Audio Engineering Society

Convention e-Brief 589

Presented at the 148th Convention
2020 June 2–5, Online

*This Engineering Brief was selected on the basis of a submitted synopsis. The author is solely responsible for its presentation, and the AES takes no responsibility for the contents. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Audio Engineering Society.*

# Recording first-order Ambisonics with a differential array of two dual-diaphragm microphones

Thomas Deppisch[1] and Christoph Frank[2]

[1] *University of Technology and University of Music and Performing Arts, Graz, Austria*
[2] *Austrian Audio GmbH, Vienna, Austria*

Correspondence should be addressed to Thomas Deppisch (thomas.deppisch@student.tugraz.at)

**ABSTRACT**

First-order Ambisonics consists of four signals corresponding to one omnidirectional and three figure-of-eight polar patterns aligned with the Cartesian axes. We obtain these signals from a matched pair of dual-diaphragm microphones, where each diaphragm output is accessible individually. While the signals for figure-of-eight patterns in x- and y-direction are calculated directly from the microphone outputs, the z-direction figure-of-eight signal is calculated as differential signal from omnidirectional patterns of both microphones, resulting in low-frequency attenuation and a phase shift. We equalize the frequency response using a combination of a model-based IIR and a measurement-based FIR filter and provide an open-source plugin performing the processing.

## 1 Introduction

Audio for 360° video is in most cases either recorded with the camera's internal microphones, which are mainly small MEMS capsules with rather high noise floor and limited frequency response, or with dedicated coincident first-order Ambisonics microphone arrays. Even in the latter case, frequency response and noise floor are most often inferior compared to studio microphones in the same price range, which make them a tough buying decision for sound engineers. We introduce a method, where two matched dual-diaphragm large condenser microphones can be used as a B-format microphone if they give individual access to both, the front and back diaphragm's signals. We evaluate our method exemplarily for a pair of Austrian Audio OC818 microphones, which are particularly suited for this method as they are calibrated within 0.5 dB sensitivity and can be brought closely together due to their flat-top form factor.

## 2 Microphone array processing

To obtain the omnidirectional and the three figure-of-eight polar patterns needed for first-order Ambisonics, we arrange two dual-cardioid-diaphragm microphones on top of each other, such that they are aligned with respect to their x- and y-coordinates but displaced in z as shown in figure 1. Hence, an omnidirectional pattern and two figure-of-eight patterns aligned with x- and y-axis can be created by summing and subtracting the cardioid signals from the individual microphones. For the figure-of-eight pattern aligned with the z-axis, we create a differential pattern by subtracting omnidirectional patterns from both microphones (cf. section 3). After equalization of the differential figure-of-eight pattern and diffuse-field equalization of all four patterns (section 4), we apply Ambisonic SN3D normalization

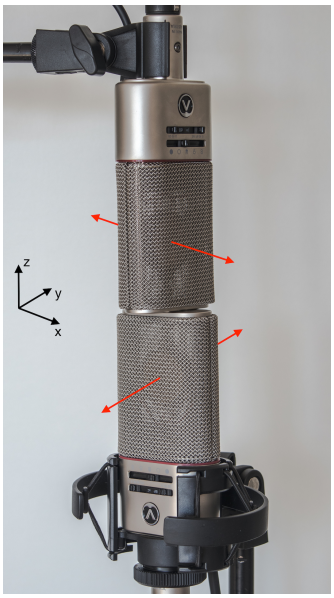



**Fig. 1:** Proposed array configuration. Red arrows show x- and y-axis alignment of the diaphragms.

factors $N_n^{|m|} = \sqrt{\frac{n-|m|!}{n+|m|!}\frac{2-\delta_m}{4\pi}}$, where $\delta_m = 1$ for $m = 0$ and $\delta_m = 0$ in all other cases [1]. Here, $n$ and $m$ are Ambisonics order and degree, which are $n = 0$, $m = 0$ for the omnidirectional and $n = 1$, $m = \{-1,0,1\}$ for figure-of-eight patterns aligned with y-, z- and x-axis, respectively.

## 3 IIR filter design

Calculation of the differential figure-of-eight array signal is done in two steps: First, both microphone outputs of each of the two dual-cardioid microphones are summed together, creating omnidirectional polar patterns for the individual microphones. Then, the signals corresponding to the omnidirectional patterns are subtracted from each other, yielding a signal corresponding to a figure-of-eight characteristic for wavelengths larger than twice the distance of the diaphragms $\lambda > 2d$. For low frequencies, the subtracted signals are nearly in phase, resulting in a low-cut frequency response including a 90° phase shift. This section derives model-based IIR filters for equalization of the low-frequency amplitude slope and phase shift.

Sources in far-field are modeled via the plane-wave pressure $p(\boldsymbol{x},\omega,t) = e^{-i\boldsymbol{k}^\mathrm{T}\boldsymbol{x}}e^{i\omega t}$ at observation point $\boldsymbol{x} = [x,y,z]^\mathrm{T}$, where the wave vector $\boldsymbol{k} = k\boldsymbol{u}$ describes the propagation from the origin in direction $\boldsymbol{u} = [u_x,u_y,u_z]^\mathrm{T}$ with wave number $k = \omega/c = 2\pi f/c$ and speed of sound $c$. We assume a single-frequency harmonic wave with angular frequency $\omega$ and unit amplitude. In the following, we focus on the frequency response and the polar pattern of the array and leave out the temporal evolution of the plane wave by setting the time to $t = 0$. To simulate the differential two-microphone array with spacing $d$, we subtract the pressure at $\boldsymbol{x}_1 = [0,0,-d]^\mathrm{T}$ created by a plane wave with direction vector $\boldsymbol{u} = [0,0,-1]^\mathrm{T}$, from the pressure at the origin,

$$H_\mathrm{p}(\omega) = 1 - e^{-i\boldsymbol{k}^\mathrm{T}\boldsymbol{x}_1} = 1 - e^{-ikd}\,. \tag{1}$$

The transfer function of the broadband continuous-time equalization filter is given by the corresponding inverse $\hat{H}(\omega) = 1/H_\mathrm{p}(\omega)$. As the inherent high-frequency comb-filters fluctuate strongly with source position and array imperfections, we focus on equalization of the low-frequency slope. The low-frequency slope is modeled by a low-pass $H_\mathrm{lp}(\omega) = 1/(1+i\omega/\omega_{\mathrm{c}_1})$, ensuring finite gain for low frequencies. We force the filter to apply constant −6 dB gain for high frequencies by adding a rising slope $H_\mathrm{hs} = (1+i\omega/\omega_{\mathrm{c}_2})$ and normalizing, yielding the continuous-time equalization $H(\omega)$ filter, effective only for low frequencies,

$$H(\omega) = \frac{\omega_{\mathrm{c}_2}}{2\omega_{\mathrm{c}_1}}\frac{1+i\omega/\omega_{\mathrm{c}_2}}{1+i\omega/\omega_{\mathrm{c}_1}}\,. \tag{2}$$

The cutoff frequencies $\omega_{\mathrm{c}_1} = 2\pi f_{\mathrm{c}_1}$, $f_{\mathrm{c}_1} = 10$ Hz and $\omega_{\mathrm{c}_2} = 2\pi f_{\mathrm{c}_2}$, $f_{\mathrm{c}_2} = c/(3.2d)$ are chosen for $H(\omega)$ to match magnitude of $\hat{H}(\omega)$ for low frequencies and phase for frequencies between 100 Hz and 1 kHz. Note that the main purpose of this IIR filter is equalization of the phase response for low frequencies. The diffuse-field magnitude responses of the array signals are equalized at a later step (cf. section 4).

The corrected impulse invariance [2][3] is used to derive a discrete-time filter from its continuous-time counterpart. By splitting the transfer function into a strictly proper low-pass and a constant throughput term, even (non-strictly) proper transfer functions can be transformed to discrete time [4] without the need of non-linear frequency warping as for the bilinear transform. The resulting discrete-time $z$-domain filter is derived as

$$\begin{aligned} H_\mathrm{d}(z) = &\frac{1}{2}\frac{1}{1-e^{-\omega_{\mathrm{c}_1}T}z^{-1}}\left(T/2\,(\omega_{\mathrm{c}_2}-\omega_{\mathrm{c}_1})+1\right.\\ &\left.-e^{-\omega_{\mathrm{c}_1}T}(1-T/2\,(\omega_{\mathrm{c}_2}-\omega_{\mathrm{c}_1}))z^{-1}\right), \end{aligned} \tag{3}$$

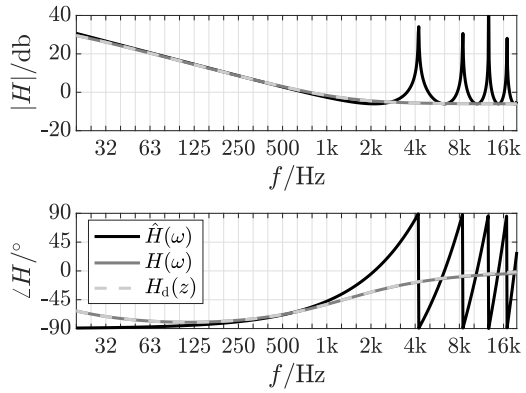



**Fig. 2:** Magnitude (top) and phase response (bottom) of broadband continuous-time equalization filter $\hat{H}(\omega)$, continuous-time low-frequency approximation $H(\omega)$ and its discrete-time counterpart $H_\mathrm{d}(z)$.

where the sampling period $T$ is the inverse of the sampling frequency $f_\mathrm{s}$. Figure 2 shows magnitude and phase response of the derived filter. The low-frequency approximation filter $H(\omega)$ follows the low-frequency slope in magnitude response very well. Due to the approximation of the slope via a low-pass characteristic, instead of an infinitely rising slope, the phase response is only modeled accurately for a frequency range between 100 Hz and 1 kHz. The discrete-time filter $H_\mathrm{d}(z)$ matches the continuous-time filter $H(\omega)$ very well in magnitude and phase over the whole frequency range.

## 4 Diffuse-field equalization FIR filter design

The IIR filter derived in section 3 equalizes magnitude and phase response of differential figure-of-eight signals based on a physical model. Real microphones exhibit sample variances, which are smaller if a matched microphone pair is used, but still non-negligible in context of differential array processing as small differences in the frequency response prevent perfect cancellation for low frequencies. Hence, a linear-phase FIR filter is designed to equalize the diffuse-field magnitude response of the signal corresponding to the differential figure-of-eight pattern aligned with the z-axis $H_{8,\mathrm{z}}$. The diffuse-field response is calculated from $4\times36$ impulse responses measured in 10° steps in a vertical circle for both diaphragms of both microphones, assuming rotationally symmetric polar patterns around the z-axis. From these measurements, the response of the differential figure-of-eight pattern is calculated for every measurement position. The diffuse-field magnitude response $|H_{\mathrm{df},8}(\omega)|$ is then calculated as a weighted sum of the responses in all directions $|H_{8,\mathrm{z},i}(\omega)|$, where the weight $w_i$ for each direction $i \in [1,36]$ is set according to the spherical segment $A_i$ on the unit sphere it corresponds to,

$$|H_\mathrm{df}(\omega)| = \sum_{i=1}^{36} \tilde{w}_i \, |H_i(\omega)| \,. \tag{4}$$

With the unit sphere $\mathbb{S}^2$, the azimuth angle $\varphi \in [0,2\pi]$ and the zenith angle $\vartheta \in [0,\pi]$, the weight $w_i$ corresponding to the spherical segment $A_i$ bounded by angles $\vartheta_i$ and $\vartheta_{i+1}$ is calculated as

$$\begin{aligned} w_i &= \frac{A_i}{\oint_{\mathbb{S}^2} dS} = \frac{\int_0^{2\pi} \int_{\vartheta_i}^{\vartheta_{i+1}} \sin\vartheta \, d\vartheta \, d\varphi}{4\pi} \\ &= 1/2\,(-\cos\vartheta_{i+1} + \cos\vartheta_i) \,. \end{aligned} \tag{5}$$

If there are two different measurement angles in azimuth, as for our full 360° measurement, we need to half the weights for measurement positions with the same zenith angle,

$$\tilde{w}_i = \begin{cases} w_i & \text{if } i \in \{1,18\}\,, \\ 1/2\, w_i & \text{else}\,. \end{cases} \tag{6}$$

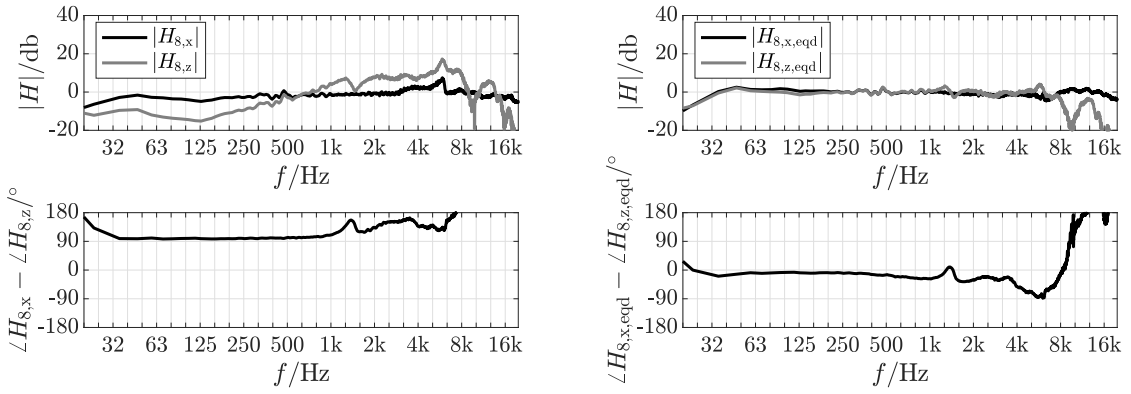
Altogether, the magnitude response of the diffuse-field equalization filter for the differential figure-of-eight pattern is calculated from the magnitude responses $|H_{8,\mathrm{z},i}(\omega)|$ in directions $i$ after applying the IIR equalization filter $H_\mathrm{d}(\omega)$,

$$|H_{8,\mathrm{z,df}}(\omega)| = 1/\left( \sqrt{3} \sum_{i=1}^{36} \tilde{w}_i \, |H_{8,\mathrm{z},i}(\omega)\, H_\mathrm{d}(\omega)| \right), \tag{7}$$

such that the diffuse-field response of the IIR-filtered signal is equalized. The factor $\sqrt{3}$ ensures that the squared diffuse-field response is not equalized to 0 dB but to $10\log_{10}(1/3) = -4.8$ dB as a figure-of-eight pattern only picks up $1/3$ of the energy an omnidirectional pattern picks up,

$$\frac{\int_0^{2\pi} \int_0^{\pi} (\cos\vartheta)^2 \sin\vartheta \, d\vartheta \, d\varphi}{\oint_{\mathbb{S}^2} dS} = \frac{1}{3} \,. \tag{8}$$

After octave-band smoothing and regularization for very low and very high frequencies, a 2048-tap linear-phase filter impulse response is calculated via inverse discrete Fourier transform, exhibiting a delay of 1024 samples.




**(a)** Before equalization.

**(b)** After equalization.

**Fig. 3:** Top: Magnitude response of figure-of-eight pattern aligned with the x-axis $H_{8,\mathrm{x}}$ and of the differential figure-of-eight aligned with the z-axis $H_{8,\mathrm{z}}$. Bottom: Phase difference $\angle H_{8,\mathrm{x}} - \angle H_{8,\mathrm{z}}$.

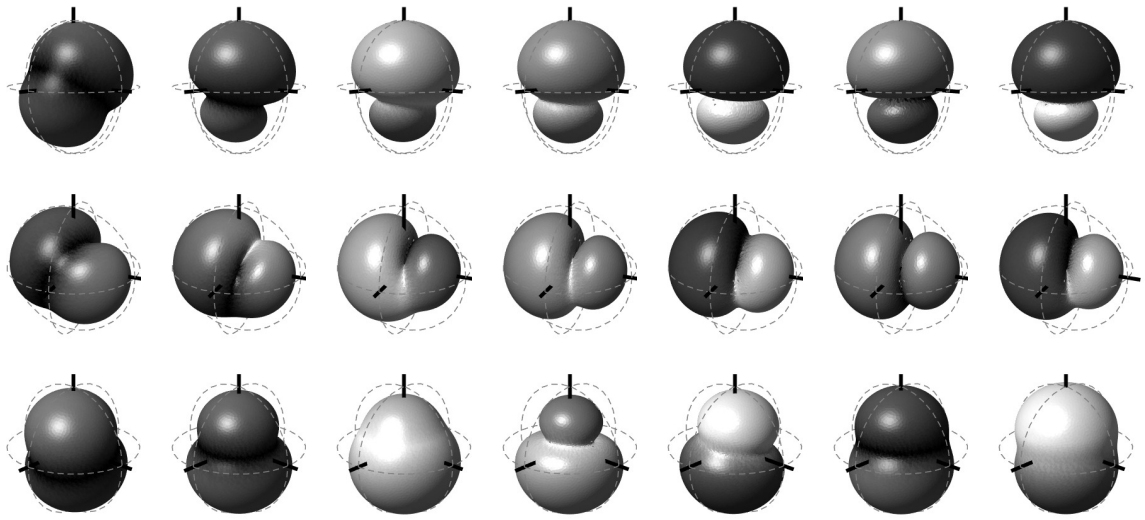
Figure 3a and b show on-axis magnitude responses of the figure-of-eight pattern in x-direction $|H_{8,\mathrm{x}}|$, and the differential figure-of-eight pattern in z-direction $|H_{8,\mathrm{z}}|$, in the upper plots, before and after equalization. In the lower plots, they show the difference of the phase responses $\angle H_{8,\mathrm{x}} - \angle H_{8,\mathrm{z}}$. After equalization, the magnitude responses match each other well until 8 kHz, where the effect of regularization is observable. The 90° phase difference for frequencies between 32 Hz and 1 kHz in figure 3a is successfully equalized by the IIR filter as shown in figure 3b.

## 5 Analysis of resulting directivity patterns

To analyze the effectiveness of the equalization filters, we evaluate the directivity the microphone array is able to reproduce in an Ambisonic context. For that, the equalization filters are applied to the measured array impulse responses and an Ambisonic impulse response $\boldsymbol{\chi}(t) = [\chi_w(t), \chi_x(t), \chi_y(t), \chi_z(t)]^{\mathrm{T}}$ is created by weighting the omnidirectional and the three figure-of-eight signals with the corresponding Ambisonic normalization factors as described in section 2. To plot the directivity pattern the array reproduces for a single source, the Ambisonic impulse response is decoded to the dense grid of a 5200-point 100-design [5] by evaluating the spherical harmonics in vector $\boldsymbol{y}$ at t-design positions $\boldsymbol{\theta} = [\theta_x, \theta_y, \theta_z]^{\mathrm{T}}$ and multiplying them with $\boldsymbol{\chi}(t)$,

$$g(\boldsymbol{\theta}) = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta})\boldsymbol{\chi}(t)\,. \tag{9}$$

Figure 5 shows an ideal super-cardioid first-order Ambisonics directivity pattern for a source at the horizon ($\varphi = 0°$, $\vartheta = 90°$) as a reference. The pattern is plotted logarithmically with a dynamic of 40 dB between the center and the surface of the surrounding sphere. Figure 4 shows directivity patterns reproduced by the microphone array for six different frequencies between $32 - 2000$ Hz and the source being either at zenith ($\varphi = 0°, \vartheta = 0°$), at the horizon ($\varphi = 0°, \vartheta = 90°$) or at nadir ($\varphi = 0°, \vartheta = 180°$). As the microphone array is coincident in x- and y-directions, sources at the horizon (middle row) are reproduced almost ideally over the whole frequency range. Only for very low frequencies (32 Hz) the plots show a slight miss-orientation and degradation of the directivity pattern, which is due to the figure-of-eight patterns of the individual microphones becoming more omnidirectional. This observation is also made for the other source directions (top and bottom rows). Sources at zenith (top row) are also reproduced very well for all the analyzed frequencies, while for sources at nadir (bottom row) one can observe the effect of spatial aliasing for high frequencies, where the the condition $\lambda > 2d$ does not hold anymore. Thus, for high frequencies $f > c/(2d)$ spatial aliasing prevents signals from top and bottom to be reproduced correctly. Here, reproduction of a source at zenith is less sensitive to spatial aliasing than at nadir as the omnidirectional Ambisonics component $\chi_w(t)$ is calculated from the upper microphone. In real-live playback scenarios reproduction errors in z-components might not be a big problem as playback environments often do not have loudspeakers placed directly at the bottom and sources tend not to be at nadir in many recording scenarios. If an assumption of no sources coming from nadir can be made, this problem can also solved by encoding the created signal to zenith, neglecting nadir directions.



**(a)** 32 Hz **(b)** 63 Hz **(c)** 125 Hz **(d)** 250 Hz **(e)** 500 Hz **(f)** 1 kHz **(g)** 2 kHz

**Fig. 4:** Reproduced directivity of the microphone array after equalization for frequencies between 32 – 2000 Hz and a single source at zenith (top row), the horizon (middle row) and nadir (bottom row). Phase is depicted via gray scales from 0 (white) to $\pm\pi$ (black) as shown in figure 5.

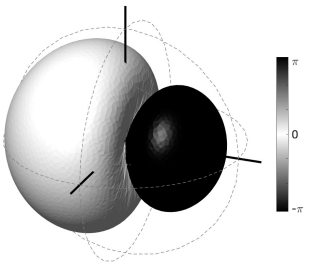



**Fig. 5:** Ideal first-order Ambisonics directivity plot for a single source at the horizon ($\varphi = 0°, \vartheta = 90°$). Phase is depicted via gray scales.

## 6 Conclusion

We showed that first-order Ambisonics can be recorded with a differential array of two dual-diaphragm microphones. Equalization of the differential figure-of-eight pattern aligned with the z-axis is necessary to compensate for the high-pass characteristic and phase shift of the differential signals. Equalization using a combination of an IIR filter targeting the phase-shift and FIR filters equalizing the diffuse-field magnitude responses of the array, is shown to be effective for low frequencies where no spatial aliasing occurs. For high frequencies, aliasing for strongly elevated sources cannot be eliminated. Still, this might not be a big problem in many real-life recording and playback scenarios. The described filters and corresponding processing is implemented in an open-source plugin, provided at `https://github.com/AustrianAudio/AmbiCreator`.

## References

[1] Zotter, F. and Frank, M., *Ambisonics, A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, Springer, 2019, doi:10.1007/978-3-030-17207-7.

[2] Jackson, L. B., "Correction to impulse invariance," *IEEE Signal Processing Letters*, 7(10), pp. 273–275, 2000.

[3] Mecklenbräuker, W. F., "Remarks on and correction to the impulse invariant method for the design of IIR digital filters," *Signal Processing*, 80(8), pp. 1687–1690, 2000.

[4] Eitelberg, E., "Convolution invariance and corrected impulse invariance," *Signal Processing*, 86(5), pp. 1116–1120, 2006.

[5] Gräf, M. and Potts, D., "On the computation of spherical designs by a new optimization approach based on fast spherical Fourier transforms," *Numerische Mathematik*, 119(4), pp. 699–724, 2011.