A METHOD FOR CONVERTING BETWEEN CYLINDRICAL AND SPHERICAL HARMONIC REPRESENTATIONS OF SOUND FIELDS

Mark R. P. Thomas[⋆]*, Jens Ahrens*[†]*, Ivan Tashev*[⋆]

[⋆]Microsoft Research, One Microsoft Way, Redmond, WA, USA
[†]University of Technology Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany
{*markth, ivantash*}*@microsoft.com, jens.ahrens@tu-berlin.de*

ABSTRACT

Spherical microphone and circular microphone arrays are useful for sampling sound fields that may be resynthesized with loudspeaker arrays. Spherical microphone arrays are desirable because of their ability to capture three-dimensional sound fields, however it is often more practical to construct loudspeaker arrays in the form of a closed circle located in the horizontal plane. This leads to a spatial undersampling as such a circular sampling can only yield a perfect representation of a height-invariant sound field. This paper investigates the consequences of such spatial undersampling by converting between cylindrical and spherical harmonic decompositions of solutions to the wave equation. We show analytically and via numerical simulations that 1) the result of the spatial undersampling is a purely horizontally propagating sound field, and 2) the ratio of travelling and standing components in the undersampled sound field varies depending on the incidence colatitude. The conversion is also used in a beamforming scenario and shows that the beamformer response becomes increasingly omnidirectional as the source moves away from the horizontal plane.

Index Terms— Cylindrical harmonics, spherical harmonics, Ambisonics, microphone arrays, loudspeaker arrays.

1. INTRODUCTION

It has been shown in the context of Ambisonics that orthogonal decompositions are powerful tools that enable the storage and transmission of captured sound fields [1]. This process is termed *spatial encoding*. The encoded fields can be re-synthesized (*decoded*), for example, using loudspeaker arrays of various geometries [2, 3]. Surface spherical harmonics are a convenient orthogonal basis for three-dimensional sound fields. Expansion coefficients can be obtained for real fields by processing a spherical microphone array recording [4].

Circular microphone arrays are not capable of capturing three-dimensional information but they provide a number of advantages over spherical ones: 1) Their physical construction is less cumbersome, 2) the required number of microphones for a given azimuthal resolution is significantly lower, and 3) the processing of the captured microphone signals is more convenient because no Associated Legendre Polynomials or factorials of relatively high numbers need to be evaluated. Circular microphone arrays are often implemented using a cylindrical scattering object and are therefore termed cylindrical arrays [5, 6].

Most large-scale loudspeaker arrays are limited to horizontal-only sound field synthesis, which has not shown to be a substantial limitation in most situations. This suggests that horizontal circular microphone arrays might be sufficient for recording content for such loudspeaker setups. The performance of circular microphone arrays has been treated in the literature in [7, 6, 8, 9, 10, 11]. Some of the literature considers only height-invariant sound fields [8, 9]. The remaining works [6, 10, 11] analyze the limitations of circular arrays capturing 3D fields only from an overall error point-of-view but do not reveal the detailed properties of resulting orthogonal decompositions.

In this paper we consider the more general problem of obtaining a cylindrical harmonic representation of a sound field from a given spherical harmonic representation. We identify what information is lost and analyze the properties of the sound fields that are reconstructed from the undersampled decomposition. A perfect estimate of the plane wave coefficients is assumed in order to reveal some fundamental properties. In practice, several considerations on the distribution of microphones and the treatment of acoustic scattering are required in order to accurately estimate these plane wave coefficients [4].

The remainder of the paper is organized as follows. The problem is formulated and background theory is given in Section 2, conversions between cylindrical and spherical representations are discussed in Section 3, experimentation with cylindrical representations of sources outside the horizontal plane is presented in Section 4 and conclusions are given in Section 5.

2. PROBLEM FORMULATION

Consider an interior 3D sound field [12] $S(\theta', \phi', r, \omega)$ in a spherical coordinate system at angular frequency ω, where (θ', ϕ', r) are colatitude angle, azimuth angle and radius respectively. This may be expanded into a continuum of propagating plane waves with respect to the surface of a notional unit sphere as

$$S(\theta', \phi', r, \omega) = \int_0^{2\pi} \int_0^{\pi} \bar{S}(\theta, \phi, \omega) e^{i\frac{\omega}{c} r \cos\Theta} \sin\theta d\theta d\phi, \quad (1)$$

where $\bar{S}(\theta, \phi, \omega)$ is termed the *signature function* [13] describing the amplitude of propagating plane waves with incidence angles (θ, ϕ), and Θ is the solid angle between (θ', ϕ') and (θ, ϕ). It is often convenient to expand $\bar{S}(\theta, \phi, \omega)$ in terms of the spherical basis functions $Y_n^m(\theta, \phi)$:

$$\bar{S}(\theta, \phi, \omega) \simeq \sum_{n=0}^{N} \sum_{m=-n}^{n} \breve{S}_n^m(\omega) Y_n^m(\theta, \phi), \quad (2)$$

where $\breve{S}_n^m(\omega)$ are the coefficients of expansion of order $m = \{-N \dots N\}$ and degree $n = \{0 \dots N\}$, and $Y_n^m(\theta, \phi)$ are

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

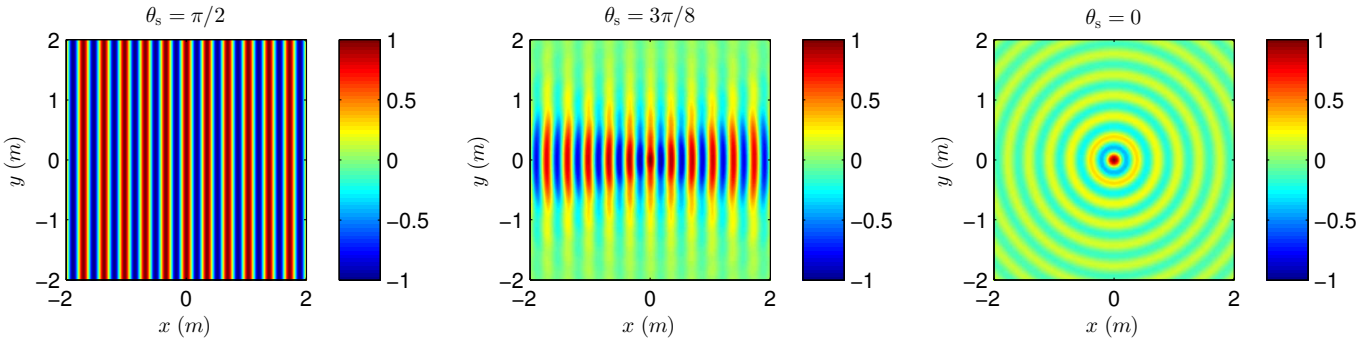

Fig. 1. 3D plane waves synthesized with cylindrical harmonic coefficients (17) and (7) for colatitude $\theta_s = \{\pi/2,\ 3\pi/8,\ 0\}$. As θ_s moves away from the horizontal plane, the travelling component decreases in magnitude and spatial support, leaving only a radial standing wave component when $\theta_s = \{0, \pi\}$.

the *spherical harmonics*

$$Y_n^m(\theta,\phi) = (-1)^m \sqrt{\frac{2n+1}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos\theta) e^{im\phi}. \quad (3)$$

The equality (2) becomes exact when $N = \infty$. A general sound field is expanded on the spherical harmonics as [13]

$$S(\theta',\phi',r,\omega) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} 4\pi i^n j_n\left(\frac{\omega}{c}r\right)\breve{S}_n^m(\omega) Y_n^m(\theta',\phi'), \quad (4)$$

where $j_n(x)$ is a spherical Bessel function describing the radial dependence of the field. In the special case where $\breve{S}_n^m(\omega) = Y_n^{-m}(\theta_s,\phi_s)$, (4) represents plane wave with incidence angle (θ_s,ϕ_s).

Let us now consider a 2D (height-invariant) field. The 2D analog of the signature function in (1) is

$$S(\phi',r,\omega) = \frac{1}{2\pi}\int_0^{2\pi} \bar{S}(\phi,\omega) e^{i\frac{\omega}{c}r\cos(\phi'-\phi)} d\phi, \quad (5)$$

where $e^{im\phi}$ are the *cylindrical harmonics*. The cylindrical form of the harmonic expansion (2) is

$$\bar{S}(\phi,\omega) \simeq \sum_{m=-N}^{N} \mathring{S}_m(\omega) e^{im\phi}, \quad (6)$$

where $\mathring{S}_m(\omega)$ are the cylindrical expansion coefficients that are a function of $m = \{-N \dots N\}$ only, and the equality is again exact when $N = \infty$. A height-invariant sound field is reconstructed from its cylindrical harmonic coefficients

$$S(\phi',r,\omega) = \sum_{m=-N}^{N} i^m J_m\left(\frac{\omega}{c}r\right)\mathring{S}_m(\omega) e^{im\phi'}, \quad (7)$$

where $J_m(x)$ is a cylindrical Bessel function. In the special case where $\mathring{S}_m(\omega) = e^{-im\phi_s}$, (7) forms the Jacobi-Anger Expansion for a plane wave [13] with incidence azimuth θ_s.

The remainder of this paper will consider the case when a 3D sound field is to be represented with a 2D decomposition, such as when a sound field captured with a spherical microphone array is to be reproduced with a horizontal-only loudspeaker arrangement. Estimating a cylindrical harmonic representation from a spherical harmonic encoded field inherently produces sound fields that propagate parallel to the horizontal plane. In order to convert $\breve{S}_n^m(\omega)$ $\big((N+1)^2$ terms$\big)$ to $\mathring{S}_m(\omega)$ $(2N+1$ terms) it is necessary to discard some information. The aim is to define a method for performing such a conversion and to understand some of its properties.

3. CONVERSIONS BETWEEN REPRESENTATIONS

3.1. Method

We will begin by deriving the spherical coefficients in terms of the cylindrical coefficients. The cylindrical and spherical spatial Fourier transforms are defined respectively for a field propagating in the horizontal plane as

$$\mathring{S}_m(\theta_s,\omega) = \frac{1}{2\pi}\int_0^{2\pi} \bar{S}(\pi/2,\phi,\omega) e^{-im\phi} d\phi \quad (8)$$

$$\breve{S}_n^m(\omega) = \int_0^{2\pi}\int_0^{\pi} \bar{S}(\pi/2,\phi,\omega) Y_n^{-m}(\theta,\phi)\sin\theta d\theta d\phi. \quad (9)$$

The colatitude angle $\theta_s = \pi/2$ is the only angle for which the integration contours over azimuth angle in (8) and (9) are mutually compatible. While information is inevitably lost when deriving $\mathring{S}_m(\omega)$ from $\breve{S}_n^m(\omega)$, it is desirable to exploit this compatibility to ensure no loss of information for components propagating in the horizontal plane. The behaviour of the conversion for fields outside this plane is considered in Section 3.2.

The cylindrical decomposition in (8) can be formulated as a spherical decomposition by applying the sifting property of the Dirac delta function [14] and noting that $Y_n^{-m}(\theta_s,\phi) = Y_n^m(\theta_s,0)e^{-im\phi}$:

$$\breve{S}_n^m(\omega) = \int_0^{2\pi}\int_0^{\pi} \delta(\theta-\pi/2)\bar{S}(\pi/2,\phi,\omega) Y_n^{-m}(\theta,\phi)\sin\theta d\theta d\phi \quad (10)$$

$$= 2\pi \mathring{S}_m(\pi/2,\omega) Y_n^m(\pi/2,0). \quad (11)$$

This represents a spatial undersampling in θ as only the values of $\bar{S}(\theta,\phi,\omega)$ for $\theta = \pi/2$ are considered. Eq. (11) can be rearranged for $\mathring{S}_m(\pi/2,\omega)$:

$$\mathring{S}_m(\pi/2,\omega) = \frac{\breve{S}_n^m(\omega)}{2\pi Y_n^m(\pi/2,0)}. \quad (12)$$

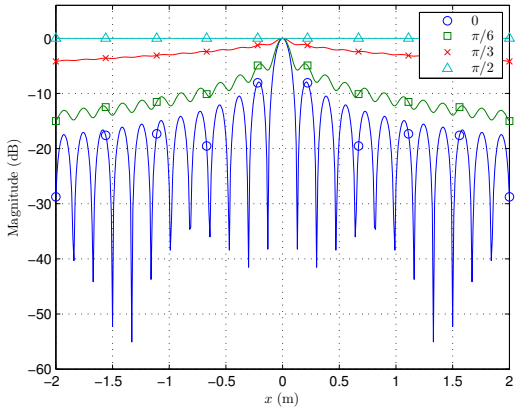

Fig. 2. Cross-section through Fig. 1 along the direction of propagation for differing incidence colatitude.

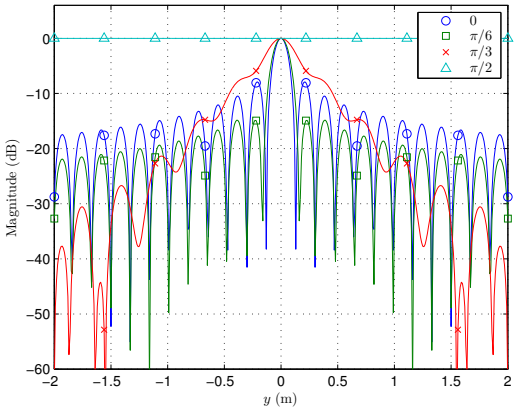

Fig. 3. Cross-section through Fig. 1 perpendicular to the direction of propagation for differing incidence colatitude.

Here an ambiguity arises as n only appears on the right hand side. In order to derive $\mathring{S}_m(\omega)$ from $\breve{S}_n^m(\omega)$, consider again the case of a plane wave arriving in the plane $\theta_\mathrm{s} = \pi/2$. There is a wide choice of coefficients in $\breve{S}_n^m(\omega)$ from which to derive $\mathring{S}_m(\omega)$ as (11) reveals that several terms differ only by a scaling factor $Y_n^m(\pi/2,0)/Y_{n'}^m(\pi/2,0)$. One possible solution is obtained (letting $\mathring{S}_m(\omega) \equiv \mathring{S}_m(\theta_\mathrm{s},\omega)$ for simplicity),

$$\mathring{S}_m(\omega) = \frac{\breve{S}_{|m|}^m(\omega)}{2\pi Y_{|m|}^m(\pi/2,0)}. \tag{13}$$

The motivation for this choice of coefficients is twofold: 1) the spherical harmonics $Y_{|m|}^m(\theta,\phi)$ are the *sectoral* harmonics [13], which have no zero crossings in colatitude, and 2) the terms $Y_{|m|}^m(\pi/2,0)$ are also necessarily nonzero for all values of m; conversely, all terms $Y_{|m|}^m(\theta,\phi)$ for which $n+|m|$ is odd are necessarily 0 for plane waves propagating in $\theta_\mathrm{s} = \pi/2$ [3]. Several other choices exist but are outside the scope of this paper.

3.2. Analysis

Generalizing to arbitrary angles of arrival $(\theta_\mathrm{s},\phi_\mathrm{s})$, the plane-wave spherical harmonic coefficients become

$$\breve{S}_n^m(\omega) = Y_n^{-m}(\theta_\mathrm{s},\phi_\mathrm{s}), \tag{14}$$

which, combining with (13), produces

$$\mathring{S}_m(\omega) = \frac{Y_{|m|}^{-m}(\theta_\mathrm{s},\phi_\mathrm{s})}{2\pi Y_{|m|}^m(\pi/2,0)}. \tag{15}$$

It is natural to ask how this relation varies as a function of arrival angle θ_s. Using the identity [13]

$$P_n^n(\cos\theta) = (-1)^n(2n-1)!!\sin^n\theta, \tag{16}$$

and recalling (3) and (15), it is straightforward to show that

$$\mathring{S}_m(\omega) = \frac{\sin^{|m|}(\theta_\mathrm{s})e^{-im\phi_\mathrm{s}}}{2\pi}. \tag{17}$$

Qualitatively, the orders are weighted with a sinusoidal envelope that is a function of the colatitude angle of arrival. Plane waves incident on $\theta_\mathrm{s} = \pi/2$ are unaffected. Plane waves incident in $\theta_\mathrm{s} = \{0,\pi\}$ only have a component in $m = 0$, thereby losing all angular dependence. Recalling (7), a standing wave is produced

$$S(\phi',r,\omega) = J_0\left(\frac{\omega}{c}r\right) \text{ when } \theta_\mathrm{s} = 0, \tag{18}$$

since the result is purely real.

4. EXPERIMENTATION

The cylindrical coefficients of a 1 kHz plane wave propagating with incidence angle $(\theta_\mathrm{s},0)$ were simulated with order $N = 60$ using (17) and evaluated with (7).

Fig. 1 shows the real part of the reconstructed field on the horizontal plane for incidence colatitude $\theta_\mathrm{s} = \{\pi/2, 3\pi/8, 0\}$. As θ_s moves away from the horizontal plane, the energy becomes concentrated in a corridor along the direction of propagation. As θ_s moves further towards the pole, the travelling plane wave components become progressively less pronounced until only a standing cylindrical wave remains at $\theta_\mathrm{s} = 0$ or $\theta_\mathrm{s} = \pi$.

Fig. 2 shows the magnitude of the field in Fig. 1 through a cross-section along the direction of incidence, revealing that the magnitude of the field is a) unity when $\theta_\mathrm{s} = \pi/2$ and, b) a 0th order Bessel function when $\theta_\mathrm{s} = 0$ as predicted in (18).

Fig. 3 shows the magnitude of the field in Fig. 1 through a cross-section perpendicular to the direction of incidence. This shows similar behaviour for sources on the horizontal plane and at the pole, but with more rapid decrease in magnitude as the source moves away from the horizontal plane. Similar phenomena are seen with the order-limited reconstruction of plane waves [15]; here there is a soft limitation caused by the coefficient extraction (13) that leads to a sinusoidal weighting (17), which may be interpreted as spatial aliasing in this context due to spatial undersampling in colatitude. The consequence of this order limitation is the aforementioned concentration of energy.

Further insight into the asymptotic behaviour between unit magnitude and the 0th order Bessel function is given by evaluating the ratio

$$\zeta(\theta_\mathrm{s}) = 10\log_{10}\left(\frac{|\mathring{S}_0(\omega|\theta_\mathrm{s})|^2}{\sum_{m\neq 0}|\mathring{S}_m(\omega|\theta_\mathrm{s})|^2}\right). \tag{19}$$

This gives a quantitative measure of the energy in the 0th order component (the standing wave component) and the energy in the

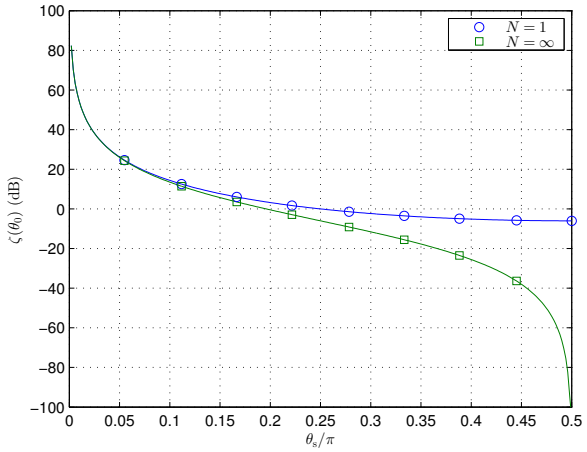

Fig. 4. Log ratio of the 0th order component to the remaining components (19) as a function of source colatitude. As θ_s tends to one of the poles, the energy becomes increasingly concentrated in the 0th order component.

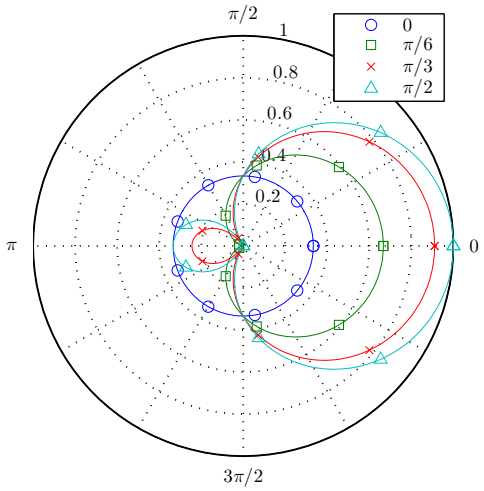

Fig. 5. Azimuth-only phase mode beamformer directivity patterns (20) as a function of source incidence colatitude. As θ_s tends to one of the poles, the response tends to omnidirectional.

non-zeroth order components. Fig. 4 shows $\zeta(\theta_s)$ for $N = 1$ and $N = \infty$. In the latter case, this measure necessarily asymptotes to ∞ as θ_s tends to one of the poles and to $-\infty$ as $\theta_s \to \pi/2$ (curves for all other values of N lie between those for $N = 1$ and $N = \infty$). Qualitatively, as the source moves towards the pole the order of the field is effectively limited, leading to the phenomena depicted in Figs 1–3.

4.1. Beamforming Example

An additional practical use of the spherical to cylindrical conversion is in the field of modal beamforming [4]. The output of a cylindrical *phase mode* beamformer is given by

$$y(\phi) = \frac{1}{2N+1} \sum_{m=-N}^{N} \mathring{S}_m e^{im\phi}. \qquad (20)$$

Note that spectral independence is assumed and recall that $\mathring{S}_m$ are plane wave cylindrical coefficients. The 1st order ($N = 1$) cylindrical coefficients of a 1 kHz plane wave propagating with incidence angle $(\theta_s, 0)$ were simulated. The beamformer directivity pattern $|y(\phi)|$ is shown in Fig. 5. When $\theta_s = \pi/2$, the result is a classic 1st order hypercardioid. As θ_s moves away from the horizontal plane the response becomes increasingly omnidirectional as the magnitudes of the components corresponding to $m \neq 0$ are reduced. With the exception of a scaling factor, this process is identical to the synthesis of the 1st order directivity patters (cardioid, hypercardioid etc.) [16]. As $N \to \infty$, the beamformer response tends to 0 for all but $\theta_s = \pi/2$ due to the normalization factor $1/(N+1)$.

5. CONCLUSIONS

The conversion between cylindrical and spherical basis expansions of 3D sound fields has been investigated. Converting from a spherical representation to a cylindrical representation inevitably involves a loss of information if the direction of arrival is outside the horizontal plane. An analysis reveals that the chosen approach leads to sinusoidally weighting the components as a function of the colatitude angle of incidence. This effectively produces an order-limited sound field and is manifest as increased spatial concentration of energy and the reduced energy of propagating components with increasing incidence angle from the horizontal plane.

6. REFERENCES

[1] J. Daniel, J-B. Rault, and J-D. Polack, "Ambisonics encoding of other audio frormats for multiple listening conditions," in *105th AES Convention*, Sept. 1998.

[2] J. Ahrens and S. Spors, "An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions," *Acta Acustica utd. with Acustica*, vol. 94, no. 6, pp. 988–999, Nov. 2008.

[3] J. Ahrens and S. Spors, "Wave field synthesis of a sound field described by sphercial harmonics expansion coefficients," *J. Acoust. Soc. Am.*, vol. 3, no. 3, pp. 2190–2199, Mar. 2012.

[4] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 1, pp. 135–143, Jan. 2005.

[5] H. Teutsch, *Modal Array Signal Processing: Principles and Applications of Acoustic Wavefield Decomposition*, Springer, 2007.

[6] D. N. Zotkin, R. Duraiswami, and N. A. Gumerov, "Plane-wave decomposition of acoustical scenes via spherical and cylindrical microphone arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 1, pp. 2–16, 2010.

[7] T. D. Abhayapala and A. Gupta, "Alternative to spherical microphone arrays: Hybrid geometries," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 81–84.

[8] E. Tiana-Roig, F. Jacobsen, and E. Fernandez-Grande, "Beamforming with a circular array of microphones mounted on a rigid sphere," *J. Acoust. Soc. Am.*, vol. 130, no. 3, pp. 1095–1098, Sept. 2011.

[9] A. Parthy, C. Jin, and A. van Shaik, "Measured and theoretical performance comparsion of a broadband circular microphone array," in *Proc. Audio Eng. Soc. Conventions*, London, England, June 2007, pp. 25–27.

[10] J. Meyer, "Beamforming for a circular microphone array mounted on spherically shaped objects," *J. Acoust. Soc. Am.*, vol. 109, no. 1, pp. 185–193, Jan. 2001.

[11] E. Hulsebos, *Auralization Using Wave Field Synthesis*, Ph.D. thesis, TU Delft, 2004.

[12] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London, first edition, 1999.

[13] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*, Elsevier, 2004.

[14] B. Girod, R. Rabenstein, and A. Stenger, Eds., *Signals and Systems*, Wiley, 2001.

[15] J. Ahrens, *Analytic Methods of Sound Field Synthesis*, Springer, Berlin, Heidelberg, 2012.

[16] I. Tashev, *Sound Capture and Processing: Practical Approaches*, Wiley, 2009.