

ANALYTICAL DRIVING FUNCTIONS FOR HIGHER ORDER AMBISONICS

Jens Ahrens, Sascha Spors

Deutsche Telekom Laboratories, Berlin University of Technology,
Ernst-Reuter-Platz 7, 10587 Berlin, Germany
{jens.ahrens, sascha.spors}@telekom.de

ABSTRACT

In this paper, we present the derivation and investigation of analytical expressions for the loudspeaker driving signals for higher order Ambisonics. The approach relies on the assumption of a continuous distribution of secondary sources on which sampling is performed to yield the actual loudspeaker signals for real-world implementations. For source-free volumes enclosed by the secondary source distribution, this formulation of Ambisonics leads to what is known as simple source approach. Since the simple source approach is theoretically well documented, we will depart from it and concentrate on the special case of a circular distribution of secondary point sources and derive analytical expressions for the driving signals. We furthermore derive a closed-form expression for the actual reproduced wave field for the circular secondary source distribution.

Index Terms— Spatial audio, Ambisonics, simple source approach, spherical harmonics, audio reproduction in a plane

1. INTRODUCTION

Higher order Ambisonics (HOA) is a sound reproduction technique that utilizes a large number of loudspeakers to physically recreate a wave field in a specific listening area. Since the so-called near-field correction has been introduced in [1] accurate versatile reproduction is possible. The desired wave field is typically described via its spatial harmonics expansion coefficients [2]. These can be yielded either from appropriate microphone recording techniques which utilize a Fourier series representation of the recorded signals [3] (data-based rendering) or virtual sound scenes may be composed of virtual sound sources whose spatial harmonics expansion coefficients are derived from analytical source models (model-based rendering). In this paper we will concentrate on the latter case.

The typical Ambisonics approach is based on the assumption of a finite number of discrete loudspeakers whose emitted wave fields superpose to an approximation of the desired one. The concept of modeling a continuous loudspeaker distribution has been briefly presented by the authors in [4] for the investigation of spatial aliasing in two-dimensional higher order Ambisonics and will be further extended in this paper. The theoretical basis is the so-called simple source approach [5] which has gained only little attention in conjunction with spatial audio reproduction. We will limit our derivations to the rendering of virtual plane waves since a suitable superposition of plane waves can be used to represent arbitrary wave fields.

1.1. Nomenclature

Our Ambisonics approach implicitly includes the near-field correction [1]. When we speak of Ambisonics in the remainder of this paper, we thus implicitly mean what is typically referred to as near-field corrected higher order Ambisonics (NFC-HOA). In the continuous case we will not refer to loudspeakers but rather to secondary

sources, and also to secondary source driving functions rather than to loudspeaker signals.

The following notational conventions are used: For scalar variables lower case denotes the time domain, upper case the temporal frequency domain. Vectors are denoted by lower case boldface. The three-dimensional position vector in Cartesian coordinates is given as $\mathbf{x} = [x \ y \ z]^T$. The Cartesian coordinates are linked to the spherical coordinates via $x = r \cos \alpha \sin \beta$, $y = r \sin \alpha \sin \beta$, and $z = r \cos \beta$. α denotes the azimuthal angle, β the elevation. Confer also to figure 1. The acoustic wavenumber is denoted by k . It is related to the temporal frequency by $k = |\frac{\omega}{c}|$ with ω being the radial frequency and c the speed of sound.

Outgoing monochromatic plane and spherical waves are denoted by $e^{-i\mathbf{k}_{pw}^T \mathbf{x}}$ and e^{-ikr} respectively, with $\mathbf{k}_{pw}^T = [\cos \theta_{pw} \sin \phi_{pw} \sin \theta_{pw} \sin \phi_{pw} \cos \phi_{pw}]$ and (θ_{pw}, ϕ_{pw}) being the propagation direction of the plane wave. i is the imaginary unit ($i = \sqrt{-1}$).

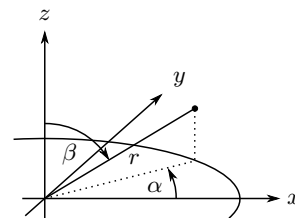


Fig. 1. The (spatial) coordinate system used in this paper. In the wave number domain k corresponds to r , the angle θ corresponds to α , and ϕ corresponds to β .

1.2. Mathematical preliminaries

This section provides a summary and reference of the mathematical tools employed in the remainder of this paper. We refer the reader to [5, 6] for a more thorough treatment.

One of the basic tools we employ is the spherical harmonics expansion

$$A(\alpha, \beta, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \hat{A}_n^m(\omega) Y_n^m(\alpha, \beta), \quad (1)$$

whereby $\hat{A}_n^m(\omega)$ denote the spherical harmonics expansion coefficients of the function $A(\alpha, \beta, \omega)$. The spherical harmonics $Y_n^m(\alpha, \beta)$ are defined as

$$Y_n^m(\alpha, \beta) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} \cdot P_n^m(\cos \beta) \cdot e^{im\alpha}, \quad (2)$$

with $P_n^m(\cdot)$ denoting the m -th order associated Legendre polynomial of n -th degree. Convolution on the surface of a sphere is de-

defined as

$$A(\alpha) = B(\alpha) *_{\text{sph}} C(\alpha) = \int_0^{2\pi} \int_0^\pi B(\alpha_0) C(\alpha - \alpha_0) \sin \beta_0 d\beta_0 d\alpha_0. \quad (3)$$

For a spherical convolution like equation (3) applies

$$\hat{A}_n^m = 2\pi \sqrt{\frac{4\pi}{2n+1}} \hat{B}_n^m \cdot \hat{C}_n^0. \quad (4)$$

In two dimensions, the analog to (1) is the Fourier series expansion

$$A(\alpha, \omega) = \sum_{m=-\infty}^{\infty} \hat{A}_m(\omega) e^{im\alpha}. \quad (5)$$

Circular convolution is defined as [7]

$$A(\alpha) = B(\alpha) *_{\text{circ}} C(\alpha) = \int_0^{2\pi} B(\alpha_0) C(\alpha - \alpha_0) d\alpha_0. \quad (6)$$

In that case applies

$$\hat{A}_m = 2\pi \hat{B}_m \cdot \hat{C}_m. \quad (7)$$

2. THE AMBISONICS APPROACH

In the basic three-dimensional Ambisonics approach, the loudspeakers of the respective reproduction system are located on a sphere around the listening area. Both the desired wave field as well as the sound fields emitted by the loudspeakers are expanded into series of orthogonal basis functions [8, 2]. This results in an equation system that is solved for the optimal loudspeaker driving signals. These drive the loudspeakers such that their superposed wave fields best approximate the desired one in a given sense:

$$P(\mathbf{x}, \omega) = \sum_{n=0}^{N-1} D(\mathbf{x}_n, r_0, \omega) \cdot G(\mathbf{x}, \mathbf{x}_n, \omega), \quad (8)$$

where $P(\mathbf{x}, \omega)$ denotes the desired wave field, $D(\mathbf{x}_n, r_0, \omega)$ the driving signal of the loudspeaker located at the position $\mathbf{x}_n = r_0 \cdot [\cos \alpha_n \sin \beta_n \sin \alpha_n \sin \beta_n \cos \beta_n]^T$, and $G(\mathbf{x}, \mathbf{x}_n, \omega)$ its spatio-temporal transfer function. Typically, numerical algorithms are employed to find the appropriate loudspeaker driving signals. These algorithms tend to be computationally costly and only little insight into the properties of the actual reproduced wave-field is gained. The Ambisonics approach is usually divided into an encoding and a decoding stage to allow for storing and transmission of content independently from the loudspeaker setup. For ease of illustration we will skip the encoding/decoding procedure and directly derive the loudspeaker driving signals from the initial virtual wave field description. Note that the extension of our approach for the encoding and decoding of wave fields is straightforward. However, if only the encoding procedure is considered, the same stability issues as in [1] arise for which workarounds are given ibidem.

3. FROM AMBISONICS TO THE SIMPLE SOURCE APPROACH

3.1. General outline

The formulation of the basic Ambisonics equation (8) for a continuous secondary source distribution on a sphere whose center resides in the coordinate origin reads

$$P(\mathbf{x}, \omega) = \int_{\Omega_0 \in S_R^2} D(\Omega_0, \omega) \cdot G(\mathbf{x}, \Omega_0, \omega) d\Omega_0, \quad (9)$$

whereby Ω_0 denotes the surface of the sphere with radius r_0 on which the secondary sources are located. The explicit integration operation is $\int_{\Omega_0} d\Omega_0 = \int_0^{2\pi} \int_0^\pi r_0 \sin \beta_0 d\beta_0 d\alpha_0$. When the spatial transfer function of the loudspeakers is modeled as a spherical wave with flat temporal frequency response (thus making $G(\mathbf{x}, \Omega_0, \omega)$ the free-field Green's function) equation (9) essentially constitutes the simple source approach for an interior problem in a spherical volume. The simple source approach for interior problems states that the acoustic field generated by events outside a volume can also be uniquely generated by a continuous distribution of monopole sources replacing these events and enclosing the respective volume [5]. The proof that an arbitrary source-free wave field can be recreated inside the sphere ($|\mathbf{x}| < r_0$) according to (9) is given in [3]. Note that the simple source approach does not pose any restrictions on the wave field in locations outside the sphere ($|\mathbf{x}| > r_0$).

3.2. Derivation of the driving function for a virtual plane wave

In this section we illustrate how wave field reproduction according to (9) and thus according to the simple source approach can be accomplished. We outline the procedure that yields the secondary source driving function for rendering a virtual plane wave. Note that the driving functions for any other kind of virtual source (point sources, complex sources, etc.) can be derived accordingly. Equation (9) can be interpreted as a convolution along the surface of a sphere. From equation (4) we can thus deduct that

$$\hat{D}_n^m(\omega) = \frac{1}{2\pi} \sqrt{\frac{2n+1}{4\pi}} \frac{\hat{P}_n^m(r, \omega)}{r_0 \cdot \hat{G}_n^0(r, \omega)}. \quad (10)$$

The wave field of a plane wave with propagating direction $(\theta_{\text{pw}}, \phi_{\text{pw}})$ can be expanded into its radial and angular dependencies around the origin of the coordinate system as [9]

$$P(\mathbf{x}, \omega) = \hat{S}(\omega) \cdot e^{-i\mathbf{k}_{\text{pw}}^T \mathbf{x}} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{\hat{S}(\omega) \cdot 4\pi (-i)^n j_n(kr) Y_n^m(\theta_{\text{pw}}, \phi_{\text{pw}})^*}_{\hat{P}_n^m(r, \omega)} Y_n^m(\alpha, \beta), \quad (11)$$

whereby $\hat{S}(\omega)$ denotes the temporal spectrum of the plane wave and $j_n(\cdot)$ the n -th order spherical bessel function. Modeling the spatial transfer function of a loudspeaker at position $\mathbf{x}_0 \in \Omega_0$ as the free-field Green's function and expanding it into its radial and angular dependencies leads to [9]

$$G(\mathbf{x}, \mathbf{x}_0, \omega) = \frac{1}{4\pi} \frac{e^{-ik|\mathbf{x}-\mathbf{x}_0|}}{|\mathbf{x}-\mathbf{x}_0|} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{(-ik) j_n(kr) h_n^{(2)}(kr_0) Y_n^m(\alpha_0, \beta_0)^*}_{\hat{G}_n^m(r, \omega)} Y_n^m(\alpha, \beta), \quad (12)$$

with $h_n^{(2)}(\cdot)$ being the n -th order spherical hankel function of second kind. Note that (12) is only valid for $r \leq r_0$.

Combining (1), (10), (11), and (12) yields

$$D_{\text{pw}}(\alpha_0, \beta_0, r_0, \omega) = \hat{S}(\omega) \times 4\pi \frac{i}{kr_0} \sum_{n=0}^{\infty} \sum_{m=-n}^n \frac{(-i)^n}{h_n^{(2)}(kr_0)} Y_n^m(\theta_{\text{pw}}, \phi_{\text{pw}})^* Y_n^m(\alpha_0, \beta_0). \quad (13)$$

Equation (13) can be verified by inserting it into (9). After interchanging the order of summation and integration and exploiting the orthogonality of the spherical harmonics one arrives at (11) which completes the proof.

4. REPRODUCTION IN A PLANE

Ambisonics systems are frequently restricted to reproduction in the horizontal plane. The secondary sources are arranged on a circle. In this case, the propagation directions of the virtual plane waves as well as the listening positions are bounded to the horizontal plane. For this two-dimensional setup the free-field Green's function required by the simple source approach can be interpreted as the spatial transfer function of a line source [5]. Implementations of Ambisonics systems usually employ loudspeakers with closed cabinets whose spatial transfer function can be approximated by that of a point source. This secondary source mismatch prevents us from perfectly recreating any source-free wave field inside the secondary source array. We have to expect artefacts. This circumstance is also a well treated problem in wave field synthesis [10].

4.1. Derivation of the driving function for a virtual plane wave

For a planar circular distribution of secondary point sources equation (9) degenerates to

$$P(\mathbf{x}, \omega) = \int_0^{2\pi} D(\alpha_0, \omega) \cdot G(\mathbf{x}, \mathbf{x}_0, \omega) r_0 d\alpha_0. \quad (14)$$

To bound our area of interest to the horizontal plane we set the elevation angle β in all position vectors to $\frac{\pi}{2}$. For circular geometries the Fourier series expansion (cf. to (5)) constitutes a useful mathematical tool. We therefore adopt the expansion of the loudspeaker wave fields in (12) by exchanging the order of summations to arrive at such a Fourier series expansion:

$$\begin{aligned} G(\mathbf{x}, \mathbf{x}_0, \omega) &= -ik \sum_{m=-\infty}^{\infty} \sum_{n=|m|}^{\infty} j_n(kr) h_n^{(2)}(kr_0) \times \\ &\quad \times Y_n^m(\alpha, \frac{\pi}{2}) Y_n^m(\alpha_0, \frac{\pi}{2})^* = \sum_{m=-\infty}^{\infty} e^{im(\alpha-\alpha_0)} \times \\ &\quad \times \underbrace{\sum_{n=|m|}^{\infty} (-ik) j_n(kr) h_n^{(2)}(kr_0) \frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!} P_n^{m2}(0)}_{\Lambda_m(r, r_0, \omega)}. \end{aligned} \quad (15)$$

The according Fourier expansion of the plane wave field yields

$$\begin{aligned} P(\mathbf{x}, \omega) &= \hat{S}(\omega) \cdot \sum_{m=-\infty}^{\infty} \sum_{n=|m|}^{\infty} 4\pi (-i)^n j_n(kr) \times \\ &\quad \times Y_n^m(\alpha, \frac{\pi}{2}) Y_n^m(\theta_{pw}, \frac{\pi}{2})^* = \hat{S}(\omega) \cdot \sum_{m=-\infty}^{\infty} e^{im(\alpha-\theta_{pw})} \times \\ &\quad \times \underbrace{\sum_{n=|m|}^{\infty} 4\pi (-i)^n j_n(kr) \frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!} (P_n^m(0))^2}_{\Psi_m(r, \omega)}. \end{aligned} \quad (16)$$

Interpreting (14) as circular convolution and following the procedure outlined in section 3.2 leads to

$$D_{pw}^{2D}(\alpha, r_0, \omega) = \hat{S}(\omega) \cdot \sum_{m=-\infty}^{\infty} \frac{\Psi_m(r, \omega)}{2\pi r_0 \Lambda_m(r, r_0, \omega)} e^{im(\alpha_0 - \theta_{pw})} \quad (17)$$

for the driving function for a plane wave with propagation direction θ_{pw} . The radius r appears in the expression for the driving function suggesting that (14) can only be satisfied for a single listening position. This finding has already been derived in [11]. We thus have to reference the reproduced wave field to a point which is then the only location where the reproduction is correct. Referencing the reproduction to the center of the circular secondary source array is most favorable since this assures equal properties for all propagation directions of the virtual plane wave.

Setting $r = 0$ in (17) leads to an undefined expression of the form $\frac{0}{0}$ for $n \neq 0$ since spherical Bessel functions of argument 0 equal 0 $\forall n \neq 0$. We therefore apply de l'Hospital's rule to yield a defined expression for $r = 0$ reading

$$D_{pw}^{2D}(\alpha_0, r_0, \omega) = \hat{S}(\omega) \cdot \sum_{m=-\infty}^{\infty} \frac{2i}{kr_0} \frac{(-i)^{|m|}}{h_{|m|}^{(2)}(kr_0)} e^{im(\alpha_0 - \theta_{pw})}. \quad (18)$$

4.2. Reproduced wave field

We yield the actual reproduced wave field by inserting (18) in (14) as

$$\begin{aligned} P_{pw}(\mathbf{x}, \omega) &= 2 \cdot \hat{S}(\omega) \sum_{n=0}^{\infty} j_n(kr) h_n^{(2)}(kr_0) \times \\ &\quad \times \sum_{m=-n}^n \frac{(-i)^{|m|}}{h_{|m|}^{(2)}(kr_0)} Y_n^m(\alpha, \frac{\pi}{2}) Y_n^m(\theta_{pw}, \frac{\pi}{2})^*. \end{aligned} \quad (19)$$

The real part and the absolute value of $P_{pw}(\mathbf{x}, \omega)$ are depicted in figures 2(a) and 2(b) for a virtual plane wave of $f = 1000$ Hz and unit amplitude ($\hat{S}(\omega) = 1$) with propagation direction $\theta_{pw} = \frac{3\pi}{2}$ rendered by a continuous circular secondary source distribution with $r_0 = 1.5$ m. The angular bandwidth was limited to $n_{\max} = 40$ for the simulation. Both figures only display the wave field inside the secondary source distribution since (19) holds there only.

From figure 2(a) it can be seen that the wave fronts of $P_{pw}(\mathbf{x}, \omega)$ are indeed perfectly plane. Though, an amplitude decay of approximately 3dB per doubling of the distance is apparent when following the propagation path of the plane wave. Figure 2(b) further illustrates this amplitude decay by depicting the absolute value of the sound pressure in logarithmic scale.

4.3. Implementation

Real-world implementations of audio reproduction systems will always employ a limited number of discrete secondary sources. This discretization constitutes spatial sampling and thus potentially produces spatial aliasing. The treatment of these artefacts is beyond the scope of this paper. We refer the reader to [10] and [4]. In this section, we comment on proper handling of the expressions for the driving signals for a finite number of discrete loudspeakers.

Analysis of the dimensionality of the reproduced field reveals that the spatial bandwidth of the loudspeaker driving function can be spatially truncated [12]. The most suitable choice is to use as many

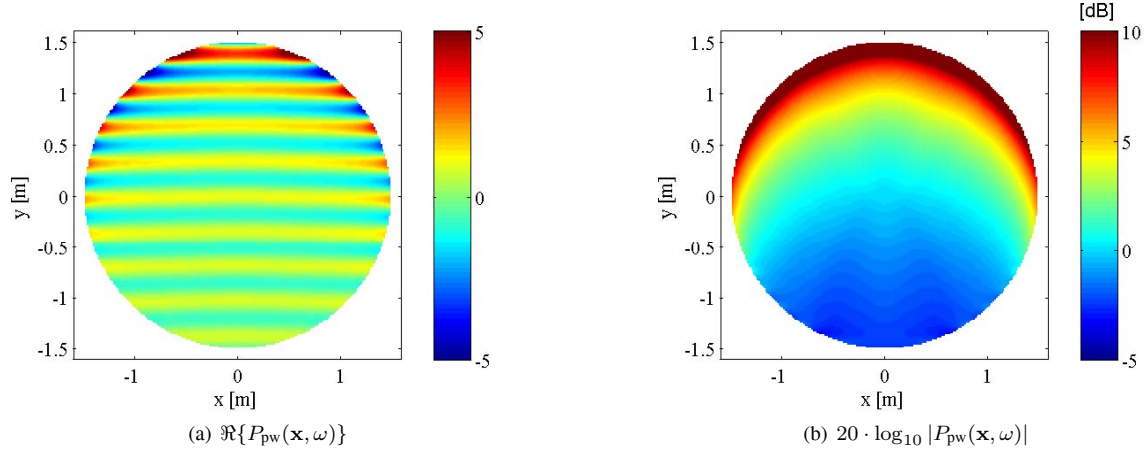


Fig. 2. Sound pressure $P_{pw}(\mathbf{x}, \omega)$ of a continuous circular distribution with radius $r_0 = 1.5$ m of secondary monopole sources rendering a virtual plane wave of $f = 1000$ Hz and unit amplitude with propagation direction $\theta_{pw} = \frac{3\pi}{2}$ referenced to the coordinate origin.

orders as secondary sources, thus

$$D_{pw}^{2D}(\alpha_0, r_0, \omega) = \hat{S}(\omega) \cdot \sum_{m=-\frac{L-1}{2}}^{\frac{L-1}{2}} \frac{2i}{kr_0} \frac{(-i)^{|m|}}{h_{|m|}^{(2)}(kr_0)} e^{im(\alpha_0 - \theta_{pw})}, \quad (20)$$

with L being an odd number of loudspeakers. For even L the limits of the summation in (20) have to be set accordingly.

The reproduced wave field of such a circular equiangular distribution of 56 secondary monopole source rendering the same virtual plane wave as in figure 2 is depicted in figure 3. The driving function is spatially limited according to (20).

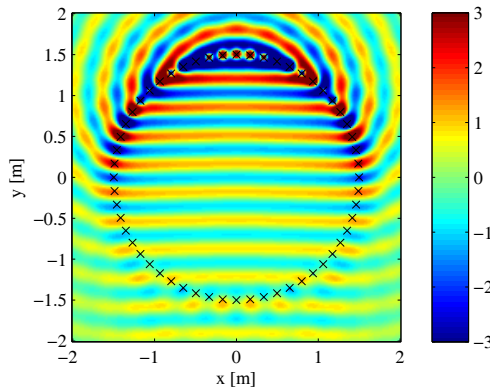


Fig. 3. $\Re\{P_{pw,s}(\mathbf{x}, \omega)\}$ of a circular distribution with radius $r_0 = 1.5$ m of 56 monopole point sources rendering the virtual plane wave from figure 2. The marks indicate the secondary source positions.

5. CONCLUSIONS

In this paper, a formulation of the Ambisonics approach assuming a continuous distribution of secondary monopole sources was presented. In the case of a volume enclosed by the secondary source distribution, this formulation lead directly to the simple source approach, thus providing the physical justification for Ambisonics to recreate arbitrary source-free wave fields. We furthermore presented analytical expressions for the secondary source driving signals of a spherical distribution when rendering a virtual plane wave. We then concentrated on the special case of a circular distribution of secondary point sources and investigated its properties when rendering a virtual plane wave. It turned out that the departure from

the physical fundament of the simple source approach indeed introduces artefacts, i.e. an amplitude decay of approximately 3dB per doubling of the distance in direction of the propagation of the virtual plane wave. Finally, we commented on proper handling of the expressions for the driving functions in real-world implementations. Contrary to the traditional Ambisonics approach, our approach provides the capability to analytically derive the actual reproduced wave field. This facilitates the investigation of the consequences of an insufficient loudspeaker layout which occurs in real world implementations (spatial sampling, incomplete spherical arrangements). Furthermore, no inverse matrix computations or solving of equation systems (as e.g. in [3, 11]) are necessary. The advantages in terms of computational complexity still have to be investigated. Finally, there are no stability issues for incomplete loudspeaker setups as in the traditional approach.

6. REFERENCES

- [1] J. Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format," in *23rd Int. Conf., May 23–25, Copenhagen, Denmark, 2003*, Audio Engineering Society (AES).
- [2] J. Daniel, "Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia," PhD thesis Université Paris 6, 2001.
- [3] M.A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *Journal of the Audio Engineering Society (AES)*, vol. 53, no. 11, pp. 1004–1025, Nov. 2005.
- [4] J. Ahrens and S. Spors, "Rendering of virtual sound sources with arbitrary directivity in higher order ambisonics," *123rd AES Conv., NY, NY, USA, 5–8 Oct., 2007*.
- [5] E.G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography*, Academic Press, London, 1999.
- [6] J.R. "Driscoll and D.M." Healy, "Computing fourier transforms and convolutions on the 2-sphere," *Advances in Applied Mathematics*, vol. 15, pp. 202–250, 1994.
- [7] B. Girod, R. Rabenstein, and A. Stenger, *Signals and Systems*, J.Wiley & Sons, 2001.
- [8] M.A. Gerzon, "With-height sound reproduction," *Journal of the Audio Engineering Society (JAES)*, vol. 21, pp. 2–10, 1973.
- [9] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer, Berlin, 1998.
- [10] S. Spors and R. Rabenstein, "Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis," *120th Convention of the AES, Paris, France, May 2006*.
- [11] D.B. Ward and T.D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans. on Sp. and Audio Proc.*, vol. 9(6), Sep. 2001.
- [12] R.A. Kennedy, P. Sadeghi, T.D. Abhayapala, and H.M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, June 2007.