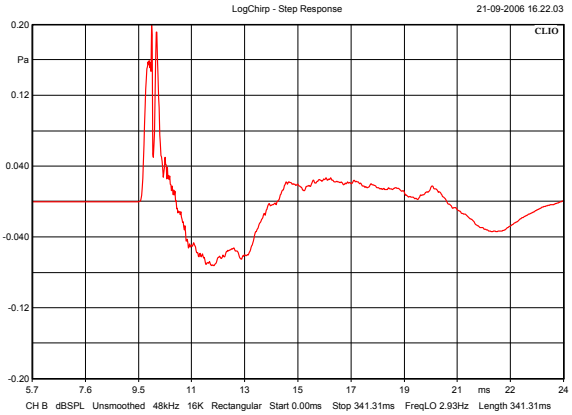

Fig. 3: Step response of a two-way loudspeaker box

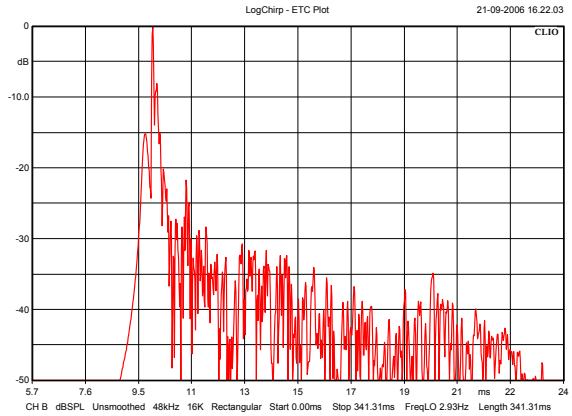

Fig. 4: ETC plot of a two-way loudspeaker box

frequency domain. The Step response of the two-way system is shown in figure 3.

The ETC is defined as the envelope of the analytic impulse. The analytic impulse is a complex quantity, with real part equal to the IR and imaginary part calculated through Hilbert transform of the IR. The ETC does not directly represent the sound energy. The ETC can be interpreted as a measure of the magnitude of the Impulse Response, since the envelope removes the negative ripples of the IR. From our point of view we can state that there are no relevant frequency information into the ETC plot. Figure 4 shows an ETC plot of the previous two-way loudspeaker system.

None of these time representations can give us clear information on the frequency domain.

2.2. Complex Frequency Response

The Complex Frequency Response $H(\omega)$ of a system is the Fourier Transform of the impulse response $h(t)$. It is another complete description of the linear behavior of the loudspeaker system. By means of the product of the transform of the input signal $X(\omega)$ and the Complex Frequency Response $H(\omega)$ it is possible to predict the output $Y(\omega)$ of the system.

Some insights on the time behavior of a system can be obtained by looking at the complex frequency response. It is common practice to check the time alignment of a loudspeaker system by looking at the behavior of its phase response.

In fact, if we define the *perfect system* as a system featuring flat magnitude and linear phase in the band of our interest, then the system will show no linear amplitude distortion as well as no time smearing and consequently no phase distortion (with the exception of the eventual total delay delay that it would remain constant as function of frequency). The Impulse Response of a system of this kind, it will appear as a well packed sinc-like function shifted in time by the system delay.

A direct simple relationship between phase and time takes place only if the system features the response of an all-pass system. In this case the negative derivative of the phase response (group delay) exactly represents the system's time delay. If a system is not an all-pass it is not possible to define a single arrival time as function of frequency. Even simple systems like it could be an high pass filter, the energy is not arriving at a single time instant [9] [10].

Generally speaking, some time spreading of the energy at the output of a system usually occurs. If we drive a system with an impulse we can measure at the system's output a modified version of the impulse: the frequency components may have modified levels and some phase distortion probably occurs resulting in the time smearing of the impulse.

One of the main difficulties that can be faced while trying to interpret the phase response in a Complex Frequency Response lies in the measurement conditions and in the assumptions we make doing the measurement as well. While measuring a loudspeaker system, we actually measure the combined response of the loudspeaker system itself and the room response. In a more general approach we need to introduce into the model of the measurement the entire audio chain from the source to the receiver. We will consider here a simplified model in which we include only the loudspeaker and the direct sound propagation, assuming other components being linear, the room being anechoic (or quasi-anechoic using IR windowing) and the air being still and quiet.

The measured response is composed of the loudspeaker system response $H_{\mathrm{DUT}}(\omega)$ and a propagation term $H_{\mathrm{propagation}}(\omega)$:

$$H_{\mathrm{meas}}(\omega) = H_{\mathrm{DUT}}(\omega) \cdot H_{\mathrm{propagation}}(\omega) \qquad (4)$$

If the distance between the loudspeaker under test and the measuring microphone is less than few meters we can neglect the frequency dependent effect of the air absorption and consider the term $H_{\mathrm{propagation}}(\omega)$ as a simple delay. Even using this assumption, still remains very difficult to recover the loudspeaker system response, since it is buried into the propagation component (Figure 6).

There are some different ways to remove the propagation part and recover the system response: however it is necessary to create a model for the measurement and make some a priori assumptions.

One of the possible solutions is to model our system as a point source with a known position (figure 5) and measure the distance d between the reference position RP of the source and the microphone.

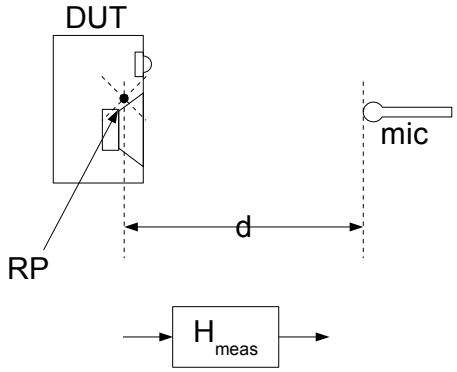

Fig. 5: Modeling the source as a point with known reference position

Once the distance d is known it is possible to calculate the delay time t_{delay}:

$$t_{\text{delay}} = \frac{d}{c} \tag{5}$$

then the delay t_{delay} can be mathematically removed from the $H_{\text{meas}}(\omega)$ by means of the relation:

$$H_{\text{DUT}}(\omega) = H_{\text{meas}}(\omega) e^{j\omega t_{\text{delay}}} \tag{6}$$

An alternative solution used by some measurement systems (particularly by those measuring the live transfer function) is to refer the delay value at the time where the impulse response reach its maximum. Figure 7 shows the Complex Frequency Response after the removal of the delay using the time of peak value of the IR as delay.

A third solution consists in examining the excess phase group delay of the system. In fact, the use of the Hilbert transform makes possible to separate the measured phase response $H_{\text{meas}}(\omega)$ into two separated contributions: a minimum phase part $H_{\text{minimum}}(\omega)$ and an excess phase part:

$$H_{\text{meas}}(\omega) = H_{\text{minimum}}(\omega) \cdot H_{\text{excess}}(\omega) \tag{7}$$

$$H_{\text{meas}}(\omega) = H_{\text{minimum}}(\omega) \cdot H_{\text{all pass}}(\omega) \cdot H_{\text{delay}}(\omega) \tag{8}$$

The group delay calculated from the excess phase part yields a time delay as a function of frequency. If it is possible to find a band where the excess phase

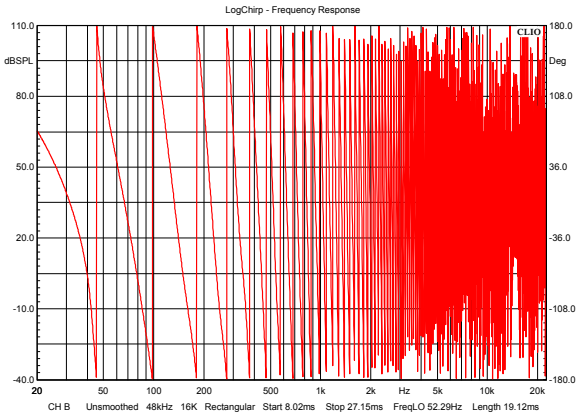

Fig. 6: Phase Response of two-way loudspeaker system, delay not removed

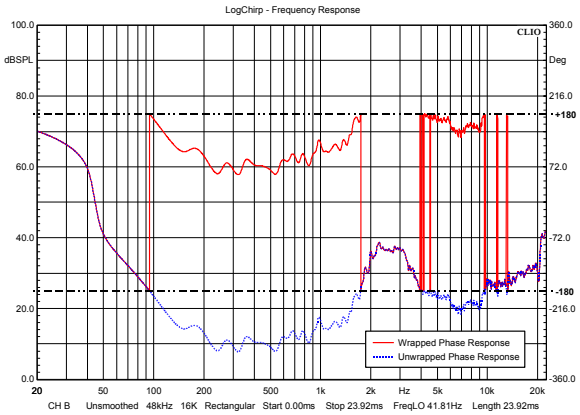

Fig. 7: Phase response two-way loudspeaker system, delay removed

group delay is constant, this means that the measured $H_{\text{meas}}(\omega)$ is minimum phase in that band. Using this constant value as delay is possible to separate the delay $H_{\text{delay}}(\omega)$ from the system response.

Luckily the exact knowledge of the phase part of the Complex Frequency Response usually is not strictly necessary. The measured responses are used into simulation and design software applications that are quite robust to the exact delay of the system. In fact, they usually can handle relative measurements that can be referred to a common reference point or delay.

Assuming that we have correctly removed the measurement delay, there is still a source of uncertainty in the reading of phase plots. The phase response, in fact, is usually represented in a log frequency scale plot because it generally share the same frequency scale used for the magnitude response. And even though this could be helpful in some cases because deviations from some specific behavior could be clearly pointed out (i.e. when the phase plot is flat), at the same time, the interpretation of the general case of phase linearity in a logarithmic scale plot can be sometime very hard. As an example in figure 9 is plotted the response of a pure delay in both scales.

It clearly appears that both the IR and the Complex Frequency Response are not suited to the analysis of the joint time-frequency characteristic of the system response, they must be thought as pictures of the same phenomena taken from two completely different angles.

2.3. Cumulative Spectral Decay

The Cumulative Spectral Decay CSD is a view of the frequency content of the energy decay of the system. The CSD is calculated by means of Fourier Transform of progressively shorter sections of the IR. Phenomena that last longer remain included in later slices. Then, looking at the evolution of the frequency response modulus, it is possible to detect long lasting components of the IR.

An example of CSD representation in a waterfall plot is in figure 10 for our two-way loudspeaker system.

2.4. Short Time Fourier Transform STFT

The idea behind the STFT consist in following the temporal evolution of the IR, by means of the sec-

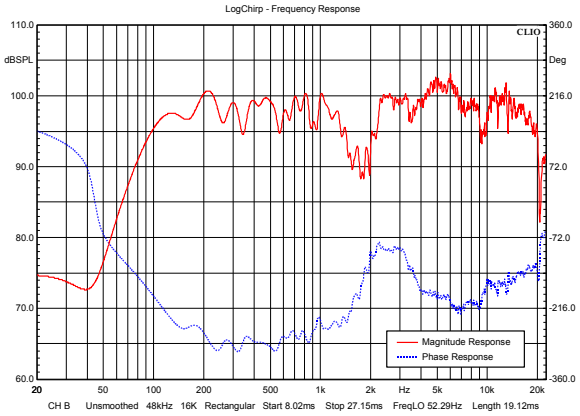

Fig. 8: Complex Frequency response two-way loudspeaker system, delay removed

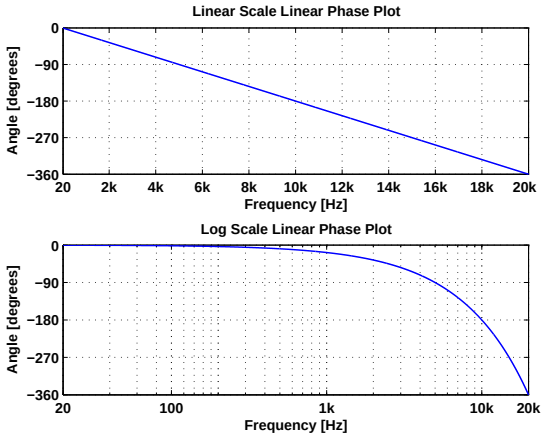

Fig. 9: Linear phase (delay) represented in linear and logarithmic scale plots

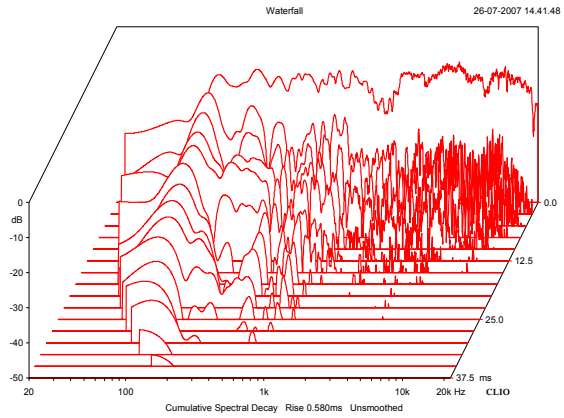

Fig. 10: CSD waterfall of two-way loudspeaker system

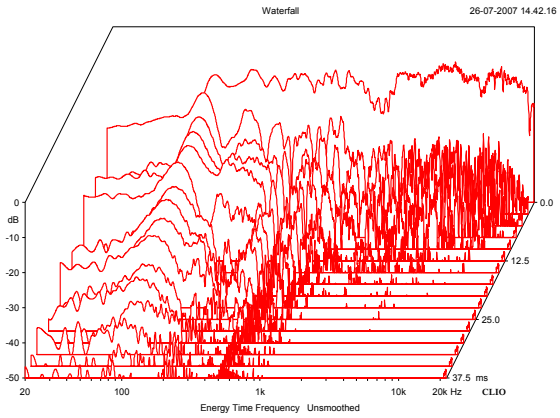

Fig. 11: STFT waterfall of two-way loudspeaker system

tioning the IR itself and the application of Fourier Transform to its pieces.

$$\mathcal{F}_h(\tau,\omega) = \int_{-\infty}^{\infty} h(t)\overline{g(t-\tau)}e^{-j\omega t}dt \qquad (9)$$

The problem that arises with the STFT is that its time-frequency resolution is fixed over the entire time-frequency plane; the same window is then applied to all frequencies. Therefore, the choice of a short FFT window yields too low frequency resolution and this it could be too coarse at low frequencies. On the other hand, the choice of a long FFT window improves the resolution but it also ends up in loosing time resolution. The STFT could be a valid tool for the analysis of either narrow band or short duration events. It is generally of little help if applied at the complete analysis of wide band and long duration signals as the overall Impulse Response of a loudspeaker system is.

To overcome the limitations of the STFT it is possible to think to its expansion as multi-resolution STFT. This can be done computing a set of STFTs for different sizes and recombining the results into the same plot. Some successful attempt in pattern recognition and musical instruments identification were recently introduced [11] but they were not applied to loudspeaker characterization.

It must be also cited another joint time-frequency distributions: the Wigner-Ville. This distribution was already used for loudspeaker system analysis [12], but it exhibits cross-components artifacts that are not present in the STFT and in the Wavelet Analysis.

3. CONTINUOUS WAVELET TRANSFORM

The Continuous Wavelet Transform CWT of a signal $h(t)$ is defined as the inner product:

$$\mathcal{W}_h(a,b) = \langle h(t), \psi_{a,b}(t)\rangle = \int_{-\infty}^{\infty} h(t)\overline{\psi_{a,b}(t)}dt \qquad (10)$$

between $h(t)$ and translated and scaled version of a function $\psi(t)$ called Mother Wavelet:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \qquad (11)$$

where the factor $1/\sqrt{|a|}$ is introduced to ensure energy normalization for every scale a.

Thus the CWT can be wrote as:

$$\mathcal{W}_h(a,b) = \frac{1}{\sqrt{|a|}}\int_{-\infty}^{\infty} h(t)\overline{\psi\left(\frac{t-b}{a}\right)}dt \qquad (12)$$

In order to ensure perfect reconstruction by means of inverse CWT trough the coefficients $\mathcal{W}_h(a,b)$ the

Mother Wavelet must satisfy the admissibility condition:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \qquad (13)$$

where $\Psi(\omega)$ is the Fourier Transform of $\psi(t)$.

As a consequence of the admissibility condition, the $\psi(t)$ must be a Bandpass Impulse Response, since this kind of signal looks like a small wave, the transform is named Wave*let* [13]. There are few different interpretation of the Wavelet Transform, basically it consist on a mapping of the one-dimensional time domain of the impulse response $h(t)$ into a two dimensional space consisting of a scale a and a translation b [14].

Since the Mother Wavelet function $\psi(t)$ is scaled and not modulated, like the STFT kernel, the Wavelet Analysis is called time-scale analysis rather than time-frequency. Variation of parameters a and b has no effect on the form of the transform kernel, but time and frequency resolution depends on a. For high analysis frequency (low scale a) we have good time localization but poor frequency resolution, on the other hand for low frequency (high scale a) we have good frequency but poor time resolution. Thus the Wavelet Analysis can be understood as a constant-Q analysis and, for this reason, it could be considered a perfectly suitable tool to inspect non-stationary wide band signals as the impulse response of a loudspeaker system.

3.1. Scalogram

The *Spectrogram* it is a well known tool for showing the energy of a signal in the time-frequency plane and is defined as the squared modulus of the STFT.

In a similar way it is possible to define the *Scalogram* as the squared modulus of the CWT. The energy of the signal can be mapped as related to the scale a and translation b:

$$E_h = \int_{-\infty}^{\infty} h^2(t)dt = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{0}^{\infty} |\mathcal{W}_h(a,b)|^2 \frac{dadb}{a^2} \qquad (14)$$

The quantity $|\mathcal{W}_h(a,b)|^2$ can be represented in a color map plot, with time in the x axis and frequency scale in the y axis, using the color gradation in order to represent in each point of the plot the associated

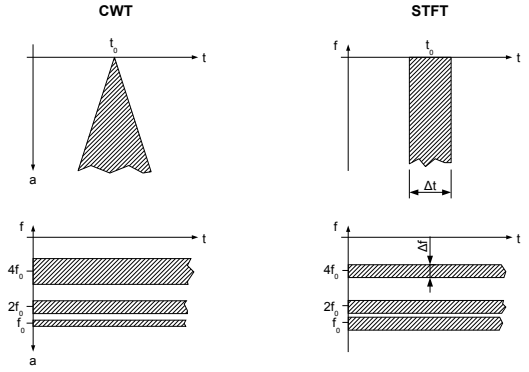

Fig. 12: Comparison of Wavelet and STFT resolutions: region of influence of a Dirac pulse (top) and three sinusoids (bottom)

magnitude. In fact these are modified versions of the parameter scale a and translation b. Once the Mother Wavelet has been defined, it is possible to define a center frequency for the Wavelet thus expressing scale a in terms of an equivalent frequency. The parameter b can be directly interpreted as time.

3.2. Time-Frequency resolution

The idea behind the CWT is to use a set of basis functions that feature an optimal time-frequency resolution. Is it possible to define the temporal t_ψ and spectral ω_ψ centers of the wavelet basis function as:

$$t_\psi = \int_{-\infty}^{\infty} t \frac{|\psi(t)|^2}{\|\psi\|^2} dt \qquad (15)$$

$$\omega_\psi = \int_{-\infty}^{\infty} \omega \frac{|\Psi(\omega)|^2}{\|\Psi\|^2} d\omega \qquad (16)$$

Is it also possible to define the temporal Δ_{t_ψ} and spectral Δ_{ω_ψ} widths as:

$$\Delta t_\psi = \sqrt{\int_{-\infty}^{\infty} (t - t_\psi)^2 \frac{|\psi(t)|^2}{\|\psi\|^2} dt} \qquad (17)$$

$$\Delta \omega_\psi = \sqrt{\int_{-\infty}^{\infty} (\omega - \omega_\psi)^2 \frac{|\Psi(\omega)|^2}{\|\Psi\|^2} d\omega} \qquad (18)$$

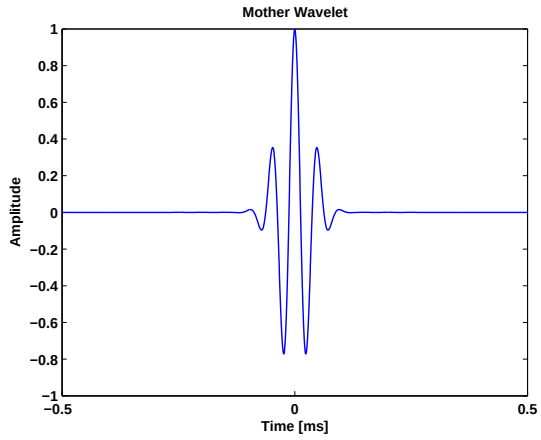

Fig. 13: Mother wavelet (Real Part) with 1/3 octave BW

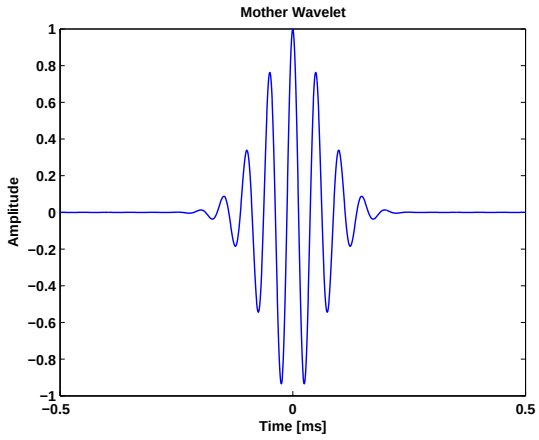

Fig. 14: Mother wavelet (Real Part) with 1/6 octave BW

The time-frequency product is bounded by the time-frequency uncertainty principle:

$$\Delta_{t_\psi}\Delta_{\omega_\psi} \geq \frac{1}{2} \tag{19}$$

The function that achieves the minimum time-bandwidth product is the Gaussian pulse. In fact its energy is reasonably concentrated both in the time and the frequency domain, but at the same time, it is not limited in both of the domains. Therefore, a function which is a modulated Gaussian pulse is definitely a good candidate as a Mother Wavelet.

3.3. Mother Wavelet

In agreement with Loutridis, we chose to use as a Mother Wavelet a modified Morlet complex wavelet:

$$\psi(t) = \frac{1}{\sqrt{\pi B}} e^{j\omega_0 t} e^{-t^2/B} \tag{20}$$

It has to be noted that this function is composed of a complex oscillating term $e^{j\omega_0 t}$ and a decaying factor $e^{-t^2/B}$, as requested by the admissibility condition. We also need to note that, in our specific application, the admissibility condition can be relaxed, since we are interested in the analysis of the property of the signal and not into its reconstruction.

The Fourier Transform of the Mother Wavelet function is:

$$\Psi(\omega) = e^{-(\omega-\omega_0)^2 B/4} \tag{21}$$

Time and frequency resolutions of the chosen Wavelet can be calculated from the previous definitions:

$$t_\psi = 0 \tag{22}$$

$$\omega_\psi = \omega_0 \tag{23}$$

$$\Delta t_\psi = \frac{\sqrt{B}}{2} \tag{24}$$

$$\Delta\omega_\psi = \frac{1}{\sqrt{B}} \tag{25}$$

Is it possible to refer the B parameter to the analysis bandwidth BW and Q:

$$\mathrm{BW} = \frac{1}{Q} = \frac{2\Delta\omega_\psi}{\omega_\psi} = \frac{2}{\omega_0\sqrt{B}} \tag{26}$$

$$B = \frac{4}{(\omega_0 \mathrm{BW})^2} = 4\left(\frac{\mathrm{Q}}{\omega_0}\right)^2 \tag{27}$$

4. COMPUTATION OF WAVELET COEFFICIENTS

The calculation of the Wavelet Coefficients directly from equation (12) is very expensive from the computational point of view.

At the beginning of this work unsatisfactory attempts were made to brute force calculate the equation in Matlab.

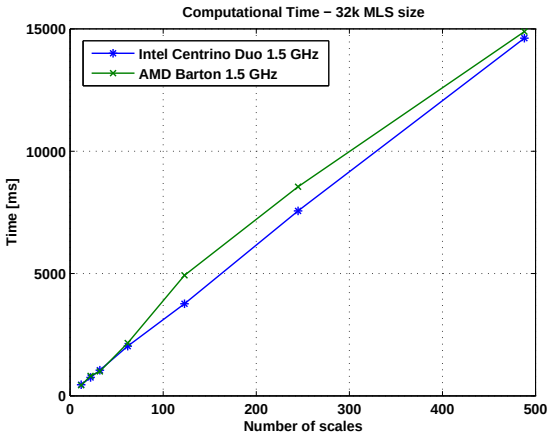

Fig. 15: Computational time - 32k MLS

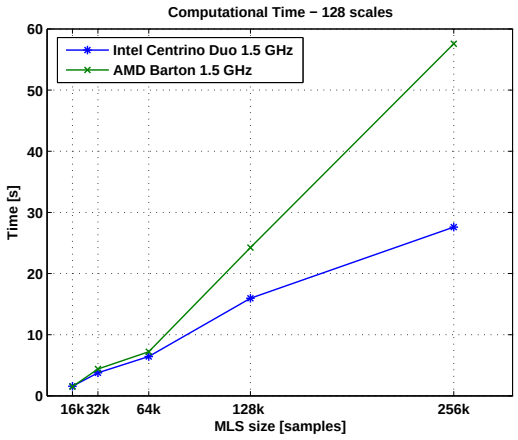

Fig. 16: Computational time - 128 scales

4.1. Calculation Algorithm

An alternative approach based on conventional Fourier Transform technique can be used. By means of the FT and its time and frequency shift properties, it is possible to obtain the FT of the Wavelet kernel function:

$$h(t) \quad \overset{\mathcal{F}}{\Longleftrightarrow} \quad H(\omega) \tag{28}$$

$$\psi(t) \quad \overset{\mathcal{F}}{\Longleftrightarrow} \quad \Psi(\omega) \tag{29}$$

$$\psi_{a,b}(t) \quad \overset{\mathcal{F}}{\Longleftrightarrow} \quad \Psi_{a,b}(\omega) \tag{30}$$

$$\frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \quad \overset{\mathcal{F}}{\Longleftrightarrow} \quad \sqrt{|a|}e^{-j\omega b}\Psi(a\omega) \tag{31}$$

From equation (10) and using Parseval's relation we get:

$$\begin{aligned}\mathcal{W}_h(a,b) &= \langle h(t), \psi_{a,b}(t)\rangle \\ &= \langle H(\omega), \Psi_{a,b}(\omega)\rangle \\ &= \sqrt{|a|}\int_{-\infty}^{\infty} H(\omega)\overline{\psi(a\omega)\, e^{-j\omega b}}d\omega \\ &= \sqrt{|a|}\int_{-\infty}^{\infty}\left[H(\omega)\overline{\psi(a\omega)}\right]e^{j\omega b}d\omega\end{aligned} \tag{32}$$

that is in the form of an inverse Fourier Transform of the product of the frequency content of the signal $H(\omega)$ and the conjugated complex of the scaled Wavelet function $\Psi(\omega)$, except for an $1/\sqrt{2\pi}$ factor.

Given the scale factor a, is it possible to calculate CWT coefficients using conventional numerical implementation of the Fourier Transform, direct Fast Fourier Transform FFT and Inverse Fast Fourier Transform IFFT:

$$\mathcal{W}_h(a,b) = \sqrt{|a|}\text{IFFT}\left[\text{FFT}\left[h(t)\right]\overline{\Psi(a\omega)}\right] \tag{33}$$

Iterating this numerical procedure for a number of a values leads to a bi-dimensional matrix of coefficients $\mathcal{W}_h(a,b)$ that can be then post-processed to produce color map plots. Since the graphical representation resolution is linked to the screen resolution of the output media (typically a computer screen) there is no need to calculate more than few hundreds of scales a.

The Wavelet Analysis tool was implemented into the CLIO measurement system, and, in order to achieve an effective functionality of the tool a set of computation speed tests were performed (Figure 15 and 16).

We found experimentally that the best result in terms of speed and analysis accuracy can be achieved using a number of calculated scales between 2^7 and 2^8, in these cases computational time is always under 60 seconds. While in theory a post-processing can theoretically run for an almost unlimited amount of time, we believe that a fairly fast response is important to the practical application of the Wavelet Analysis tool.

5. GRAPHICAL REPRESENTATION

To represent the Scalogram, as above mentioned, we chose to use a color map plot. In our opinion, this representation yields a clear view of the evolution of the energy flow into the system. While in Level vs. Time and Level vs. Frequency plots only one domain can be clearly shown, the color map plot of the Scalogram clearly shows the joint time-frequency domain. Every attempt to interpret the display in terms of only one domain will be unsuccessful.

Once the Wavelet coefficients $\mathcal{W}_h(a, b)$ are calculated for a given number of scales, i.e. for a given number of frequencies, the Scalogram can be plotted. The color map is plotted using bilinear interpolation from the available frequency-time points.

5.1. Normalization

We also applied scale normalization to the Scalogram to gain a better graphical output thus rendering it more understandable. This normalization has the same result of ignoring the energy normalization introduced into equation (11).

Figure 17 shows the Wavelet Analysis Scalogram of a Dirac impulse. We recall that this is the response of a perfect system; therefore this is the target Wavelet Analysis response of every linear loudspeaker system.

Figure 18 shows the Wavelet Analysis of the two-way professional loudspeaker system illustrated before.

If the Wavelet Analysis is applied in studying the system decay or its time alignment, the Scalogram can be normalized in level. In this case every row of the Scalogram, where the coefficient are calculated by equation (33), is normalized to its maximum value. This results in energy differences between the different scales being neglected, resulting in a clearer view of the energy decay of the system. This is very similar to the computation of a set of narrow-band ETC curves.

As an example we show in figure 19 the time-normalized Wavelet Analysis for the same two-way system.

It can also be interesting to plot the peak energy arrival time curve, i.e. the curve connecting the maximum energy points for every row of the Scalogram. Figure 20 shows the Peak Energy Arrival Time Curve as a dotted-dashed line superimposed

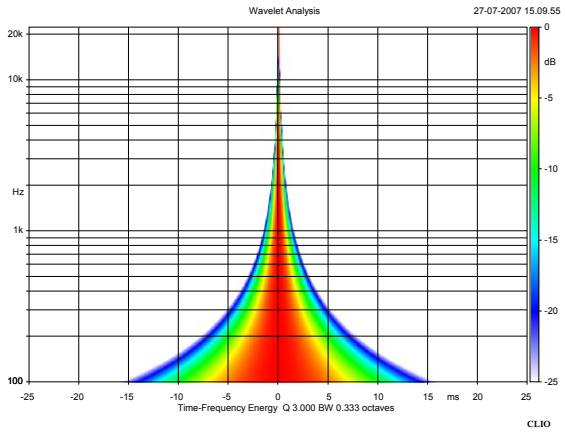

Fig. 17: Wavelet Analysis of a perfect system

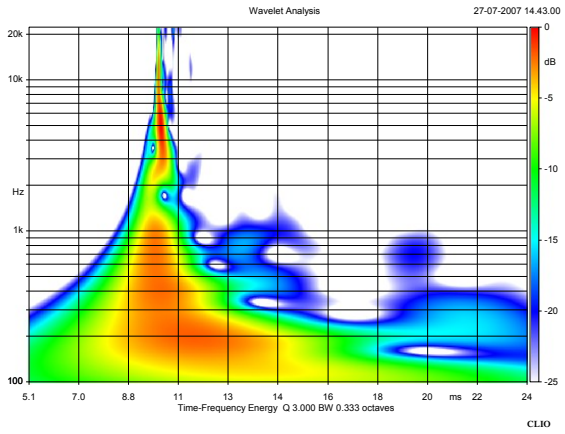

Fig. 18: Wavelet Analysis of two-way loudspeaker box, 1/3 octave BW

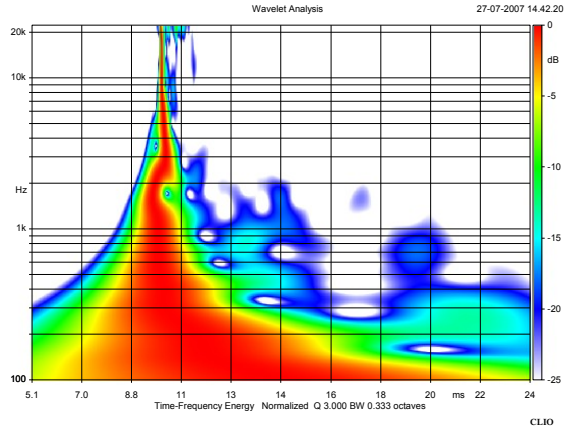

Fig. 19: Wavelet Analysis of two-way loudspeaker box, 1/3 octave BW, with level normalization

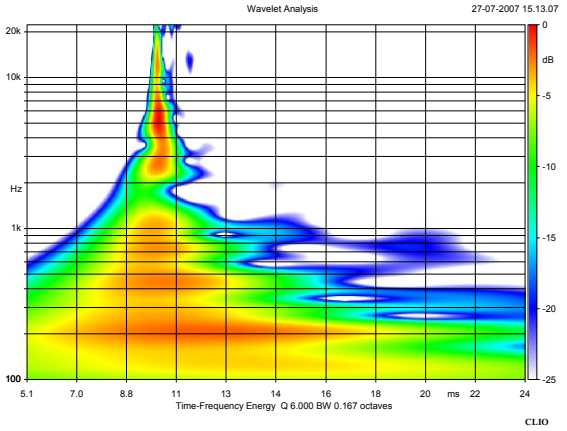

Fig. 21: Wavelet Analysis of two-way loudspeaker box, 1/6 octave BW

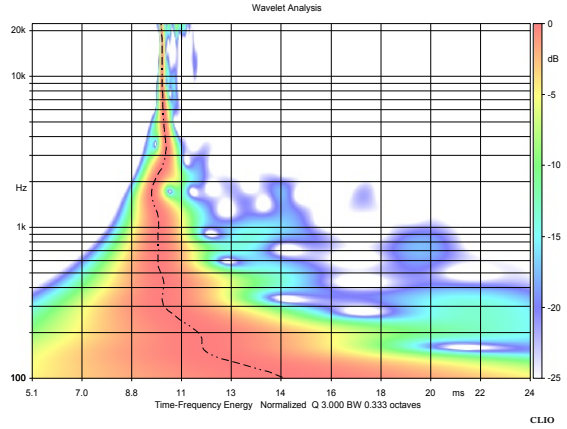

Fig. 20: Wavelet Analysis of two-way loudspeaker box, Peak Energy Arrival Time Curve

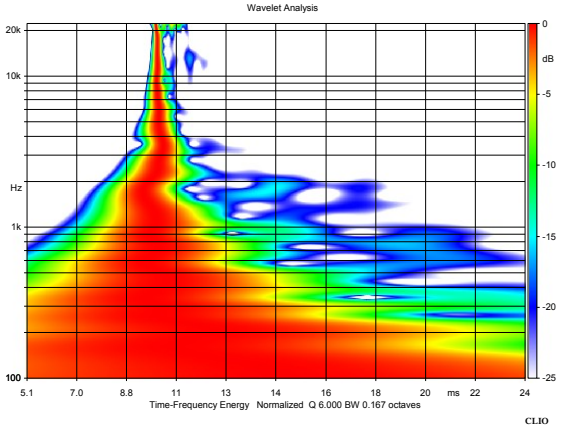

Fig. 22: Wavelet Analysis of two-way loudspeaker box, 1/6 octave BW, with level normalization

to the Scalogram plot. Into this figure the Scalogram has been intentionally faded to emphasize the curve.

5.2. Trading Frequency and Time Resolution

The product of the temporal and spectral widths is a fixed quantity; the effect of the BW parameter on the Wavelet Analysis results in trading Time Resolution with Frequency Resolution. We experienced that optimal values of BW are between of one third and one twelfth of octave.

Figure 21 show the Wavelet Analysis of the two way loudspeaker with 1/6 octave BW, notice the reduced time resolution and enhanced frequency resolution in respect to the previous 1/3 octave BW analysis.

6. APPLICATIONS

In this section we would like to show some of the real world case study where we found very useful the practical applications of the Wavelet Analysis Tool that we will call here WAT. The developing of the WAT itself have been originated from the consciousness that a joint time-frequency analysis tool was very desirable in order to be able to have a more complete sight of some time-frequency-magnitude characteristic related to specific transducers or full loudspeaker systems. The information that can be gathered analyzing an impulse response with the WAT are generally related to some specific topics: time response of a transducer by itself or eventually coupled to additional devices like waveguides, time-frequency decay of the sound energy from a loudspeaker, time alignment between drivers in a multiway loudspeaker system.

6.1. Two-way 8" Professional Loudspeaker

Lets have a rapid look to a simple two way loudspeaker system equipped with an 8" cone woofer and a 1" compression driver loaded by a proper constant coverage waveguide. This simple system could represent a quite common example of application for the WAT. We could be interested, in this case, in analyzing how two different crossover strategies can affect the time alignment between the drivers and which of the two could perform better in terms of time coherence.

It can be easily obtained, and it is exactly the case here, that the system frequency responses for both the alignments are almost identical (Figure 23). By the way, comparing the two phase response curves (Figure 24) it can be easily seen that the overall phase shift of the loudspeaker system while processed with the Linear Phase Crossover (LPC) filter (Blue Curve) exhibit a remarkably reduced phase shift if compared to the same system while processed with an All Pass Network (APN) crossover approach. In the previous paper we have shown, with more details, the time performances obtained on a loudspeaker system being processed using alternatively several different crossover strategies. In this case, looking at the results of the WAT for the two approaches, the clear evidence of improved time response that the LPC yields to the system if compared to the APN approach can be seen in Figure 26. The WAT analysis result obtained from the APN approach shown in Figure 25 demonstrates, in this case, that the energy from the high frequency unit leads that emitted from the low frequency unit by a remarkable amount of time. The main reason for the lag of low frequency information over the overall impulse response is a direct consequence of the different amount of *in-band* Group Delay associated to the Low Pass filter if compared to that of the High Pass filter. The APN approach doesn't compensate in any way this source of time spreading of the system response. The LPC approach instead, uses FIR linear phase filter that feature a fixed and constant Group Delay that has the same amount for each band; once the geometric delay is compensated, no additional time spreading is introduced in the system time response. Figure 27 shows the response of the system with LPC crossover with both the drivers connected with the proper polarity (Red Curve) and with one of the drivers polarity reversed (Blue Curve). The WAT result for the analysis of the un-properly connected system shown in figure 28 reveals a clear and evident signature of the phase cancellation phenomenon.

6.2. Three Way Professional Vertical Array Element

In order to show some more application of the effectiveness of the CLIO WAT module while studying the time alignment between drivers in a multiway loudspeaker system, an additional reference can be cited from the above mentioned previous paper. This three way vertical array loudspeaker module analyzed here is very similar to system that have

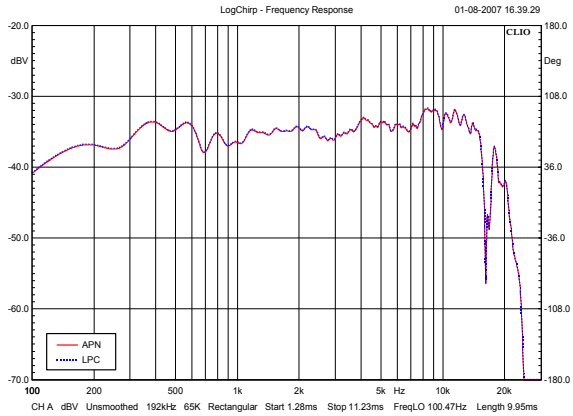

Fig. 23: Two-way 8" Professional Loudspeaker Frequency Response

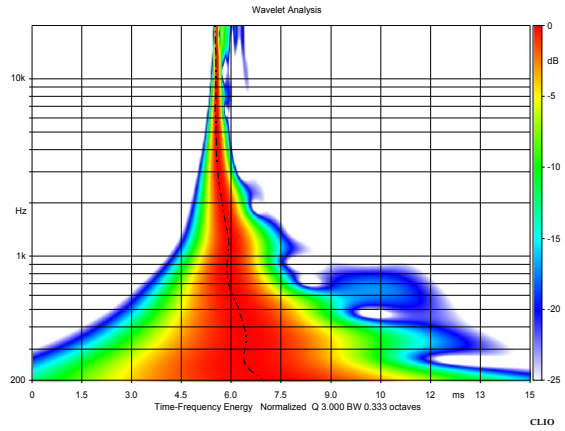

Fig. 25: Two-way 8" Professional Loudspeaker Wavelet Analysis, APN case

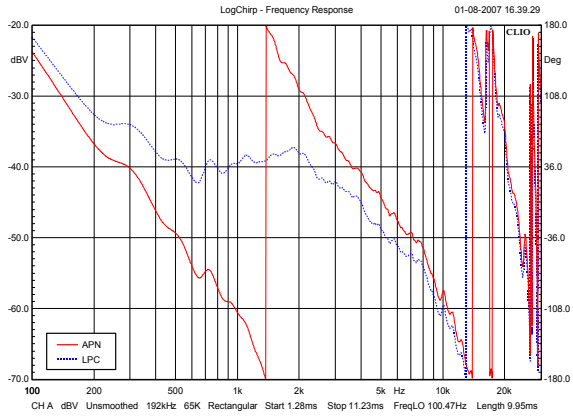

Fig. 24: Two-way 8" Professional Loudspeaker Phase Response

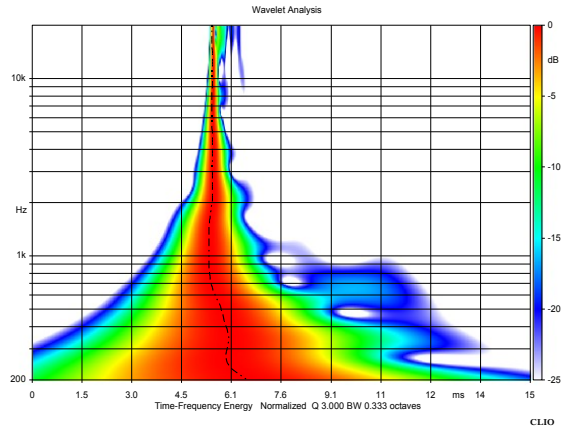

Fig. 26: Two-way 8" Professional Loudspeaker Wavelet Analysis, LPC case

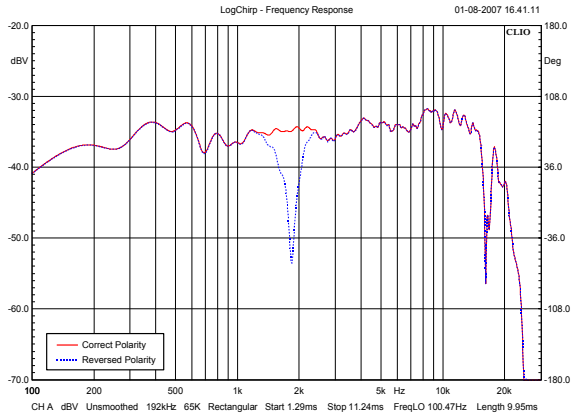

Fig. 27: Two-way 8" Professional Loudspeaker Frequency Response, LPC case

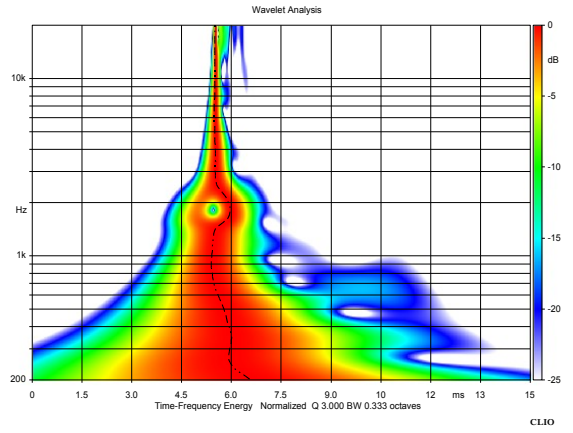

Fig. 28: Two-way 8" Professional Loudspeaker Wavelet Analysis, LPC polarity reverse case

been shown there while being of a bigger format. A quick look at the frequency response (Figure 29) plot shows that the frequency response of the system being processed by its original crossover and a crossover implemented with a FIR Linear Phase Crossover are almost identical and the small differences and not remarkable at all. The Phase Response curves are, however, remarkably different each other (Figure 30). The phase response of the original crossover is shown in the red curve while the blue curve represents the system response with the Linear Phase set. The Wavelet Analysis of IR obtained by means of the original Crossover (Figure 31) shows that the chosen crossover approach, even though has some positive effect in reducing the overall phase shift in respect to an hypothetical 3 way APN crossover, it shows some clear trace of time spreading. The midrange output, in fact, is evidently leading the high and the low frequency outputs. The Wavelet Analysis applied to the same system when processed by a Linear Phase Crossover set (Figure 32) shows unquestionably how much the time Coherence is improved and the time spreading is reduced having the most of the energy delivered at the same time through almost the full frequency span.

6.3. Compression Driver on a Constant Directivity Diffraction Horn

One of the most common features of a Constant Directivity Horn (CD) is the diffraction slot used at the horn throat. Since large format, wide coverage HF horns use compression drivers that feature an exit that's usually in the range of 1.5" or 2" of diameter, the horn cannot be entirely designed with a conical shape. The reason for this is that, because of the diameter of the drivers used, the HF directivity will be dominated by the inherent directivity of the driver itself at those frequencies. In order to maintain a wide coverage angle up to the upper end of the audio band still maintaining the use of large format HF compression driver, is common practice to couple the driver to an exponential portion of the horn that ends up in a very narrow slot that is forced to diffract in a subsequent conical section of the horn. This section can be usually followed by additional conical sections with higher expansion rate as well. The sound energy that reaches the diffraction slot is not radiated completely into the conical section. A

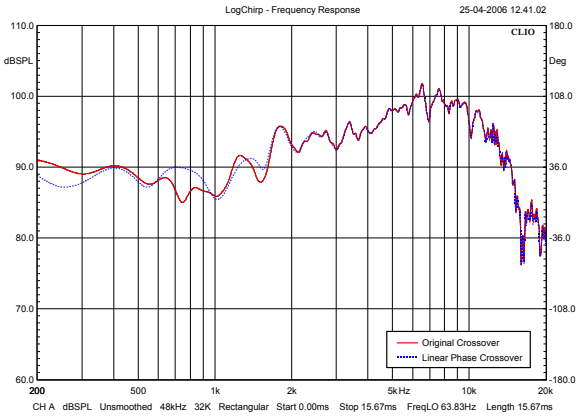

Fig. 29: Vertical Array Element Frequency Response

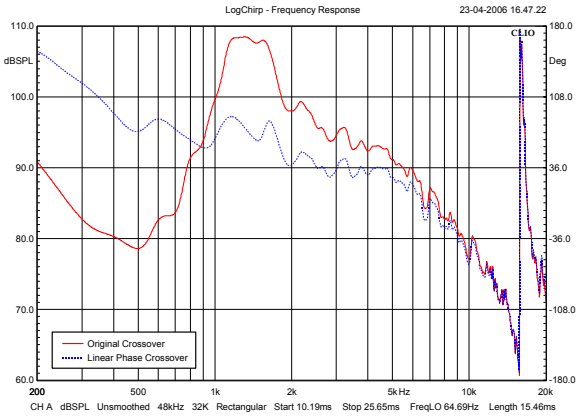

Fig. 30: Vertical Array Element Phase Response

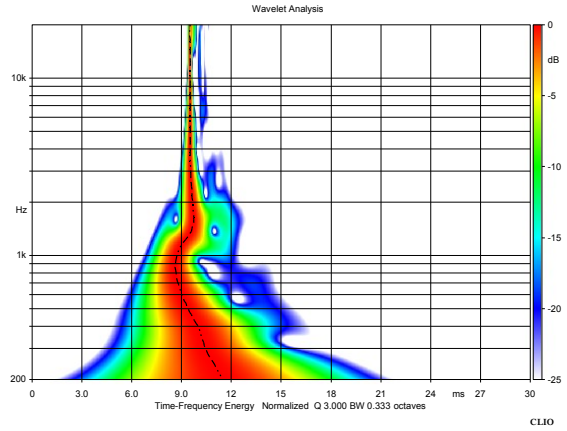

Fig. 31: Vertical Array Element Wavelet Analysis, Original Crossover

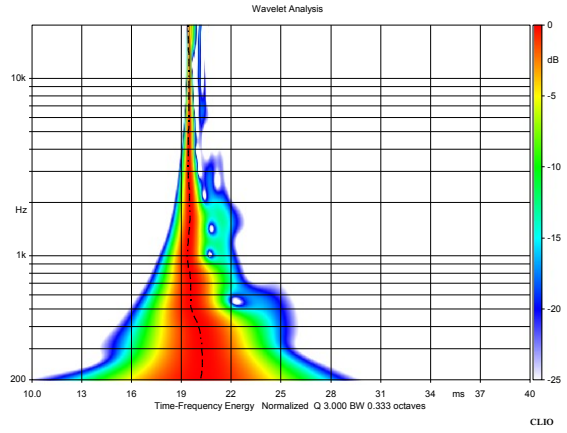

Fig. 32: Vertical Array Element Wavelet Analysis, Linear Phase Crossover

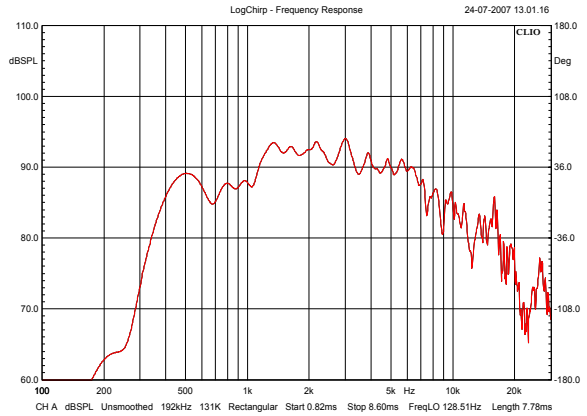

Fig. 33: Compression Driver on a CD horn Frequency Response

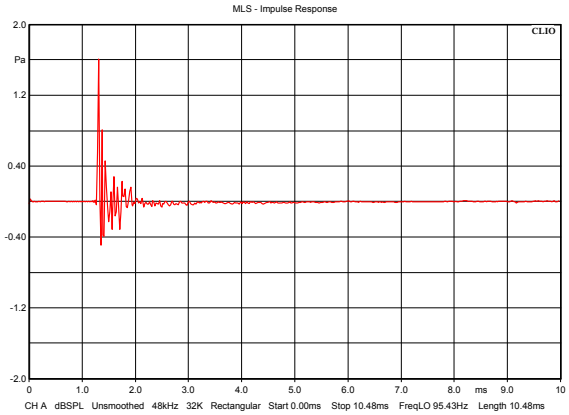

Fig. 35: Quad ESL-63 Impulse Response

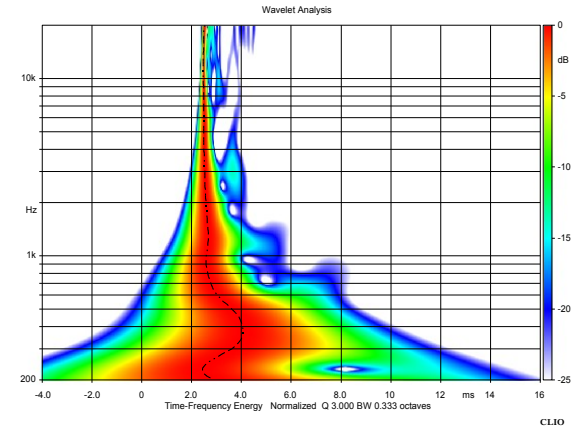

Fig. 34: Compression Driver on a CD horn Wavelet Analysis

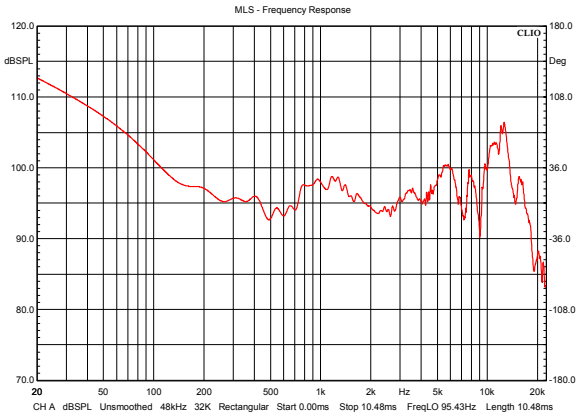

Fig. 36: Quad ESL-63 Phase Response

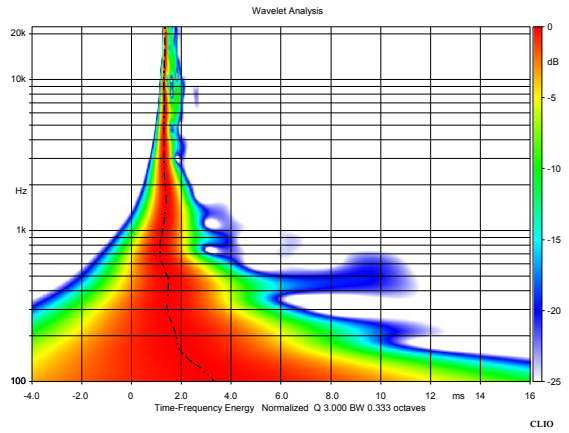

Fig. 37: Quad ESL-63 Wavelet Analysis

portion of it is reflected back to the driver throat because of the abrupt change in the surface shape that takes place at the diffraction point. This generates reflected waves. The Wavelet Analysis applied to the impulse response of such an horn can show how much energy is reflected back and forward inside of the horn and which are the frequency bands essentially affected by the phenomenon for that specific horn. In figure 33 it can be clearly seen the effect of these reflections on the frequency response and in figure 34 can be seen how the WAT results should look like. Even though this sample waveguide shows a remarkable effect of energy reflection at the diffraction slot, it is enough representative of the average quality that can be found in wide coverage diffraction horns used in the general purpose professional loudspeaker systems. Inside the extra wide coverage waveguides that are nowadays in common use in the majority of the Vertical Array elements this effect caused by the diffraction slot happens to be even worst than this.

6.4. Electrostatic Hi-Fi Loudspeaker

The IR of an electrostatic Hi-Fi loudspeaker, Quad ESL-63, has been measured (Figure 35). The system response is well time aligned due to its principle of operation. This is confirmed by the almost flat Phase Response plot (Figure 36) and the Wavelet Analysis plot (Figure 37). Beginning from 100 Hz the system is practically time perfect, except for a few milliseconds decay around 450 Hz.

7. CONCLUSIONS

The Wavelet Analysis is a powerful tool for inspecting the time response of a loudspeaker. The reason for this lies in the joint Time-Frequency characterization of the system that it can provide. A single and easy to read color map plot contains all the information of the joint time-frequency behavior of the system.

This tool can be part of the everyday work of the loudspeaker design engineer and/or the transducer engineer as well. The Wavelet Analysis tool is not going to replace the Phase Response plot into the time alignment of loudspeaker systems or the Cumulative Spectral Decay when the decay of the response must be known. It is a tool that can be used side by side along with proven techniques, it just shed a clearer light on some characteristics of the loudspeakers that are (and still remain) quite difficult to manage: time response and phase response.

As future development we would like to examine faster calculation algorithms as suggested in [15] and the integration of the Wavelet Analysis into a tool with different time-frequency distributions.

8. AKNOWLEDGEMENTS

We would like to thank Audiomatica for supporting this work and opening their CLIO measurement system to this new field, we would also like to thank Fabio Blasizzo for many useful suggestions.

9. REFERENCES

[1] M. Di Cola, M. T. Hadelich, D. Ponteggia, D. Saronni, “Linear Phase Crossover Filters Advantages in Concert Sound Reinforcement Systems: a practical approach”, Presented at the AES 121st Convention, San Francisco, CA, USA, 2006 October 5-8

[2] H. Møller, P. Minnaar, S. K. Olesen, F. Christensen, J. Plogsties, “On the Audibility of All-Pass Phase in Electroacoustical Transfer Functions”, J. Audio Eng. Soc., vol. 55, No. 3, pp. 115-134 (2007, March).

[3] D. B. Keele, “Time-Frequency Display of Electroacoustic Data Using Cycle-Octave Wavelet Transforms”, Presented at the AES 99th Convention, New York, NY, USA, 1995 October 6-9

[4] D. W. Gunness, W. R. Hoy, "A Spectrogram Display for Loudspeaker Transient Response", Presented at the AES 119th Convention, New York, NY, USA, 2006 October 7-10

[5] S. J. Loutridis, "Decomposition of Impulse Responses Using Complex Wavelets", J. Audio Eng. Soc., vol. 53, No. 9, pp. 796-811 (2005, Sept.).

[6] O. Rioul, M. Vetterli, "Wavelets and signal processing", IEEE Signal Processing Magazine, vol. 8, no. 4, Oct. 1991, pp. 14-38

[7] A. Duncan, "The Analytic Impulse", J. Audio Eng. Soc., vol. 36, pp. 315-327 (1988, May).

[8] *CLIO Software - User's Manual*, Audiomatica, `http://www.audiomatica.com`

[9] R. C. Heyser, "Loudspeaker Phase Characteristics an Time Delay Distortion, Parts I and II", J. Audio Eng. Soc., vol. 17, No. 1, pp. 30-41 (1969, Jan.); vol. 17, No. 2, pp. 130-137 (1969 Apr.).

[10] J. D'Appolito, *Testing Loudspeakers*, Audio Amateur Press, 1998

[11] A. Lukin, J. Todd, "Adaptive Time-Frequency Resolution for Analysis and Processing of Audio", Presented at the AES 120th Convention, Paris, France, 2006 May 20-23

[12] C. P. Janse, A. J. M. Kaizer, "Time-Frequency Distributions of Loudspeakers: The Application of the Wigner Distribution", J. Audio Eng. Soc., vol. 31, No. 4, pp. 198-223 (1983, Apr.).

[13] A. Mertins, *Signal Analysis: Wavelets, Time-Frequency Transforms and Applications*, Wiley, 1999

[14] R. L. Allen, D. Mills, *Signal Analysis: Time, Frequency, Scale, and Structure*, Wiley-IEEE Press, 2004

[15] O. Rioul, P. Duhamel, "Fast Algorithms for Discrete and Continuous Wavelet Transforms", IEEE Transactions on Information Theory, vol. 38, no. 2, Mar. 1992, pp. 569-586.