DSP loudspeaker 3D complex correction

Víctor Manuel Catalá Iborra[1], Francis F. Li [2]

[1] *DAS AUDIO GROUP, Fuente del Jarro, Spain*

[2] *University of Salford, Salford, Greater Manchester, UK*

Correspondence should be addressed to Víctor Catalá (vcatala@dasaudio.com)

ABSTRACT

An advantageous approach to DSP equalization of loudspeakers is proposed in this paper adopting spatial averages of complex responses acquired from 3D balloon measurements. Alignment of the off-axis impulses responses with the on-axis impulse responses are accomplished using a cross-correlation technique prior to spatial averaging to attain meaningful statistics of magnitude and phase responses. This is performed over a pre-defined listening window from the complete loudspeaker response balloons (both magnitude and phase). The resulted average of the complex response within a suitably defined listening window is used to obtain, via the least mean square adaptive technique, an inverse filter that corrects the linear behaviour of the loudspeaker.

1 Introduction

Real-time digital signal processors and associated algorithms as broadly adopted in modern audio systems for diverse purposes are a capable addition in loudspeaker electroacoustic property correction or compensation, showing a great potential and having gradually attracted more and more attention of designers. Indeed, many artifacts of loudspeakers can be corrected or compensated for in an effective and flexible manner with DSP algorithms and suitable processors that allow for real-time processing of audio signals.

There are several artifacts of loudspeakers that the designers may wish to correct, including linear and non-linear distortions, while this paper focuses on the linear ones. Level correction based on a single response is a relatively straightforward task, which is to synthesize an inverse filter from an impulse response measured at a certain distance on-axis from the loudspeaker. However, most loudspeakers are not designed for the use in a single listening point and in an anechoic condition, this simplistic approach to loudspeaker response correction is not entirely sensible; perceived quality of a loudspeaker cannot be completely characterized by a single on-axis response. Besides, level only corrections ignore the fact that the transfer function of a linear system is determined by its magnitude and phase responses. Acting on the phase response can potentially give more accurate time response, i.e. perceived transient characteristics.

The paper is reported from an attempt to equalize both magnitudes and phase responses in an arithmetic mean sense in a listen window based on loudspeaker balloon measurements in an anechoic condition. Loudspeaker 3D balloon responses are first acquired; an algorithm to average the measured impulse responses in a specific window is developed and the inverse filter derived via the least mean square (LMS) adaptive FIR filter approximation. The compensatory algorithm (inverse filter) was then implemented on a DSP chip to enable real-time operation.
Measurements showed satisfactory correction of the system and pilot listening tests confirmed this by notable improvements in sound quality.

2 Spatial Correction of Loudspeaker Responses

2.1 Complex spatial correction

Although DSP opened a new horizon for loudspeaker corrections, the objective of such equalization and compensation is a complicated and somewhat mysterious question. Equalizing based on the on-axis responses might cause unwanted artifacts (peaks or dips) off-axis, which mitigates the general perceived sound quality. Besides, as it will be seen later, not all issues can be corrected with pre-conditioning filters, and some unwise correction attempts can simply cause more problems. Taken these into consideration, several authors have proposed the use of spatial averaging to get a "representative response" to work with and suggested a number of ways to do this. A brief review of these is summarized below.

Given the fact that the major concern is loudspeaker correction not room equalization or correction, Salamouris et al. [1] advocated a consistent and representative measurement of the loudspeaker responses under minimal room effects for "intrinsic" loudspeaker equalization. Greenfield and Hawksford [2] proposed the use of the average of responses within a finite listening space instead of the on-axis response as the representative response or reference for loudspeaker correction to avoid the detrimental artifacts occur in off-axis responses. Di Cola and Ponteggia [3] advocated spatial averaging and smoothing of frequency responses of the loudspeakers to get a representative response and avoid local problems at specific angles. However, the spatial averaging method was restricted to magnitude-frequency responses. For the phases, only on-axis responses instead of spatial information were considered. Vaucher [4] took phase linearity into serious consideration and proposed the use of complex impulse responses and spatial average as the basis for the design of inverse filters for loudspeaker correction. Even so, for phase responses, he only considered the on-axis phase response, neglecting angular phase variations. Response smoothing has also been recommended by many other authors in the context of loudspeaker-room equalization e.g. [5] and [6]. Pedersen and Thomsen [7] proposed the combination of the listening position measurement with information about the 3D sound field in a room, obtained from other position measurements.

Most recently, Toole [8] showed the importance of a flat on-axis amplitude response and well behaved directional characteristics in achieving a good sounding loudspeaker by extensive research with listening tests. Some rules of thumb following [9] [10] [11] and [12], might be summarised to include that (1) Correction of loudspeakers should be based on anechoic measurements. (2) The average response within a listening window, say ± 30° horizontal ± 10° vertical, might be deemed as being representative of the direct sound received by the listeners. (3) The use of the inverse of such representative instead of a specific or the average of on-axis response as a basis for inverse filter design can mitigate specific local problems.

2.2 Importance of complex 3D measurements

Now that the importance of spatial response average has been established, it becames necessary determine how to acquire loudpeaker angular data, including level and phase. The methods including simulation and measurement techniques have evolved over decades. Early electroacoustic simulation software was based on radiation polar measurements, considering only magnitude variations but neglecting phase responses. Similar situation happens to routine measurements. It is still rare to find loudspeaker specification sheets that give phase response information. To the best of the authors' knowledge, most self-powered professional loudspeakers with built-in DSP correction only correct phases at the crossover point(s), and overall phase corrections are not always adopted.

Several authors have argued the importance and advocated the inclusion of phase data using complete 3D balloon measurements (magnitude and phase) to achieve more accurate characterization of loudspeaker behaviours. Ahnert et al. [13] showed the need of complex response acquisition for proper simulation of loudspeaker interactions, since magnitude only models cannot accurately predict interactions between loudspeakers with different phase responses. In their work, attenuation and phase balloons were used as a fundamental part of loudspeaker modelling or, a step further, a means to consider source interactions, e.g.[14]. The complex and the magnitude only models were quantitatively compared in these publications, and the complex model revealed notable advantages. One other important consideration when performing 3D measurements is the fly time, i.e. the time that it takes for sound to travel from the source to the microphone. It has to be accurately extracted from the measured transfer function to obtain an intrinsic phase response, independent from the measurement distance, e.g. [14] [15]. To further avoid small deviations between measurement set-ups due to mechanical or environmental reasons, Feistel et al. [15] suggested re-normalization when multiple redundant on-axis measurements are made.

A standardization effort was documented in [16]. This AES standard describes how polar radiation measurements shall be made and documented for sound system predictions and characterization, among other uses, and proposes 5º resolution measurements for both Φ and θ angles, a minimum distance of 4 m, typically 8 m, depending on loudspeaker dimensions, and the use of redundant data points to determine repeatability. It was reaffirmed fairly recently in 2014, therefore, was considered up to data.

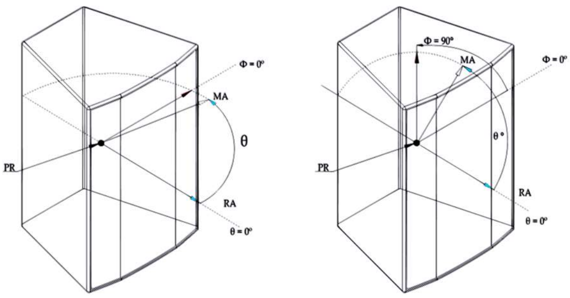


Figure 2-1: Polar system coordinates per AES, 2008 (reaffirmed 2014).

Figure 2-1 gives illustrations of $\theta=0^o$, and $\varphi=0°$ and $\varphi=90°$ according to the standard. It is assumed that measurements can be done in an anechoic environment. Nevertheless, it is difficult to find anechoic chambers large enough to allow for the distances required and with a low enough frequency cut-off to perform appropriate measurements. Some authors have proposed methods to perform quasi-anechoic measurements in non-anechoic environments. For example, Gander [17] proposed ground plane techniques to measure loudspeaker systems, although his approach it is not easily applicable to 3D balloons acquisition. Struck and Temme [18] proposed time gating of impulse responses to acquire only the direct sound before the first reflection arrives to the measurement microphone, rejecting the sound reflected by the room. The measurement low frequency extension depends on the window length. With a relatively large room, with short reverberation times and absorbers on the floor to avoid the floor reflection, good anechoic equivalent measurements can be achieved down to mid-low frequencies. If low frequency responses are needed, near field techniques weighting the radiation from the different radiators can be applied. Then, the near field response can be scaled and combined with the far field response to obtain a full range response. Other technique has been proposed by Benjamin [19] to obtain the low frequency responses of loudspeakers in non-anechoic rooms based on the assumption of a typical high pass response according to the design of the loudspeaker system being measured, and applying a post-processing filter that counteract the effect on frequency response of the window being applied to reject reflections.

More recently, several new techniques, based on initial near-field scanning and later far field prediction, have been validated for 3D radiation measurements, which allow for measurements being performed in non-anechoic conditions [20], [21] and [22].

2.3 The roles of pre-conditioning filters

The use of signal pre-conditioning filters is a typical way to apply DSP for the real-time compensation of loudspeaker responses. Before starting any design process, it is necessary to understand what phenomena are correctable and how. This paper limits its scope of applying linear and time invariant (LTI) DSP algorithms as pre-conditioning filters. Gunness [23] states the requirements for a system to be corrected with such preconditioning filters. If a system is time-variant (its response varies with time), it is nonlinear (its response varies with level) or is spatially variable (its response varies with direction), then, it cannot be corrected with an input filter. However, as discussed and will be further showed that although strict sense compensation of spatial variation of loudspeaker is not possible by a linear and time invariant filter in signal path, such a filter suitably designed to represent an inverse of a certain average of spatial variation will be beneficial. In this way a quasi or partial compensation is possible.

In order to be correctable by a LTI filter, the system has to be sufficiently linear over a wide input level range. Deviations at high levels have to be avoided with the use of limiters or with filters acting in a level dependent manner [23]. Some other authors, e.g. [24], propose the correction of the nonlinearities of the loudspeaker by using systems with feedback or using a model of the nonlinear system to predict the response of the speaker. These methods are beyond the scope of this paper.

The time invariance is another theoretical requirement for preconditioning filter correction. If some variation is produced after some use of the loudspeaker, a point in the middle of the variation range can be chosen [23]. This is also what other authors e.g. [4], adopted as the so called "aging process" before evaluating the response of the loudspeaker.

From a manufacturing point of view, unit to unit variability must be considered. Avoiding applying the inverse filter to phenomena which shows strong variations between units [23] or analysing and averaging the variations between them to get a representative response [4] and [3] are the two common tactics.

2.4 Spatial averaging

Spatial variation in responses is the intrinsic characteristics of loudspeakers which might benefit from some sort of compensation based on the inverse of the "averaged" measures. Some authors have proposed average methods that can be applied to this purpose. In the context of impulse interpolation, Gunness [25] proposed a method using the cross-correlation between different impulse responses to align them in time and eliminate excess delay. He suggested to work with complex responses and to use the geometric mean to obtain a response with level and phase which lie between the two original level and phase responses being interpolated.

The method is probably best described by an example. The impulse responses of the considered loudspeaker in this paper, off-axis at $\Phi=0^o$, and $\theta=5^o$ (red) and on-axis at $\Phi=0^o$, $\theta=0^o$ (green) are shown in Figure 2-2.

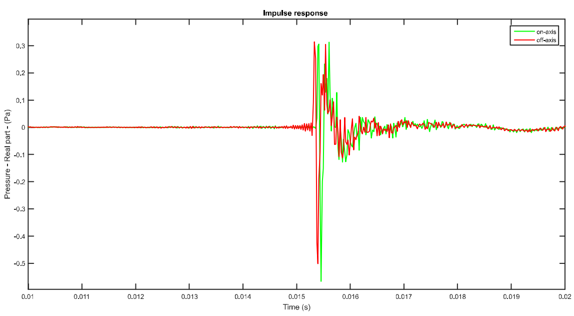


Figure 2-2: On-axis, Φ=0° θ=0° (green), and off-axis, Φ=0° θ=5° (red) impulse responses.

Different fly times between both measurements can be seen in Figure 2-2. They make straightforward averaging meaningless. Time-alignment is necessitated. To do so, the cross-correlation of the two impulse responses is calculated and shown in Figure 2-3.

When the on-axis and the off-axis impulses are similar enough the cross-correlation function shows a defined peak at the lag difference (in time). In this case, Figure 2-3 shows a lag difference of 0.0625 ms (3 samples at 48000 Hz sampling frequency) and the cross-correlation peak is close to 1. This means that the two signals are rather similar but are shifted 0.0625 ms. Following time-lag compensation the signals are visualized in Figure 2-4. The alignment of delays makes averaging meaningful.

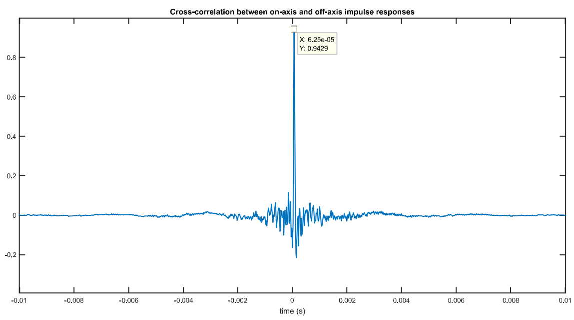


Figure 2-3: Cross-correlation function between the On-axis, Φ=0° θ=0° and off-axis Φ=0° θ=5° impulse responses.

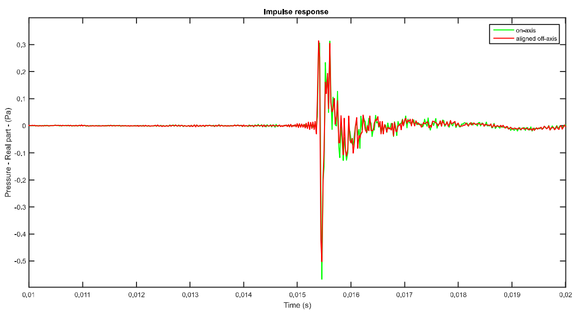


Figure 2-4: On-axis, Φ=0° θ=0° (green), and off-axis, Φ=0° θ=5° (red) impulse responses, with compensated lag time.

With the time delay compensated for, frequency responses can be obtained from them using the FFT, with phase results not being affected by different fly times. Once the complex transfer functions are obtained, the fly time can be extracted from them dividing them by the transfer function of a pure delay (e^{-jwt}), being t the fly time. In this case the fly time was considered as the time at the maximum amplitude point of the on-axis impulse response. After this procedure, average can be performed in a more meaningful way.

Figure 2-5 shows the similarities between the on-axis (green) and off-axis (red) phase responses once the lag time is compensated and the fly time is extracted. Similar phase response allows for meaningful phase interpolation or average.

Gunness [25] proposed the geometric mean of complex responses for interpolation, to get level and phase frequency responses that lie between the original ones. In this paper, the authors have worked with magnitude and phase frequency responses independently and use their arithmetic mean for averaging purposes. This is considered simpler, more intuitive when handling multiple responses and gives rather similar result to that obtained from the geometric mean with two curves. The interpolation of magnitude and phase separately was proposed by Panzer and Ferekidis [26]. With the two considered signals in the example, the averaged level runs in the middle of the two original ones in dB, and the averaged phase stays in the middle of the two original phases, in degrees (Figure 2-5, magenta curves).

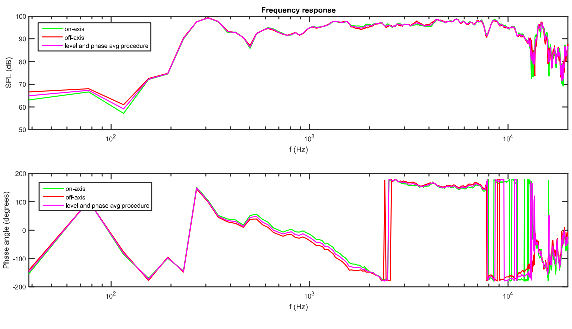


Figure 2-5: On-axis, Φ=0° θ=0° (green), off-axis, Φ=0° θ=5° (red) frequency responses and average response with the proposed method (magenta).

A Matlab script was used to perform alignment of the impulse responses of complete 3D balloons as described previously and to transform them into frequency responses. Phase and level responses were averaged independently within a specified listening window and the averaged complex response was created from the independent averages. Going back to the time domain using IFFT, the average impulse response was created and exported to be used as the base for the inverse filter calculation. It has to be noticed that the averaged impulse response obtained this way neglects the angular fly time variations because the average algorithm performs all calculations after aligning impulse responses and no fly time averaging is applied. This is not relevant in our case because the average impulse response is used as the base for corrections and the corrections are independent of the fly time. Only the shape of the average impulse is relevant, not its absolute location in time.

In addition, phase and attenuation balloons were calculated. The lag time difference obtained by cross-correlation between on-axis and off-axis impulses that was extracted to align them for average purposes was added to off-axis responses here to get real meaningful phase balloons.

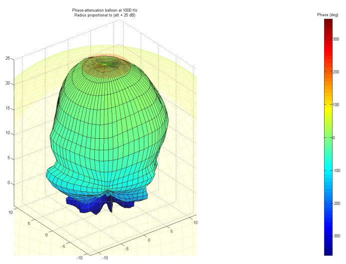


Figure 2-6: Combined phase-attenuation balloon at 1 kHz. The average window is shown over the reference sphere in darker colour.

It is observed that redundant on-axis measurements are not exactly identical. To avoid this uncertainty re-normalization was considered necessary and used. This means that the on-axis reference (θ=0) was varied with the Φ corresponding with the set of measurements being processed. For example, all Φ=n, θ=m measurements are referenced to the on-axis measurement Φ=n, θ=0.

3 A design paradigm for a PA loudspeaker

3.1 Loudspeaker system

A professional public address (PA) loudspeaker system has been chosen to identify the improvements that can be achieved with the proposed method. In particular, the correction has been applied to the mid-high frequency section of the loudspeaker system shown in Figure 3-1: from top to bottom, the high frequency horn, the mid frequency horn and mid speaker with its phase plug, and the open port of the low frequency unit. Only the mid-high arrangement has been studied due to measurement constrictions. The loudspeaker system has an 8 inches loudspeaker loaded with a horn with phase plug arrangement as the mid unit, and a 3 inches diaphragm compression driver coupled to a 40° x 30° horn as the high frequency unit. Mid and high frequency sections are combined with a passive crossover filter at 1650 Hz. The use of passive crossover, cone radiator and compression driver with large horns and phase plugs makes it a good candidate for LTI preconditioning correction.

Figure 3-1: Loudspeaker system under study.

3.2 3D measurements

Following the standard [16], the Clio Software application note AN-002, [27], describes a practical procedure to acquire 3D balloon measurements. The spherical coordinates system is defined Figure 3-2, and the method for simulated free field measurements by [18] is adopted for this paper. Figure 3-3 shows the measurement system with two turntables. The polar turntable is placed with its rotation axis perpendicular to the floor and provides the θ angular increments. The azimuth turntable has its rotation angle parallel to the floor and provides Φ variations. The point of rotation (POR) is defined by the intersection of the two-rotation axis.

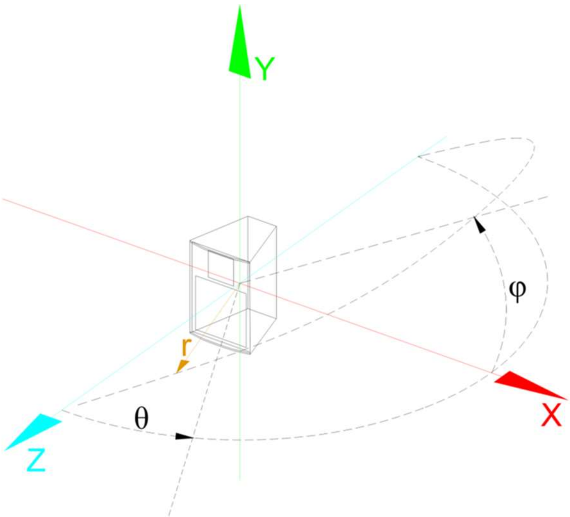


Figure 3-2: Graph following Polar coordinates at Clio Software application note AN-002.

With a 5° angular resolution, a complete balloon is achieved with 72 Φ positions (a complete turn) x 37 θ positions (half turn). This gives a total of 2664 measurements. All $\theta = 0°$ measurements are redundant on-axis measurements and all $\theta = 180°$ measurements are redundant measurements at the back of the loudspeaker system.

Although redundant measurements should theoretically be the same, the finite measurement distance, the different locations of the components referred to the POR, small mechanical deviations during measurements and the asymmetrical measurement space, all contribute to slightly different measurement results. Re-normalization is recommended for each set of Φ data.

Figure 3-4 specifies the Point of Rotation, Mechanical Reference and measurement distance used for 3D balloon data acquisition. It can be seen that the mid frequency unit is very close to the P.O.R. while the high frequency unit is about 19 cm away from the P.O.R.

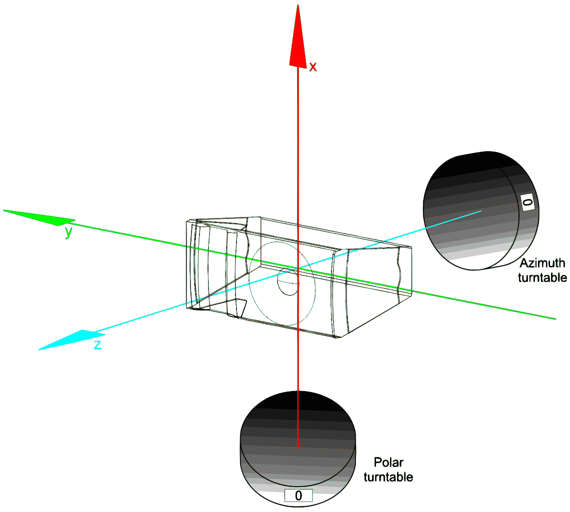


Figure 3-3: Graph following 3D acquisition system description at Clio Software application note AN-002.

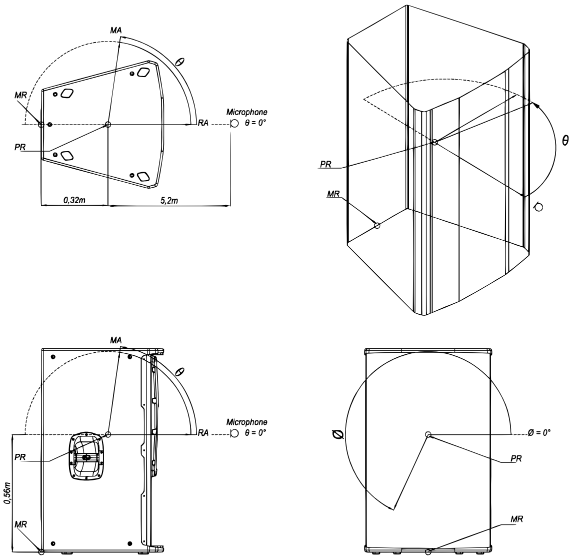


Figure 3-4: Plan and perspective views showing the Point of Rotation (PR) and Mechanical Reference (MR) – top. Side and front views showing the Point of Rotation (PR) and Mechanical Reference (MR) – bottom.

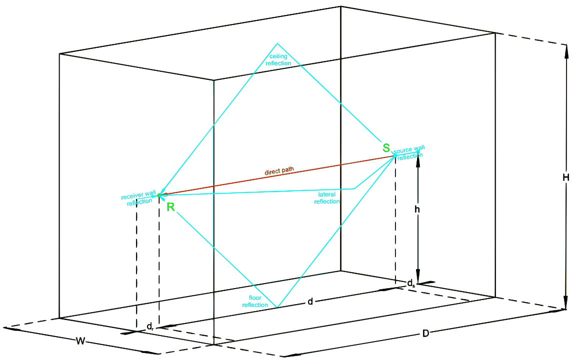


Figure 3-5: Graph following the shoe box model for reflection analysis at Clio Software application note AN-002.

Figure 3-6: Measurement set-up at DAS Audio Auditorium.

The measurement space was taken in a room with a low RT (RT_{60} about 0.3 s) with large enough dimensions to allow for anechoic equivalent measurements down to a useful frequency. The ceiling is about 5 m high at the first ceiling reflection point and the first floor reflection is absorbed with appropriate absorber wedges. A picture can be seen at Figure 3-6.

$$H = \sqrt{\left(\frac{\frac{c}{f_{\mathrm{LOW}}}+d}{2}\right)^2 - \left(\frac{d}{2}\right)^2} + h \qquad (1)$$

$$W = 2\sqrt{\left(\frac{\frac{c}{f_{\mathrm{LOW}}}+d}{2}\right)^2 - \left(\frac{d}{2}\right)^2} \qquad (2)$$

$$d_s = \frac{c}{2f_{\mathrm{LOW}}} \qquad (3)$$

$$D = d + 2d_s \qquad (4)$$

Based on [27], equations (1) (2) (3) (4) and Figure 3-5, the minimum dimensions to get anechoic measurements down to 100 Hz (f_{LOW}), with a source-microphone distance d = 5 m (approximately the available maximum) and the microphone placed at h = 1.5 m, are

H = 4.88 m
W = 6.75 m

With the described 3D measurement process and the already explained alignment method, complete balloons, both level and phase, were acquired to analyse the loudspeaker system performance and to be used for averaging within the specified listening window (10º per side cone, 20º total). A set of balloons can be seen at Appendix A.

Figure 3-7 shows the on-axis frequency response at $\Phi=0°$, $\theta=0°$ (red) compared with the 10° cone (20º total) averaged response (green). It can be seen that several strong cancellations at high frequencies get smoothed when averaging, revealing that they are localized mainly on-axis. Narrow corrections based only on the on-axis response would produce unwanted peaks off-axis, probably worsening the perceived quality of the loudspeaker.

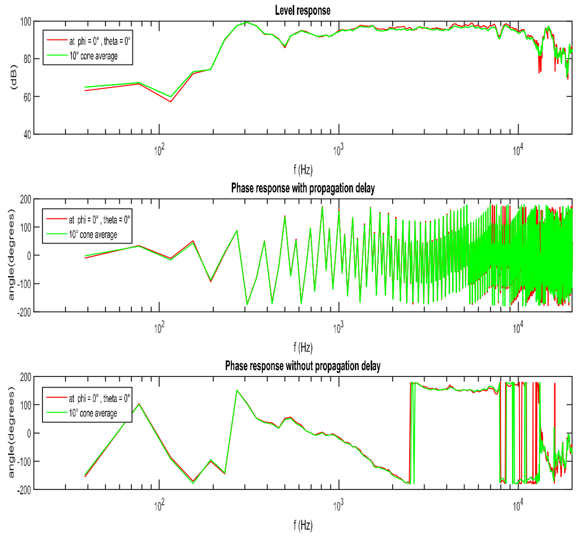


Figure 3-7: On-axis and averaged frequency responses.

The propagation delay was taken as the time at maximum amplitude at the on-axis impulse response. Once extracted, the intrinsic loudspeaker phase response can be seen at the bottom of Figure 3-7. It can also be seen the 180º phase shift from 1 kHz to 3 kHz introduced by the crossover. As the level response remains mainly flat, it causes the non-minimum phase behaviour of the loudspeaker.

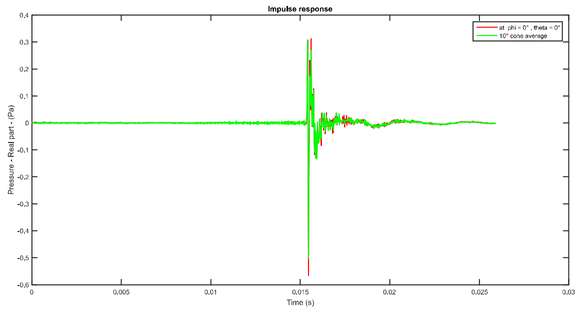


Figure 3-8: On-axis and averaged impulse responses.

The on-axis impulse response (red) and the average impulse response (green) can be seen at Figure 3-8. Notice that the average impulse response has been calculated from aligned responses using the cross-correlation technique. It is, therefore, aligned with the on-axis impulse. A "true" average impulse should be moved in time according with the average lag between the off-axis responses on the on-axis response. In this case is not necessary because the average impulse response has been obtained to calculate the correction inverse filter and only its shape matters. The obtained impulses are the result of the combination of the impulse responses of the mid and high units passed through the crossover network. Because they are not perfectly physically aligned, neither on-axis nor off-axis, some phase variation is also introduced in the phase responses. Narrow and quick peaks and deeps are due to the high frequency unit while slower oscillations are due to the mid unit.

3.3 Inverse filter calculation

Individual loudspeaker drivers are often approximated as minimum phase systems, but the 2-way passive loudspeaker system deviates from a minimum phase considerably. This needs to be considered when synthesizing the inverse filter.

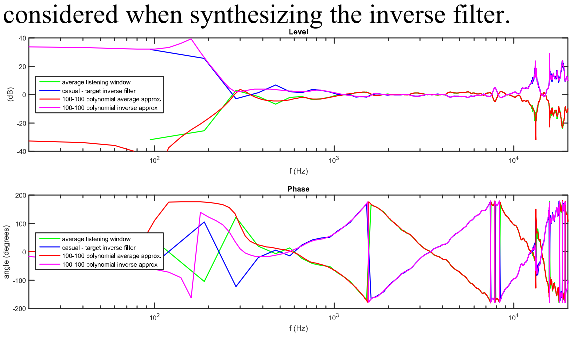


Figure 3-9: Average response, its casual inverse and their approximations.

Figure 3-9 shows the average response within the selected 10º cone listening window, with the green track showing the phase shift caused by the crossover starting at about 1 kHz. Because of the non-minimum phase behaviour of the original response, an appropriate delay has to be added to the inverse filter to make it casual, that is, to move the main part of its impulse to positive times, or equivalently, working with the z-transform, moving the poles of the inverse filter inside the unit circle. In this case, the added delay has been optimized to be minimum and to provide the best possible correction filter. The frequency response of the casual inverse and its approximation with 100 order polynomials can be seen at Figure 3-9 (blue and magenta curves).

Once the optimum delay is found, it is applied to the theoretical inverse filter to make it casual: this is the target response. From the target impulse response, the approximated inverse filter is calculated using the LMS algorithm.

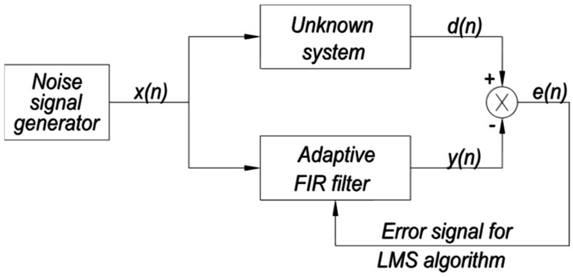


Figure 3-10: Graph following the block diagram of system identification or system modelling problem from [28].

The LMS algorithm compares the output signal of the unknown system, in this case the theoretical inverse plus the delay, with the FIR filter output, and modifies the FIR filter coefficients to minimize the error signal. 400 taps were used in this case.

A Matlab script was used to import the average impulse response of the loudspeaker system, process it and calculate inverse filters that correct the system's level and phase.

Figure 3-11 shows the frequency responses of the average listening window, the target correction filter, the calculated filter and the calculated corrected system. No smoothing has been applied to show the accuracy that the used algorithm can provide although some smoothing of the response to be corrected can help to reduce inverse filter pre-ringing artifacts and high Q. The allowed maximum correction was chosen to 6 dB and the minimum frequency for correction was chosen at 200 Hz. The optimum calculated delay was 1.1 ms. It can be seen how the inverse filter flattens both, level and phase responses. Where the required correction exceeds 6 dB, the boost is limited to this value to avoid driver overload.

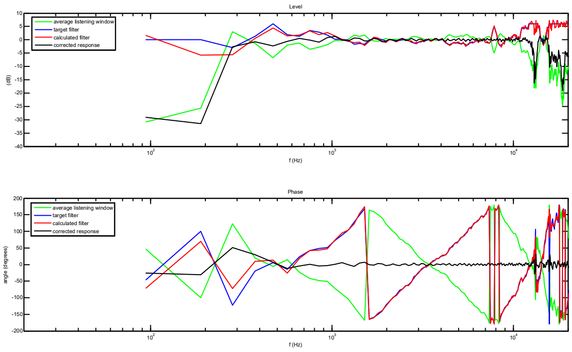


Figure 3-11: Freq. resp.: Average listening window, target filter, calculated filter and corrected system.

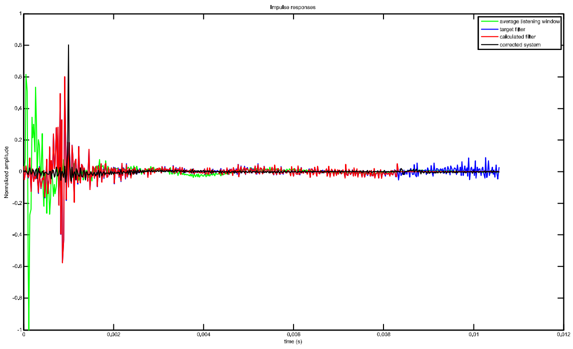


Figure 3-12: Impulse response: Average listening window, target filter, calculated filter and corrected system.

4 Results

In data processing and inverse filter design were done in Matlab environment. The coefficients of the synthesised inverse FIR filter as describe in the previous section were then exported to an FIR block in Analog Device's SigmaStudio for further implementation on the ADAU 1701 based DSP platform as the real-time input filter for the DUT. Objective and subjective assessments were carried out to compare the performance with and without the proposed correction. Results are presented in this section.

4.1 Objective Assessments

3D balloons were acquired, and average was performed again with the preconditioning filter being applied. The 3D balloons are not included again because, as it was said before, they are representations of attenuation and phase variation relative to the on-axis values. Because the inverse filter has been applied as input filter, it affects the response at all angles and, therefore, relative variations remain intact. See Appendix A.

Level and phase frequency responses with the inverse filter applied have been clearly improved both on-axis and on average within the listening cone, Figure 4-1. Average level response lies between ±1.5 dB from 500 Hz to 12 kHz while phase response lies between +27º to -5º in the same frequency range. The agreement with the calculated corrected response at Figure 3-11 is good.

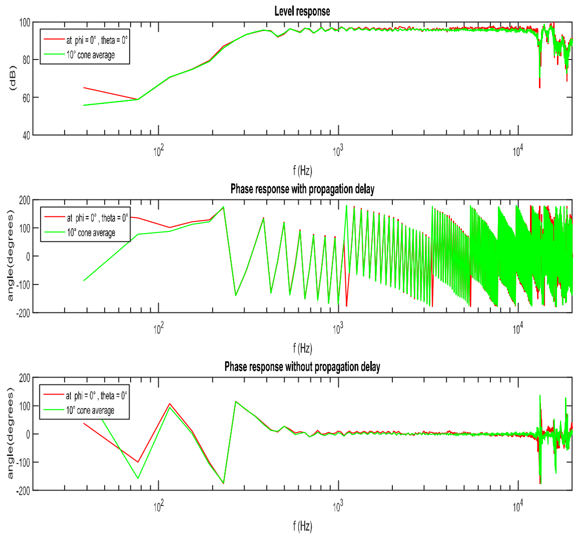


Figure 4-1: Corrected system's Frequency responses

Impulse responses, Figure 4-2, approximate the desirable delta shape of an ideal system and shows similar behaviour to the calculated corrected system impulse response at Figure 3-12. The corrected system impulse response is close to ideal, with some small pre and post ringing added, and some noise. The delay added to the system is mainly due to the filter delay introduced to make it casual, plus the delay produced during AD/DA conversions. The total added delay is about 2.3 ms, allowing the system to be used in real time applications, even in monitoring applications [4].

4.2 Listening test

To further evaluate the correction method and determine how the evident improvements found in objective parameters translate into perceivable improvement in sound quality, pilot listening tests were conducted using a small panel of specialists in house.

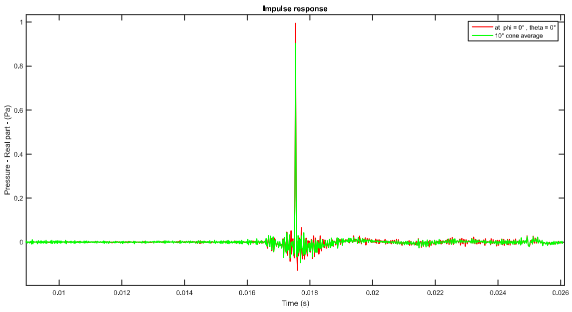


Figure 4-2: Corrected system's impulse response. On-axis and listening cone average.

4.2.1 Set-up

A SigmaStudio project was created including two different paths for: (1) no correction processing, and (2) level and phase FIR filter correction. A multiplexer cell allowed the listeners to choose between the two different signal paths.

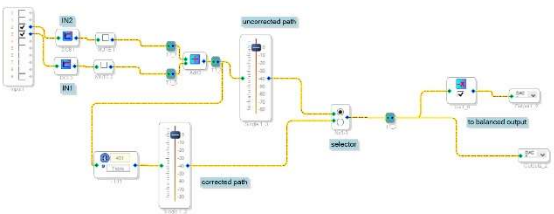


Figure 4-3: Block diagram of the DSP project for the listening test

Two blind options were available, and each subject was free to listen to each of them as many times as needed to identify differences. Switching between filters was instantaneous and subjects could switch between two paths at any time they wanted. They were asked to rank the sound quality at the two switch positions. The initial position was randomized between subjects.

4.2.2 Materials

Given all subjects were chosen from experienced loudspeaker engineers and audio engineers; they all have accumulated a list of preferred materials for subjective listening test. It is arguable that the use of experts' own choices means the opportunities for them to apply the best suited probe stimuli to their listening experiences. The materials used included pop and rock music, with vocal songs, and were different for each subject.

4.2.3 Listening room

The listening space was the same DAS auditorium where the measurements were made and with a similar set-up. Wedges on the floor avoid floor reflections and the first reflections arrived from the ceiling after 11 ms. The listening position was on-axis with the speaker, at 5.2 m from it, and was the same for all listeners. Comfortable levels were chosen for each subject. Due to the high dynamic range of the loudspeaker and the rest of the equipment being used, all test was run in their linear zone. Several measurements of the reverberation time around the listening position can be seen at Figure 4-4.

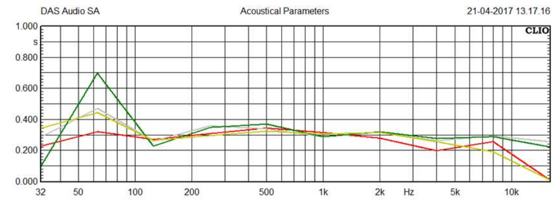


Figure 4-4: DAS Auditorium RT_{60}.

4.2.4 Subjects

Five subjects participate in the test. All are audio equipment engineers, or audio applications engineers that usually make subjective product evaluations or product tuning for demos, installations or tours. According to [29] "Experienced acoustic experts and audio equipment engineers" show "high" reliability at subjective audio evaluations and are suitable for “high precision measurements”. Therefore, the listening test results, although informal, can be taken as a reliable initial result.

4.2.5 Results

All five subjects perceived the processed systems as clearly superior to the un-processed system regarding to spectral balance, smoothness, clarity and attack.
The processed option was appreciated as correctly equalized and nothing odd or artificial was recognized.

5 Discussion

5.1 Selection of appropriate listening cone for averaging

Two different correction targets may be considered namely flat and non-flat ones according to chosen listening cones.
A 20º (-10º to +10º) listening cone was chosen for the averaged response to be used to calculate the inverse filter in this paradigm. This was to include the angles of major radiation but avoid the coverage where the radiation level falls below -6 dB. Within a so-chosen listen cone, responses should be generally flat and thus the design objective would be a flat response. With this approach and in general, the listening cone should be carefully determined according to the loudspeaker’s actual spatial characteristics. The measured balloons here play an important role, since they help accurately decide the averaged angle(s), avoiding unreliable nominal data.

A different approach could be the selection of wider average listening windows that include radiation beyond the -6dB limit and to correct the system to a non-flat target, for example an average response with some roll-off at mid-high frequencies to be discussed later.

5.2 Distribution of the measurements over the sphere

It has to be noticed that with the measurement distribution specified in the A.E.S. Standard as used in this project, when considering all measurements within the $\theta=+/-10^o$ cone, 72 $\theta=0^o$ on-axis measurements are taken, i.e. at the same angular position. This gives more weight for the on-axis responses than the off-axis ones. It has not been considered a bad option because some authors stressed and confirmed the importance of a flat on-axis response to get good sounding speakers. If this extra weighting effect for the on-axis responses was not desirable, it could be avoided by calculating the average of on-axis responses and treating it as a single measurement or assigning a desired weighting factor to it.

Other measurement distributions over the sphere can be considered to balance the weighting factors applied on different angles or angular sectors or to select different horizontal and vertical angles to define the window widths according to the loudspeaker coverage.

5.3 Configurable target responses

As mentioned previously, a narrow listening cone has been selected for response averaging, and consequently, a flat target response has been chosen. A wider listening window, including responses close to the coverage of the loudspeaker would have required an average target response with some roll-off at high frequencies. Otherwise, a flat average response would produce excessive high frequency content on-axis and at angles close to on-axis. The roll-off slope would depend on the loudspeaker directivity response against frequency. A wide constant directivity speaker would require a different target response for correction than a long throw high Q speaker.

5.4 DSP resources and latency

The inverse FIR filters were calculated and implemented at a 48 kHz sampling rate. 400 coefficients were needed to achieve a good match with the target response down to 500 Hz, in the case of level and phase correction. A total of 450 instructions were used, out of a possible 1024 instruction cycles per audio sample, mainly dedicated to the FIR filter. The delay introduced to get a casual correcting filter at calculation stage was 1.1 ms. The total introduced delay including AD/DA conversions and processing with the FIR filter, was about 2.3 ms. Such short delay is not perceivable, even in close monitoring applications.

5.5 Limitations

The method presented in this paper attempts to correct the linear behaviour of the loudspeaker only. The use of windowing (time domain gating) during data acquisition because of the non-anechoic environment imposed a low frequency limit for measurements. A large anechoic room or the use of near field techniques would be necessary to obtain usable full range measurements.

All measurement sessions were run in normal working days and the measurement room, in a loudspeaker factory without special noise isolation. Although the signal to noise ratio was generally good, the average impulse response calculated from the measurements and used to derive the correcting filter was not completely noise free as desirable.

Some of the recommendations about loudspeaker aging or average between production samples made at the literature review section were not followed, because the project was to focus on data acquisition, averaging and processing and was developed with only 1 unit as a proof of concept. Of course, if implemented industrially, these recommendations should be followed.

The average response obtained from the 3D measurements has been taken as reference and the calculated inverse filter tries to correct all issues as linear phenomena. However, some of the narrow notches that appear at high frequencies are related to break-up modes of the diaphragm and to geometrical constrains of the design and, therefore, cannot be fixed with equalisation. This can be seen from measurements of the corrected system, where some notches were not improved although equalisation was applied. In this case, further investigation is required to identify and handle this issue and exclude them from the target correction.

The proposed method is based on spatial average, and this average is further based on impulse alignment using cross-correlation. In well off-axis scenarios, the cross-correlation between the on and off impulses falls, these impulse responses cannot be aligned. Therefore, the method becomes unsuitable when trying to average over very wide listening windows for narrow coverage loudspeakers.

6 Concluding remarks

A full procedure for loudspeaker correction based on 3D average of complex responses within a listening window has been proposed, developed and evaluated. The procedure includes acoustic data acquisition, data processing, averaging, and inverse filter synthesis for complex correction. From the presented example, the measurements obtained after correction showed clear improvements of the loudspeaker system in both its level and phase frequency responses, and in its transient response. The average corrected level response lies between ±1.5 dB from 500 Hz to 12 kHz, while average phase response lies between +27º to -5º in the same frequency range. The average impulse response showed a clear reduction of resonances, approximating an ideal delta function. Therefore, measurements showed that the proposed method can provide proper equalization from an objective measurement point of view.

The total delay introduced by the use of DSP on a common audio processing chip to implement the proposed pre-conditioning correction is short enough to be used in real time applications, even in near field monitoring speakers, provides an effective FIR filter pre-conditioning correction method suitable for self or externally-powered loudspeakers for applications where perceptible latency is not acceptable.
Pilot listening tests with a small number of experts revealed clearly improved sound quality overall. Given the objective and subjective assessment results, the proposed method is thought a valid one for loudspeaker equalization.

Further research about the presented method compared to other equalization procedures based on only on-axis, only polar or only level is needed. This work is being developed and will be proposed as a journal paper.

Acknowledgments
Thanks to DAS Audio for the use of its facilities and the collaboration of its engineers during the listening tests.
Thanks to Jonathan Hargreaves (The Salford University), Francisco Peiró (DAS Audio) and Gonzalo Arroyo (DAS Audio) who, independently, agreed on the idea of showing attenuation and phase variations on the same balloon.
Thanks to Monica Cayuela (DAS Audio) for her help with drawings.

References

[1] S. Salamouris, K. Politopoulos, V. Tsakiris, and J. Mourjopoulos, "Digital System for Loudspeaker and Room Equalisation," presented at the 98th A.E.S. Convention, Paris, France, 1995.

[2] R. Greenfield and M. Hawksford, "Efficient filter design for loudspeaker equalisation," presented at the 86th A.E.S. Convention, Hamburg, 1989.

[3] M. Di Cola and D. Ponteggia, "Application of Optimized Inverse Filtering to improve Time Response and Phase Linearization in Multi-way Loudspeaker Systems.," presented at the 129th A.E.S. Convention, San Francisco, CA, USA, 2010.

[4] R. Vaucher, "Linear phase implementation in loudspeaker systems; measurements, processing methods and application benefits," presented at the 135th A.E.S. Convention, New York, NY, USA, 2013.

[5] P. D. Hatziantoniou and J. N. Mourjopoulos, "Results for Room Acoustics Equalisation Based on Smoothed Responses," presented at the 114th A.E.S. Convention, Amsterdam, The Netherlands, 2003.

[6] P. Rubak and L. G. Johansen, "Design and Evaluation of Digital Filters Applied to Loudspeaker/Room Equalization," presented at the 108th A.E.S. Convention, Paris, France, 2000.

[7] J. A. Pedersen and K. Thomsen, "Fully automatic loudspeaker-room adpatation," presented at the AES 32nd International Conference, Hillerłd, Denmark, 2007 September 21-23, 2007.

[8] F. E. Toole, "The Measurement and Calibration of Sound Reproducing Systems " A.E.S. Journal, vol. 63, No. 7/8, 2015.

[9] F. E. Toole, "The Measurement and Calibration of Sound Reproducing Systems " A.E.S. Journal, vol. 63, no. 7/8, pp. 512-541, 2015.

[10] L. A. GEDEMER, "PREDICTING THE IN-ROOM RESPONSE OF CINEMAS FROM ANECHOIC LOUDSPEAKER DATA," presented at the AES 57th International Conference, Hollywood, CA, USA, 2015.

[11] B. Bank, "Combined quasi-anechoic and in-room equalization of loudspeaker responses," presented at the 134th A.E.S. Convention Rome, Italy, 2013.

[12] P. Newell, K. Holland, J. Newell, and B. Neskov, "New Proposals for the Calibration of Sound in Cinema Rooms," presented at the 130th A.E.S. Convention, London, UK, 2011.

[13] W. Ahnert, J. Baird, S. Feistel, and P. Meyer, "Accurate Electroacoustic Prediction utilizing the Complex Frequency Response of Far-Field Polar Measurements," presented at the 108th A.E.S. Convention, Paris, 2000.

[14] S. Feistel and W. Ahnert, "The Significance of Phase Data for the Acoustic Prediction of Combinations of Sound Sources," presented at the 119th A.E.S. Convention, New York, 2005.

[15] S. Feistel, W. Ahnert, C. Hughes, and B. Olson, "Simulating the Directivity Behavior of Loudspeakers with Crossover Filters," presented at the 123th A.E.S. Convention, New York, 2007.

[16] Standard on acoustics - Sound source modeling - Loudspeaker polar radiation measurements, 2008 (reaffirmed 2014).

[17] M. R. Gander, "Ground-PlaneAcoustic Measurement of Loudspeaker Systems," J. Audio Eng. Soc., vol. 30, no. 10, October 1982 1982.

[18] C. J. STRUCK and S. F. TEMME, "Simulated FreeField Measurements," J. Audio Eng.Soc., vol. 42, no. 6, 1994.

[19] E. Benjamin, "Extending Quasi-Anechoic ElectroacousticMeasurements to Low Frequencies," presented at the 117th A.E.S. Convention, San Francisco, CA, USA, 2004.

[20] W. Klippel and C. Bellmann, "Holographic Nearfield Measurement of Loudspeaker Directivity," presented at the 141st A.E.S. Convention, Los Angeles, USA, 2016.

[21] M. Sanalatii, P. Herzog, M. Melon, R. Guillermin, J.-C. Le Roux, and N. Poulain, "Measurement of the frequency and angular responses of loudspeaker systems using radiation modes," presented at the 141st A.E.S. Convention, Los Angeles, CA, USA, 2016.

[22] M. Sanalatii et al., "Assessment of the radiation mode method for in situ measurements of loudspeaker systems," presented at the 142nd A.E.S. Convention, Berlin, Germany, 2017.

[23] D. W. Gunness, "Improving Loudspeaker Transient Response with Digital Signal Processing," presented at the 119th A.E.S. Convention, New York, New York USA 2005.

[24] W. Klippel, "Active Compensation of Transducer Nonlinearities," presented at the AES 23rd International Conference, Copenhagen, Denmark, 2003.

[25] D. W. Gunness, "Loudspeaker Transfer Function Averaging and Interpolation," presented at the 111th A.E.S. Convention, New York, NY, USA, 2001.

[26] J. Panzer and L. Ferekidis, "The use of continuous phase for interpolation, smoothing and forming mean values of complex frequency response curves," presented at the AES 116th Convention, Berlin, Germany, 2004 May 8–11, 2004.

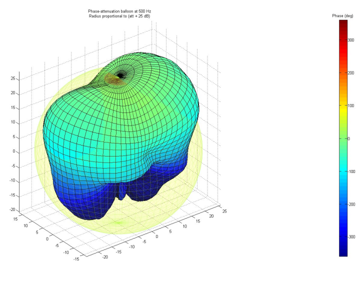
[27] D. Ponteggia, "AUTOMATED BALLOON MEASUREMENTS WITH CLIO 10," Audiomatica, 2009.

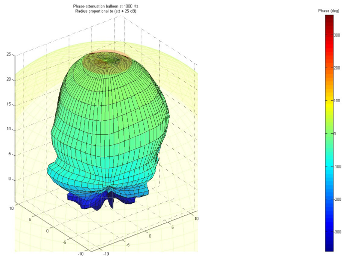
[28] V. K. Ingle, Digital signal processing using MATLAB, 3rd ed. ed. Stamford, Conn.: Stamford, Conn. : Cengage Learning, 2012.

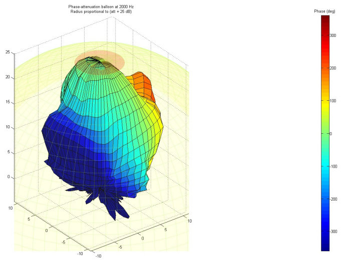
[29] F. Toole, Sound reproduction : the acoustics and psychoacoustics of loudspeakers and rooms. Oxford: Oxford : Focal, 2008.

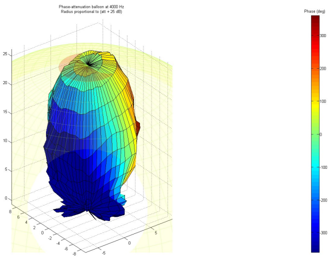
Appendix A.

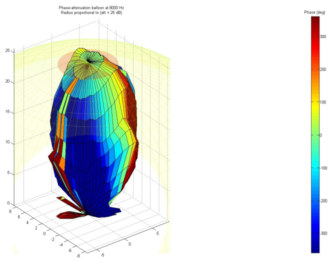
Combined phase-attenuation balloons.




Phase-attenuation balloon at 1000 Hz
Radius proportional to (att + 25 dB)
Phase (deg)


Phase-attenuation balloon at 2000 Hz
Radius proportional to (att + 25 dB)
Phase (deg)


Phase-attenuation balloon at 4000 Hz
Radius proportional to (att + 25 dB)
Phase (deg)


Phase-attenuation balloon at 8000 Hz
Radius proportional to (att + 25 dB)
Phase (deg)