Quality evaluation of natural Ambisonics recordings

Johann-Markus Batke
Thomson Corporate Research, Karl-Wiechert-Allee 74, 30625 Hannover, Germany, Email: Jan-Mark.Batke@thomson.net

Introduction

Higher Order Ambisonics (HOA) is a sound field description techniques that is capable to represent a sound field with all three spatial dimensions. The coefficients for such a HOA representation are usually acquired by a spherical microphone array when natural sounds should be captured.

The mechanical size of such an array and the used number of microphone capsules limit the accuracy of the acquired HOA coefficients. Furthermore, the spatial extension of the microphone diaphragm can be taken into account. In this contribution, these influences are discussed with regard to the spatial quality of the sound field representation.

The theoretically calculated representation of a plane wave is compared to the output of a simulated microphone array under ideal conditions. As the major result it should be pointed out that the quality of HOA coefficients acquired by a spherical microphone array shows a strong dependency from the angle of incidence. Variations of the array size, the number of capsules and the microphone capsules' size exhibit possibilities for improvements. The simulations made are confirmed by measurements using a commercially available sound field microphone.

Higher Order Ambisonics

The Ambisonics representation is a sound field description method employing a mathematical approximation of the sound field in one location. The pressure at point $\boldsymbol{r}$ in space is described by

$$p(\boldsymbol{r},k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} A_n^m(k)\, j_n(kr)\, Y_n^m(\theta,\phi). \tag{1}$$

This series is sometimes also regarded as "Fourier-Bessel-Series" [1] or "Multipole Expansion" [4]. Normally n runs to a finite order N. The coefficients $A_n^m(k)$ of the series describe the sound field (assuming sources outside the region of validity [12]), $j_n(kr)$ is the spherical Bessel function of first kind and $Y_n^m(\theta,\phi)$ denote the spherical harmonics. The latter only depend on the angles and describe a function on the unity sphere [12]. Coefficients $A_n^m(k)$ are regarded as Ambisonics signals in this context.

B-Format

Ambisonics signals of first order are also known as B-Format signals. More precisely, the mapping between the commonly used notation for the B-Format and the Ambisonics coefficients is

$$W = A_0^0/\sqrt{2}, \quad X = A_1^1, \quad Y = A_1^0, \quad Z = A_1^{-1}, \tag{2}$$

where W denotes the pressure signal and X, Y, Z are describing the directional signals.

Ambisonics signals with order higher than one are summarised under the denomination Higher Order Ambisonics (HOA), often using a notation such as in equation (1).

To describe the characteristics of a microphone array, the free field response is evaluated here. The free field response of a microphone is the output signal when the microphone is exposed to a plane wave. Therefore, the theoretically expected values of Ambisonics coefficients describing a plane wave are derived now.

Plane waves

A plane wave impinging from direction $\boldsymbol{k}_i$ is written as

$$p_i(\boldsymbol{r},k_i) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}_i^T \boldsymbol{r}} \tag{3}$$

in the frequency domain [12]. The term $\mathrm{e}^{\mathrm{i}\boldsymbol{k}_i^T \boldsymbol{r}}$ can be also expressed as [12]

$$\mathrm{e}^{\mathrm{i}\boldsymbol{k}_i^T \boldsymbol{r}} = 4\pi \sum_{n=0}^{\infty} \mathrm{i}^n j_n(kr) \sum_{m=-n}^{n} Y_n^m(\theta,\phi) Y_n^m(\theta_i,\phi_i)^*. \tag{4}$$

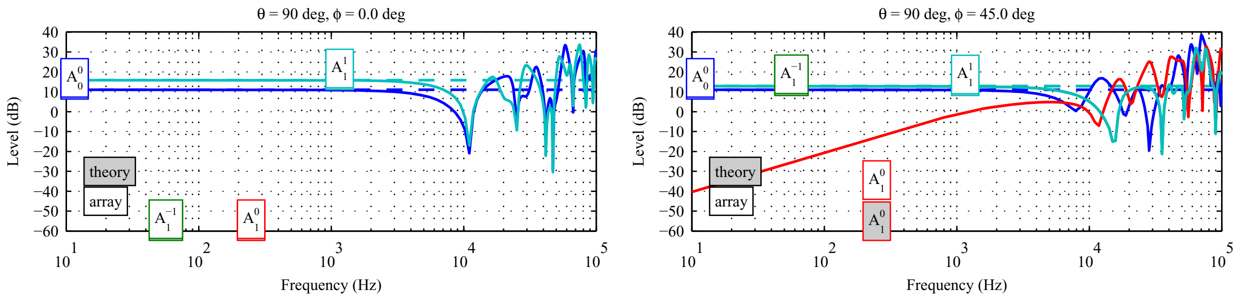
Comparing this to equation (1), the Ambisonics coefficients describing such a plane wave are [9]

$$A^m_{n,\mathrm{plane}}(\theta_i,\phi_i) = 4\pi\, \mathrm{i}^n\, Y_n^m(\theta_i,\phi_i)^*, \tag{5}$$

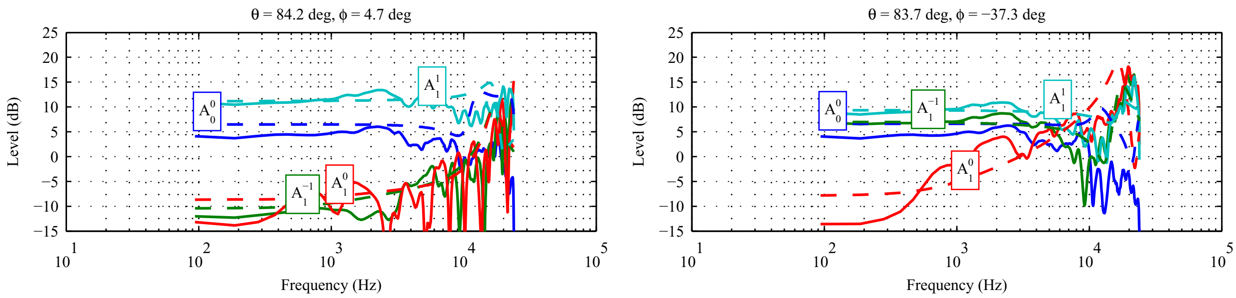
and it is easily seen that they are constant over frequency. A perfect microphone array would output signals matching such coefficients.

Microphone arrays

In general, microphone arrays are required when it is necessary to synthesise higher directivities that are available from first-order microphones [9]. Different microphone array geometries are in use, in case of HOA signal acquisition spherical arrays are often used. These spherical arrays can be subdivided in free-field sphere and solid sphere designs. A signal processing unit is necessary to calculate the Ambisonics coefficients from the capsules' signals.
The "soundfield microphone" DSF-1 is a commerical available microphone array for music recording with processing unit for processing of the B-Format [11]. It is a free-field sphere array using four microphone capsules in a tetrahedral setup with a array radius of 1.47 cm [2]. For the (music) acquisition at higher orders only different research prototypes exist (e. g. [5, 7, 8]) but no commercially available version. In this contribution the DSF-1 is used to verify some soundfield simulations, which in turn are used to evaluate microphone arrays of higher orders. The polar pattern of the capsules used



(a) Theoretical expected coefficients (– –) and simulations (—) for $\phi = 0°$ and $\phi = 45°$.



(b) Simulated (– –) and measured (–) coefficient signals. Angular deviations are due the positioning of the DSF-1 on the turntable.

Figure 1: Comparison of theoretically expected coefficients, a simulation of a first order microphone array and measurements of the DSF-1.

in the DSF-1 is sub-cardioid, the diameter of the diaphragm is about 12 mm. Smaller microphone capsules cause less distortion of the sound field, but also show typically a decreased signal to noise ratio. The size of the capsule causes a decay of level when its size gets in the same dimension as half of the wavelength and the diaphragm is hit sideways from the impinging wave.

Quality Evaluation

In the follwing, a simulated plane wave recording is compared with a measured sound field. The Ambisonics coefficients of the simulated sound field are also referred here as output of a simulated microphone. All measurements described here were carried out using a loudspeaker system (JBL LSR 6325), a measurement microphone (Neumann KM 100) and the reference device mentioned above, the DSF-1 microphone with a digital control unit. The distance between microphone and loudspeaker is about 1 m. The microphone is rotated in steps of 45° using a turntable.

First order signals

As described above, first order signals are also denoted with W containing the pressure signal, and X, Y, Z for the directional signals, which are proportional to the velocity of air in direction of the respective Cartesian coordinates. Accordingly, a wave front impinging for instance in direction of the X-axis should result in a pressure component W and an X-component, whereas the Y- and Z-components should vanish. This szenario is examined next in a simulation of the DSF-1 exposed to a plane wave.

In figure 1(a) the first order Ambisonics coefficients over the frequency are shown for a plane wave impinging from $\Omega = (\theta, \phi)$ with $(90°, 0°)$ (left) and $(90°, 45°)$ (right). Because of the array's symmetry these two angles characterise also all further rotation positions. Labels in figure 2(a) located on the margins are denoting signals that vanish. In case of $(90°, 0°)$ the simulation results fit to the theoretically expected results up to 10 kHz. Above this frequency comb filtering like effects are visible due to the finite size of the capsules' diaphragm. Interestingly, in case of $(90°, 45°)$ an unwanted A_1^0 or Z-component becomes visible. Several reasons for errors can be taken into account. Spatial aliasing occurs because of infinite order of the sound field, whereas the microphone is order one [10]. Moreover, for higher frequencies the dimension of the array causes aliasing errors when its size is more half of the wavelength. For the first order array under test, obviously also the geometric arrangement of the capsules comes into play. The analysis of these errors is proposed for further study.

Now some measurements results are shown to confirm the simulation findings. Figure 1(b) depicts the output of the simulated microphone together with the B-Format output of the DSF-1. Since the directional calibration of the microphone standing on the turntable was done by sight, an angular deviation of about ±5° from the desired values was produced. The actual angles were calculated using the sound intensity [6]. The coefficients of the simulated microphone take these actual angles into account and match the measured values in a wide range. As the B-Format signal contains a diffuse field correction, deviations from the simulated microphone output in the higher frequency range become visible.

In principle, the tetrahedral array could be improved with regard to the mechanical parameters. A smaller diameter of the microphone's diaphragm improves the high frequency behaviour. On the other hand, the maximum frequency producing reasonable results for Ambisonics coefficients from a spherical microphone array is affected by the radius of the array. Further simulations show that an array radius of 1 cm is a reasonable choice for an upper frequency limit of 20 kHz. This in turn would require smaller microphone capsules because of the necessary closer placement.

Higher order signals

In the following, results of a simulation of a fifth order microphone array are presented. To provide a realistic scenario for that case, an array radius of 5 cm is assumed, the capsules' diameter is chosen to 6 mm. The positions on the sphere are taken from [3].

In figure 2 the frequency responses of coefficients from simulated microphone arrays (solid line, white labels) are compared with theoretically expected coefficients from equation (5) (dashed line, Gray labels), using $\theta = 90°, \phi = 0°$ for all figures. To contrast the differences between the various outputs, a first order array with $L = 4$ spatial sampling points (upper plots) is shown opposed to an array with order $N = 5$ using $L = 36$ sampling points. The components of each order ($n = 0, 1$ for the first order array and $n = 0 \ldots 5$ for the fifth order array) are separated to give a better overview. The ordinate values range from -60 dB up to 60 dB in all plots for the sake of comparability, labels located on the margins are denoting signal components whether lower or higher than these limits.
All coefficients from the both microphone arrays have an upper frequency limit where the result compares reasonable to the theoretical values. In figure 2(a) for $n = 0$, the A_0^0 component of the fifth order array in figure 2(b) fits to theoretical value up to 10 kHz, whereas the first order array works well up to 2 kHz. Moreover, the first order array coefficients show a significant level of the A_1^0 component, which actually should vanish. The fifth order array sufficiently suppresses this component up to 2 kHz. Summarising the first order case, it is obvious that the coefficients of the fifth order array show up better quality than the coefficients of the first order array. This is due to the fact that the fifth order array is considerably overdetermined for firstet order signals.
For $n = 2 \ldots 5$ in figure 2(b) the coefficients of the fifth order array are compared to the theoretically expected values only. In case of $n = 2$ the unwanted components A_2^{-2}, A_2^{-1}, and A_2^1, are sufficiently suppressed up to 1 kHz. In case of $n = 3$ only a maximum level distance of about 40 dB is achieved for theoretically vanishing components. For $n = 4$ and $n = 5$ the coefficient signals of the array even show high pass character, and the quality of the signals is questionable for most purposes.

Conclusion

This contribution evaluated the quality of Ambisonics coefficients. First, simulations and measurements of Ambisonics coefficients using a tetrahedral microphone array of first order were compared with the theoretically expected result for a plane wave. It turned out that the quality of frequency responses is directivity depend. Spatial aliasing is considered to be the main reason for errors. Furthermore, simulations of a first order and a fifth order array were compared. The quality of the first order Ambisonics components of the fifth order array are significantly better than the output of the first order array. However, for the higher order components the deviation from the theoretically expected results increases.

References

[1] Jérôme Daniel, Rozenn Nicol, and Sébastien Moreau. Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging. In *Audio Engineering Society Preprints*, March 2003. Paper 4795 presented at the 114th Convention.

[2] Ken Farrar. Soundfield microphone — design and developement of microphone and control unit. *Wireless World*, pages 48–50, 1979.

[3] Jörg Fliege and Ulrike Maier. A two-stage approach for computing cubarure formulae for the sphere. Technical report, Fachbereich Mathematik, Universität Dortmund, 1999. Node numbers are found at `http://www.mathematik.uni-dortmund.de/lsx/research/projects/fliege/nodes/nodes.html`.

[4] Nail A. Gumerov and Ramani Duraiswami. *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Elsevier, first edition edition, 2004.

[5] Arnaud Laborie, Rémy Bruno, and Sébastien Montoya. A new comprehensive approach of surround sound recording. In *Audio Engineering Society Preprints*, volume 51, March 2003. Paper 5717 presented at the 114th Convention.

[6] Juha Merimaa and Ville Pulkki. Spatial impulse response rendering I: Analysis and synthesis. *J. Acoust. Soc. Am.*, 53(12):1115–1127, December 2005.

[7] Jens Meyer and Tony Agnello. Spherical microphone array for spatial sound recording. In *Audio Engineering Society Preprints*. AES, September 2003. Paper 5975 presented at the 155th Convention.

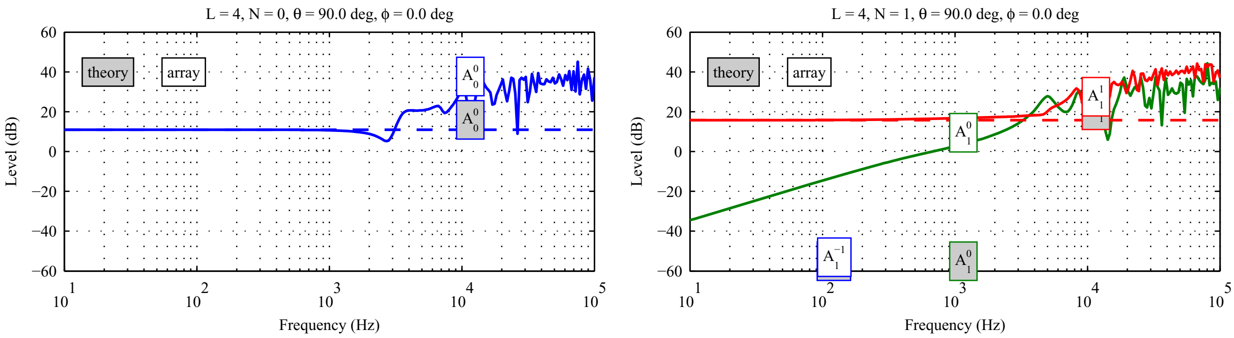
[8] Sébastien Moreau, Jérôme Daniel, and Stéphanie Bertet. 3d sound field recording with higher order ambisonics – objective measurements and validation of spherical microphone. In *Audio Engineering Society Preprints*, May 2006.

[9] M. A. Poletti. Three-dimensional surround sound systems based on spherical harmonics. *J. Audio Eng. Soc.*, 53(11):1004–1025, November 2005.

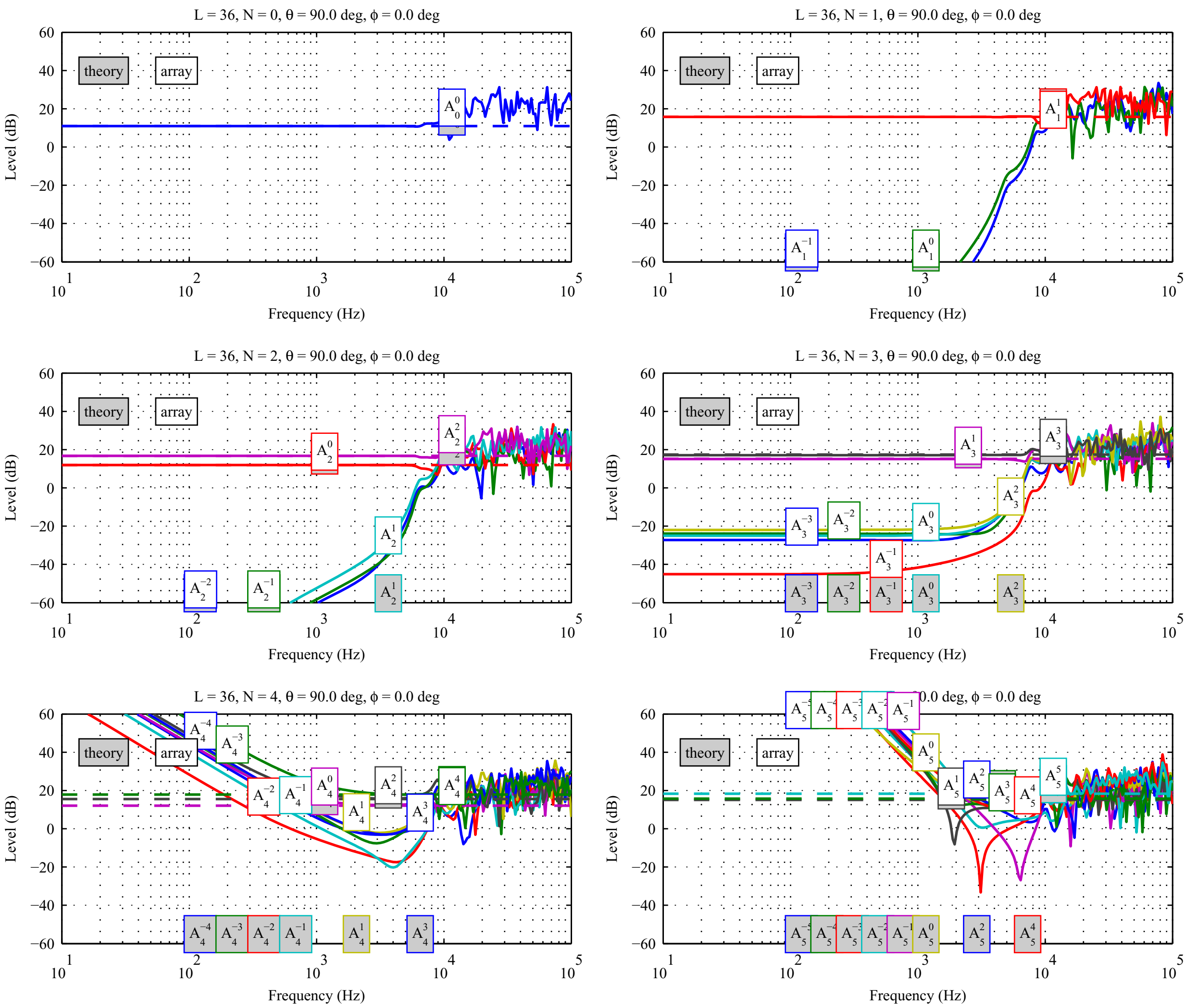
[10] Boaz Rafaely. Spatial aliasing in spherical microphone arrays. *IEEE Transactions on Signal Processing*, 55(3): 1003–1010, March 2007.

[11] *DSF-1 Performance Microphone System – User Guide*. Soundfield Ltd., West Yorkshire, England, 1.0 edition.

[12] Earl G. Williams. *Fourier Acoustics*. Academic Press, 1999.



(a) Simulation of a first order array, using $L = 4$ sampling points



(b) Simulation of a fifth order array, using $L = 36$ sampling points

Figure 2: Comparison of Ambisonics coefficients from a first order and a fifth order spherical microphone array.