AURALISATION LEVEL CALIBRATOIN

S Di Rosario Buro Happold
A Meachin Buro Happold

1 INTRODUCTION

This paper describes the study we have performed to establish a practical procedure for sound level and frequency response calibration of an auralisation reproduction system for the purpose of business presentation of acoustic issues.

1.1 The Problem of Auralisation Accuracy

Auralisations are demonstrations of simulated acoustic environments in a form that can be heard, and may include combinations of recorded and synthesized sounds.

It is a regular requirement that auralisation demonstrations have to be conducted in less than ideal situations. Necessary compromises may include:

- Demonstration in a room with poorly controlled acoustic conditions, such as high levels of reverberation and background noise (e.g. client's meeting rooms)
- The requirement to simultaneously demonstrate to an audience of many people
- Limited space meaning that loudspeakers have to be placed close to reflecting boundaries and to audience members
- Set-up time constraints limiting system adjustments
- The need to use compact and portable equipment

These constraints will give rise to inaccuracies in frequency response linearity, sound level, time domain response and background noise level.

1.2 Our Approach to Auralisation Accuracy

In our experience, although the constraints identified above lead to poorer reproduction quality than could be achieved in good conditions, a sub-optimal system can still be a useful communication tool if its limitations are understood.

Careful consideration must be given to whether the accuracy of a demonstration will be adequate for it to fulfill its purpose. This consideration does not automatically exclude an auralisation from being used for any given purpose (e.g. choice of reverberation time design target or predicted sound insulation performance), but does require analysis of whether the important features of the demonstration can be adequately experienced despite its faults. For example, it would be pointless to attempt to experience an NR20 noise level in a meeting room with NR35 background noise level.

This paper principally addresses the issue of frequency response linearity.

1.3 Analysis of an Example Auralisation System

The system under consideration is a first order horizontal only Ambisonic four loudspeaker portable rig that could easily be set up in a standard meeting room for the purpose of presentation to client or design team. Figure 1 illustrates the system configuration (described in more detail in section 3.1), and seats and loudspeakers are labeled with references used later in this paper. We consider the use of this reproduction system in a typical meeting room without any deliberate treatment to control room acoustics.

Horizontal first order Ambisonic reproduction has been selected as it is a flexible technique with the ability to reproduce 2D sound fields using a small number of loudspeakers compared to other techniques. It also provides a mathematical encapsulation of most of the scientifically recognized auditory localization models except for the pinna coloration and the impulsive inter-aural time delay models [1]. Ambisonics represents an extensively studied surround reproduction technique with plenty of scientific literature issued and likely to be developed in the future.

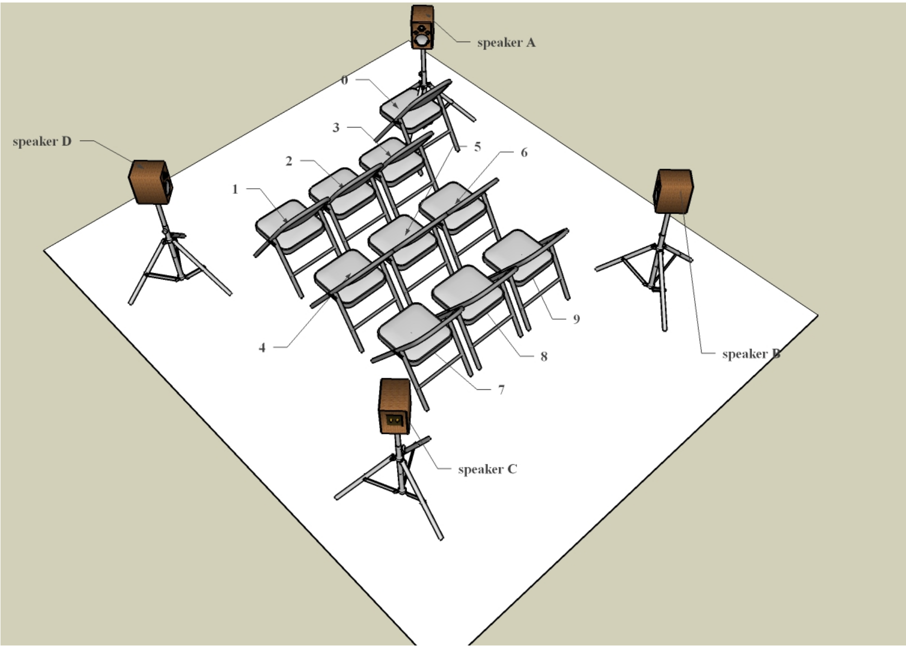

Figure 1 - Sound reproduction system under consideration

2 REVIEW OF AMBISONICS

2.1 Ambisonic Theory

The Ambisonics theory is based on a spherical harmonic decomposition of the sound field whose elements correspond to the instantaneous sound pressure (W, omni-directional channel) and the three components of its gradient (X, Y, and Z channels with axes in the direction of the Cartesian axes), which are related to the particle velocity at a point in space [2].

The loudspeaker signals are derived by using a linear combination of these four channels (also called B-Format), where each signal is dependent on the actual position of the loudspeaker in relation to the centre of an imaginary sphere the surface of which passes through all available loudspeakers. For two-dimensional first order reproduction only the W, X and Y channels are used.

The decomposition of the sound field in only four components is called first order Ambisonics and defines the minimum number of components to reproduce three dimensional sound field information. Higher order Ambisonics has been studied and developed since the early years of the theory with the aim of accurately reproducing the sound field over a larger area than can be achieved by first order (first order only reproduces the sound field over a very small area normally identified around the geometric centre of the loudspeaker rig).

Second order Ambisonics requires five additional B-format signal (R, S, T, U, V) while third order contains a further seven (K, L, M, N, O, P, Q) and so on, introducing the need for a more complex system for the reproduction of all the components. This complexity is the reason why most practical applications are only first order reproductions, as is the one we are presenting in this paper.

M. Gerzon, who is considered the father of Ambisonic meta-theory, defines two primitive models for hearing perception that are characterized by the velocity localization vector $\mathbf{r_V}$ and the energy localization vector $\mathbf{r_E}$ [3]. These mathematically encapsulate the Inter-aural Time Difference (ITD) and Inter-aural Level Difference (ILD) hearing mechanisms respectively. The direction of these vectors indicates the direction of the localization perception and the magnitude indicates the quality of the localization; in the real case of natural hearing their magnitude equals one and their direction always matches the real source directional. In the case of Ambisonic reproduction this cannot be reached due to various factors beyond the scope of this paper.

According to Gerzon's model [4], several different strategies for Ambisonic reproduction can be used:

- "Velocity decoding": maintaining the magnitude of the velocity vector near unity for all frequencies. This optimizes perception by the ITD (Inter-aural Time Difference) mechanism. The decoding process reproduces the exact pressure and velocity information over large area at low frequencies (<700 Hz) but reconstructs the sound field perfectly only in a very small area for high frequencies (700 Hz to 4 kHz).

- "Energy decoding": maximizing the energy vector magnitude over the entire frequency range. This optimizes perception by the ILD (Inter-aural Level Difference) mechanism. The decoding process reproduces the pressure exactly but the velocity component of the sound field is reduced by a factor of $\sqrt{2}$, which enlarges the listening area at mid to high frequencies but gives poorer source localization results at the array centre. This is the best chose if shelf filters are not employed, and is appropriate for large listening areas.

- "Velocity and Energy decoding": Velocity vector magnitude is near unity in the low frequency range (<700 Hz) and energy vector magnitude is maximized in the mid to high frequencies range (700 Hz to 4 kHz). This decoding process is the one proposed for designing Ambisonics decoders that comply with the theoretical conditions for small audience area [5].

2.2 Ambisonic Decoder Filters

Standard Ambisonic decoders perform a number of functions to covert B-format signals into signals for amplification and reproduction by loudspeakers. These are listed below, and our approach to each function is stated.

1. Adjust the ratio of pressure (W) and velocity (X, Y & Z) signals to optimize sound source direction perception according to psycho-acoustic models described in section 2.1. Required ratios may be different at low and high frequencies, and if so 'shelf filters' are used to achieve this [6].

   Velocity decoding (ITD optimization) requires the signals sent to some loudspeakers to have opposite polarity to those in the simulated source direction, which for a large listening *area can cause incorrect directional cues. As is a common solution for large listening areas [6,7,8], we have chosen for our example configuration to use energy decoding (optimize ILD) at all frequencies, and therefore not use shelf filters.*

2. Adjust signal output levels for all B-format signals above and below the transition frequency of the shelf filters to match perceived output in the two frequency ranges. Gains for this adjustment may be based on assumptions of either coherent or energy summation in the regions either side of the transition frequency. This is discussed further in section 2.3.

   This adjustment may be made by the above shelf filters if the assumption is made that the ITD / ILD perception transition is at the same frequency as the coherent / energy transition. Our approach is not to make this correction analytically, but according to measurements at listening positions.

3. Adjust the ratio and phase relationship of pressure and velocity signals to compensate for the low-frequency acoustic impedance changes when listeners are in the near field of loudspeakers [9, 10].

   Our implementation does not include near field correction filters. Such filters are normally used to 'fine tune' the velocity decoding and are not appropriate for our energy decoding at all frequencies. Also, there would be a difficulty in determining appropriate filter values due to the variation in listener distance.

4. Add 3dB gain to the W channel to compensate for the 3dB gain reduction inherent in the Ambisonic signal definition [11].

   Our processing method corrects the W channel gain as required.

5. Apply a gain to all signals if required.

   We apply a calibrated gain to allow reproduction at the required absolute sound level.

2.3 Low-Frequency Coherent Summation

At the exact centre of the loudspeaker array, the signals from all loudspeakers will sum coherently at all frequencies. At positions away from the centre of the array, low frequencies will sum approximately coherently and high frequencies may be considered to sum energetically; the transition frequency between these two frequency ranges is influenced by the distance from the array centre. The implications of this are that close to the array centre low frequencies will have greater emphasis than at listening positions farther from the centre.

The sound levels at the extreme low and high frequency conditions can be expressed as:

$$\mathrm{L\ (low\ frequency)} = 10\log\left(\sum_{i=1}^{\mathrm{N}} \mathrm{g_i}\right)^2$$

$$\mathrm{L\ (high\ frequency)} = 10\log\left(\sum_{i=1}^{\mathrm{N}} \mathrm{g_i}^2\right)$$

Evaluation of these equations using first order Ambisonic energy decoding shows that at low frequencies coherent summation will mean that the signals are 6dB higher than those from a single loudspeaker. At high frequencies, energetically averaged signals are 1.76dB higher than for a single loudspeaker. Note that these values are particular to a two dimensional four loudspeaker first order array.

The phenomenon is illustrated in Figure 2. Each line shows the calculated averaged energy over a range of receiver positions at the stated distance from the centre of an array due to a source reproduced by a first-order horizontal only square Ambisonic rig. In calculating these curves, we have assumed that loudspeakers are a long way from the listeners and hence plane waves reach listening positions. The curves are ⅓-octave smoothed. Values are relative to a single loudspeaker with the same signal as an Ambisonic source in the direction of that loudspeaker.

If an Ambisonic sound reproduction system includes a range of listening positions corresponding to those in Figure 2, an appropriate frequency response equalization curve must be chosen to give the best listening conditions at these positions. The line labeled 'mean' is the arithmetic average of the decibel values of the curves shown. The inverse of this curve could be considered the best equalization choice to use, as a compromise between equalizing response in any of the various listening positions.

The line labeled 'filter' shows the inverse of a simpler shelf-type filter response that might be used in an attempt to correct the difference between low and high frequencies. Note that this filter is not the same as the shelf filters often applied to optimize Ambisonic direction perception (which have different values for W and X/Y channels) [6], and is applied to all B-format signals.

It is clear that whatever filter is used the 'corrected' response will have significant errors at some listening positions and frequencies, of the order of ±4dB.

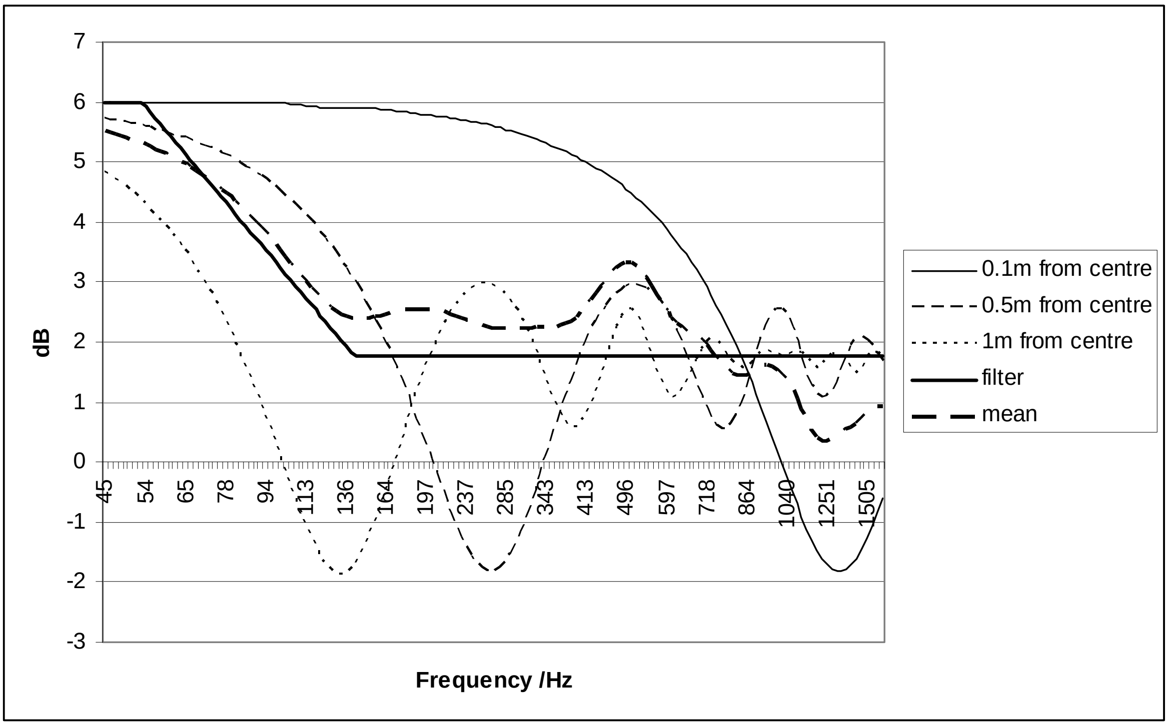

Figure 2 - Coherent and energetic summation regions at various distances from array centre

The filter shown has a slope of -3dB per octave. It approximately matches the response at a distance d from the array centre when it has a centre frequency equal to 0.2c/d, where c is the speed of sound. We have determined this slope and centre frequency empirically to approximately match the response variations calculated.

2.4 Absolute Level Information

There is no standard method to control or document sound levels represented by Ambisonic systems, so we have developed the following system of recording this information.

- Ambisonic microphone sensitivity information is obtained from the microphone supplier, based on tests in anechoic conditions. Pistonphone calibrators are used to confirm the sensitivity of omni-directional microphones

- Electronically generated calibration tones (for Ambisonic microphones) or pistonphone reference tones (for omni-directional microphones) are recorded at the start of each recording session.

- A 'recording record sheet' is kept for every recording, noting calibration levels, equipment used, environmental details and any other necessary information

- Sound files are stored electronically, and file names indicate the sound level represented by full scale signals, for example, *recording_FSD108dB.wav*

- A record is kept of any sound file manipulation carried out, and output files are named to indicate sound levels represented by full scales

- Where sound files are used for which no level calibration information is available, assumptions must be made to estimate the sound levels represented. The assumptions used are recorded and files are named to indicate full scale level representations.

These measures systematically track what sound levels files represent, and allow auralisations to be calibrated to absolute sound levels.

3 EXPERIMENTAL METHOD

3.1 Test Set-up

The main components of the sound reproduction system under test were:

- Four *Genelec 8030A* powered loudspeakers mounted on microphone stands with tweeters at a height of 1.2m

- *Edirol UA-101* sound card

- Laptop running:
    - *VVmic for Tetramic* software to perform Ambisonic encoding [12]
    - *Sonar* software for reproduction control
    - *Audacity* software for audio file manipulation [13]

Loudspeakers were located at the corners of a square with 4m sides, in the arrangement illustrated in Figure 1.

Seats rows were separated by 0.72m, and seats within a row by 0.52m. The centre of the loudspeakers was mid-way between ear positions at seat 5. Ear separation distance was assumed to be 0.15m, and ear height 1.2m.

The system used for reproduction system measurements was:

- *Earthworks M30* omni-directional microphone on microphone stand at height of 1.2m
- *Edirol UA-101* sound card
- *Dell* laptop running *Easera* software
- *Dummy Head* fitted with *AKG C462CM* microphones at each ear canal entrance position

3.2 Measurement Methods

Impulse responses were measured at each ear position at each of seats 0, 3, 5 and 6. All seats were unoccupied during tests. Four repeats of 5.5s swept sine stimulus signals were averaged at each measurement position.

3.3 Equalisation Procedure

System equalization was performed in two stages:

1. Equalize the sound pressure response of individual loudspeakers over the usable frequency range
2. Apply a single filter to all loudspeakers to compensate for the low-frequency response increase when using the loudspeakers in an Ambisonic arrangement (due to more that one loudspeaker being active for a given source direction).

To allow frequency response equalization, the apparent sound level at each seat had to be derived from the measurements made at two ear positions. This was done by calculating average energy at the two positions. However, it is not obvious that this is the best method, and further study may help clarify the issue.

A ⅓-octave graphic equalizer in *Audacity* software was used to adjust the frequency response of each loudspeaker individually. The gain of each band was determined by calculating the value that would make that arithmetic average of decibel frequency responses at each seat zero. This equalizer is limited to whole decibel values for gain in each band.

Following initial measurement of impulse responses, all further manipulation and analysis was performed on the impulse responses, without further acoustic measurements.

3.4 Audience Influence Measurement

An additional set of measurements was made to determine typical effects on frequency response caused by other members of the audience.

These tests were conducted in a slightly smaller meeting room, again with no deliberate treatment to control reflections, shown in Figure 3. Loudspeakers were placed at the corners of a 2m x 2m square. The main listening position was in the centre of the loudspeaker array (shown occupied in Figure 3) and a further three audience seats located behind, to the left and behind-left of the main listening seat. Note that this arrangement includes a total audience of four (including dummy), and a smaller total audience area used for other measurements..

Impulse response measurements were made using the method described in sections 3.1 and 3.2, and frequency response analyzed in ⅓-octave bands using *Easera* software. However, for this test a dummy head fitted with microphones at ear positions was used. This was a shop dummy head fitted with *AKG C462CM* microphones rather than an acoustically correct head and torso simulator, but was nonetheless considered of adequate quality to indicate the magnitude of response changes at ear positions. The dummy head was placed so that microphone positions were 1.2m above floor level. The effective frequency response was approximated by calculating the mean energy in ⅓-octave bands at each ear.

Frequency responses were measured for each loudspeaker individually for the following two situations:

- Dummy head on centre seat, all other seats unoccupied
- Dummy head on centre seat, human listeners in the three other seats

The difference between the two sets of measurements was then calculated.

Figure 3 - Room used for measurements of audience influence

4 EXPERIMENTAL RESULTS

4.1 Impulse Responses

Figure 4 shows the early part of a typical measured impulse response. Significant floor and ceiling reflections can be seen.
To illustrate room acoustic conditions, Figure 5 shows T_{20} reverberation time calculated from this impulse response. Mid-frequency reverberation time was approximately 1.0s.

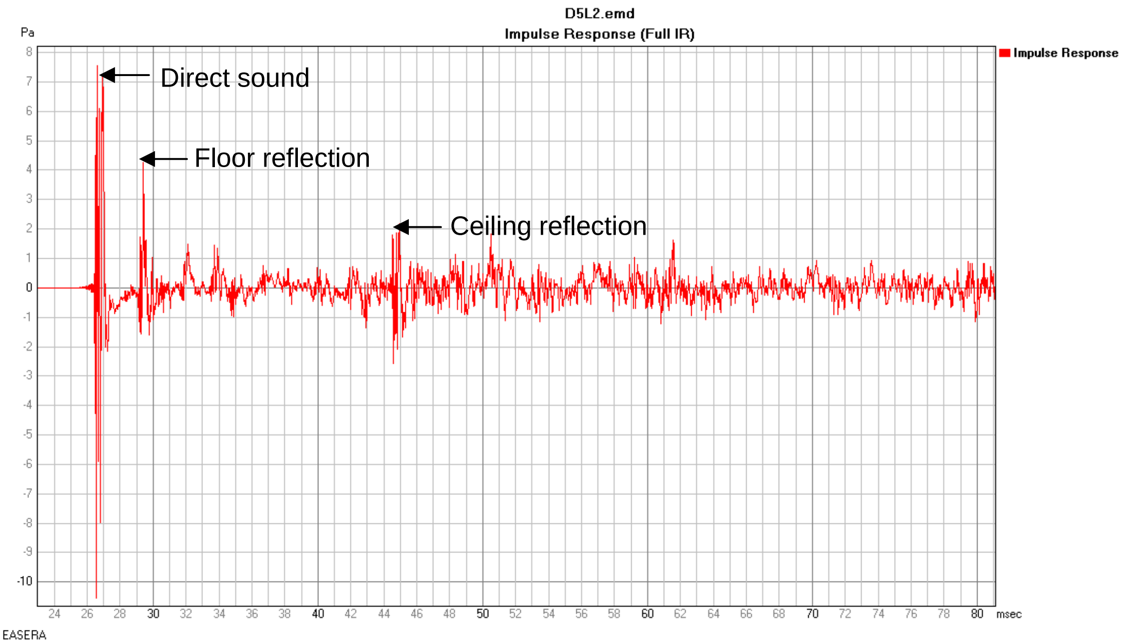

Figure 4 – Early part of typical loudspeaker impulse response (loudspeaker D, seat 5, left ear)

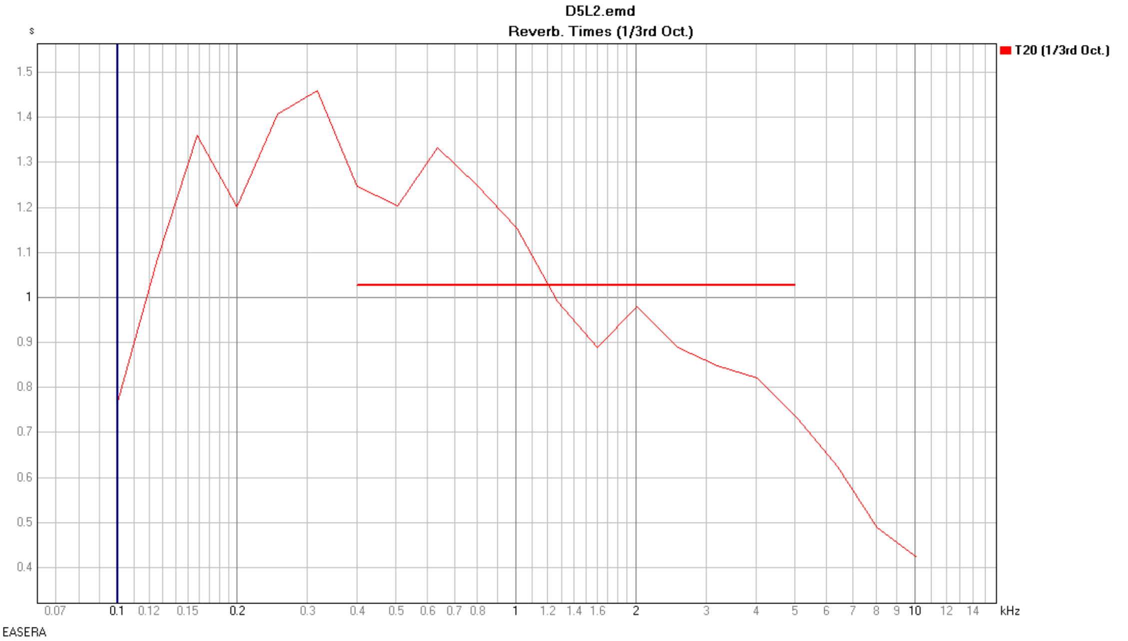

Figure 5 - Measured room reverberation time T_{20}

4.2 Loudspeaker Frequency Response Equalisation

Figure 6 shows the measured frequency response from loudspeaker A to all ear positions before any equalization modifications were made. Frequency response was calculated by Easera software, and averaged over each ⅓-octave band. Similar results were obtained for the other three loudspeakers.

Note that the loudspeaker has a low output below 50Hz, so it was decided not to attempt to control response below this frequency.

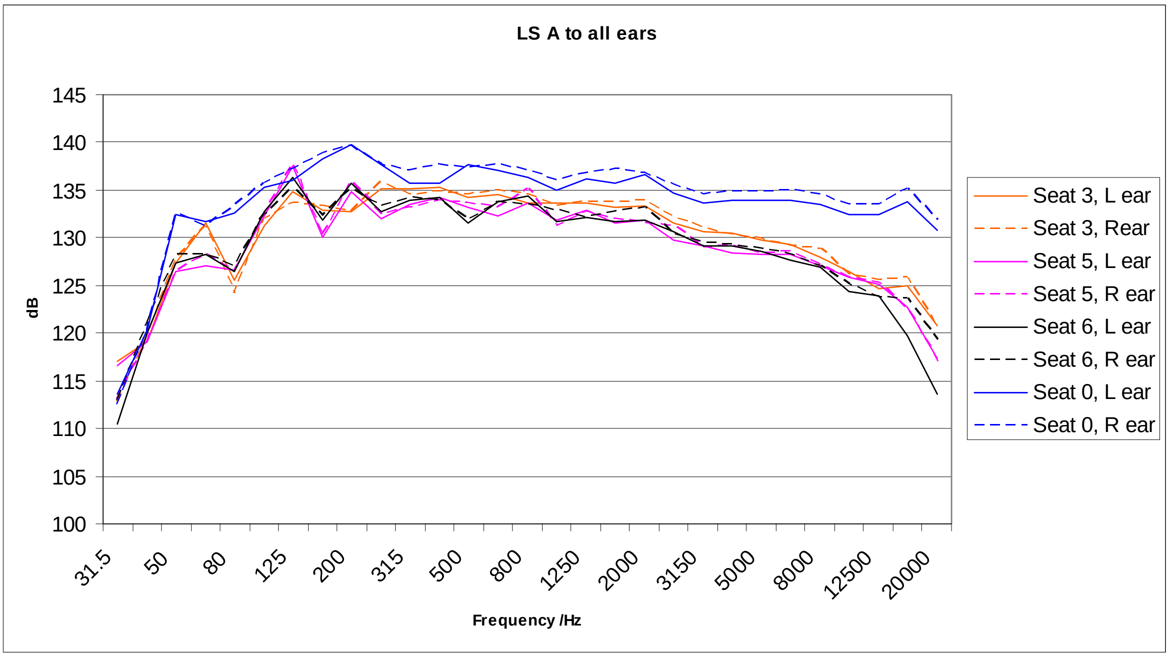

Figure 6 - Measured frequency response of loudspeaker A at all ear positions

Graphic equalizer band gains were calculated as described in section 3.3. Figure 7 shows the calculated graphic equalizer gains for each ⅓-octave band for each loudspeaker. Note that the exact response of the filter differs from this plot, as the filter control settings are illustrated rather than the detailed filter response to these settings. Measured impulse responses were processed using these filters to simulate the effect of applying the filters to sound files.

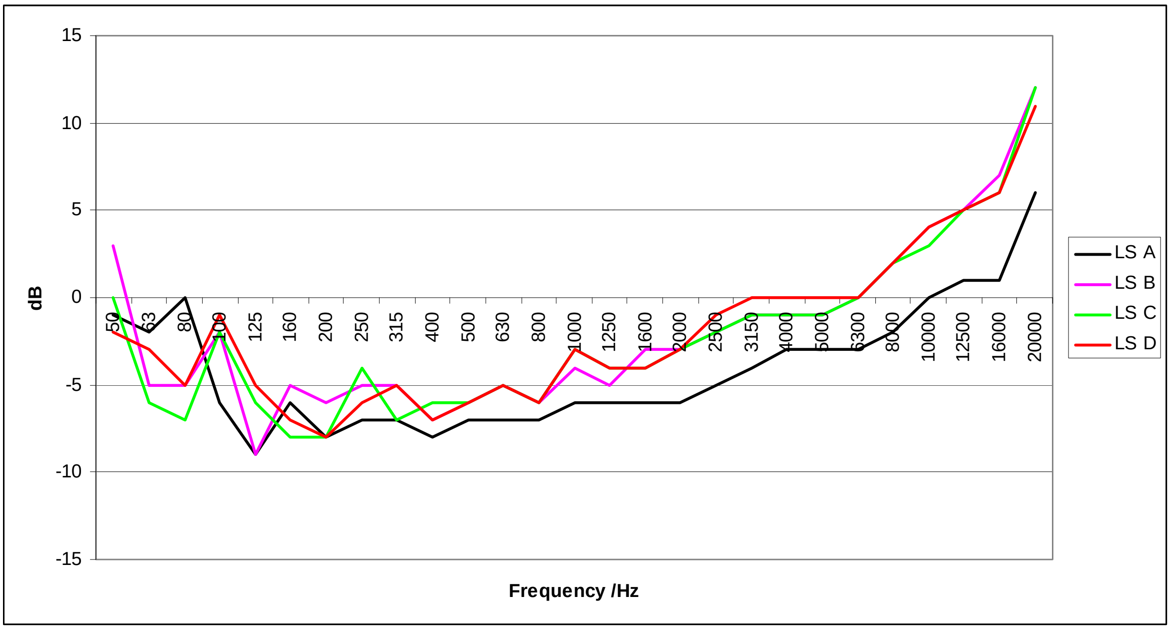

Figure 7 - ⅓-octave equalization applied to each loudspeaker

4.3 Ambisonic Source Frequency Response Equalisation

Impulse responses were then combined with appropriate individual gain adjustments to simulate the impulse responses of Ambisonic sources reproduced by the loudspeaker rig at each ear position.

Simulated source directions were:

- Loudspeaker A
- Loudspeaker C
- Mid-way between loudspeakers A and B
- Mid-way between loudspeakers C and D

All simulated sources were plane wave sources.

To determine the effective frequency response at each seat, the arithmetic mean of the simulated sources (in decibels) to that seat were calculated. The results are shown in Figure 8, along with the mean of these values. Values are relative to a single loudspeaker with the same signal as an Ambisonic source in the direction of that loudspeaker.

It can be seen that there is a significant low-frequency level increase of a similar form to that predicted in section 2.3. The low-frequency boost is higher for seats close to the array centre as expected. The values at high frequencies are close to the predicted value of +1.76dB. However, a significant peak is seen at 125Hz for most seats, which is above the range of frequencies calculated to suffer from significant low-frequency coherent summation error. The cause of this is not clear, but may be reflections from room boundaries.

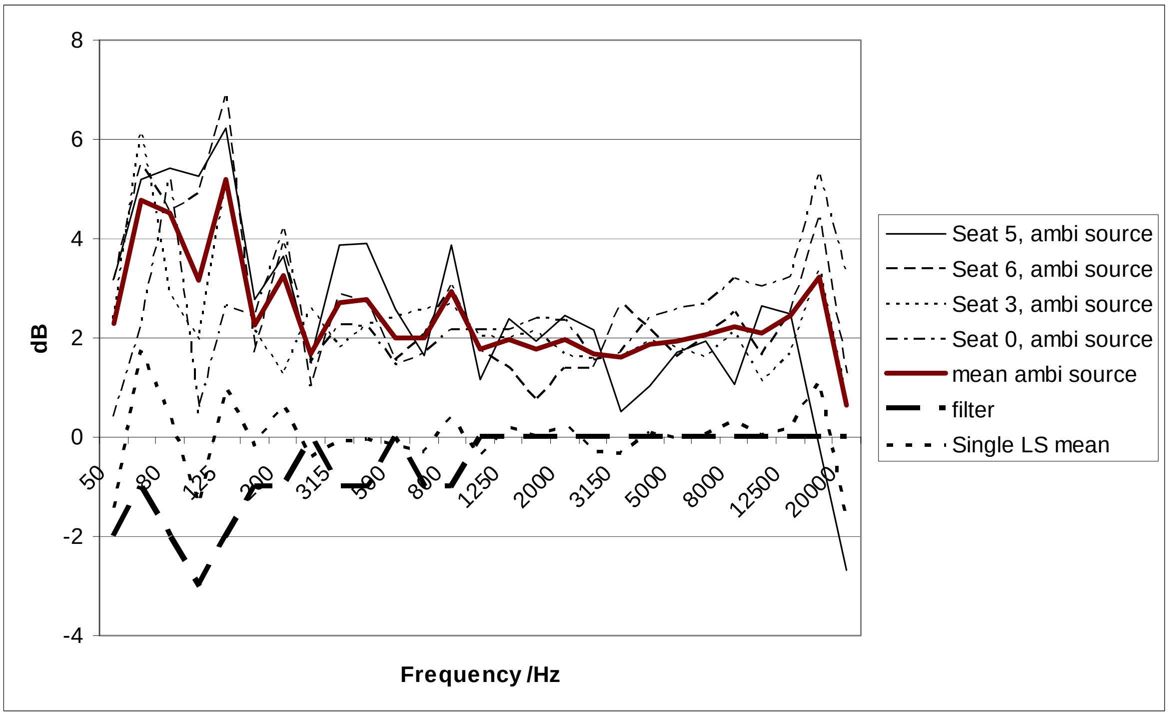

Figure 8 - Mean frequency response at each seat for simulated Ambisonic sources before the application of low frequency correction filters. Also shown are mean response from single loudspeakers and low-frequency correction filter band-centre values

A filter to compensate for the difference in frequency response between a single loudspeaker and a simulated Ambisonic source was derived. Again, the Audacity graphic equalizer was used, and ⅓-octave band settings are shown as 'filter in Figure 8. This filter was applied to all impulse responses for simulated Ambisonic sources.

4.4 Residual Errors

Following the application of the low-frequency compensation filter, the mean of ⅓-octave frequency responses over all Ambisonic source directions and seats was calculated. As before, this was based on an energy average over each pair of ear positions. Figure 9 shows this result, along with the standard deviation from this 'average frequency response' curve at all seats, standard deviation from 0dB and minimum and maximum values calculated. The frequency response is shown normalized to a mean value of 0dB.

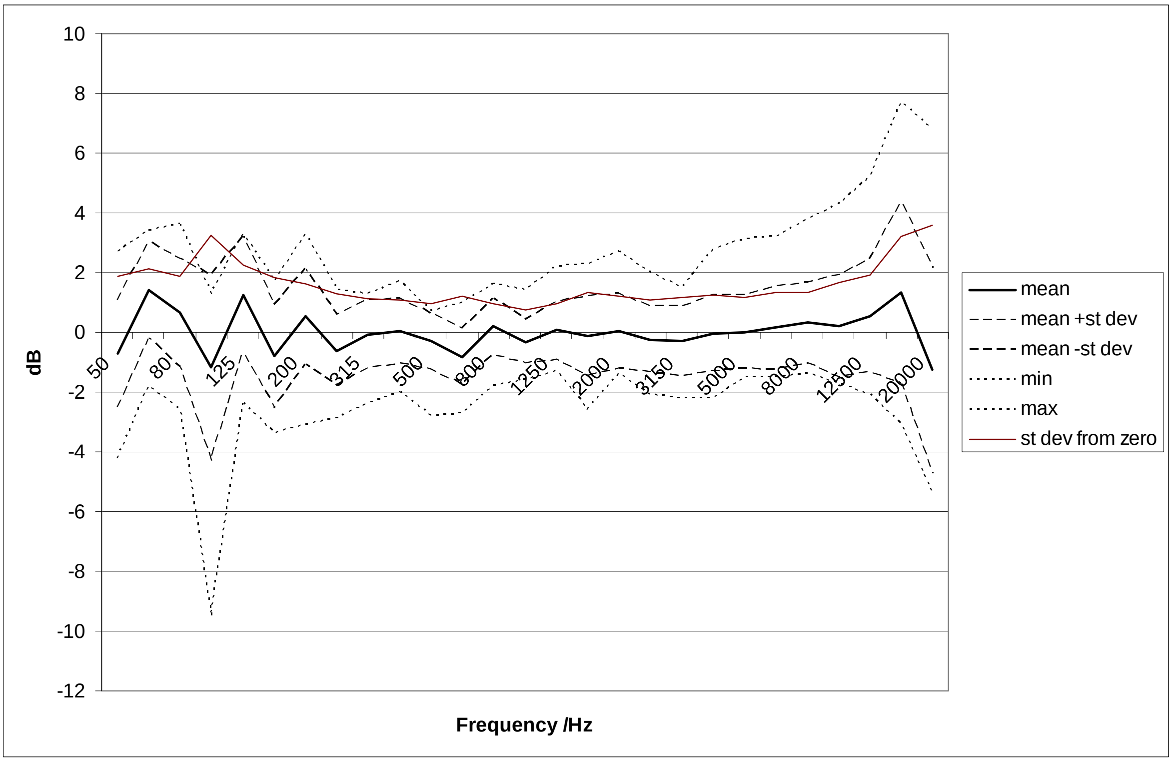

Figure 9 - Mean frequency response for all simulated Ambisonic sources at all seats

We could summarize the frequency response accuracy of the system as:

- ⅓-octave standard deviation in 50Hz to 20kHz bands: <4dB
- ⅓-octave maximum deviations at any seat in 50Hz to 20kHz bands: +8dB, -10dB
- ⅓-octave standard deviation in 250Hz to 8kHz bands: <1.5dB
- ⅓-octave maximum deviations at any seat in 250Hz to 20kHz bands: +4dB, -3dB

If the sounds to be auralised were known, it would be possible and useful to calculate the accuracy of the system for their specific spectra. For example, if pink noise band limited to 50Hz to 10kHz was of interest, we could state:

- Standard deviation of A-weighted level of pink noise 50Hz to 10kHz: <1dB
- Maximum deviation of A- weighted level of pink noise 50Hz to 10kHz: +3dB, -2dB

This approach would permit a simple description of accuracy for a given sound file to be stated, which would be likely to be better understood by users of the demonstration.

4.5 Sound Level Calibration

Although sound level calibration of the Ambisonic sound reproduction system was not carried out experimentally as part of this investigation, the results above have implications for the best method and accuracy of the process.

Using the method described in section 2.4, a sound file representing a known sound level may be reproduced by the loudspeaker system. The overall system gain can then be adjusted to match the required level.

Narrowband signals are not suitable for level calibration, as they can vary significantly from one measurement position to another. Also, wideband noise would be subject to large variations as low frequencies (due to loudspeaker interference) and high frequencies (due to variations in direct / reverberant ratio). Therefore, we suggest that in this case analysis of pink noise over the range 250Hz to 2kHz would be a suitable method to adjust overall reproduction level.

4.6 Subjective Assessment

Reproduction quality was briefly and informally assessed by the experimenters, comparing quality with and without the use of all equalization and low-frequency compensation filters. A recording of a talker moving around the Ambisonic microphone was assessed. Key observations were:

- Perceived source direction quality was fairly good, despite room acoustic conditions
- Apparent source position differed slightly between signals processed by the loudspeaker equalization filters and those without this processing

4.7 Audience Influence Measurements

The change in frequency response due to the presence of three audience members was measured as described in section 3.4, and the results are shown in Figure 10.

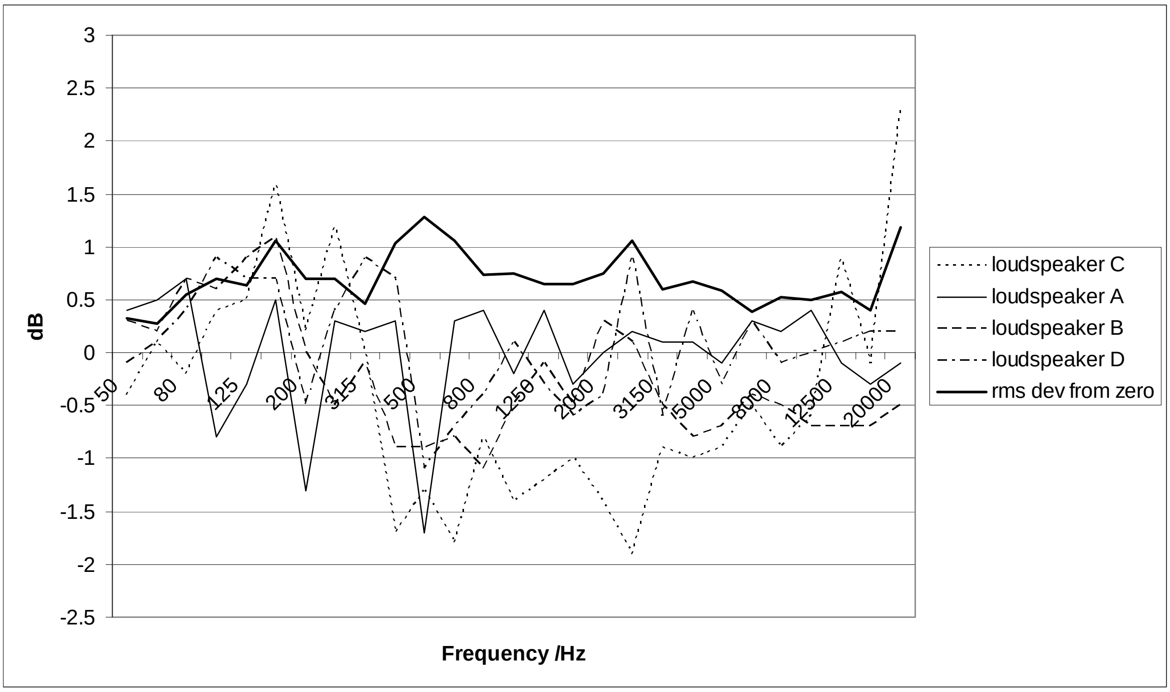

Figure 10 - Change in frequency response due to presence of audience

A complicated set of variations can be seen, as would be expected to result from the reflections from audience members.

It can be seen that typical change in level for a ⅓-octave band was of the order of ±1dB.

Note that the repeatability of this measurement was not checked, and some response changes may be due to slight movements of the dummy head during cable changes. To confirm the validity of these results this aspect should be checked.

These errors can be summarized as:

- Standard deviation of ⅓-octave band level due to audience members: <1.5dB
- Maximum deviation of ⅓-octave band level due to audience members: +2.5dB, -2dB

5 DISCUSSION

5.1 Narrowband Errors

Analysis above has shown how the accuracy of the reproduction system can be quantified by reference to ⅓-octave band average response. It should be remembered though that narrow-band defects greater than those implied by Figure 8 may exist. For some source material this may be significant.

5.2 Low-Frequency Corrections

We have shown how the frequency response of the system under test was not the same for single loudspeakers as for simulated Ambisonic sources. This can be thought of as a result of the number of loudspeaker simultaneously active to simulate a source represented by Ambisonic signals. The difference calculated is particular to first order horizontal only square loudspeaker layouts, and to the assumption of an extended listening area. If an array with more loudspeakers or a three-dimensional array was used, this difference would be larger. Higher order Ambisonic reproduction would be expected to show smaller differences between single loudspeakers and simulated Ambisonic sources [14, 15].

This of course has implications for 'mixed-order' auralisations, which may include for example a mix of a first-order recording, a second-order synthesized sound field and a direct sound source represented by a single loudspeaker. Each of these would need a different low-frequency correction filter to compensate for the number of loudspeakers simultaneously active.

A reproduction system with more extended low-frequency response would also exhibit larger differences than those measured here.

5.3 Improvements to Method

During this investigation, some aspects in which the methods applied to the example system could be improved have become apparent: -

1. The selection of listening positions (seats 0, 3, 5 and 6) was not symmetrical with respect to the loudspeaker array. Seat 0 was very much closer to loudspeaker A than any of the other seat-loudspeaker combinations. This led to the responses averaged over all seat positions incorrectly implying that loudspeaker A had a higher output than the other loudspeakers, and incorrectly making a 'correction' for this in the loudspeaker equalization.

   To prevent this undesirable effect, the sample listening positions should be selected to be symmetrical with respect to the loudspeaker layout.

2. Measured sound levels from loudspeakers at very high frequencies were significantly higher at seat 0 than at other seats. This was because the loudspeakers have higher directivity at high frequencies than at low frequencies, and hence lower reverberant levels at high frequencies. The higher direct / reverberant ratio at seat 0 therefore meant that high frequencies were emphasized.

   It should be noted that the semi-reverberant conditions led to low variation in total sound level over the listening area; for some demonstrations this advantage could outweigh the disadvantage of low time domain accuracy in these conditions.

   To avoid this problem, loudspeakers with less directivity variation could be used, although this is at odds with our requirement for a portable system. An alternative approach is to avoid locating any listening positions too close to any loudspeakers. If this is unavoidable then the best approach may be to simply quantify and understand the errors at these positions.

3. Brief subjective evaluation of reproduction quality suggested apparent source positions changed slightly with signals processed by the loudspeaker equalization filters. This may be due to the filters, which were different for each loudspeaker, introducing phase response differences between loudspeakers. Although in a semi-reverberant environment phase relationships are somewhat chaotic, it may be that maintaining equal direct sound phase responses between the loudspeakers would be beneficial. Further study may clarify this point.

   The use of a single set of filter settings to process all loudspeaker signals may be beneficial. However, this advantage would have to be weighed against the loss of ability to individually correct loudspeaker frequency responses. Linear phase filters may also offer a useful approach.

6 CONCULSIONS

This paper has described a practical method of optimizing frequency and level response of a first order horizontal only Ambisonic reproduction system, and of quantifying system accuracy

Key conclusions of the study are:

- A practical method is presented to quantify the accuracy of a sound reproduction system's sound level response. This gives auralisation users the information they need to understand demonstration accuracy, and hence what degree of trust they can place in the demonstration representing the simulated environment.

- Analysis of the performance of a first order horizontal only four loudspeaker Ambisonic sound reproduction system in a typical meeting room indicated that the standard deviation of 1/3-octave reproduced levels over typical listening positions was <4dB in the range 50Hz to 20kHz. In the range 250Hz to 8kHz standard deviations were <1.5dB.

- Maintaining level calibration and an understanding of error magnitude throughout the production of an auralisation is not a trivial task, and requires careful record keeping.

- For large listening areas, minimizing average frequency response error requires a low frequency correction filter to be applied in addition to single loudspeaker equalization. This correction is dependant on Ambisonic order, number of loudspeakers and number of dimensions of the sound field reproduced.

- Care must be taken in selecting sample listening positions for equalization, and if individual loudspeaker corrections are to be applied they should be symmetrical with respect to the loudspeaker layout.

7 FURTHER WORK

Further work which may help advance understanding and methods of level calibration for auralisations includes:

- Development of a method to quantify the time domain accuracy of a system

- Investigating appropriate methods of combining pairs of responses at ear positions to estimate perceived response

- Investigation into the degree of perceptual degradation caused by the use of dissimilar equalization filters for individual loudspeaker frequency response equalization

8 REFERENCES

1. Lee, R., A.J. Heller and E.M. Benjamin, “Is my decoder Ambisonic?”, presented at the 125th AES convention, San Francisco, USA, 2-5 October 2008;

2. http://en.wikipedia.org/wiki/Ambisonic;

3. Gerzon, M.A., “Pratical periphony: the reproduction of full sphere sound”, Preprint no. 1571 (A6) of the 65th AES Convention, 25th to 28th February 1980;

4. Gerzon, M. A., “Surround Sound Psychoacoustics”, Wireless World, December 1974;

5. Lee, R., E. Benjamin and A. Heller, “Localization in Horizontal-Only Ambisonic Systems”, presented at the 121st AES convention, San Francisco, USA, 8th October 2006, revised and corrected version by the authors;

6. Lee, R., “SHELF FILTERS for Ambisonic Decoders”, http://www.ambisonia.com/Members/ricardo

7. Lee, R. and A. Heller, “Ambisonic Localisation – Part 2”, presented at the 14th International Congress on Sound and Vibration, Cairns, Australia, 9-12 July 2007;

8. Mahlam, D.G., Experience with Large area 3-d Ambisonic sound systems, Proceedings of the institute of Acoustics Autumn Conference on Reproduced Sound 8, 1992;

9. Jerome Daniel, “Spatial Sound Encoding including Near Field Effect: Introducing distance coding filters and a viable, new Ambisonic format”, Preprints 23rd AES International Conference, Copenhagen, Denmark, 2003;

10. Adriansen, F., “Near Field filters for High Order Ambisonics”, http://www.kokkinizita.net/papers/hoafilt.pdf;

11. Farina, A. and E. Ugolotti, “Software Implementation of B-Format Encoding and Decoding, Pre-prints of the 104th AES Convention, Amsterdam, 16-19 May 1998;

12. VVmic for Tetramic software: http://www.core-sound.com/TetraMic/3.php and http://mcgriffy.com/audio/ambisonic/vvmic/;

13. Audacity, Free audio editor: http://audacity.sourceforge.net/;

14. Malham, D., “Higher Order Ambisonic Systems”, abstracted from “Space in Music - Music in Space”, an Mphil thesis by Dave Malham, submitted to the University of York in April 2003, revised and passed in December 2003;

15. Daniel, J., Rozenn N. and Moreu S., “Further Investigations of High Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging”, presented at 114th AES Convention, Amsterdam, The Netherlands, 22-25 March 2003.