Audio Engineering Society

Convention Paper

Presented at the 126th Convention
2009 May 7–10 Munich, Germany

Design and Limitations of Non-Coincidence Correction Filters for Soundfield Microphones

Christof Faller[1] and Mihailo Kolundžija[2]

[1] *Illusonic LLC, Chemin du Trabandan 28A, 1006 Lausanne, Switzerland*

[2] *Audiovisual Communications Laboratory, EPFL, 1015 Lausanne, Switzerland*

Correspondence should be addressed to Christof Faller (christof.faller@illusonic.com)

ABSTRACT

The tetrahedral microphone capsule arrangement in a Soundfield microphone captures a so-called A-format signal which is then converted to a corresponding B-format signal. The phase differences between the A-format signal channels due to non-coincidence of the microphone capsules cause errors in the corresponding B-format signals and linear combinations thereof. Various strategies for designing B-format non-coincidence correction filters are compared and limitations are discussed.

1. INTRODUCTION

A Soundfield microphone [1,2] captures four signals with sub-cardioid microphone capsules arranged as a tetrahedral microphone array [3]. These sub-cardioid signals (often denoted A-format) are then converted to the well known B-format, which corresponds to signals with an omnidirectional characteristic (W signal) and three signals with dipole characteristic (X, Y, and Z signals). The sub-cardioid microphone capsules in a Soundfield microphone are arranged as coincidentally as possible, but nevertheless gain and directional response errors result from the phase differences between the microphone capsule signals when converted to B-format.

It has been claimed previously that errors due to non-coincidence can be corrected [1–4]. While a 4-by-4 filter matrix could be optimized for the task, it has been argued that a 4-by-4 scalar matrix followed by an equalization filter for each of the B-format signals (W, X, Y, and Z) is sufficient and simpler to determine [3].

In this paper, we derive the theory of the non-coincidence correction proposed by Gerzon [3]. Based on simulation of a Soundfield microphone, the effect of the non-coincidence is illustrated and

least mean squares optimal non-coincidence correction filters are proposed. Practical considerations are discussed and it is shown that nearly optimal non-coincidence correction filters can be computed based on a small number of microphone measurements.

The theory and simulation data presented in this paper also clarify the limitations of non-coincidence correction for Soundfield microphones. It is shown that while in the $x = 0$, $y = 0$, and $z = 0$ planes errors due to non-coincidence can be corrected up to high frequencies, in other planes the performance of the microphone is worse despite of non-coincidence correction.

The paper is organized as follows. Section 2 describes a Soundfield microphone, A-format, and B-format signals in detail. It also derives the A-format (microphone capsule signals) to B-format scalar conversion matrix assuming perfect coincidence. Non-coincidence correction is discussed from various points of view in Section 3, including theoretical results, simulations, and practical ways of determining non-coincidence correction filters. The conclusions are given in Section 4.

2. SOUNDFIELD MICROPHONE, A-FORMAT, AND B-FORMAT

A Soundfield microphone uses four cardioid capsules whose membrane centers are arranged at the tips of a tetrahedron pointing outwards [1], as is illustrated in Figure 1. According to [1], the cardioid capsules are configured to yield sub-cardioid polar pattern ($\frac{1}{3}(2+\cos\phi)$), which has the advantage that off axis response is better than for the cardioid case.

Figure 2 illustrates the microphone capsule center positions relative to the x, y, and z axes. Note that the axes are defined such that the B-format X signal corresponds to a dipole pointing towards the positive x-axis. Similarly, the B-format Y and Z signals are oriented towards the positive y and z axes, respectively.

The coordinates of the microphone capsule center positions are

$$\begin{aligned} \mathbf{p}_{\mathrm{LF}} &= d\mathbf{v}_{\mathrm{LF}} \\ \mathbf{p}_{\mathrm{RB}} &= d\mathbf{v}_{\mathrm{RB}} \\ \mathbf{p}_{\mathrm{LB}} &= d\mathbf{v}_{\mathrm{LB}} \\ \mathbf{p}_{\mathrm{RF}} &= d\mathbf{v}_{\mathrm{RF}}\,, \end{aligned} \tag{1}$$

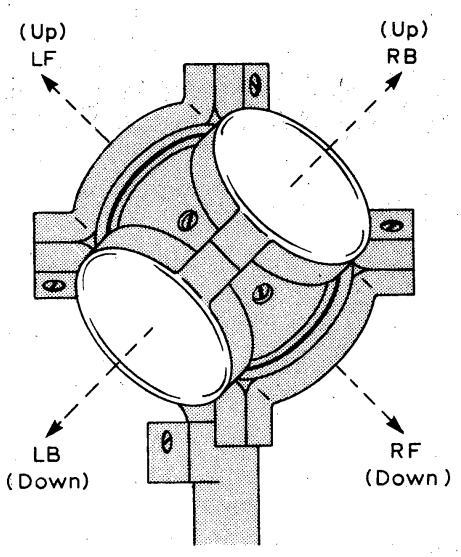

Fig. 1: Sub-cardioid microphone capsule arrangement in a Soundfield microphone (picture source: [1]).

where d in meter is the distance of the origin to the edges of the tetrahedron,

$$\begin{aligned} \mathbf{v}_{\mathrm{LF}} &= \frac{1}{\sqrt{3}}[\,1 \;\; 1 \;\; 1]' \\ \mathbf{v}_{\mathrm{RB}} &= \frac{1}{\sqrt{3}}[-1 \;\; -1 \;\; 1]' \\ \mathbf{v}_{\mathrm{LB}} &= \frac{1}{\sqrt{3}}[-1 \;\; 1 \;\; -1]' \\ \mathbf{v}_{\mathrm{RF}} &= \frac{1}{\sqrt{3}}[\,1 \;\; -1 \;\; -1]'\,, \end{aligned} \tag{2}$$

are unit vectors defining the look directions of the microphone capsules, and $'$ is the transpose operation. Throughout this paper we use $d = 1.47$ cm [3].

The A-format is converted to B-format by

$$\begin{bmatrix} W(t) \\ X(t) \\ Y(t) \\ Z(t) \end{bmatrix} = \mathbf{M}_{\mathrm{AB}} \begin{bmatrix} S_{\mathrm{LF}}(t) \\ S_{\mathrm{RB}}(t) \\ S_{\mathrm{LB}}(t) \\ S_{\mathrm{RF}}(t) \end{bmatrix}, \tag{3}$$

where t is the time variable and $S_{\mathrm{LF}}(t)$, $S_{\mathrm{RB}}(t)$, $S_{\mathrm{LB}}(t)$, and $S_{\mathrm{RF}}(t)$ are the microphone capsule signals. $\mathbf{M}_{\mathrm{AB}}$ is a 4-by-4 matrix which is derived in the following for the coincident case ($d = 0$).

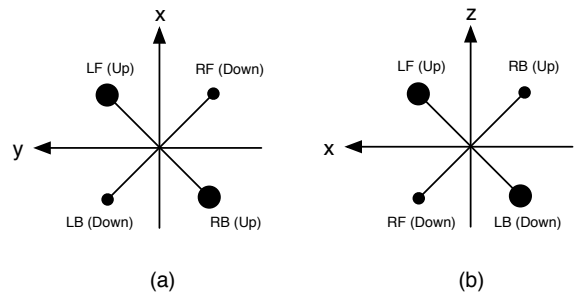

Fig. 2: Microphone capsule center positions: (a) top view and (b) side view. A large dot indicates a position above and a small dot a position below the planes defined by the shown axes.

A signal with a directional response of $a+(1-a)\cos\phi$ pointing towards azimuth θ_0 and elevation ϕ_0 can be obtained by combining the B-format signals:

$$S(t) = aW(t) + \frac{1-a}{\sqrt{2}}\mathbf{v}'_{\theta_0,\ \phi_0}\begin{bmatrix} X(t) \\ Y(t) \\ Z(t) \end{bmatrix}, \quad (4)$$

where $\mathbf{v}_{\theta_0,\ \phi_0}$ is a unit vector pointing towards azimuth θ_0 and elevation ϕ_0, i.e.

$$\mathbf{v}_{\theta_0,\ \phi_0} = [\cos\theta_0\cos\phi_0 \ \ \sin\theta_0\cos\phi_0 \ \ \sin\phi_0]'. \quad (5)$$

From (4) it follows that the matrix converting B-format signals to the microphone capsule signals (A-format) is

$$\mathbf{M}_{\mathrm{BA}} = \begin{bmatrix} a & \frac{(1-a)}{\sqrt{2}}\mathbf{v}'_{\mathrm{LF}} \\ a & \frac{(1-a)}{\sqrt{2}}\mathbf{v}'_{\mathrm{RB}} \\ a & \frac{(1-a)}{\sqrt{2}}\mathbf{v}'_{\mathrm{LB}} \\ a & \frac{(1-a)}{\sqrt{2}}\mathbf{v}'_{\mathrm{RF}} \end{bmatrix} \quad (6)$$

$$= \begin{bmatrix} a & \frac{1-a}{\sqrt{6}} & \frac{1-a}{\sqrt{6}} & \frac{1-a}{\sqrt{6}} \\ a & -\frac{1-a}{\sqrt{6}} & -\frac{1-a}{\sqrt{6}} & \frac{1-a}{\sqrt{6}} \\ a & -\frac{1-a}{\sqrt{6}} & \frac{1-a}{\sqrt{6}} & -\frac{1-a}{\sqrt{6}} \\ a & \frac{1-a}{\sqrt{6}} & -\frac{1-a}{\sqrt{6}} & -\frac{1-a}{\sqrt{6}} \end{bmatrix},$$

where the constant a is chosen according to the specification of the microphone capsules. The original Soundfield microphone uses $a = \frac{2}{3}$ [1], which corresponds to the sub-cardioid directional response shown in linear scale in Figure 3.

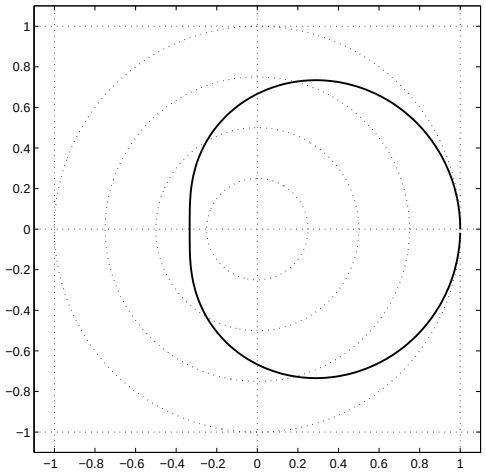

Fig. 3: Sub-cardioid directional response in linear scale.

The matrix for converting the microphone capsule signals (A-format) to B-format signals (3) is the inverse of (6), i.e.

$$\mathbf{M}_{\mathrm{AB}} = \mathbf{M}_{\mathrm{BA}}^{-1} = \begin{bmatrix} \frac{1}{4a} & \frac{1}{4a} & \frac{1}{4a} & \frac{1}{4a} \\ b & -b & -b & b \\ b & -b & b & -b \\ b & b & -b & -b \end{bmatrix}, \quad (7)$$

with

$$b = \frac{\sqrt{6}}{4(1-a)}. \quad (8)$$

Figure 4 shows ideal directional responses of the W and X signals. Note that all dipole directional responses shown in this paper are normalized such that the maximum gain is 0 dB (1 in the linear scale which is shown) as opposed to 3 dB according to B-format definition. In the following, it is shown that non-coincidence of the microphone capsules will impair the directional responses of the B-format signals.

3. NON-COINCIDENCE CORRECTION

Due to the non-coincidence the directional responses of the B-format signals deviate from the ideal case shown in Figure 4. Note that from the symmetry of the setup (microphone capsule positions, coordinate system) it follows that:

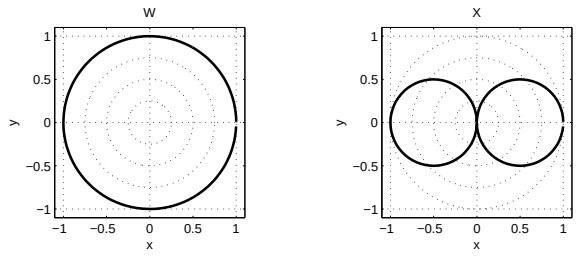

Fig. 4: Polar plots of the directional responses of the omni-directional W and dipole X signals in linear scale.

- The directional responses of W measured in the $x = 0$ and $y = 0$ planes are identical to the directional response of W measured in the horizontal ($z = 0$) plane.
- The directional response of X measured in the $y = 0$ plane is identical to the directional response of X measured in the $z = 0$ plane.
- The directional response of Y measured in the $x = 0$ plane is identical to the directional response of Y measured in the $z = 0$ plane.
- The directional response of Z measured in the $x = 0$ plane is identical to the directional response of Z measured in the $y = 0$ plane.
- The 3D directional responses of Y and Z are identical to the directional response of X, if Y and Z are rotated accordingly.

Gerzon proposed [3] applying a scalar matrix (7) for A-format to B-format conversion, followed by a filter for each B-format channels, i.e.

$$\begin{aligned} \tilde{W}(t) &= h_W(t) \star W(t) \\ \tilde{X}(t) &= h_X(t) \star X(t) \\ \tilde{Y}(t) &= h_Y(t) \star Y(t) \\ \tilde{Z}(t) &= h_Z(t) \star Z(t)\,, \end{aligned} \tag{9}$$

where $\star$ denotes linear convolution and $h_W(t)$, $h_X(t)$, $h_Y(t)$, and $h_Z(t)$ are the non-coincidence correction filters. Assuming ideal identical microphone capsules positioned precisely, Gerzon showed that $h_Z(t) = h_Y(t) = h_X(t)$ holds, which also follows from the previously mentioned B-format properties due to symmetry.

Section 3.1 gives a detailed derivation of Gerzon's proposed non-coincidence correction filters in a generalized form (not only for cardioid capsules) and discusses the results. In Section 3.2, a Soundfield microphone is simulated and least mean squares optimal non-coincidence correction filters are estimated. Considerations on what is a good approach to use in practice, based on microphone measurement, are discussed in Section 3.3.

3.1. Theoretical derivation of non-coincidence correction

A monochromatic far-field sound with temporal frequency ω, propagating from the direction defined by a unit vector $\mathbf{u} = [\cos\theta\cos\phi \;\; \sin\theta\cos\phi \;\; \sin\phi]'$, can be described by a plane wave

$$p(\mathbf{r}, t) = Ae^{-j(\omega t - \mathbf{k}\cdot\mathbf{r})}\,, \tag{10}$$

where A is a complex amplitude (conveying the information about both the real amplitude and the initial phase), and $\mathbf{k} = k\mathbf{u}$ is the wave vector that points toward the direction of sound propagation with intensity $k = \frac{\omega}{c}$, where c is the speed of sound.

The signals captured by the four microphone capsules of a Soundfield microphone can be expressed as

$$\begin{aligned} S_{\mathrm{LF}}(t) &= (a + (1-a)\mathbf{v}_{\mathrm{LF}} \cdot \mathbf{u})\, p(\mathbf{p}_{\mathrm{LF}}, t) \\ S_{\mathrm{RB}}(t) &= (a + (1-a)\mathbf{v}_{\mathrm{RB}} \cdot \mathbf{u})\, p(\mathbf{p}_{\mathrm{RB}}, t) \\ S_{\mathrm{LB}}(t) &= (a + (1-a)\mathbf{v}_{\mathrm{LB}} \cdot \mathbf{u})\, p(\mathbf{p}_{\mathrm{LB}}, t) \\ S_{\mathrm{RF}}(t) &= (a + (1-a)\mathbf{v}_{\mathrm{RF}} \cdot \mathbf{u})\, p(\mathbf{p}_{\mathrm{RF}}, t)\,. \end{aligned} \tag{11}$$

Using the decomposition of a plane wave into a sum of spherical waves (e.g., see [3])

$$e^{jkd(\mathbf{u}\cdot\mathbf{v})} = \sum_{n=0}^{\infty} j^n (2n+1) P_n(\mathbf{u}\cdot\mathbf{v})\, j_n(kd)\,, \tag{12}$$

where $P_n(x)$ is the nth-degree Legendre polynomial and $j_n(x)$ is the nth-order spherical Bessel function of the first kind, together with the A-format to B-format conversion formula given in (3), one can express signals W and X as

$$\begin{aligned} W &= Ae^{j\omega t} \sum_{n=0}^{\infty} O_n^W(kd) \left(aM_n^W(\mathbf{u}) + (1-a) N_n^W(\mathbf{u}) \right) \\ X &= Ae^{j\omega t} \sum_{n=0}^{\infty} O_n^X(kd) \left(aM_n^X(\mathbf{u}) + (1-a) N_n^X(\mathbf{u}) \right)\,, \end{aligned} \tag{13}$$

n	$M_n^W(\mathbf{u})$	$N_n^W(\mathbf{u})$	$M_n^X(\mathbf{u})$	$N_n^X(\mathbf{u})$
0	4	0	0	$\frac{4}{\sqrt{3}}\cos\theta\cos\phi$
1	0	$\frac{4}{3}$	$\frac{4}{\sqrt{3}}\cos\theta\cos\phi$	$\frac{8}{3}\sin\theta\cos\phi\sin\phi$
2	0	$\frac{6}{\sqrt{3}}\sin 2\theta\cos^2\phi\sin\phi$	$4\sin\theta\cos\phi\sin\phi$	$\frac{8}{5\sqrt{3}}\cos\theta\cos\phi - \frac{1}{\sqrt{3}}\cos 3\theta\cos^3\phi + \frac{\sqrt{3}}{5}\cos\theta\cos\phi(5\sin^2\phi - 1)$

Table 1: Values of functions $M_n^W(\mathbf{u})$, $N_n^W(\mathbf{u})$, $M_n^X(\mathbf{u})$ and $N_n^X(\mathbf{u})$ up to the second order, expressed as linear combinations of spherical harmonics.

where

$$O_n^W(x) = j^n \frac{2n+1}{4a}\, j_n(x)$$
$$O_n^X(x) = j^n \frac{\sqrt{6}(2n+1)}{4(1-a)}\, j_n(x)\,,$$

and

$$\begin{aligned}
M_n^W(\mathbf{u}) &= P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{LF}}) + P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{RB}}) \\
&\quad + P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{LB}}) + P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{RF}}) \\
N_n^W(\mathbf{u}) &= \mathbf{u}\cdot\mathbf{v}_{\mathrm{LF}}\, P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{LF}}) + \mathbf{u}\cdot\mathbf{v}_{\mathrm{RB}}\, P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{RB}}) \\
&\quad + \mathbf{u}\cdot\mathbf{v}_{\mathrm{LB}}\, P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{LB}}) + \mathbf{u}\cdot\mathbf{v}_{\mathrm{RF}}\, P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{RF}}) \\
M_n^X(\mathbf{u}) &= P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{LF}}) - P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{RB}}) \\
&\quad - P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{LB}}) + P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{RF}}) \\
N_n^X(\mathbf{u}) &= \mathbf{u}\cdot\mathbf{v}_{\mathrm{LF}}\, P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{LF}}) - \mathbf{u}\cdot\mathbf{v}_{\mathrm{RB}}\, P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{RB}}) \\
&\quad - \mathbf{u}\cdot\mathbf{v}_{\mathrm{LB}}\, P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{LB}}) + \mathbf{u}\cdot\mathbf{v}_{\mathrm{RF}}\, P_n(\mathbf{u}\cdot\mathbf{v}_{\mathrm{RF}})\,.
\end{aligned}$$

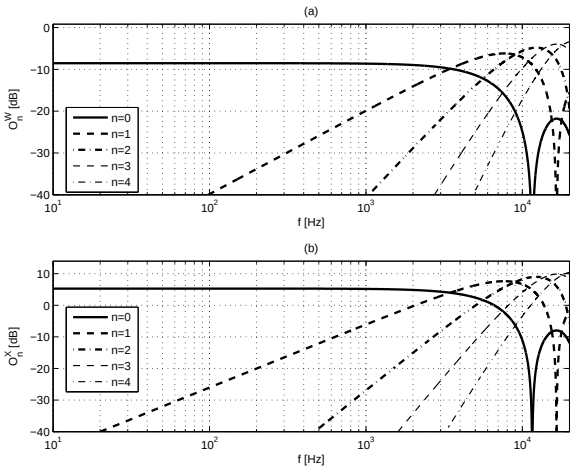

Fig. 5: Magnitude frequency responses of functions $O_n^W(kd)$ (a) and $O_n^X(kd)$ (b) for different orders n.

The directional responses of W and X are direction dependent linear combinations of $O_n^W(kd)$ and $O_n^X(kd)$, respectively. Figure 5 shows the magnitude frequency characteristics of $O_n^W(kd)$ and $O_n^X(kd)$ up to order $n = 4$. The data shown in the figure indicate that the signals W and X are mostly determined by the terms of orders n up to two for frequencies up to about 8 kHz. Furthermore, for orders higher than two, the functions $M_n^W(\mathbf{u})$, $N_n^W(\mathbf{u})$, $M_n^X(\mathbf{u})$, and $N_n^X(\mathbf{u})$, whose values are given in Table 1, contain only spherical harmonics of orders higher than one[1]. Since the signals W and X should be proportional to zeroth (constant) and first order ($\cos\theta\cos\phi$) spherical harmonics, respectively, higher than first order spherical harmonics can be considered as aliasing distortions. Note also that, due to orthogonality, high-order spherical harmonic terms do not affect the non-coincidence correction of the zeroth and first order spherical harmonic responses of signals W and X.

Following these observations, the signals W and X for a monochromatic far-field sound of frequency ω that propagates from the direction defined by unit vector $\mathbf{u}$ can be simplified as follows:

$$\begin{aligned}
W(t) &= (F_W(\omega) + R_W(\omega,\theta,\phi))p_0(t) \\
X(t) &= (\sqrt{2}\cos\theta\cos\phi F_X(\omega) + R_X(\omega,\theta,\phi))p_0(t)\,,
\end{aligned} \tag{14}$$

[1] Tables of spherical harmonics can be found in [5].

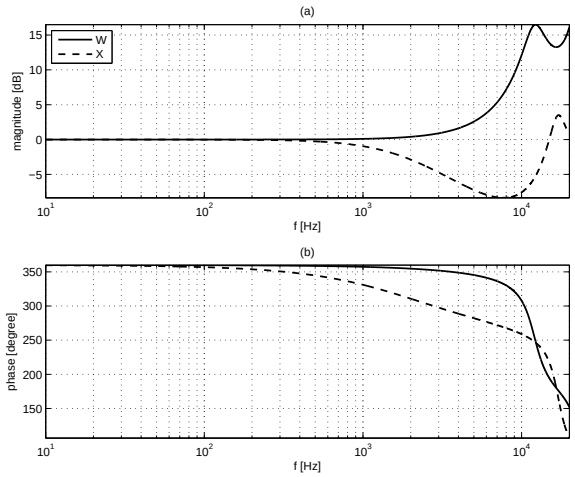

Fig. 6: Magnitude (a) and phase (b) frequency characteristics of non-coincidence correction filters H_W and H_X according to Gerzon's theory.

where

$$F_W(\omega) = j_0\left(\frac{\omega d}{c}\right) + j\, j_1\left(\frac{\omega d}{c}\right)$$
$$F_X(\omega) = j_0\left(\frac{\omega d}{c}\right) + j\frac{3a}{1-a}\, j_1\left(\frac{\omega d}{c}\right) - 2\, j_2\left(\frac{\omega d}{c}\right) \quad (15)$$

are the Soundfield microphone's frequency responses of zeroth- and first-order harmonics, respectively, and the remaining terms $R_W(\omega, \theta, \phi)$ and $R_X(\omega, \theta, \phi)$ contain only higher-order spherical harmonic terms.

For non-coincidence correction, one needs to invert the frequency characteristics $F_W(\omega)$ and $F_X(\omega)$. The non-coincidence correction filters $h_W(t)$ and $h_X(t)$ are the inverse Fourier transform of

$$\begin{aligned} H_W(\omega) &= F_W^{-1}(\omega) \\ H_X(\omega) &= F_X^{-1}(\omega)\,. \end{aligned} \quad (16)$$

$H_W(\omega)$ and $H_X(\omega)$ are shown for $a = \frac{2}{3}$ in Figure 6.

The responses given in (15) are in accordance with the frequency responses of zeroth- and first-order spherical harmonics of a Soundfield microphone using cardioid capsules ($a = \frac{1}{2}$), given by Gerzon in [3]. It should also be noted that non-coincidence correction by filters $h_W(t)$ and $h_X(t)$ is effective up to some limiting frequency at which aliasing terms, i.e. high-order spherical harmonics, start to dominate the directional responses of the signals W and X. In [3], Gerzon gives an approximation of the limiting frequency in the form $f_l = 10/d$ kHz (where d is expressed in cm), which for $d = 1.47$ cm gives a limiting frequency $f_l = 7.35$ kHz. Above the limiting frequency f_l, Gerzon suggests correcting nominal zeroth- and first-order harmonic signals, W and X, to have flat responses to diffuse random sound fields.

It is also worth noting that even though the signal X contains strong second-order spherical harmonic terms $N_1^X(\mathbf{u}) = \frac{8}{3}\cos\theta\cos\phi\sin\phi$ and $M_2^X(\mathbf{u}) = 4\cos\theta\cos\phi\sin\phi$ (as observed from the last two rows of Table 1), its directional response in the plane $z = 0$ is not affected by them, since the spherical harmonic $\cos\theta\cos\phi\sin\phi$ equals zero in this plane. Due to symmetry, the same holds for the $y = 0$ plane. Also, the directional response of W is better in the $x = 0$, $y = 0$, and $z = 0$ planes, since the large third order spherical harmonic term $N_2^W(\mathbf{u})$ is zero in these planes. Simulations shown in the following section confirm these observations.

3.2. Non-Coincidence correction determined by simulation

In this section, it is shown how to compute non-coincidence correction filters based on simulating the Soundfield microphone described in Section 2. Ideal sub-cardioid capsules are assumed with axially symmetric directional responses. Directional responses are simulated considering far field sound. The non-coincidence correction filters are determined such that the B-format directional response errors are minimized in a least mean squares sense.

The top two panels of Figure 7 show the directional responses of the B-format signals W and X, respectively, measured in the horizontal plane ($z = 0$). There is no need to show the directional responses of Y and Z because they are identical up to a rotation to the X directional response. Note that the omnidirectional W gain is decreasing and the dipole gains are increasing as frequency increases, which is quantitatively in line with the theoretical analysis, the data shown in [6], and the Soundfield microphone measurements given in [7]. Note that the gain changes relatively quickly as frequency increases, but the shape of the directional responses is rather correct up to about 10 kHz.

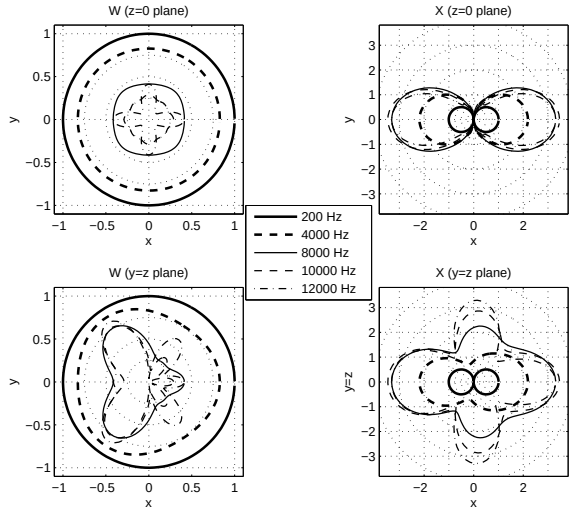

Fig. 7: Polar plots at various frequencies of directional responses of W and X in the $z = 0$ (top two panels) and $y = z$ (bottom two panels) planes.

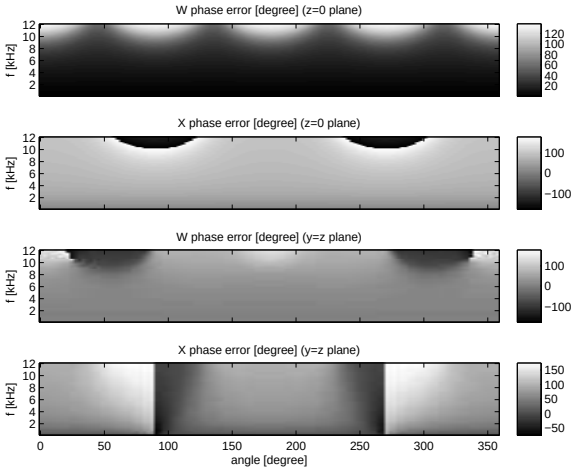

Fig. 8: The phase error of X and W in the $z = 0$ (top two panels) and the $y = z$ (bottom two panels) planes.

The bottom two panels of Figure 7 show the directional responses of the same B-format signals but measured in the $y = z$ plane. Note that in the $y = z$ plane the shape of the directional responses of the W and X signals degrade much more quickly than in the horizontal plane as frequency increases. This is in line with the previously derived theoretical results, which showed that large third and second order spherical harmonic terms in the W and X signals, respectively, vanish in the $z = 0$ plane.

The phase differences between the simulated W and X signals and ideal corresponding signals are shown in Figure 8. Except at the lowest frequencies, there is a significant phase error.

Figure 9 shows examples of B-format decodings measured in the horizontal ($z = 0$) plane where a cardioid is steered to different directions in the horizontal plane. The data shown imply that the gain and directional response are highly frequency-dependent at and above 4 kHz.

As mentioned previously, the data shown in the top panels of Figure 7 indicate that the shapes of the B-format magnitude responses in the horizontal plane are close to desired, except for the highest frequencies. Thus, by means of a frequency-dependent gain

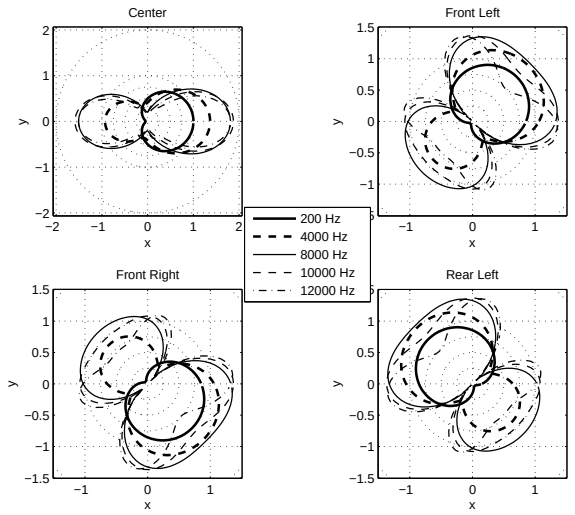

Fig. 9: Polar plots at various frequencies of directional responses of horizontal ($z = 0$) B-format decodings.

correction one can obtain B-format signals with a magnitude response much more similar to the desired one.

However, as has been shown, the directional responses measured in other planes can degrade much more quickly as frequency increases. Non-coincidence correction by means of B-format frequency-dependent gain correction will only result in directional responses approximating the ideal responses at those frequencies and in those planes where the shape of the B-format directional response is correct.

A good magnitude response alone is not enough, since phase differences between the B-format channels will result in errors in the magnitude responses of B-format decoded signals. Thus, phase differences between W and the X, Y, and Z dipoles should be minimized if possible. The data shown in Figure 8 indicate that the phase error of the B-format signals has usually a significant non-zero mean for each frequency, which if subtracted will reduce the average phase error.

The Fourier transform of the filters $h_W(t)$ and $h_X(t)$, which minimize the mean square error between the desired and actual directional responses, are

$$\begin{aligned} H_W(\omega) &= \frac{\int_{-\pi}^{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d_W(\theta,\phi) D_W^\star(\theta,\phi,\omega)\cos\phi \mathrm{d}\theta\mathrm{d}\phi}{\int_{-\pi}^{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} D_W(\theta,\phi,\omega) D_W^\star(\theta,\phi,\omega)\cos\phi \mathrm{d}\theta\mathrm{d}\phi} \\ H_X(\omega) &= \frac{\int_{-\pi}^{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d_X(\theta,\phi) D_X^\star(\theta,\phi,\omega)\cos\phi \mathrm{d}\theta\mathrm{d}\phi}{\int_{-\pi}^{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} D_X(\theta,\phi,\omega) D_X^\star(\theta,\phi,\omega)\cos\phi \mathrm{d}\theta\mathrm{d}\phi}, \end{aligned} \tag{17}$$

respectively, where θ is azimuth, ϕ is elevation, $d_W(\theta,\phi)$ and $d_X(\theta,\phi)$ are the desired omnidirectional (W) and dipole (X) directional responses

$$\begin{aligned} d_W(\theta,\phi) &= 1 \\ d_X(\theta,\phi) &= \cos\theta\cos\phi, \end{aligned} \tag{18}$$

and $D_W(\theta,\phi,\omega)$ and $D_X(\theta,\phi,\omega)$ are the actual frequency dependent complex directional responses.

The magnitude and phase responses of $H_W(\omega)$ and $H_X(\omega)$ are shown in Panels (a) and (b) of Figure 10, respectively. Comparison with Figure 6 indicates that above 3 kHz Gerzon's filters start to deviate from the least mean squares filters. The magnitude

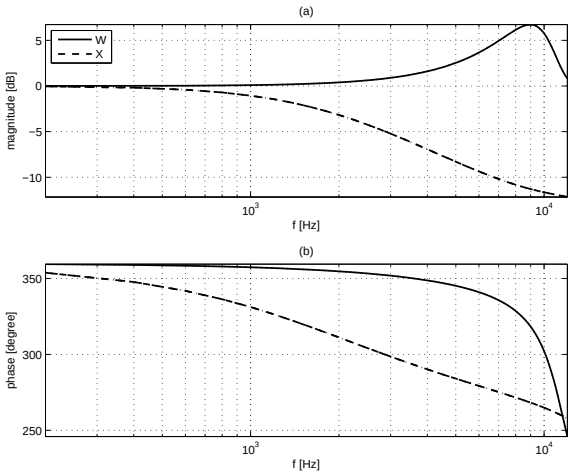

Fig. 10: Magnitude (a) and phase (b) frequency characteristics of the least mean squares non-coincidence correction filters H_W and H_X.

of $H_W(\omega)$ at high frequencies of Gerzon's filter is significantly larger than that of the corresponding least mean squares filter. The deviation of Gerzon's filters from the least mean squares filters above 3 kHz implies that the former are sub-optimal[2].

Figure 11 shows the magnitude of the corrected B-format directional responses corresponding to the responses shown in Figure 7. The directional responses in the horizontal ($z = 0$) plane are similar to the desired responses (shown in Figure 4) up to high frequencies. The directional responses in the $y = z$ plane also improve in terms of gain.

Figure 12 shows the phase error of the corrected B-format signals. The data shown indicate that the phase correction eliminates the phase errors nearly ideally up to more than 8 kHz in the horizontal plane. The average phase error is also reduced in the $y = z$ plane.

The example decodings in the horizontal plane of the corrected B-format, shown in Figure 13, show that the directional response gains and shapes are greatly

[2] As mentioned earlier in the paper, Gerzon proposed correction of B-format at higher frequencies (> 7.35 kHz) to have flat response to diffuse random sound fields, which corresponds to filters which deviate less from the least mean squares filters at these frequencies.

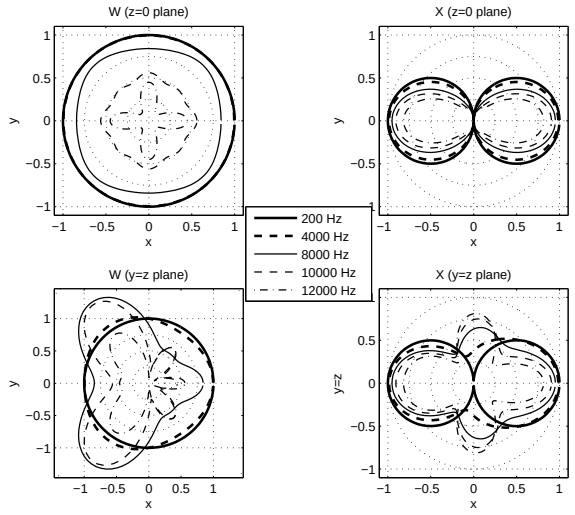

Fig. 11: Polar plots at various frequencies of directional responses of non-coincidence corrected W and X in the $z = 0$ (top two panels) and $y = z$ (bottom two panels) planes.

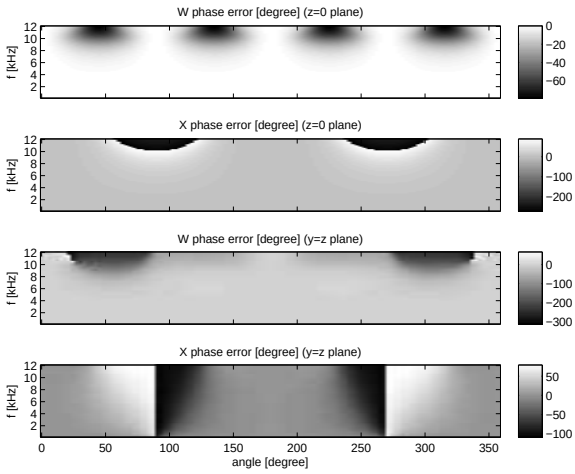

Fig. 12: The phase error of non-coincidence corrected X and W in the $z = 0$ (top two panels) and the $y = z$ (bottom two panels) planes.

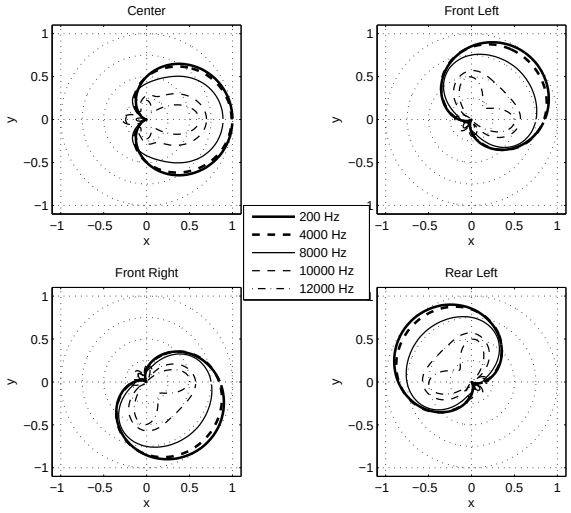

Fig. 13: Polar plots at various frequencies of directional responses of non-coincidence corrected horizontal ($z = 0$) B-format decodings.

improved compared to the non-corrected case shown in Figure 9.

Similar decodings, but with 45 degrees elevation ($\phi = \frac{\pi}{4}$), are shown in Figure 14, indicating that decoded B-format signals are worse in planes other than the $x = 0$, $y = 0$, or $z = 0$ planes. But note that the least mean squares non-coincidence correction filter also improves this case compared to not using non-coincidence correction filters.

3.3. Practical Considerations

Determination of B-format correction filters based on theory is problematic for various reasons. Practical constraints such as non-ideal microphone capsules, non-matching microphone capsules, limited positioning precision, and diffraction effects make filters determined under ideal assumptions deviate substantially from the desired ideal filters.

An advantage of the proposed least mean squares filters (17) is that they can be computed from actual measurements, i.e. $D_W(\theta, \phi, \omega)$ and $D_X(\theta, \phi, \omega)$. Perfect symmetry can not be assumed in practice and it is favorable to compute separate non-coincidence correction filters for Y and Z, $h_Y(t)$ and $h_Z(t)$, based on measurement of $D_Y(\theta, \phi, \omega)$ and $D_Z(\theta, \phi, \omega)$.

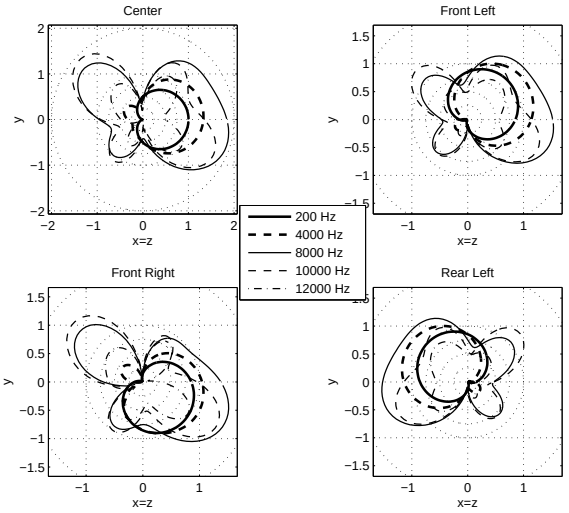

Fig. 14: Polar plots at various frequencies of directional responses of non-coincidence corrected elevated ($z = x$) B-format decodings.

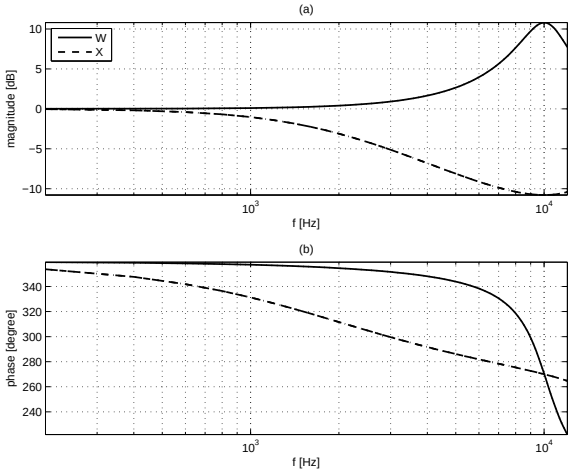

Fig. 15: Magnitude (a) and phase (b) frequency characteristics of the single measurement based non-coincidence correction filters H_W and H_X.

Farina proposed computation of B-format non-coincidence correction filters based on a single on-axis forward measurement per capsule [8]. This approach is much simpler in terms of measurement effort. The procedure is as follows:

- Measure magnitude and phase response of W relative to sound arriving from the positive x-direction.
- Measure magnitude and phase responses of X, Y, and Z relative to sound arriving from the positive x, y, and z directions, respectively.
- Design the B-format correction filters such that their magnitude and phase responses are the inverse of the measured magnitude and phase responses and multiply the X, Y, and Z filters with $\sqrt{2}$.

Figure 15 shows the magnitude and phase responses of the non-coincidence correction filters obtained with a single measurement per filter. Comparison with Figure 10 indicates that up to 7 kHz these filters are very similar to the least mean squares filters. Above 7 kHz the magnitude correction of the single measurement based filter for W is larger than the magnitude correction of the least mean squares filter. Thus, it is favorable to reduce the single measurement based filter magnitude at high frequencies to prevent excessive high frequency gain.

Using the same Soundfield microphone simulation as previously, Figure 16 shows the directional responses of W and X at various frequencies when using single measurement based non-coincidence correction. Corresponding B-format decodings in the $z = 0$ and $z = x$ planes are shown in Figures 17 and 18, respectively. Comparison of these B-format decodings with the corresponding ones using least mean squares correction filters, shown in Figures 13 and 14, confirms too high gains at high frequencies (in the $z = x$ plane). As mentioned, this effect may be reduced by reducing the gain of the W correction filter at high frequencies.

4. CONCLUSIONS

Non-coincidence correction for Soundfield microphones was investigated in this paper. B-format non-coincidence correction filters determined based on Gerzon's theory were compared to filters determined by simulation and optimization. The best correction filters minimize the mean square error between the achieved and desired B-format directional responses. The practically most useful approach is, as proposed by Farina, correction based on a single

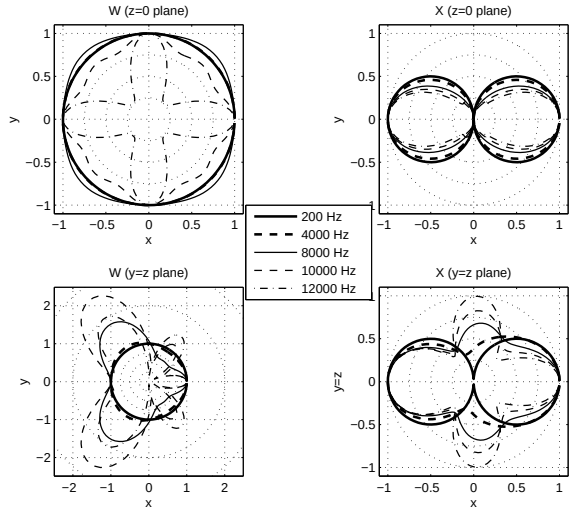

Fig. 16: Polar plots at various frequencies of directional responses of non-coincidence corrected W and X in the $z = 0$ (top two panels) and $y = z$ (bottom two panels) planes.

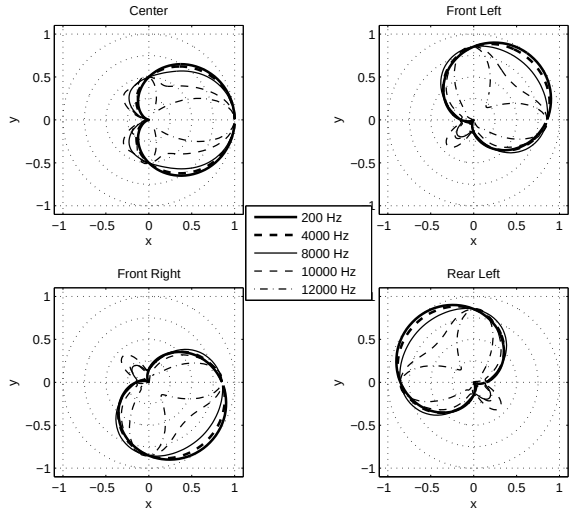

Fig. 17: Polar plots at various frequencies of directional responses of non-coincidence corrected horizontal ($z = 0$) B-format decodings.

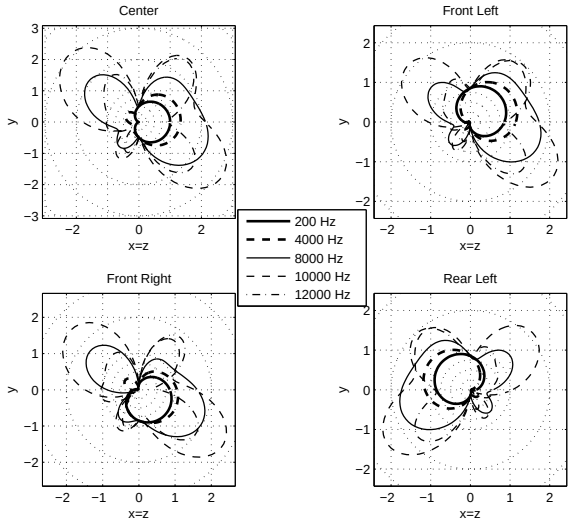

Fig. 18: Polar plots at various frequencies of directional responses of non-coincidence corrected elevated ($z = x$) B-format decodings.

measurement per B-format channel. This simplified correction filter is as good as the least mean squares filter up to about 7 kHz.

Analysis and simulations presented in this paper clarify fundamental limitations of non-coincidence correction. The non-coincidence corrected B-format W signal measured in the $x = 0$, $y = 0$, and $z = 0$ planes yields nearly ideal directional responses up to about 10 kHz. Similarly, the X signal yields very good results in the $z = 0$ and $y = 0$ planes, the Y signal in the $x = 0$ and $z = 0$ planes, and the Z signal in the $x = 0$ and $y = 0$ planes. In other planes, the errors of the directional responses of the B-format signals are higher. This can be explained theoretically by the absence of strong second and third order spherical harmonic terms in these planes.

For B-format decoding, a Soundfield microphone performs best for sources in the $z = 0$ plane and simultaneously decoding in the same plane. Similar good performance is achieved for sources in the $x = 0$ and $y = 0$ planes when decoding in these planes. The results are worse for sources in other planes or when using elevated decoding (e.g. $z = x$ plane).

5. REFERENCES

[1] K. Farrar, "Soundfield microphone: Design and development of microphone and control unit," *Wireless World*, pp. 48–50, Oct. 1979.

[2] K. Farrar, "Soundfield microphone - 2: Detailed functioning of control unit," *Wireless World*, pp. 99–103, Nov. 1979.

[3] M. A. Gerzon, "The design of precisely coincident microphone arrays for stereo and surround sound," in *Preprint 50th Conv. Aud. Eng. Soc.*, Mar. 1975.

[4] P. G. Craven and M. A. Gerzon, "Coincident microphone simulation covering three dimensional space and yielding various directional outputs," *U.S. Patent Specification 4042779*, Aug. 1977.

[5] Eric W. Weisstein, "Spherical harmonic," http://mathworld.wolfram.com.

[6] M. A. Gerzon, "Ambisonics in multichannel broadcasting and video," *J. Aud. Eng. Soc.*, vol. 33, no. 11, pp. 859–871, Nov. 1985.

[7] A. Farina, "Anechoic measurement of the polar plots of a soundfield st-250 b-format microphone," *http://pcfarina.eng.unipr.it/Public/B-format/Test-Soundfield-Microphones/Probe_Comparison_3.PDF*, undated.

[8] A. Farina, "A-format to b-format conversion," *http://pcfarina.eng.unipr.it/Public/B-format/A2B-conversion/A2B.htm*, Oct. 2006.