



**QUEEN'S
UNIVERSITY
BELFAST**

Spectral impairment of enclosed higher order ambisonics

Stitt, P., Bertet, S., & Van Walstijn, M. (2012). Spectral impairment of enclosed higher order ambisonics. In *45th Audio Engineering Society UK Conference 2012* Curran Associates, Inc.

Published in:
45th Audio Engineering Society UK Conference 2012

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access

This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: <http://go.qub.ac.uk/oa-feedback>

SPECTRAL IMPAIRMENT OF ENCLOSED HIGHER ORDER AMBISONICS

PETER STITT, STÉPHANIE BERTET, MAARTEN VAN WALSTIJN

Sonic Arts Research Centre, Queen's University, Belfast, UK

Ambisonics is spatial audio technique that attempts to recreate a physical sound field over as large an area as possible. Higher Order Ambisonic systems modelled with near field loudspeakers in free field as well as in a simulated room are investigated. The influence of reflections on the image quality is analysed objectively for both a studio-sized and large reproduction environment using the relative intensity of the reproduced sound field. The results of a simulated enclosed HOA system in the studio-sized room are compared to sound field measurements in the reproduced area.

1 - INTRODUCTION

Michael Gerzon [1] first proposed the theory behind Ambisonics in the 1970s as an alternative to the then prevalent quadrasonic systems. It is a multichannel reproduction technique that attempts to recreate a physical sound field over as large a listening area as possible. It is a scalable technique where higher orders allow a larger listening area, but also require more loudspeakers to recreate the sound field [2].

Since the recreated sound field is not reproduced perfectly, different measures of error have been used to describe its accuracy. Three possible measures of the error are pressure field error, D-error [3] and relative intensity. The pressure field and D-error are related to the difference in pressure between the recreated and ideal sound field. The relative intensity gives a ratio of the recreated field intensity to the ideal field intensity. Solvang [4] proposed the use of the relative intensity as measure of error because at higher frequencies the inter-aural level difference (ILD) is used for localisation instead of the inter-aural time difference (ITD). It can also give an indication of any colouration i.e. spectral impairment to the recreated sound image.

Much of the theory of Ambisonics investigates it in anechoic conditions but, when considering a system that is enclosed in a room, ignoring the boundary reflections influence on the recreated image is no longer ideal. The influence of the boundary reflections may become important for systems enclosed in reverberant

environments, as they will influence the character of the image and possibly the listener's ability to localise it. In this situation the assumption that the loudspeakers can be modelled as emitting plane waves can no longer be used to simulate the recreated pressure field, as the wave emitted must be considered in all directions, not just at the listener position. Instead, a finite distance model of the loudspeakers must be used.

This paper begins with a brief overview of Ambisonic theory, followed by an account of Ambisonic reproduction with finite distance loudspeakers. The next section places the Ambisonic system with finite distance loudspeakers in two simple rooms, one studio-size and one large, using the image source method. In the final section studio measurements are taken to verify the model.

2 - 2D HIGHER ORDER AMBISONICS

In its most basic form Ambisonics [1] is used to reconstruct a plane wave by decomposing the sound field using spherical harmonic functions. It is based on the decomposition of an acoustic pressure field using the Fourier-Bessel series

$$p(kr, \theta, \phi) = \sum_{m=0}^{+\infty} i^m j_m(kr) \sum_{n=0}^m \sum_{s=\pm 1} \hat{a}_{mn}^s B_{mn}^s Y_{mn}^s(\theta, \phi), \quad (1)$$

where k is the wave number, θ and ϕ are the azimuth and elevation angles, i is the imaginary unit, $j_m(kr)$ are the spherical Bessel functions and $Y_{mn}^s(\theta, \phi)$ are the

spherical harmonic functions. B_{mn}^s are the coefficients of the spherical harmonic functions that are used to describe the field being reconstructed. In theory an infinite number of spherical harmonics must be used to recreate the sound field but in practice the series must be limited to a finite order N . For a plane wave of incident azimuth θ_s and elevation ϕ_s the spherical harmonic coefficients are

$$B_{mn}^s = S Y_{mn}^s(q_s, f_s), \quad (2)$$

where S is the signal being reconstructed. Ambisonics works by encoding the sound field using spherical harmonics and decoding the encoded signals to the appropriate loudspeaker signals needed to reconstruct the field.

2.1 - Encoding

Since its initial development Ambisonics has been expanded to higher orders [2, 5] as Higher Order Ambisonics (HOA). To encode a signal for 2D higher order reconstruction the equations for the spherical harmonic components are

$$\begin{aligned} B_0^1 &= S, \\ B_n^1 &= S\sqrt{2} \cos(nq_s), \\ B_n^{-1} &= S\sqrt{2} \sin(nq_s), \end{aligned} \quad (3)$$

where n is the HOA order. These equations produce the fully normalised (N2D) components. For a reconstruction of order N the encoding process will generate $2N + 1$ components that are transmitted as separate channels to the decoder.

2.2 - Decoding

The decoding process calculates the appropriate loudspeakers gains needed to recreate the sound field. The gain $G_m(\theta_s)$ of the m -th loudspeaker is a summation of the encoded channels weighted by the appropriate spherical harmonics for the loudspeaker position.

$$G_m(q_s) = \frac{1}{M} \sum_{\hat{e}} B_0^1 + \sum_{n=1}^N \sum_{s=\pm 1} \hat{a}_n^s B_n^s Y_n^s(q_m) \quad (4)$$

where θ_m is the angle of loudspeaker and M is the total number of loudspeakers in the array. This is known as basic decoding but there are options for decoding over large areas, known as in-phase decoding [6], or to focus the concentration of energy, known as max r_E [7].

3 - ANECHOIC HOA REPRODUCTION WITH FINITE DISTANCE LOUSPEAKERS

The theory of HOA is based on recreating a plane wave using a regular loudspeaker array. Much of the literature assumes the loudspeakers are far enough from the listener to be assumed to be emitting plane waves. In

practice the loudspeakers may be placed at a distance where the waves are still spherical in nature across the listening area. Consider the case where $M = 2N + 1$ loudspeakers are used in a regular array and the image is placed directly at the position of one of the loudspeakers. In this case only a single loudspeaker is used to recreate the sound field, which will therefore have the same radiation pattern as the loudspeaker. Daniel [8] compensated for the finite distance of loudspeakers to allow the reproduction of real plane waves or near field sources within the listening area. Furthermore, if the room is to be taken into account then the plane wave model of loudspeakers can no longer be used, as it does not allow boundary reflections to be included. The pressure field create by a finite distance loudspeaker, modelled as a point source, is

$$p(kr, q, q_m) = A_m \frac{e^{-jk|\mathbf{r}_m - \mathbf{r}|}}{|\mathbf{r}_m - \mathbf{r}|}, \quad (5)$$

where A_m is the amplitude of the loudspeaker, r is the radius from the coordinate system origin, k is the wavenumber, \mathbf{r}_m is the position vector of the loudspeaker and \mathbf{r} is a position vector from the origin of the coordinate system. A 2D HOA array with M loudspeakers will therefore have a pressure field described as a summation of the fields created by each of the loudspeakers,

$$p(kr, q, q_m)_A = \sum_{m=1}^M G_m(q_s) \frac{e^{-jk|\mathbf{r}_m - \mathbf{r}|}}{|\mathbf{r}_m - \mathbf{r}|}, \quad (6)$$

where $G_m(\theta_s)$ and \mathbf{r}_m are the HOA gains and position vector of the m -th loudspeaker respectively. θ_s is the angle of the reproduced sound source. The recreated sound field will not be completely accurate due to the truncation of reproduction order, and therefore there will be an error in the reproduced sound field.

3.1 - Error in the Reproduced Sound Field

There are several measures of error to describe the difference between the recreated and ideal sound fields. The pressure field error E_p can be defined as the absolute normalised difference between the recreated p_{ambi} and reference p_{ref} sound fields.

$$E_p = \left| \frac{p_{ref} - p_{ambi}}{p_{ref}} \right|, \quad (7)$$

A second measure of error is the D-error [3], which is defined as

$$\text{D-error} = \frac{1}{2\rho|S|} \int_0^{2\rho} |p_{ambi}(\mathbf{r}(q)) - p_{ref}(\mathbf{r}(q))| dq, \quad (8)$$

where $\mathbf{r}(\theta) = r \cdot \mathbf{u}_\theta$. The D-error is an integral error over a circle centred on the array centre, giving a measure of error at a particular radius. For finite distance sources the $|S|$ term must be replaced by the pressure of the

reference field at the centre of the array to compensate for the decay in amplitude from the spherical waves, therefore making sure the normalisation is correct. This gives the D-error for finite distance loudspeakers as

$$\text{D-error} = \frac{|r_0 e^{jk_0 r_0}|^{2\rho}}{2\rho |S|_0} \int_0^{2\rho} |p_{\text{ambi}}(\mathbf{r}(q)) - p_{\text{ref}}(\mathbf{r}(q))| dq. \quad (9)$$

where r_0 is the radius of the loudspeaker array. For both E_p and D-error, a well reproduced area can be defined as any region with less than 20% (-14dB) error. The size of the well reproduced region depends on the frequency of the signal being recreated and the recreation order. Lowering the frequency or increasing the order will increase the size of the well reproduced area. For a halving of the frequency or a doubling of the reproduction order N the size of the well reproduced region will approximately double.

In the case of finite distance loudspeakers, the radius of the well reproduced region given by plane wave loudspeaker theory could be larger than the radius r_0 of the loudspeaker array. Figure 1 shows the increase in size for the well reproduced region size with decreasing frequency at 3rd order with 8 loudspeakers at a radius $r_0 = 1.6\text{m}$, for both plane wave emitting and finite distance

loudspeakers. The growth of the well reproduced region can be seen for the plane wave loudspeakers but for finite distance sources the well reproduced region does not continue to increase with decreasing frequency. Table 1 shows the radius below which each of the frequencies is reproduced with less than 20% D-error, for both plane wave and finite distance arrays. For the plane wave loudspeakers the well reproduced area continues to increase as the frequency decreases but for smaller, finite distance arrays the radius of the well reproduced region converges to a limit lower than r_0 . For example, with plane wave loudspeakers the radius of the well reproduced region grows from 0.4762m at 320Hz to 1.986m at 80Hz – a 317% increase. Over the same frequencies the region only increases from 0.4664m to 0.7139m – a 53% increase – for an array of radius $r_0 = 1\text{m}$. As the array radius increases the frequency below which the well reproduced region is limited decreases. Increasing the order of the reproduction will raise the frequency limit below which the well reproduced area is limited but will allow the limit to move closer to the loudspeakers.

As shown by Table 1 and Figure 1, there is a limit to the size of the well reproduced area that depends on the radius of the array and the order of the reconstruction.

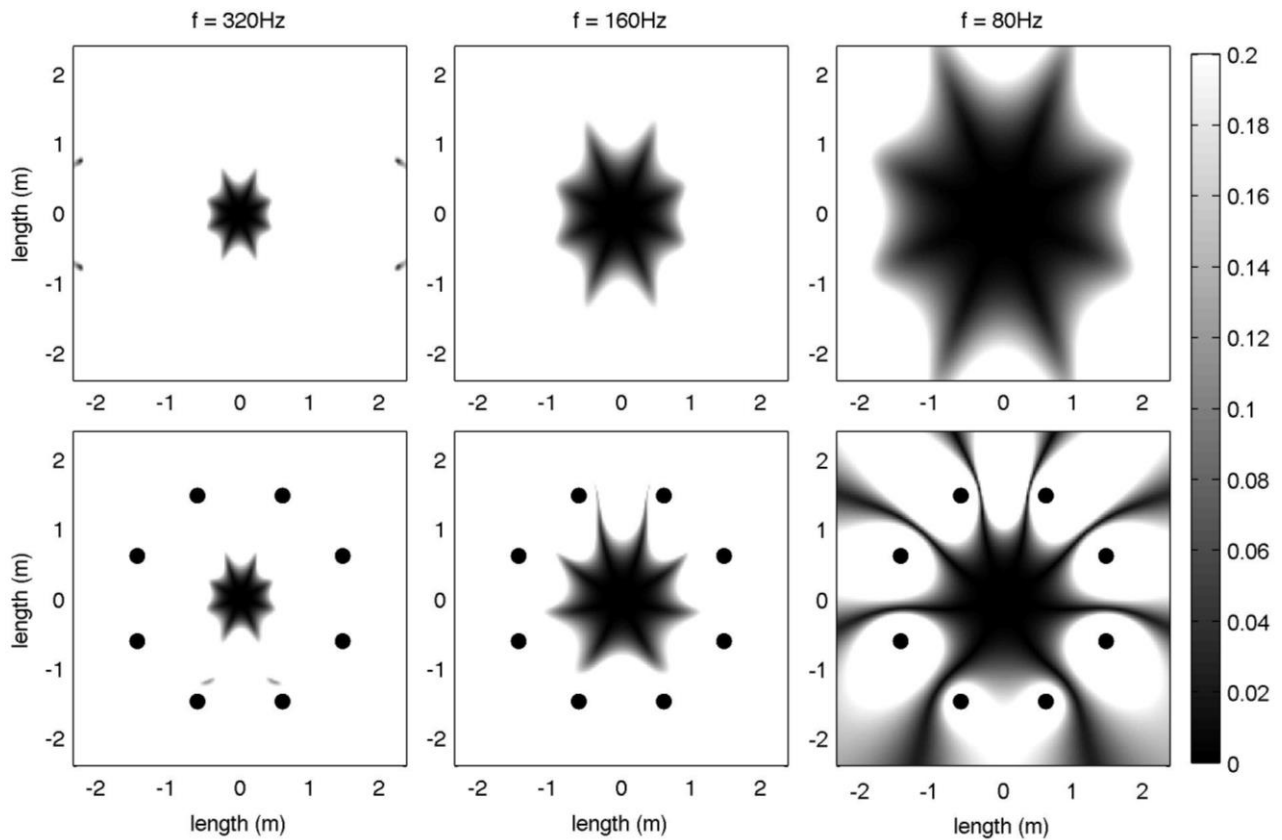


Figure 1 - E_p in the well reproduced area for signals of 320, 160 and 80 Hz with both finite distance and plane wave loudspeakers. The reproduction is 3rd order with 8 loudspeakers and an image at 0 degrees. The error scale is clipped above 20% error.

However, although the well reproduced region may be limited by the array radius, it is possible that in practice a listener will be able to be placed close enough to the loudspeakers before other perceptual cues, such as the precedence effect [9], become significant.

Frequency	320 Hz	160 Hz	80 Hz	40 Hz	20 Hz
Plane wave	0.4967m	0.9936m	1.9874m	3.9748m	7.9496m
$r_0 = 1\text{m}$	0.4664m	0.6730m	0.7139m	0.7215m	0.7233m
$r_0 = 2\text{m}$	0.4904m	0.9328m	1.3461m	1.4278m	1.4431m
$r_0 = 20\text{m}$	0.4967m	0.9931m	1.9836m	3.9448m	7.6679m

Table 1 - Radius of the well reproduced area (D-error < 20%) for frequencies from 20 Hz to 320 Hz. Shows both finite distance and plane wave arrays of 3th order using 8 loudspeakers and the image placed at 0 degree. r_0 is the radius of the loudspeaker array.

Daniel et al [10] use the D-error to find upper frequency limits of the reproduced sound field that are considered well reproduced on a circle just larger than the human head, $r = 8.5\text{cm}$. Table 2 shows the frequency limit over a circle with $r = 8.5\text{cm}$, below which the D-error is less than 20% for order 1 to 5 for both plane wave and finite distance arrays. Each array has $2N + 2$ loudspeakers and the reproduced sound image placed at an angle of 0° . It shows that the frequency limits are approximately equal for both plane wave and finite distance loudspeaker assumptions. The finite distance sources have a limit that is only marginally lower. As the array radius increases the frequency limit converges to the plane wave limit, as would be expected for far field loudspeakers.

Order	1	2	3	4	5
Plane wave	712 Hz	1282 Hz	1871 Hz	2470 Hz	3076 Hz
$r_0 = 1\text{m}$	707 Hz	1268 Hz	1849 Hz	2443 Hz	3049 Hz
$r_0 = 2\text{m}$	709 Hz	1279 Hz	1868 Hz	2467 Hz	3073 Hz
$r_0 = 20\text{m}$	712 Hz	1281 Hz	1869 Hz	2468 Hz	3074 Hz

Table 2 - Upper frequency limit given by D-error for a field reconstruction with less than 20% error on a circle radius 8.5cm for orders 1 to 5.

3.2 - Relative Intensity for Finite Distance Loudspeakers

Another measure of error in the sound field, proposed by Solvang [4], is the relative intensity between the HOA pressure field recreated and the reference signal. For finite distance arrays the reference source will appear to be placed along the circumference of the loudspeaker array. The intensity is proportional to the pressure field multiplied by its complex conjugate. For the reference source placed at the radius of the loudspeaker array this is

$$p(kr, q, q_s)p(kr, q, q_s)^* = |p(kr, q, q_s)|^2 = \frac{|S|^2}{|\mathbf{r}_s - \mathbf{r}|^2}, \quad (10)$$

and for the HOA pressure field it is

$$|p(kr, q, q_s)_A|^2 = |S|^2 \prod_{m=1}^M \prod_{p=1}^M G_m G_p \frac{e^{-jk(|\mathbf{r}_m - \mathbf{r}| + |\mathbf{r}_p - \mathbf{r}|)}}{|\mathbf{r}_m - \mathbf{r}| |\mathbf{r}_p - \mathbf{r}|}. \quad (11)$$

The relative intensity of the reconstructed sound field is therefore

$$I_{rel}(kr, q, q_s) = |\mathbf{r}_s - \mathbf{r}|^2 \prod_{m=1}^M \prod_{p=1}^M G_m G_p \frac{e^{-jk(|\mathbf{r}_m - \mathbf{r}| + |\mathbf{r}_p - \mathbf{r}|)}}{|\mathbf{r}_m - \mathbf{r}| |\mathbf{r}_p - \mathbf{r}|}. \quad (12)$$

This equation is similar to the relative intensity equation given by Solvang [4] for plane wave emitting loudspeakers. The mean relative intensity for each frequency and an image rotated around the array can be used as a measure of the error in the frequency spectrum. It can be defined as

$$\bar{I}_{rel}(kr, q) = \frac{1}{2\rho} \int_0^{2\rho} I_{rel}(kr, q, q_s) dq_s. \quad (13)$$

This is the same definition as Solvang's mean relative intensity but is integrated with respect to the image angle rather than the spatial angle.

4 - ENCLOSED HOA WITH FINITE DISTANCE LOUSPEAKERS

The finite distance loudspeaker arrays can then be modelled inside a room. As a first approximation, the room investigated will be assumed to be shoebox-type rooms (i.e. cuboidal) with dimensions (L_x, L_y, L_z)m. By assuming the room is a shoebox the image source method can be used to approximate the pressure field in the room due to direct and boundary reflections. The pressure at a receiver point \mathbf{r}_R from a single source placed at a position \mathbf{r}_{s0} is [11]

$$p(k, \mathbf{r}_{s0}, \mathbf{r}_R) = \sum_{i=-\infty}^{+\infty} G_i(\mathcal{W}) \frac{e^{-jk|\mathbf{r}_i - \mathbf{r}_R|}}{|\mathbf{r}_i - \mathbf{r}_R|}, \quad (14)$$

where \mathbf{i} is the integer vector triplet (i_x, i_y, i_z) and \mathbf{r}_{si} is the position vector of the image source (i_x, i_y, i_z) given by

$$\mathbf{r}_{si} = (i_x L_x + (-1)^{i_x} x_{s0}, i_y L_y + (-1)^{i_y} y_{s0}, i_z L_z + (-1)^{i_z} z_{s0}) \quad (15)$$

The relative gain $G_i(\theta_s)$ of image source (i_x, i_y, i_z) is

$$G_i(\mathcal{W}) = R_{x^-}^{\lfloor i_x/2 \rfloor} R_{x^+}^{\lceil i_x/2 \rceil} R_{y^-}^{\lfloor i_y/2 \rfloor} R_{y^+}^{\lceil i_y/2 \rceil} R_{z^-}^{\lfloor i_z/2 \rfloor} R_{z^+}^{\lceil i_z/2 \rceil}, \quad (16)$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor and ceiling operators respectively. R_{χ^-} and R_{χ^+} are the frequency dependent, possibly complex, reflection coefficients in the negative and positive χ -directions respectively. In practice an infinite sum cannot be carried out and the reflections are therefore limited to a chosen order.

The anechoic HOA pressure field equation (6) can be expanded using the image source equation (14) to give the pressure field in the room due to the loudspeaker array and the boundary reflections. The pressure in the

room can therefore be written as

$$p(k, \mathbf{r}_{s0}, \mathbf{r}_R)_A = \hat{\mathbf{a}} \hat{\mathbf{a}} \hat{\mathbf{a}} \hat{\mathbf{a}} G_{mi}(q_s, \mathcal{W}) \frac{e^{-jk|\mathbf{r}_{m,si} - \mathbf{r}|}}{|\mathbf{r}_{m,si} - \mathbf{r}|}, \quad (17)$$

where $\mathbf{r}_{m,si}$ is the position vector and $G_{mi}(\theta_s, \omega)$ is the relative gain of the image source (i_x, i_y, i_z) of the m -th loudspeaker. The relative gain can be written as the product of the HOA loudspeaker gain and the relative gain of the reflection images

$$G_{mi} = G_m(q_s)G_i(\mathcal{W}). \quad (18)$$

Equation (17) therefore gives the field in a shoebox room and will collapse back to the anechoic HOA pressure field equation if \mathbf{i} is set to $(0, 0, 0)$ i.e. there are no images.

4.1 - Relative Intensity for Enclosed HOA

The relative intensity for the enclosed HOA system can be found in the same manner as equation (12), by dividing the enclosed HOA intensity field by that of a free field reference source placed at the radius of the loudspeaker array. The reference signal is taken as a free field source and therefore the relative intensity will include the intensity differences from the HOA as well as those from the room. The relative intensity of the enclosed HOA field to the free field source is

$$I(k, \mathbf{r}_R, q_s) = |\mathbf{r}_s - \mathbf{r}|^2 \hat{\mathbf{a}} \hat{\mathbf{a}} \hat{\mathbf{a}} \hat{\mathbf{a}} G_{mi}(q_s, \mathcal{W}) G_{pj}(q_s, \mathcal{W}) \frac{e^{-jk(|\mathbf{r}_{mi} - \mathbf{r}| + |\mathbf{r}_{pj} - \mathbf{r}|)}}{|\mathbf{r}_{mi} - \mathbf{r}| |\mathbf{r}_{pj} - \mathbf{r}|}. \quad (19)$$

4.2 - Simulations of HOA Enclosed in a Studio-size Room

A shoebox approximation of a room with dimensions 5.44m x 6.15m x 2.53m (dimensions corresponding to one of the studios at SARC) was used to investigate increased reverberation time on the spectrum of the signal reproduced by the HOA system enclosed. The centre of the array is placed at a position of $(0.1800, -0.3150, -0.0650)$ m, where the origin of the coordinate system is the centre of the room. The array has a radius $r_0 = 2.3$ m.

Figure 2 shows the relative intensity at a receiver point displaced by 8.5cm (radius of the human head) in the x -direction for increasingly reflective boundaries with average reflection coefficients of $R = 0, 0.3, 0.6$ and 0.9 . The simulation uses a 3rd order system with 8 loudspeakers. The anechoic simulation shows the spectral impairment solely from the HOA system and the other three demonstrate the addition of the influence of boundary reflections. For $t_{60} = 0.1551$ s the spectral

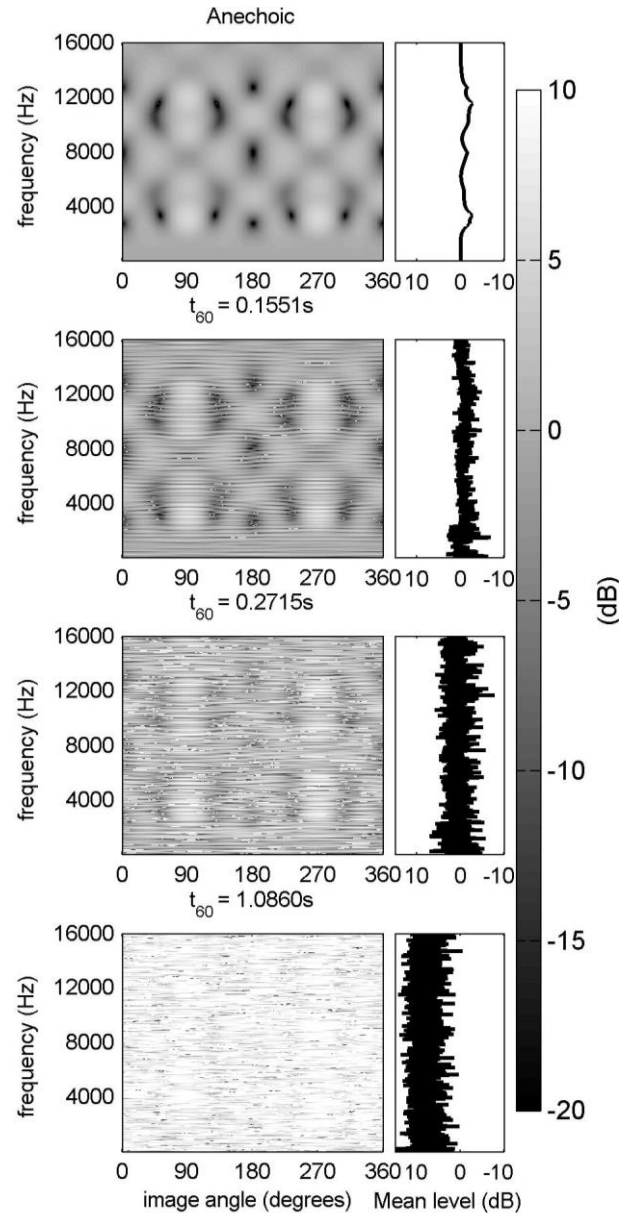


Figure 2 – Relative intensity of the recreated sound image for the studio-sized image source model, simulated with increasingly reflective boundaries. The observation point is displaced by 8.5cm from the array centre in the x -direction. Image source model uses 15th order reflections.

colouration from the HOA system is still visible, with the noisier influence of the room reflections superimposed on top. As the reverberation time increases the room reflections become dominant and the anechoic character is no longer evident from inspection of the relative intensity. As the reverberation time increases the depth of the troughs decreases but there are more fluctuations across the rest of the frequency range. This is shown in the mean relative intensity,

which becomes noisier and flatter as the reverberation time increases.

Figure 3 shows the relative intensity for a receiver point displaced 8.5cm from the origin in the x -direction, for both an anechoic room and a room of the dimensions listed above with a $t_{60} = 0.22s$ (this corresponds to the t_{60} of the studio used for measurements in section 5). The relative intensity is given for three different Ambisonic combinations:

- o1spk4 - 1st order using 4 loudspeakers from an 8 loudspeaker array,
- o1spk8 - 1st order using 8 loudspeaker and
- o3spk8 - 3rd order using 8 loudspeakers.

For all three systems, the deep troughs from the HOA system's influence in the relative intensity are still very much apparent, but not as deep as in the anechoic case. The relative intensity for system o1spk8 exhibits the largest divergence from anechoic case but the HOA spectral impairment is still visible. Solvang [4] found, in anechoic conditions, that the use of more loudspeakers than is needed leads to more spectral impairment above the well reproduced frequency limit and his conclusions are verified in a simulated room. This is likely to be because the system uses more loudspeakers than needed for first order and therefore there are more unwanted boundary reflections.

Inspecting the mean relative intensity of the enclosed system o1spk8 shows the deepest troughs are shallower than in the anechoic case, with increased fluctuation over the whole of the frequency range. In anechoic conditions, systems o1spk4 and o3spk8 have flatter mean relative intensities than o1spk8. When enclosed there is a noisier fluctuation each system. System o1spk4 has an average of 1.27dB and a standard deviation of 3.18dB in mean relative intensity, while system o3spk8 has an average of 0.44dB with a standard deviation of 2.05dB. This may be because the gains for the loudspeakers away from the image position are lower for higher orders, therefore their reflections will contribute less to the overall spectral impairment. This suggests that when using a higher reproduction order the noisy fluctuation added by the room boundary reflections may be lower.

4.3 - Simulations of HOA Enclosed in a Large Room

Again using the image source model, a rough approximation of the Sonic Lab at SARC [12] was modelled to investigate the spectral impairment in larger rooms. The room has dimensions of 17m x 13m x 14m and the loudspeaker array is placed at (0, 0, -3)m, where the origin is at the centre of the room. The Sonic Lab is a variable acoustic space and the reverberation time can

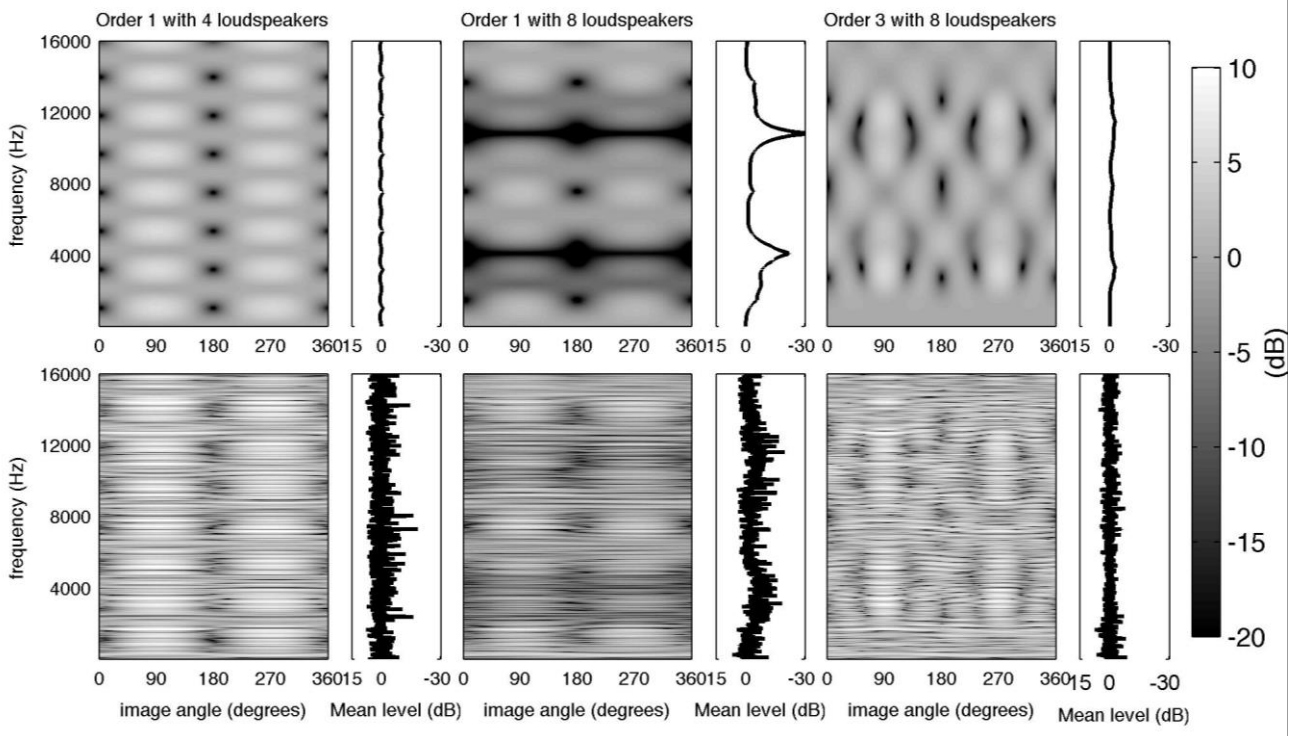


Figure 3 – The relative intensity for three Ambisonic arrangements using a studio-size image source model with 15th order reflections. The room's reflective properties are such that and (top) no boundary reflections and (bottom) $t_{60} = 0.22s$. Observation point displaced 8.5cm from the centre of the array in the x -direction.

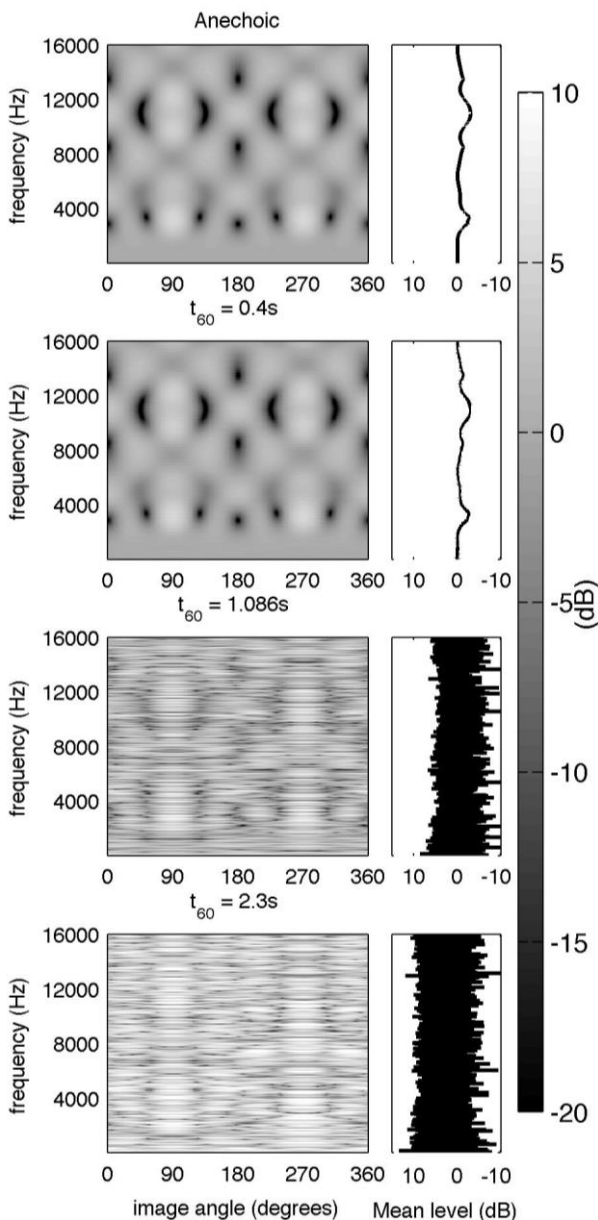


Figure 4 – Relative intensity of the recreated sound image for the large room image source model, simulated with increasingly reflective boundaries. The observation point is displaced by 8.5cm from the array centre in the x-direction. Image source model uses 15th order reflections.

be set from 0.4s to 2.3s. Figure 4 shows the spectral impairment for reverberation times of 0.4, 1.086s (the same reverberation time as the reflective studio-size model in the section 4.2) and 2.3s.

Figure 4 shows that for $t_{60} = 0.4$ s the anechoic character of the system is still very much present in the spectrum of the recreated sound with very little fluctuation. The

mean relative intensity for this reverberation time is also almost exactly the same as for the anechoic case. As with the studio-size room, increasing the reverberation in the room generally makes the troughs shallower and makes the rest of the spectrum more diffuse.

Comparing the large room and studio-size room both with $t_{60} = 1.086$ s shows that the HOA system's characteristic colouration is more visible for the large room and the average level of the mean relative intensity is also lower for the larger room. Even with a relatively long reverberation time of 2.3s the impairment is similar to that of a small room with $t_{60} = 1.086$ s but the average mean relative intensity is still lower. This suggests that the larger the room the longer the reverberation time that can be tolerated before the room becomes the dominant influence on the spectrum of the sound. A 2.3s reverberation time in the studio-size room would certainly lead to complete domination of the room reflections on the recreated image spectrum. The radius of the array is larger in the large room and this may have an influence on the overall impairment from the boundary reflections, since the placement of the loudspeakers will have an influence on reflection times.

5 – STUDIO MEASUREMENT COMPARISONS

For verification of the model results, measurements were taken in one of the 8-channel studios in the Sonic Arts Research Centre, Belfast. The room measurements are 5.44m x 6.15m x 2.53m with small indents in two of the corners. The room has t_{60} time of approximately 0.22s. The centre of the array is placed at a position of (0.1800, -0.3150, -0.0650)m, where the origin of the coordinate system is the centre of the room. The array has a radius $r_0 = 2.3$ m. The individual impulse response of each of the loudspeakers was taken at measurement point 8.5cm from the array centre in the x direction. The impulse responses were taken using a 5 second logarithmic sine sweep at a sampling rate of 88.2 kHz.

The measured impulse responses from the studio are weighted by the Ambisonic gains, given by equation (4), and summed to obtain an “Ambisonic impulse response” at the measurement point. This combined impulse response can be used to obtain the frequency response of the system and room at the measurement point by taking the Fourier transform. Using the gains from equation (4) the frequency response can be found as a reproduced image is panned around the array.

Figure 5 shows the frequency responses of the same Ambisonic systems, o1spk4, o1spk8 and o3spk8, used for the simulations in section 4.2. For system o1spk4 the spectral impairment is still very visible in the measurements. For system o3spk8 the pattern is still

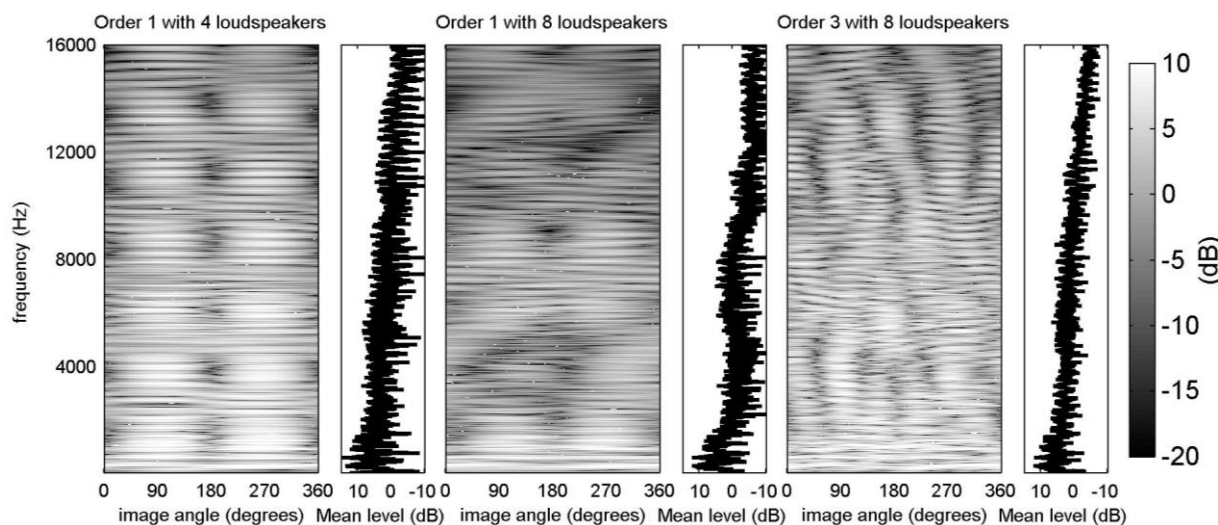


Figure 5 - Relative intensity for three Ambisonic arrangements from studio measurements at a position displaced 8.5cm from the centre of the array in the x-direction.

visible, though it is slightly skewed for image positions between 180 and 360 degrees for the measurement point displaced 8.5cm in the x direction. This may have been introduced by an error in the measurement position. The measurements for system o1spk8 do not have the same obvious pattern, but there are troughs that correspond to the anechoic spectral impairment.

The measurements for systems o1spk4 and o3spk8 correspond to the prediction from the simulation of a room with the same t_{60} , shown in Figure 3, that the impairment from the anechoic HOA system will still be prominent for a room of this size and reflectivity. The measurements for system o1spk8 do not exhibit the spectral impairment of the anechoic HOA system as obviously as the other two systems but was also the case of the model.

There is an overall drop in intensity as the frequency increases for all three systems, seen in the mean relative intensity. This may be due to a combination of the loudspeaker frequency response, its directivity pattern at higher frequencies not being omnidirectional and the frequency dependence of the studio boundaries.

6 – CONCLUSIONS

In order to simulate a HOA system inside a room the loudspeaker needs to be simulated and the assumption that loudspeakers can be modelled as emitting plane waves no longer holds. The loudspeakers were modelled as point sources. The size of the well reproduced area and, to a smaller extent, the well reproduced frequency range of the recreated sound field were shown to depend on the radius of the loudspeaker array used.

Using an image source model of a shoebox room the

spectral impairment in HOA systems was extended to include room boundary reflections. It was shown that for a studio-type room the spectral impairment from the anechoic HOA system is still evident for an off centre listening position at the radius of a human head. Increasing the boundary reflections has the effect of removing the deepest troughs from the HOA system relative intensity, at the cost of additional spectral fluctuation across the rest of the frequency range. In a comparison between large rooms and studio-size rooms, the large room, with a larger loudspeaker array radius, was shown to be able to tolerate much longer reverberation times than smaller rooms before the spectral influence on the sound becomes prominent.

Further work could include extending the model to use directional sources and frequency dependent boundaries to better align the simulation results with physical measurements. It could also include measurements in larger reverberant environments to investigate the spectral characteristics. The spectral impairment could also be expanded to investigate enclosed 3D HOA systems.

REFERENCES

- [1] M. Gerzon, “Periphery: With-height surround reproduction,” *Journal of the Audio Engineering Society*, vol. 21, no. 1, pp. 2-10, 1973.
- [2] J. S. Bamford and J. Vanderkooy, “Ambisonic Sound for Us,” in *99th Audio Engineering Society Convention*, 1995.
- [3] A. Solvang, “Spectral impairment for two-dimensional higher order Ambisonics,” *Journal*

- of the *Audio Engineering Society*, vol. 56, no. 4, pp. 267-279, 2008.
- [4] D. Malham, *Higher order Ambisonic systems*, Doctoral thesis, University of York, 2003.
 - [5] J. Daniel, Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia, PhD thesis, Université Pierre et Marie Curie (Paris VI), France, 2000.
 - [6] D. Malham, "Experience with larger area 3-D Ambisonic sound systems," in *Proceedings of Institute of Acoustics, 14*, 1992.
 - [7] M. Gerzon, "Design of Ambisonic decoders for multispeakers surround sound," in *58th Audio Engineering Society Convention*, New York, 1977.
 - [8] J. Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format," in *Audio Engineering Society 23rd International Conference*, Copenhagen, 2003.
 - [9] H. Wallach, E. Newman and M. Rosenzweig, "The precedence effect in sound localization," *The American Journal of Psychology*, vol. 62, no. 3, pp. 315-336, 1949.
 - [10] J. Daniel, J. B. Rault and J. D. Polack, "Ambisonics encoding of other audio formats for multiple listening conditions," in *105th Audio Engineering Society Convention*, San Francisco, California, 1998.
 - [11] J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," *Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943-950, 1979.
 - [12] "SARC Sonic Lab specifications," [Online]. Available: <http://www.sarc.qub.ac.uk/sites/sarc/AboutUs/TheSARCBuildingandFacilities/TheSonicLab/SonicLabSpecs/>. [Accessed 21 February 2012].