FUNCTIONAL ANALYSIS GUIDED APPROACH FOR SOUND FIELD REPRODUCTION WITH FLEXIBLE LOUDSPEAKER LAYOUTS

Wen Zhang and Thushara D. Abhayapala[*]

The Australian National University
Research School of Engineering
Canberra, 2601, Australia
{wen.zhang, thushara.abhayapala}@anu.edu.au

Filippo M. Fazi

University of Southampton
Institute of Sound and Vibration Research
Southampton, SO17 1BJ, UK
ff1@isvr.soton.ac.uk

ABSTRACT

This paper proposes a design of multiple circular and partial circular loudspeaker arrays for reproducing sound fields originated from a limited spatial region. We apply a functional analysis framework to formulate the sound field reproduction problem in closed form. Analytical solutions are derived for a circular secondary source arrangement, from which circular arc layouts are investigated and the design of placing multiple loudspeaker arrays over the limited region of interest is proposed. Such a design allows for non-spherical and non-uniform loudspeaker placement and thus provides flexibility to suit reproduction in real audio environments. The reproduction using the proposed method are illustrated by numerical simulations in comparison with the Least-squares based schemes.

***Index Terms*—** Circular loudspeaker arrays, partial layouts, sound field reproduction, 3D audio

1. INTRODUCTION

Ambisonics is based on representation of a sound field as a sum of angular modes (cylindrical/spherical harmonics) and by mode-matching equations to derive loudspeaker driving signals for a natural reproduction of sound over a large listening area [1]. Even though a few 3D reproduction systems were implemented in higher order Ambisonics [2, 3], there is a critical problem in terms of loudspeaker placement. To perfectly reconstruct incident sound fields, the mode-matching approach requires the placement of loudspeakers on a sphere that surrounds the target reproduction region [1, 4]. The deployment of a spherical loudspeaker array is impractical in reality. Non-spherical loudspeaker arrays have been deployed, such as a partial-sphere or a hemisphere [5] and multiple circular arrays [6], based on the Least-squares (LS) approach to match pressures recorded at multiple points in the target region and decomposed by the spherical harmonics. The problem in most cases is ill-posed and Tikhonov regularization is the common method for obtaining loudspeaker weights with limited energy [7]. Using the compressive sensing idea, the LS formulation of the sound field reproduction problem is regularized with the ℓ_1 norm and solved using the Least-absolute shrinkage and selection operator (Lasso) [8]. The assumption here is that the desired sound field can be reproduced by a few loudspeakers, which are placed close to the direction of incidence and sparsely distributed in space.

The functional analysis framework was firstly introduced into sound field reproduction by Fazi [2, 4]. A continuous distribution of secondary sources and spatial sound fields are interrelated by an integral operator, from which a self-adjoint operator is constructed and singular value decomposition can be applied to modal decompose the source distributions and sound fields with two sets of orthonormal functions (called as modes or modal basis functions). The construction of source distributions allows for a perfectly accurate reproduction if the full sets of source modes and sound field modes are included. As expected, the modal basis functions for source distributions and sound fields arranged on two concentric circles and spheres are cylindrical/spherical harmonics, respectively.

This paper presents the design of a flexible scheme of placing multiple circular or partial circular loudspeaker arrays for sound field reproduction. We apply the functional analysis framework to analytically derive modal basis functions for each individual array at a particular colatitude and to determine the loudspeaker driving signals from the spherical spectrum of the desired sound field. The flexible geometry would suit sound field reproduction in real rooms and auditoria; in addition, by analyzing the reproduction efficiency of each array, we can design systems to focus on sound fields originated from a limited region of interest, i.e., only loudspeakers close to the virtual source positions are active—a feature imposed as constraints and achieved implicitly in the LS-based schemes.

2. PRELIMINARIES AND PROBLEM STATEMENT

Consider a continuous secondary source distribution on the surface Λ_0 that fully encloses a bounded region Λ, the generated sound field $S(\boldsymbol{x})$ can be represented with an integral operator,

$$S(\boldsymbol{x}) = (\mathcal{A}\rho)(\boldsymbol{x}) = \int_{\Lambda_0} \rho(\boldsymbol{y}) G(\boldsymbol{x}|\boldsymbol{y}) d\boldsymbol{y},\ \boldsymbol{y} \in \Lambda_0, \tag{1}$$

where $\boldsymbol{x}$ is the observation point and $\boldsymbol{y}$ denotes a secondary source position. The control region considered here are the surface enclosing the entire reproduction region V[1], i.e., $\boldsymbol{x} \in V_0$. The reproduction region V is contained within the region Λ. $G(\boldsymbol{x}|\boldsymbol{y})$ is a Green's function that represents a spatial-temporal transfer function from $\boldsymbol{y}$ to $\boldsymbol{x}$; in free-field conditions for an omnidirectional point source it is defined as [9],

$$G(\boldsymbol{x}|\boldsymbol{y}) = \frac{e^{-ik|\boldsymbol{y}-\boldsymbol{x}|}}{4\pi|\boldsymbol{y}-\boldsymbol{x}|}, \tag{2}$$

[*]This work was supported under the Australian Research Councils Discovery Projects funding scheme (project no. DP110103369).

[1]When the wave number is not one of the Dirichlet eigenvalues of the reproduction region, reproducing a sound field on the surface enclosing the region of interest ensures the exact sound field reproduction within the region [9, 10].

where $k = 2\pi f/c$ is the wave number with f the frequency and c the speed of wave propagation. Note that here only a single frequency f is considered and for simplicity the explicit dependence on k has been suppressed in the notation.

Applying the Hilbert-space mathematical concepts [11], we can formulate the space of secondary source distributions, $\mathcal{D} \triangleq \{\rho(\boldsymbol{y}) : \|\rho\|_{\mathcal{D}} < \infty, \boldsymbol{y} \in \Lambda_0\}$ and the space of spatial sound fields, $\mathcal{M} \triangleq \{S(\boldsymbol{x}) : \|S\|_{\mathcal{M}} < \infty, \boldsymbol{x} \in V_0\}$. The operator $\mathcal{A} : \mathcal{D} \to \mathcal{M}$ defines the mapping from a secondary source distribution to a spatial sound field. An adjoint operator is defined to represent the inverse mapping $\mathcal{A}^* : \mathcal{M} \to \mathcal{D}$,

$$(\mathcal{A}^* S)(\boldsymbol{y}) = \int_{V_0} S(\boldsymbol{x}) G^*(\boldsymbol{y}|\boldsymbol{x}) d\boldsymbol{x},\ \boldsymbol{x} \in V_0, \tag{3}$$

where $G^*(\boldsymbol{y}|\boldsymbol{x})$ denotes the complex conjugate of (2). The physical meaning of the adjoint operator $(\mathcal{A}^* S)(\boldsymbol{y}; k)$ is to treat the observation points as monopole-like sources on V_0 and to represent the sound field generated by these sources on the boundary of the secondary source distributions Λ_0 [2].

Now consider the operator $\mathcal{A}^*\mathcal{A} : \mathcal{D} \to \mathcal{D}$. Given $(\mathcal{A}^*\mathcal{A})^* = \mathcal{A}^*\mathcal{A}$ and $\mathcal{A}$ is compact, $\mathcal{A}^*\mathcal{A}$ is a self-adjoint compact operator [11]. Singular valued decomposition can be applied to modal decompose the source distributions and sound fields with two sets of orthonormal functions. That is every element $\rho \in \mathcal{D}$ has a unique representation in the form,

$$\rho(\boldsymbol{y}) = \sum_{\ell=1}^{\infty} \langle \rho, u_\ell \rangle_{\mathcal{D}} u_\ell(\boldsymbol{y}) + u(\boldsymbol{y}), \tag{4}$$

where $\langle \cdot, \cdot \rangle$ represents an inner product of two functions on a defined space. The set of functions $u_\ell(\boldsymbol{y})$ are called the source distribution modes, and $u(\boldsymbol{y})$ satisfies the equation $\mathcal{A}u = 0$.

Based on the eigenvalue equation $\xi_\ell u_\ell = \mathcal{A}^*\mathcal{A}u_\ell$, a spatial sound field can also be represented by a set of sound field modes, i.e.,

$$S(\boldsymbol{x}) = \sum_{\ell=1}^{\infty} \langle S, V_\ell \rangle_{\mathcal{M}} V_\ell + V(\boldsymbol{x}), \tag{5}$$

where $\mathcal{A}^* V = 0$. The relationship between the source distribution modes, sound field modes, and corresponding coefficients can be found as

$$V_\ell = \frac{1}{\sqrt{\xi_\ell}} \mathcal{A} u_\ell,\ \langle \rho, u_\ell \rangle_{\mathcal{D}} = \frac{1}{\sqrt{\xi_\ell}} \langle S, V_\ell \rangle_{\mathcal{M}}. \tag{6}$$

This implies that for a desired sound field $S_d(\boldsymbol{x})$ on V_0 that can be completely represented as a linear combination of the orthonormal functions $V_\ell(\boldsymbol{x})$ (i.e., $V(\boldsymbol{x}) = 0$), a corresponding source distribution $\rho(\boldsymbol{y})$ can be identified such that $(\mathcal{A}\rho)(\boldsymbol{x})$ allows the exact reproduction of the desired sound field.

3. DESIGN OF FLEXIBLE LAYOUTS

In this work, we assume that the control region is the surface of a sphere enclosing the entire target region, i.e. $V_0 = \{r_{\boldsymbol{x}} = r, \theta_{\boldsymbol{x}} \in [0, \pi], \phi_{\boldsymbol{x}} \in [0, 2\pi)\}$. A secondary source distribution deployed as a circular array or a fragment of a circular array at a particular colatitude is analyzed, from which multiple arrays with flexible geometry are designed[2].

[2]The current setup is fundamental different from the conventional study of the circular source distribution for two-dimensional horizontal Ambisonic system [12].

3.1. Circular Secondary Source Distribution

The free-field Green function (2) can be expanded as [9]

$$\frac{e^{-ik|\boldsymbol{y}-\boldsymbol{x}|}}{4\pi|\boldsymbol{y}-\boldsymbol{x}|} = \sum_{n=0}^{\infty} \sum_{\ell=-n}^{n} -ikh_n^{(2)}(kr_{\boldsymbol{y}}) j_n(kr) \overline{Y_n^\ell(\hat{\boldsymbol{y}})} Y_n^\ell(\hat{\boldsymbol{x}}), \tag{7}$$

where $j_n(\cdot)$ and $h_n^{(2)}(\cdot)$ are the first kind spherical Bessel function and second kind spherical Hankel function of the nth degree, respectively. The unit vectors are denoted by $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$. $\overline{(\cdot)}$ denotes the complex conjugate operation.

The spherical harmonics are defined as [9]

$$Y_n^\ell(\hat{\boldsymbol{x}}) = \mathcal{P}_{n\ell}(\cos\theta_x) E_\ell(\phi_x), \tag{8}$$

where

$$\mathcal{P}_{n\ell}(\cos\theta_x) = \sqrt{\frac{2n+1}{2} \frac{(n-|\ell|)!}{(n+|\ell|)!}} P_n^{|\ell|}(\cos\theta_x) \tag{9}$$

is the normalized associate Legendre function and $E_\ell(\phi_x) = e^{i\ell\phi_x}/\sqrt{2\pi}$ is the normalized complex exponential function.

For a circular secondary source arrangement, i.e., $\Lambda_0 \equiv \{r_y = R, \theta_y = \theta, \phi_y \in [0, 2\pi)\}$ with radius $R > r$, by substituting (7) into (1) and (3) and based on orthogonality of the complex exponential functions, we can analytically derive the modes and eigenvalues as

$$\begin{aligned}
\xi_\ell &= \sum_{n=|\ell|}^{N} k^2 |h_n^{(2)}(kR) j_n(kr) \mathcal{P}_{n\ell}(\cos\theta)|^2 \\
u_\ell(\hat{\boldsymbol{y}}) &= E_\ell(\phi_{\boldsymbol{y}}) \\
V_\ell(\hat{\boldsymbol{x}}) &= \frac{\sum_{n=|\ell|}^{N} -ikh_n^{(2)}(kR) j_n(kr) \mathcal{P}_{n\ell}(\cos\theta) Y_n^\ell(\hat{\boldsymbol{x}})}{\sqrt{\sum_{n=|\ell|}^{N} k^2 |h_n^{(2)}(kR) j_n(kr) \mathcal{P}_{n\ell}(\cos\theta)|^2}}.
\end{aligned} \tag{10}$$

The spaces $\mathcal{D}$ and $\mathcal{M}$ have infinite number of dimensions in the strict sense. Here, we adopt a rule of thumb given in [13] to represent a sound field within the reproduction region using the spherical harmonic expansion up to the degree $N = \lceil ekr/2 \rceil$.

That is, the circular secondary source distribution is approximated as a summation of $L = 2N + 1$ modes,

$$\rho(\boldsymbol{y}) = \sum_{\ell=-N}^{N} \gamma_\ell E_\ell(\phi_y), \tag{11}$$

where the coefficients can be determined directly from the desired sound field decomposition with the derived modes or inferred from the spherical spectrum of the desired sound field,

$$\begin{aligned}
\gamma_\ell &= \langle \rho, u_\ell \rangle_{\mathcal{D}} = \frac{1}{\sqrt{\xi_\ell}} \langle S_d, V_\ell \rangle_{\mathcal{M}} \\
&= \frac{\sum_{n=|\ell|}^{N} ikh_n^{(1)}(kR) j_n(kr) \mathcal{P}_{n\ell}(\cos\theta) \beta_{n\ell}}{\sum_{n=|\ell|}^{N} k^2 |h_n^{(2)}(kR) j_n(kr) \mathcal{P}_{n\ell}(\cos\theta)|^2},
\end{aligned} \tag{12}$$

as $\beta_{n\ell} = \int_0^{2\pi} \int_0^{\pi} S_d(\hat{\boldsymbol{x}}) \overline{Y_n^\ell(\hat{\boldsymbol{x}})} \sin\theta_{\boldsymbol{x}} d\theta_{\boldsymbol{x}} d\phi_{\boldsymbol{x}}$.

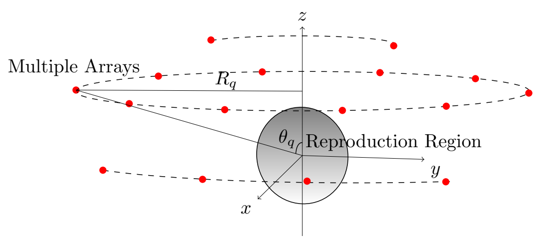


Figure 1: Scheme of flexible loudspeaker layouts for sound field reproduction.

3.2. Secondary Sources on Bounded Circle

The sum in (11) can be re-written as a scalar vector product,

$$\rho = \boldsymbol{\Gamma} \boldsymbol{E}. \tag{13}$$

For a partial circular array layout $\tilde{\Lambda}_0 \in \Lambda_0$, a new set of orthonormal basis functions can be derived by formulating the Gram-matrix [14],

$$\tilde{\boldsymbol{G}} = \langle \tilde{\boldsymbol{E}}, \tilde{\boldsymbol{E}}^T \rangle_{\tilde{\Lambda}_0} = R \langle \boldsymbol{E}, \boldsymbol{E}^T \rangle_{\tilde{\Lambda}_0} R^T \equiv I, \tag{14}$$

with $\tilde{\boldsymbol{E}} = R\boldsymbol{E}$. Eigen decomposition is applied to factorize $\langle \boldsymbol{E}, \boldsymbol{E}^T \rangle_{\tilde{\Lambda}_0}$ and using the eigenvectors $\boldsymbol{V}_b$ associated with non-zero eigenvalues $\boldsymbol{D}_b = \text{diag}\{d_1, d_2, \ldots, d_b\}$ yields the new set of basis functions, i.e., $R = \boldsymbol{D}_b^{-1/2} \boldsymbol{V}_b^T$.

Substituting the new basis functions into (13) then leads to

$$\tilde{\rho} = \tilde{\boldsymbol{\Gamma}} \tilde{\boldsymbol{E}}, \ \tilde{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma} R^{\dagger}, \tag{15}$$

where $(\cdot)^{\dagger}$ represents the pseudo-inverse operation and $\tilde{\rho}$ represents the secondary source distribution on the bounded circle.

3.3. Multiple Array Configuration

For synthesis of three-dimensional sound fields, a practical implementation is to place multiple loudspeaker arrays over the entire range of virtual source colatitudes, i.e., θ_q, $q = 1, 2, \ldots, Q$. We adopt the equiangular grid for sampling of a spherical band-limited signal up to the degree of N to place the loudspeaker arrays [15], that is the interval between two adjacent arrays is $\Delta\theta = \frac{\pi}{2(N+1)}$. In terms of the array geometry, as shown in Fig.1, different radii can be assigned to each array as long as the requirement $R_q > r$ is satisfied. This configuration would provide significantly improved flexibility for loudspeaker placement in real acoustic environments.

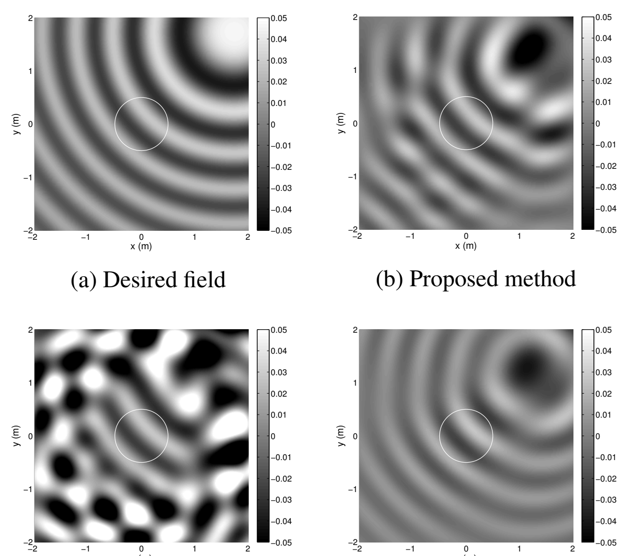
Since each array is designed separately, this configuration also provides a flexible design to active only the arrays with high reproduction efficiency based on the ℓ_2 norm of the sound field decomposed coefficients,

$$\|\eta_q\| = \sqrt{\sum_{\ell=-N}^{N} |\langle S_d, V_\ell^{(q)} \rangle_{\mathcal{M}}|^2}. \tag{16}$$

We calculate the efficiency ratio, i.e. $\|\eta_q\| / \max(\|\eta_1\|, \ldots, \|\eta_Q\|)$, and our simulation suggests that a criteria for loudspeaker array activation is the efficacy ratio larger than 90%. Loudspeaker weights are obtained from discretization of the continuous secondary source distributions [16]. In this setup, the loudspeaker weights also need to be scaled by dividing by the number of activated arrays.

4. SIMULATION

We assess and compare the performance of the proposed method with the LS and Lasso based techniques in a simulation setup of multiple circular/partial circular arrays of loudspeakers. Free-field source conditions are assumed; the desired sound field over a 3D reproduction region of radius 0.5 m is resulted from a single virtual source located at $r_{\text{vs}} = 3$ m, θ_{vs}=55°, and ϕ_{vs}=45°. The operating frequency is 600 Hz; this gives $kr = 5.54$, and the rule of thumb [13] suggests using $N = \lceil ekr/2 \rceil = 8$, thus requiring an interval between two adjacent arrays as 10° and at least 17 loudspeakers in a full set of circular array. In the simulation, we place three layers of loudspeaker arrays of different radii $\{2, 3, 2\}$ m at colatitudes of $40^\circ, 50^\circ, 60^\circ$, respectively. The loudspeakers are equiangularly placed on each array with an interval of 20°, i.e., 18 on the circular array at colatitude 40°, 10 within the azimuthal range $[0^\circ, 180^\circ]$ at colatitude 50° and 6 within the azimuthal range $[20^\circ, 120^\circ]$ at colatitude 60°. The reproduction performance is examined through two measures, one is relative mean square error (MSE) of the reproduction over the desired reproduction region and the other is the loudspeaker weight energy (LWE), i.e. the square of loudspeaker weights.

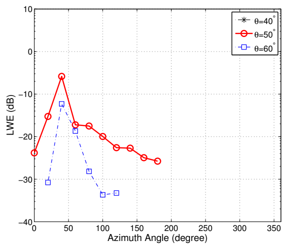


(a) Desired field (b) Proposed method

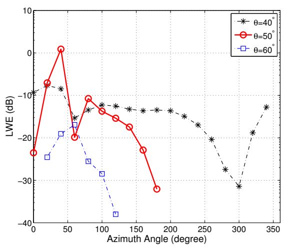
(c) LS-based (d) Lasso-based

Figure 2: Sound field reproduction at frequency of 600 Hz. The encircled region is the reproduction region.

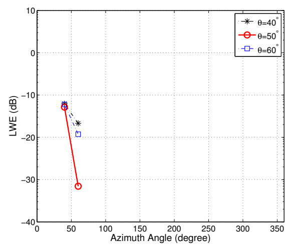
To illustrate, we plot the real parts of the desired sound field and reproduced sound fields within a square slice of the region of interest in the horizontal plane in Fig. 2. These figures are displayed as density plots, where the display is clipped to the maximum value of the reconstructed field within the reproduction region. As demonstrated, the LS method produces the most accurate reproduction (MSE around -30 dB) but with largest levels of sound outside the reproduction zone. The Lasso-based technique produces the smallest level of sound outside the reproduction region, but its accuracy is the lowest (MSE larger than -20 dB) among the three methods. Small reproduction errors (MSE around -25 dB) are obtained using the proposed functional analysis guided approach in the meantime low levels of sound field outside the reproduction region are maintained.



(a) Proposed method



(b) LS-based



(c) Lasso-based

Figure 3: Comparison of loudspeaker weight energy of three methods for reproducing a virtual source from $r_{\mathrm{vs}} = 3$ m and $\theta_{\mathrm{vs}} = 55^\circ, \phi_{\mathrm{vs}} = 45^\circ$. Marks in the plots denote the positions of activated loudspeakers.

Figure 3 compares the loudspeaker weight energy in the proposed method and the other two numerical methods. As expected, the LS technique activates all loudspeakers and has the highest LWE. Lasso activates the least number of loudspeakers due to its inherent sparsity assumption. Even though the virtual source is located at colatitude of 55°, both numerical methods assign larger weights to loudspeakers at 40° rather than loudspeakers at 60°. The proposed method on the other hand actives only loudspeakers (at colatitudes of 50° and 60°) in close proximity to the virtual source. In addition, the proposed method has lower computation complexity as the matrix inversion is avoided which makes it a good candidate for synthesis of real time 3D moving sound.

5. CONCLUSION

A practically realizable design of multiple circular and partial circular loudspeaker arrays is proposed for reproducing 3D sound fields guided by the functional analysis framework. The proposed design provides a flexible loudspeaker layout with the ability to focus on the sound field incident from a limited spatial region. We show design examples of sound field reproduction using the proposed multiple array scheme, in comparison with the performance of the numerical LS and Lasso based techniques. The method developed for a single frequency signal can be repeated and applied in parallel over a subband decomposition for wideband sources.

6. REFERENCES

[1] M. A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Society*, vol. 53, no. 11, pp. 1004–1025, 2005.

[2] F. M. Fazi, P. A. Nelson, J. E. N. Christensen, and J. Seo, "Surround system based on three-dimensional sound field reconstruction," in *125th Convention of the Audio Engineering Society*, San Francisco, CA, USA, Oct. 2008, p. 21 pages.

[3] X. Amatriain, J. Kuchera-Morin, T. Höllerer, and S. T. Pope, "The alloSphere: Immersve multimedia for scientific discovery and artistic exploration," *IEEE Multimedia*, vol. 16, no. 2, pp. 64–75, 2009.

[4] F. M. Fazi, "Sound field reproduction," Ph.D. dissertation, University of Southampton, 2010.

[5] J. Hannemann and K. D. Donohue, "Virtual sound source rendering using a multipole-expansion and method-of-moments approach," *Journal of the Audio Engineering Society*, vol. 56, no. 6, pp. 473–481, 2008.

[6] A. Gupta and T. D. Abhayapala, "Three-dimensional sound field reproduction using multiple circular loudspeaker arrays," *IEEE Trans. Audio, Speech, and Language Process.*, vol. 19, no. 5, pp. 1149–1159, 2011.

[7] T. Betlehem and C. Withers, "Sound field reproduction with energy constraint on loudspeaker weights," *IEEE Trans. Audio, Speech, and Language Process.*, vol. 20, no. 8, pp. 2388–2392, 2012.

[8] G. N. Lilis, D. Angelosante, and G. B. Giannakis, "Sound field reproduction using the Lasso," *IEEE Trans. Audio, Speech, and Language Process.*, vol. 18, no. 8, pp. 1902–1912, 2010.

[9] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. San Diego, CA: Academic Press, 1999.

[10] F. M. Fazi and P. A. Nelson, "Nonuniqueness of the solution of the sound field reproduction problem with boundary pressure control," *Acta Acoustica United with Acoustica*, vol. 98, pp. 1–14, 2012.

[11] L. Debnath and P. Mikusiński, *Introduction to Hilbert Spaces with Applications*, 3rd ed. San Diego, CA: Academic Press, 2005.

[12] J. Ahrens and S. Spors, "An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions," *Acta Acustica united with Acustica*, vol. 94, pp. 988–999, 2008.

[13] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2542–2556, 2007.

[14] H. Pomberger, F. Zotter, and A. Sontacchi, "An ambisonics format for flexible playback layouts," in *Proc. 1st Ambisonics Symposium*, Graz, Austria, June 2009, p. 8 pages.

[15] J. R. Driscoll and D. M. Healy, "Computing Fourier transforms and convolution on the 2-sphere," *Advances in Applied Mathematics*, vol. 15, no. 2, pp. 202–250, 1994.

[16] Y. Wu and T. D. Abhayapala, "Theory and design of soundfield reproducion using continuous loudspeakers concept," *IEEE Trans. Audio, Speech, and Language Process.*, vol. 17, no. 1, pp. 107–116, 2009.