inter.noise 2014
MELBOURNE AUSTRALIA
16-19 NOVEMBER

Creation of a single sound field for multiple listeners

Mark Alister POLETTI[1]; Terence BETLEHEM[1]

[1] Callaghan Innovation, New Zealand

ABSTRACT

Sound reproduction systems are limited in their ability to create sound fields over large areas and over wide bandwidths, because the required number of loudspeakers becomes prohibitive. This limits the ability of surround systems to create sound fields large enough to be heard by multiple listeners.

This paper considers an alternative approach to the creation of a single sound field for multiple listeners in which reproduction is accurate in a number of small zones of sufficient size for each listener. This results in a significant reduction in loudspeaker requirements compared to the creation of a single zone that encompasses all listeners.

The approach is based on conversion of the sound field coefficients of the global field into localized coefficients within each zone and a mode-matching method for simultaneously reproducing the local fields in each zone. The theory is developed for the 2D case, for simplicity, and simulations are presented to demonstrate the effectiveness of the approach.

Such a system would require a tracking system to determine the position of each listener, but the tracking requirements would be less stringent than are required for binaural reproduction using a cross talk canceller.

Keywords: Sound, Reproduction, Loudspeakers I-INCE Classification of Subjects Number(s): 38.2

1. INTRODUCTION

Sound reproduction systems aim to reproduce a desired sound field using a finite number of loudspeakers placed in some geometry around a desired reproduction volume. There are two common approaches (1). Wave Field Synthesis (WFS) is based on the implementation of the Kirchhoff-Helmholtz integral or – the simplified 2D form – Rayleigh integrals (2). Higher order Ambisonics (HOA) is based on the representation of sound fields in cylindrical or spherical coordinates and the recording and reproduction of the associated modes (3).

In an ideal sound reproduction system the sound field can be produced throughout the interior of the loudspeaker array allowing multiple listeners to hear the same program. However, because there are a finite number of speakers, this is only possible up to the spatial Nyquist frequency where the loudspeakers are spaced half a wavelength apart. For typical installations the spatial Nyquist frequency is low and so wide band sound fields cannot be generated over large areas for multiple listeners.

One way to overcome this problem is to identify the listener positions within the loudspeaker array and to reproduce the desired sound field only at the listener positions. If the listeners are moving around, then their positions must be tracked, in a similar manner to what is required for binaural reproduction over loudspeakers (4). However, because the sound field is reproduced over a region around each listener, instead of only at their ear canals, the tracking requirements are less stringent than for the binaural case.

This paper introduces a mode-matching approach to generating a sound field accurately at a number of separate regions within a loudspeaker array. For simplicity we consider the 2D case. The sound field at each listener location is represented with respect to a global origin using an addition theorem and the loudspeaker weights are then determined to match the local fields at each listener location. The reproduction at multiple regions may be termed multi-zone reproduction, but in the case considered here, the same sound field is produced in every region (5)–(8).

[1] mark.poletti@callaghaninnovation.govt.nz

2. THEORY

2.1 Local and global coordinates

The solution to the wave equation in 2D at a positive frequency ω for a region in which there are no sound sources (the interior solution) has the general form (9)

$$p(R,\phi) = \sum_{m=-M}^{M} J_m(kR) A_m(k) e^{im\phi} \tag{1}$$

Where $k = \omega / c$ is the wave number, c the speed of sound, $J_m(.)$ the *mth* Bessel function and $A_m(k)$ is the *mth* sound field coefficient. If the sound field is described over a limited region and bandwidth, then kR is finite and the maximum mode order M required to represent the field is given approximately by $M \approx \lceil kR \rceil$, where $\lceil . \rceil$ denotes rounding up to the next highest integer (10). Reproducing this spatially and band limited field thus requires that approximately $2kR+1$ modes are accurately reproduced. If reproduction is carried out using a regular circular array of L monopole loudspeakers, positioned at radius R_L, over the entire region within the loudspeaker array, then the number of loudspeakers must exceed the required number of modes, $L \geq 2kR_L + 1$. Setting $L = 2kR_L + 1$ produces the modal form of the spatial Nyquist frequency

$$f_{Nyq} = \frac{c(L-1)}{4\pi R_L} \tag{2}$$

For example, with 50 loudspeakers at a radius of 5 m, the Nyquist frequency is 265 Hz. If reproduction is only required for a single listener at the origin, and the zone radius is R_Z, then the required number of modes reduces to approximately $2kR_Z + 1$ and the spatial Nyquist frequency rises to

$$f_{Nyq,Z} = \frac{c(L-1)}{4\pi R_Z} \tag{3}$$

Consider now the case where the reproduction region is at the arbitrary location $\vec{R}_q = (R_q, \phi_q)$ (Fig. 1). The sound field with respect to the local origin with polar coordinates (ρ_q, θ_q) has the form

$$p_q(\rho_q,\theta_q) = \sum_{m=-M}^{M} J_m(k\rho_q) B_m(k) e^{im\theta_q} \tag{4}$$

Using the addition theorem (11) each term in Eq. (4) can be written in terms of global coordinates as

$$J_m(k\rho_q) e^{im\theta_q} = \sum_{n=-\infty}^{\infty} J_n(kR) J_{n-m}(kR_q) e^{in\phi} e^{-i(n-m)\phi_q} \tag{5}$$

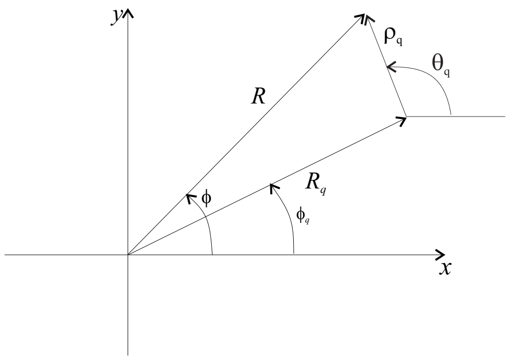

Figure 1: Coordinate systems

Hence the global expansion of the pressure in Eq. (4) is

$$p_q(R,\phi) = \sum_{n=-\infty}^{\infty} J_n(kR) e^{in\phi} \sum_{m=-\infty}^{\infty} B_m(k) J_{n-m}(kR_q) e^{-i(n-m)\phi_q} \tag{6}$$

Comparing this with Eq. (1), the global coefficients are obtained from the local coordinates as

$$A_m(k) = \sum_{m=-\infty}^{\infty} B_m(k) J_{n-m}(kR_q) e^{-i(n-m)\phi_q} \tag{7}$$

Using the identity

$$\sum_m J_{m-l}(z) J_{m-n}(z) = \delta_{n-l} \tag{8}$$

the local coefficients can be related to the global coefficients as

$$B_n(k) = \sum_{m=-\infty}^{\infty} A_m(k) J_{m-n}(kR_q) e^{im\phi_q} \tag{9}$$

2.2 Sound reproduction in local regions

Consider now the reproduction of a general sound field (Eq. (1)) at a location $\vec{R}_q$ using L monopole sources positioned in a regular circular array with source positions $\vec{R}_l = (R_L, \phi_l)$ where $\phi_l = l2\pi / L$ and $l \in [0, L-1]$. The sound field due to a single source is (11)

$$p_l(R,\phi) = \sum_{m=-\infty}^{\infty} J_m(kR) e^{im\phi} H_m(kR_L) e^{-im\phi_l} \tag{10}$$

where $H_m(.)$ is the cylindrical Hankel function of second order. We require the sum of the fields due to the L sources, weighted by coefficients w_l, to closely approximate Eq. (1)

$$\hat{p}(R,\phi) = \sum_{m=-\infty}^{\infty} J_m(kR) e^{im\phi} H_m(kR_L) \sum_{l=0}^{L-1} w_l e^{-im\phi_l} \approx p(R,\phi) \tag{11}$$

The approximate sound field can be written in terms of the local coordinates using Eq. (9)

$$\hat{p}(\rho_q, \theta_q) = \sum_{n=-N}^{N} J_n(k\rho_q) e^{in\theta_q} \left[\sum_{m=-\infty}^{\infty} J_{m-n}(kR_q) H_m(kR_L) e^{im\phi_q} \sum_{l=0}^{L-1} w_l e^{-im\phi_l} \right] \tag{12}$$

where the term in square parentheses is the approximate local coefficient $\hat{B}_n(k)$ and where we have assumed that the reproduction is required for a maximum $N = \lceil k\rho_q \rceil$. Equating these to the ideal coefficients in local coordinates (Eq. (9)), yields the mode matching equations

$$\sum_{m=-\infty}^{\infty} J_{m-n}(kR_q) e^{im\phi_q} H_m(kR_L) \sum_{l=0}^{L-1} w_l e^{-im\phi_l} = \sum_{m=-\infty}^{\infty} J_{m-n}(kR_q) A_m(k) e^{im\phi_q} \tag{13}$$

which must be satisfied for $n \in (-N, N)$. In practice the sum in m can be truncated to a maximum order $M = \lceil N + kR_q \rceil$.

If there are Q listeners then there are Q separate regions, each with its local mode matching requirement having the same form as Eq. (13). Matching the sound fields in all Q zones produces a set of mode matching equations which can be written in matrix form

$$\mathbf{H}w = d \tag{14}$$

where $\mathbf{H}$ is a $Q(2N+1)$ by L matrix, w is an L by 1 vector of loudspeaker weights and d is a $Q(2N+1)$ vector of desired local coefficients.

For sufficiently large L, $Q(2N+1) < L$ and there are infinitely many solutions to Eq. (14). The solution with the minimum weight energy is

$$w = \mathbf{H}\left(\mathbf{H}^H \mathbf{H}\right)^{-1} d \tag{15}$$

where superscript H denotes the conjugate transpose. In practice, this solution may produce large weight magnitudes which can produce non-robust solutions in the presence of perturbations from ideal loudspeaker performance (12). Therefore, Eq. (15) may be regularized as

$$w = \mathbf{H}\left(\mathbf{H}^H \mathbf{H} + \lambda \mathbf{I}\right)^{-1} d \tag{16}$$

where λ is a regularisation parameter, which in the simulations to follow was set to a fraction $\varepsilon = 0.001$ of the maximum squared singular value of $\mathbf{H}$.

A simple measure of robustness is the total loudspeaker weight energy (LWE) (12)(13),

$$LWE = \sum_{l=0}^{L-1} \left| w_l \right|^2 \tag{17}$$

2.3 Spatial Nyquist frequency

The spatial Nyquist frequency of the multi-zone reproduction system may be determined in the same manner as used to derive Eq. (2). Reproduction in a zone of radius R_Z requires $2kR_Z + 1$ modes. There are Q of these zones and therefore approximately $Q(2kR_Z + 1)$ local modes must be controlled by L loudspeakers. The spatial Nyquist frequency is then

$$f_{Nyq,Z} = \frac{c(L/Q - 1)}{4\pi R_Z} \approx f_{Nyq} \frac{R_L}{QR_Z} \tag{18}$$

Hence if QR_Z is less than the loudspeaker radius R_L, then the multi-zone system provides a higher Nyquist frequency. This equation generalizes the spatial Nyquist frequency for a single listener at the centre of the reproduction region, with zone radius R_Z (Eq. (2)), to the case for Q listeners, where the Nyquist frequency reduces by approximately $1/Q$. The minimum number of resources are used for a given number of listeners.

The spatial Nyquist frequency is shown in Fig. 2 for a circular array of 50 loudspeakers at radius 5m and up to 8 listeners, each requiring a zone of radius 0.2 m. The spatial Nyquist frequency is 6.6 kHz for a single listener, 3.2 kHz for two and 2.1 kHz for three listeners. The Nyquist frequency without tracking would require reproduction throughout the interior, yielding a value of 265 Hz.

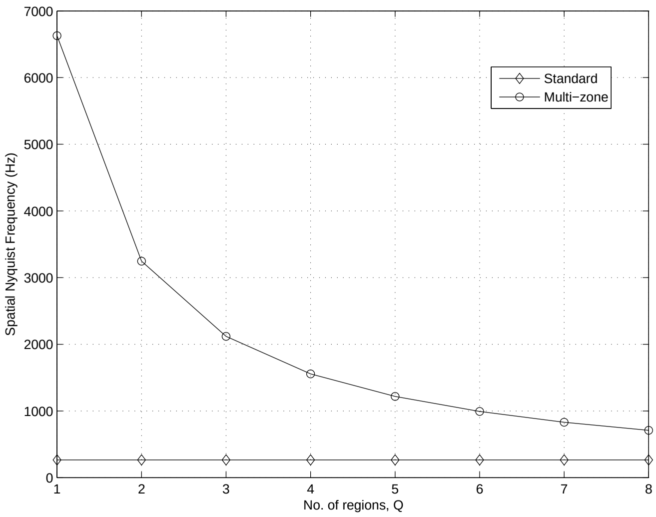

Figure 2: Spatial Nyquist frequency vs no. of listeners, R_Z = 0.2 m

3. SIMULATIONS

Numerical simulations were carried out to evaluate the mode matching technique. A circular array of 50 loudspeakers was used, unless otherwise stated, at a radius of 5m with listener zone radii of 0.2 m. The desired source was a single line source at position $\vec{R}_S = (R_S, \phi_S)$, which has the modal expansion in Eq. (10) with $\vec{R}_l$ replaced by $\vec{R}_S$. A source angle of $\phi_S = 30$ degrees was used for all simulations.

For brevity we typically plot only the normalized squared error

$$\varepsilon(R,\phi) = \frac{\left| p(R,\phi) - \hat{p}(R,\phi) \right|^2}{\left| p(R,\phi) \right|^2} \tag{19}$$

This is well defined at all points since $p(R,\phi)$ is complex.

We first consider a single zone of radius 1 metre and investigate the effect of moving the zone from the centre. Figure 3 shows the real part of the sound field and Figure 4 the reproduction error in dB for a zone at the center of the array and a desired source at a radius of 5.1 m, and at an angle of 30 degrees, radiating at 1 kHz. The LWE is 0.76. The number of sources exceeds the required number of modes

($Q\left(2\lceil kR_z \rceil+1\right)=39$) and the error is below -30 dB within the zone. This performance is attained for any source angle and for any source radius greater than the loudspeaker radius.

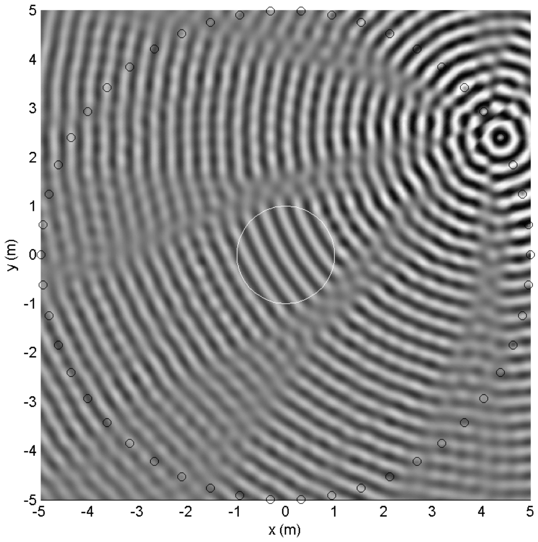

Figure 3: Sound field at 1 kHz, R_S = 5.1 m R_Z=1m at origin

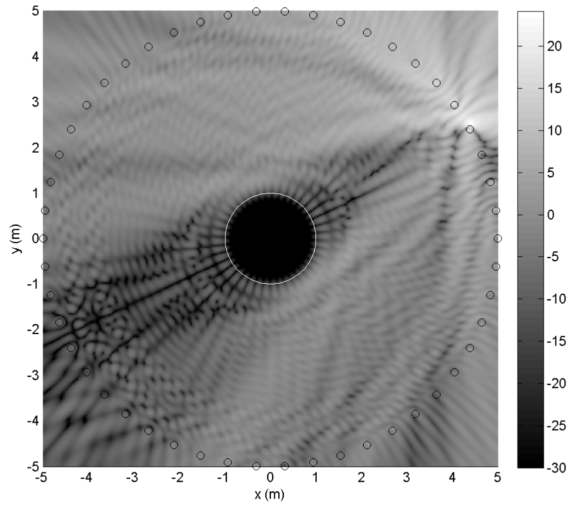

Figure 4: Reproduction error in dB at 1 kHz, R_S = 5.1 m, R_Z=1m at origin. LWE = 0.76

Figure 5 shows the error for a zone at 3m from the origin, for the same source position. The error is higher than that in Figure 4, and the LWE is 1.3 which is around twice the value for reproduction at the origin. It is harder to create a sound field near the array sources than at the origin due to the attenuation of sound from the far sources and the increased wave front curvature from the near sources.

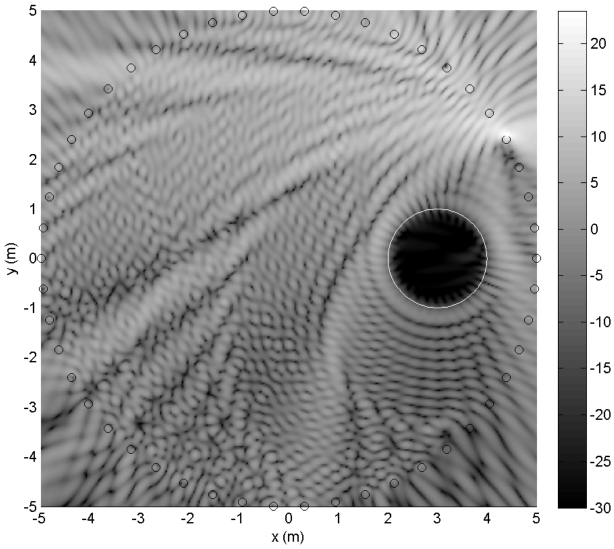

Figure 5: Reproduction error in dB at 1 kHz, R_S = 5.1 m, R_Z=1m. LWE = 1.3

Figure 6 shows the error for a source radius of 10m and a zone angle of 30 degrees, equal to the source angle. The error is higher since the weight solution must not only cope with the relative asymmetry of the sources but must also alter the wave front curvature more significantly than in the previous example. While the condition number of the matrix **H** was similar, the LWE is 8.7 which is over six times that of the previous example.

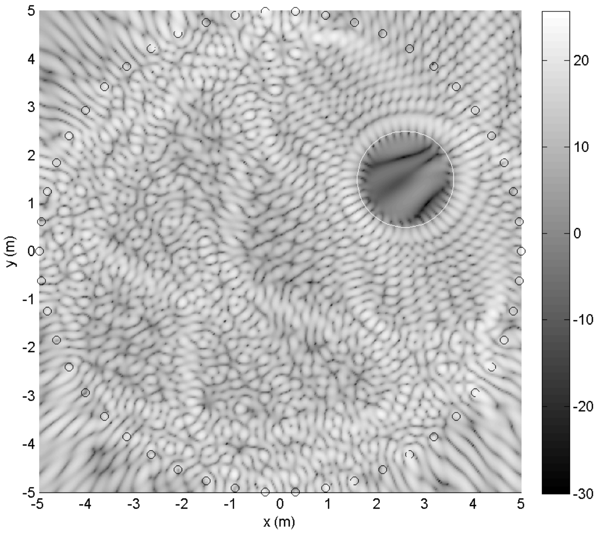

Figure 6: Reproduction error in dB at 1 kHz, R_S = 10 m, R_Z=1m, φ_q=30º. LWE = 8.7

Figure 7 shows the error for twice the number of sources (L=100) and 10 reproduction zones positioned in a spiral configuration, with each zone having a radius of 0.2 m. The number of sources just exceeds the required number of modes (90), the corresponding Nyquist frequency is 1.3 kHz and the sound field is accurately produced in all 10 zones. The LWE is 1.2.

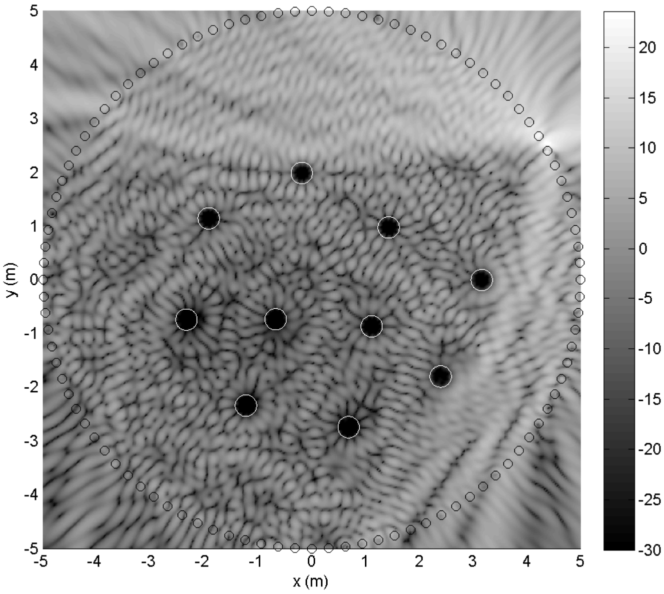

Figure 7: Reproduction error in dB at 1 kHz, f_{Nyq} = 1.3 kHz, 10 zones, R_S = 5.1 m, L = 100, LWE=1.2

Finally, the technique can be used to create a large single zone of non-circular shape. For example, Figure 8 shows the error for 5 zones, each of radius 0.2 m, placed adjacent to each other, which creates a single, approximately elliptical, region which would allow 5 people to sit in a line. The LWE is 0.53.

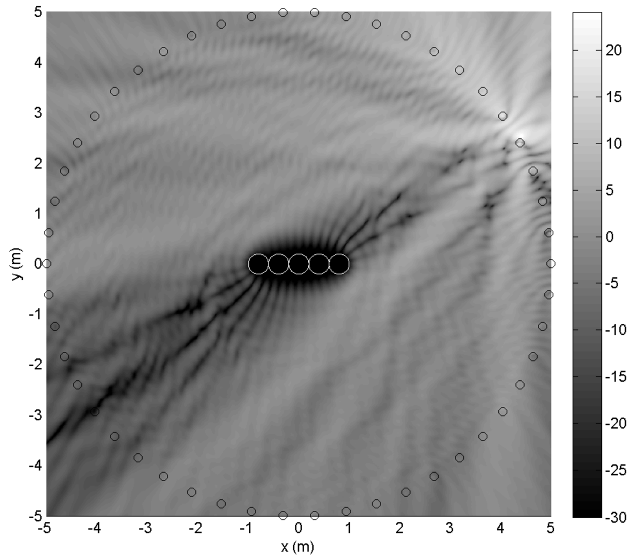

Figure 8: Reproduction in dB at 1 kHz, f_{Nyq} = 1.2 kHz, 5 adjacent zones, R_S = 5.1 m. LWE = 0.53

4. CONCLUSIONS

This paper has introduced a method for reproducing a single global sound field in multiple local zones. This allows the effort required by the array to be minimized as the sound field need only be accurate within the zones. Simulations have shown that the technique works well in the 2D case, and the weight energy values suggest that reproduction using practical arrays would be reasonably robust. However, the technique produces larger errors for zones near the loudspeakers and when significant alteration of wave front curvature is required.

REFERENCES

1. Spors, S., H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter, “Spatial Sound With Loudspeakers and Its Perception: A Review of the Current State,” *Proc. IEEE*, vol. 101, no. 9, pp. 1920–1938, 2013.

2. Berkhout, A. J., D. de Vries, and P. Vogel, “Acoustic control by wave field synthesis,” *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, 1993.

3. Gerzon, M. A., “Ambisonics in multichannel broadcasting and video,” *J. Audio Eng. Soc.*, vol. 33, no. 11, pp. 859–871, 1985.

4. Gardner, W. G., *3-D Audio Using Loudspeakers*. Dordrecht: Kluwer Academic Publishers, 1998.

5. Choi, J.-W. and Y.-H. Kim, “Generation of an acoustically bright zone with an illuminated region using multiple sources,” *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1695–1700, 2002.

6. Elliott, S. J., J. Cheer, J.-W. Choi, and Y. Kim, “Robustness and Regularization of personal audio systems,” *IEEE Trans. Audio Speech Lang. Process.*, vol. 20, no. 7, pp. 2123–2133, 2012.

7. Poletti, M. A., “An Investigation of 2D Multizone Surround Sound Systems,” in *AES 125th Convention*, 2008.

8. Wu, Y. J. and T. D. Abhayapala, “Spatial Multizone Soundfield Reproduction: Theory and Design,” *IEEE Trans. Audio Speech Lang. Process.*, vol. 19, no. 6, pp. 1711–1720, 2011.

9. Williams, E. G., *Fourier Acoustics*. San Deigo: Academic Press, 1999, pp. 115–234.

10. Ward, D. B. and T. D. Abhayapala, “Reproduction of a plane-wave sound field using an array of loudspeakers,” *IEEE Trans. Speech Audio Process.*, vol. 9, no. 6, pp. 697 – 707, 2001.

11. Watson, G. N., *A treatise on the Theory of Bessel Functions*, 2nd ed. 40 West 20th street, New York: Cambridge Univ. Press, 1995, p. chap. 7, pp 198.

12. Poletti, M. A., “Robust two-dimensional surround sound reproduction for nonuniform loudspeaker layouts,” *J. Audio Eng. Soc.*, vol. 55, no. 7/8, pp. 598–610, 2007.

13. Betlehem, T. and M. A. Poletti, “Two dimensional sound field reproduction using higher order sources to exploit room reflections,” *J. Acoust. Soc. Am.*, vol. 135, no. 4, pp. 1820–1833, 2014.