RADIAL FILTERS FOR NEAR FIELD SOURCE SEPARATION IN SPHERICAL HARMONIC DOMAIN

Isha Agrawal and Rajesh M. Hegde

Indian Institute of Technology, Kanpur
Email: {isha, rhegde}@iitk.ac.in

ABSTRACT

Radial filter design for processing near field speech sources over a spherical microphone array is challenging. Polynomial based radial filter design procedures have been proposed in earlier work. In this paper we address the issue of radial filter design using a family of orthogonal polynomials called the Gegenbauer polynomials. The radial filters designed using this approach indicate an improved radial response and greater efficiency in attenuating distant sources. Improved white noise gain and directivity index are also noted from experimental evaluations. The radial filters hence designed are used to separate directionally co-incident near field speech sources. Subjective evaluation is conducted on separated sources using measures like LSD, PESQ, and SDR. The subjective evaluation scores are motivating enough to be considered for practical speech and audio applications.

Index Terms— Radial Filters, Source Separation, Spherical Microphone Array.

1. INTRODUCTION

Spherical Microphone Arrays have been actively used for sound field analysis, room acoustic measurements and spatial filtering applications. This is due to the ease of performing array processing in spherical harmonic domain [1]. Sources can also be resolved without spatial ambiguity using a spherical microphone array. Near field spherical microphone array processing has received a lot of attention from researchers in the past decade. In [2], a robust adaptive beamformer for near-field sources has been formulated. This work is based on the idea of far-field LCMV (Linear Constrained Minimum Variance) beamformer. A near-field source localization method based on SH-MUSIC (Spherical Harmonic MUltiple SIgnal Classification) appeared in [3]. Methods for constructing directional beamformers at varying radial distances have been proposed in [4], [5] and [6] but they do not talk about radial filters explicitly. The radial filtering capabilities of a spherical array for near-field sources has been comprehensively studied in [7]. The near-field criterion in terms of array order, frequency and location has been systematically defined herein. In [8], this idea was extended to the far-field Dolph-Chebyshev beamformer to develop a Dolph-Chebyshev radial filter, capable of attenuating far-field interferences given a source close to the array surface. In [9], the capability of this radial filter was tested for real speech signals. Some additional methods for radial filter design have also been presented in [10]. The radial filters presented in most studies are polynomial-based as originally proposed in [10].

In [10], the authors have worked with low-order polynomials and the Chebyshev polynomial. However, this paper presents a design methodology for radial filters using a family of orthogonal polynomials called the Gegenbauer polynomials [11]. Design of these filters is presented along with their performance analysis based on the White Noise Gain (WNG) and Directivity Index (DI). These radial filters are also used in the separation of directionally co-incident sources. Subjective quality of the separated sources is assessed on the basis of measures such as LSD, PESQ and Signal-to-Distortion Ratio (SDR).

This work was funded by the DST under project number EE/SERB/20130277 and by the BITCOE, IIT Kanpur.

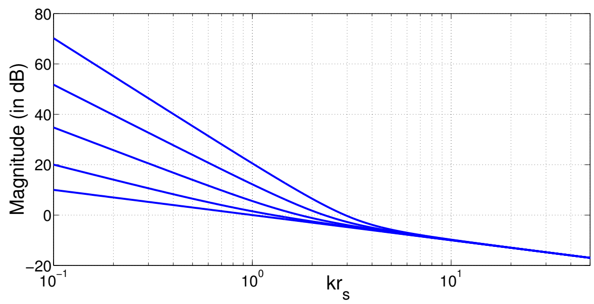

Fig. 1. Magnitude of the spherical Hankel function, $h_n(kr_s)$ for orders $n = 0$ (bottom) to $n = 4$ (top).

The paper is organized as follows. In section 2, the near-field array processing method is reviewed. The near-field criteria is stated next. In section 3, the polynomial-based radial filter design method using the Gegenbauer polynomials is detailed. This is followed by an analysis of radial response using WNG and DI. Section 4 presents experiments with source separation. Section 6 concludes the paper.

2. NEAR-FIELD SPHERICAL MICROPHONE ARRAY PROCESSING

Array processing is performed in the spherical harmonic domain (SH-Domain) because of its ability to decouple angle-dependent and frequency-dependent components [12]. Additionally, spherical wave fronts can provide important spatial cues which can be exploited in the design of efficient spatial filters. Let, the pressure at any point in a sound field be represented by $p(k, r, \Omega)$, where k is the wave number, r is the radial distance from the origin and $\Omega = (\theta, \phi)$ is the angular position of the point of interest. For a point source located at (r_s, Ω_s), pressure at (r, Ω) is given by [13],

$$p(k, r, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} b_n^s(k, r, r_s)[Y_n^m(\Omega_s)]^* Y_n^m(\Omega) \quad (1)$$

where, $Y_n^m(\Omega)$ is a spherical harmonic of order n and degree m and the asterisk, $[.]^*$, refers to the complex conjugate. $b_n^s(k, r, r_s)$ is the near-field mode strength factor given by,

$$b_n^s(k, r, r_s) = i^{-(n-1)} k b_n(kr) h_n(kr_s) \quad (2)$$

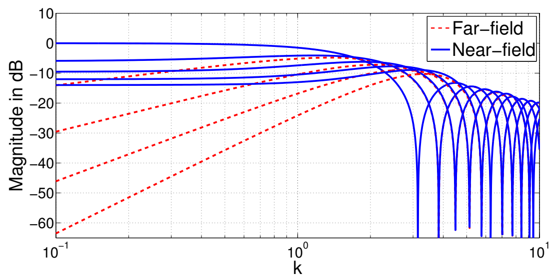

Fig. 2. Plots of the far-field mode strength $b_n(kr)$ (dashed lines) and the near-field $b_n(k, r, r_s)$ (solid lines) for $r_s = 1$m and orders $n = 0$ (top) to $n = 4$ (bottom).

Here, $b_n(kr)$ is the far-field mode strength factor, defined as,

$$b_n(kr) = \begin{cases} 4\pi i^n j_n(kr), & \text{for open sphere} \\ 4\pi i^n (j_n(kr) - \frac{j'_n(kr)}{h'_n(kr)}), & \text{for rigid sphere} \end{cases} \quad (3)$$

where, j_n is the spherical Bessel function, h_n is the spherical Hankel function and j'_n and h'_n are their respective derivatives. It can be seen from Equation 2 that the radial behaviour of the point source is determined by the spherical Hankel function as shown in Figure 1.

Taking the spherical Fourier Transform of Equation 1, sound field at the array surface in the SH-Domain is obtained as [1],

$$p_{nm}(k, a) = b_n^s(k, a, r_s)[Y_n^m(\Omega_s)]^* \quad (4)$$

where a is the radius of the SMA. Spatial sampling introduced due to the finite number of microphones on the array surface limits the array order N following the inequality $I > \beta(N+1)^2$ [14], where I is the number of microphones and β is determined by the chosen sampling scheme [15]. Having acquired the pressure field at array surface, the final array output is given by,

$$y(k) = \sum_{n=0}^{N} \sum_{m=-n}^{n} w_{nm}^*(k) p_{nm}(k) \quad (5)$$

where, w_{nm} is the spherical Fourier transform of the spatial weighting function $w(\Omega_i)$. Designing a spatial filter requires an involved computation of this weighting function [16].

2.1. Near-field Criterion in Spherical Harmonic Domain

Traditionally, Fraunhofer and Fresnel distances were used for finding the transition from the near-field to the far-field of a sensor array. A novel method for finding out the border between the near and far fields was proposed in [7]. In this work, the magnitudes of the mode strength factors of the two cases, i.e. $|b_n^s(k, r, r_s)|$ and $|b_n(kr)|$, are compared to study the equivalence. It can be observed from Figure 2 that after $kr_s \approx n$, the two functions start behaving in a similar manner. Thus for $kr_s < n$, the near-field behaviour is evident. Hence, for an array of order N, the near-field criteria r_{NF} is given by

$$r_{NF} \approx \frac{N}{k} \quad (6)$$

For an array of radius a, the highest allowable wave number to avoid spatial aliasing is given by $k_{max} = N/a$, as discussed in [17]. Thus, for a point source to be located in the near-field, the constraint on r_s is given by $a < r_s < r_{NF}$.

3. RADIAL FILTER DESIGN USING A FAMILY OF ORTHOGONAL POLYNOMIALS

The distance of the point source from the array surface is usually not known. In such cases, to obtain an array output that is spherically symmetrical about the look direction, following weighting function is used [10],

$$w_{nm}^*(k) = \frac{d_n(k)}{i^{-(n-1)} k b_n(ka)} Y_n^m(\Omega_l) \quad (7)$$

where, Ω_l is the look-up direction and d_n are called the design coefficients. Using Equations 4 and 5 the array output can be rewritten as,

$$y(k) = \sum_{n=0}^{N} d_n(k) h_n(kr_s) P_n(cos\Theta) \quad (8)$$

where, Θ is the angle between Ω_s and Ω_l and $P_n(.)$ is the Legendre function of order n. The additivity property of spherical harmonics, used to arrive at Equation 8, involves a factor of $(2n+1)/4\pi$ which is absorbed in d_n here for notational convenience. For the design of the radial filter we assume that the directional location of the point source is known. Thus assuming $\Theta = 0$, the array response in the look direction is given by,

$$y(k) = \sum_{n=0}^{N} d_n(k) h_n(kr_s) \quad (9)$$

The spherical Hankel function can be expanded using power-polynomials [18]. The array output can now be written as,

$$y(z) = h_0\left(\frac{ka}{z}\right)\left[\sum_{n=0}^{N} d_n(k) \sum_{m=0}^{n} c_m^n(k) z^m\right] \quad (10)$$

where,

$$c_m^n(k) = \frac{i^{(m-n)}}{m! 2^m} \frac{(n+m)!}{(n-m)!}(ka)^{-m} \quad \text{and} \quad z = \frac{ka}{x} \quad (11)$$

The term in square brackets is an Nth order polynomial in z. This polynomial is equated to another order-N polynomial exhibiting the desired radial filter pattern as,

$$\sum_{n=0}^{N} d_n(k) \sum_{m=0}^{n} c_m^n(k) z^m = \sum_{n=0}^{N} q_n z^n \quad (12)$$

Equation 12 can be written in matrix form as,

$$\mathbf{C}^T \mathbf{d} = \mathbf{q} \quad (13)$$

where, $\mathbf{d} = [d_0\ d_1\ ...\ d_N]^T$ and $\mathbf{q} = [q_0\ q_1\ ...\ q_N]^T$. $\mathbf{C}$ is an $(N+1) \times (N+1)$ invertible matrix with $\mathbf{C}(i, j) = c_{j-1}^{i-1}$ for $i \geq j$, zero otherwise. Finally, one can obtain the design coefficients as,

$$\mathbf{d} = (\mathbf{C}^{-1})^T \mathbf{q} \quad (14)$$

Thus, given a desired pattern, radial filter design can be obtained [10].

Chebyshev polynomials have been used for designing radial filters using the above method in [8]. In this paper, the Gegenbauer polynomial family [11] is chosen to assess the scope of this method in designing filters with varying radial response. Gegenbauer polynomial family is a class of polynomials which are orthogonal with respect to the weighing function $(1-x^2)^{\alpha-1/2}$ on the interval $[-1\ 1]$,

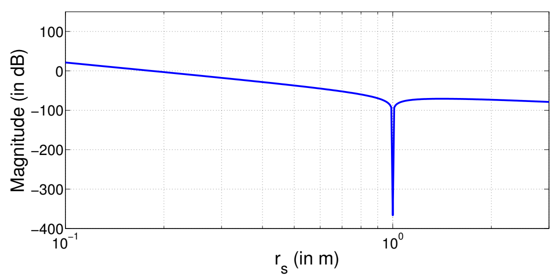

Fig. 3. Magnitude response of the radial filter for $\alpha = 0.5$, $N = 3$, $R = 1000$ and $k = 1$

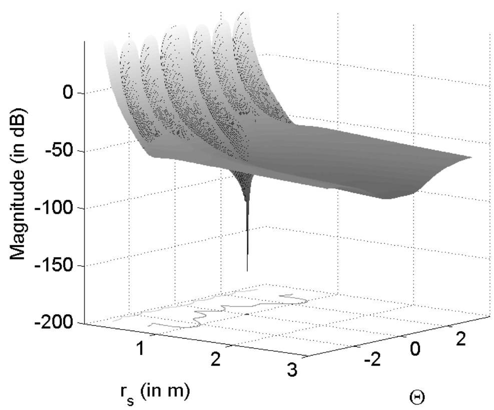

Fig. 4. Magnitude response of the radial filter with $\alpha = 0.5$, $N = 3$, $R = 1000$ and $k = 1$, for varying r_s and Θ

where α is a parameter greater than -0.5. Many standard polynomials belong to this family, like, Legendre polynomial ($\alpha = 0.5$), Chebyshev polynomial of the first kind ($\alpha = 0$) and Chebyshev polynomial of the second kind ($\alpha = 1$). As $|\alpha|$ is increased from zero, magnitudes of local maxima and minima increase along with the curvature of the polynomial. Consider a polynomial of this family given by

$$P(z) = \frac{1}{R}\sum_{n=0}^{N} p_n x_0^n z^n \tag{15}$$

where, p_n is the nth coefficient of the polynomial, x_0 is a scaling factor that decides the root positions and R controls the maxima/minima of the polynomial. Using Equations 12 and 13, we can write,

$$\mathbf{C}^T\mathbf{d} = \frac{1}{R}\mathbf{X}_0\mathbf{p} \tag{16}$$

where, $\mathbf{X}_0 = \mathrm{diag}(x_0^0, x_0^1, ..., x_0^N)$ and $\mathbf{p} = [p_0\ p_1\ ...\ p_N]^T$. Hence, the design coefficients can be computed as,

$$\mathbf{d} = \frac{1}{R}(\mathbf{C}^{-1})^T\mathbf{X}_0\mathbf{p} \tag{17}$$

Figure 3 illustrates the radial response of a filter designed using Gegenbauer polynomial with $\alpha = 0.5$ (Legendre polynomial), order $N = 3$, $R = 1000$, $k = 1$ and $a = 5$ cm. x_0 has been chosen such that the polynomial root lies at a distance of 1 m resulting in

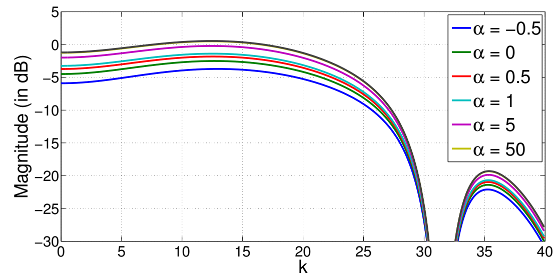

Fig. 5. Variation of WNG for different α and $N = 4$.

a notch at that position. The decrease in magnitude with increasing radial distance makes these filters suitable for attenuating distant inferences. Figure 4 shows the spectral magnitude response of the filter. It can be seen that though the notches are effective in attenuating the undesired sources the gain at the desired radial distance is not very high. Other filters can be designed in a similar manner by varying polynomial type, order and notch positions. Choosing a higher order results in more complex filter designs.

4. PERFORMANCE ANALYSIS

In this section, the radial filters constructed using different Gegenbauer polynomials have been evaluated on the basis of White Noise Gain (WNG) and Directivity Index (DI). Experiments have been performed on two sizes of spherical arrays, namely, a small size array with radius 5 cm, and a medium size array with radius 10 cm.

4.1. White Noise Gain (WNG)

WNG is the ratio of the signal-to-noise ratio (SNR) at the array output to the SNR at the array input [19]. It can be obtained using [10],

$$\mathrm{WNG} = I\frac{|\sum_{n=0}^{N} d_n(k)h_n(kr_s)(2n+1)|^2}{\sum_{n=0}^{N}|\frac{d_n(k)}{b_n(ka)}|^2(2n+1)} \times \frac{1}{\sum_{n=0}^{\infty}|b_n(ka)h_n(kr_s)|^2(2n+1)} \tag{18}$$

Simulations show that for lower order filters ($N = 1, 2$ or 3), there is no variation in WNG with change in α. This is because lower order Gegenbauer polynomials differ only in a scaling factor which is eventually normalized during processing. For higher orders ($N = 4$ and above), as α increases, WNG increases steadily and saturates after a certain value. This can be seen in Figure 5, where a medium sized array with $N = 4$ has been used. The difference between maximum (for $\alpha = 100$) and minimum (for $\alpha = -0.5$) WNG values in this case is approx 5dB. Table 1 shows such differences for other higher orders. All these radial filters typically show a low-pass behaviour. However at higher order, they tend to deviate from this behaviour at higher frequencies.

4.2. Directivity Index (DI)

The directivity index (DI) is a measure indicating the improvement in directivity of an array compared to an omni-directional microphone. Theoretically, it is the ratio of the array output in the look direction to the combined array output over all directions [20]. DI is computed in this work as,

$$\mathrm{DI} = 10\log_{10}(\mathrm{Q}) \tag{19}$$

Order N	Maximum WNG range
4	5dB
5	10dB
6	5dB

Table 1. Range of maximum WNG obtained for $-0.5 < \alpha < 50$, and varying order N.

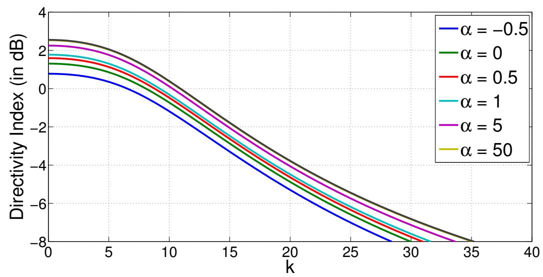

Fig. 6. Variation of DI with k and α for order $N = 4$.

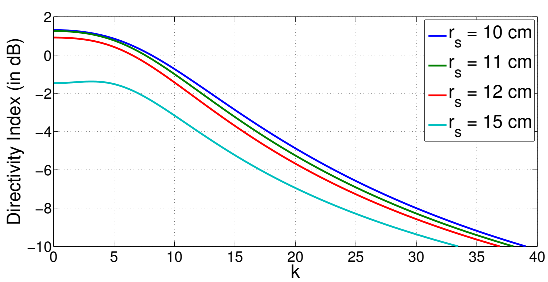

Fig. 7. Variation of DI with k and r_s for $N = 4$ and $\alpha = 0$.

where, Q is the directivity factor given by,

$$\mathrm{Q} = \frac{|\sum_{n=0}^{N} d_n(k)h_n(kr_s)(2n+1)|^2}{\sum_{n=0}^{N} |d_n(k)h_n(kr_s)|^2(2n+1)} \tag{20}$$

It can be seen from Figure 6 that, directivity improves, and ultimately saturates, with increasing α. This is in the case of polynomials with order greater than 3. It is observed that DI also exhibits low-pass behaviour which tends to get sharper with increasing polynomial order. Figure 7 shows the variation of DI for varying radial positions of the source. Directivity decreases with increasing source distance which indicates that directional response is lost in the far-field of the array.

5. EXPERIMENTS ON SOURCE SEPARATION

The radial filters presented herein are suitable for close-talk applications. They are capable of attenuating distant sources effectively in comparison to their natural $1/r_s$ decay. In this section, we assess the performance of the proposed filters, through MATLAB® simulations, in separating directionally co-incident near field sources (see Figure 8). Efficiency of the designed filters is evaluated based on three standard performance evaluation criteria, namely, Log-spectral Distance (LSD), Perceptual Evaluation of Speech Quality (PESQ), and Signal-to-Distortion Ratio (SDR).

Radial distance of Interference	LSD	PESQ	SDR
6 cm	0.54	2.52	5.11
8 cm	0.49	2.62	8.98
10 cm	0.44	2.69	12.41

Table 2. Source Separation performance for order $N = 2$.

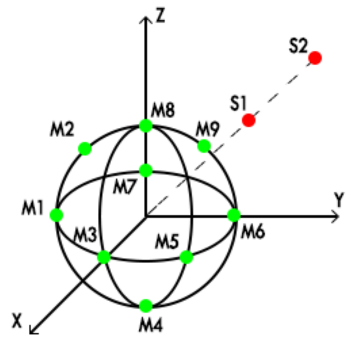

Fig. 8. Experimental setup for coincidental source separation. S_1 and S_2 are placed at different radial distances from the centre of an Eigenmike® SMA.

α	LSD	PESQ	SDR
-0.5	0.43	2.59	27.90
0	0.46	2.52	23.68
0.5	0.49	2.47	21.19
1	0.50	2.43	19.50
5	0.56	2.33	14.37
50	0.62	2.27	10.93

Table 3. Source Separation performance for order $N = 4$ with interfering source at a radial distance of 10 cm at an SIR of 0 dB.

A 32-microphone Eigenmike® [21] with radius 4.2 cm has been used for this purpose. A desired source (s_1) is placed at 5 cm from the array centre and an interfering source (s_2) is placed at different test positions beyond it. The initial signal-to-interference ratio is 1. It was verified that for order less than 4, changing the polynomial does not result in a separate filter design. Table 2 presents the evaluation outcomes for different positions of the interfering source (6 cm, 8 cm and 10 cm) for a 2nd order radial filter (any α). As the array order increases, performance improves. For an order $N = 4$, when the interfering source is placed 10 cm away from the array, the results are shown in Table 3. It can be seen that filters designed using smaller value of α exhibit better source separation. The PESQ values primarily lies between 2 and 3, which appears to be a reasonably good score.

6. CONCLUSION

In this paper, radial filter design for separating near field sources is proposed. The method uses a family of orthogonal polynomials in contrast to earlier work. It was observed that certain polynomials of the family provided better performance in terms of WNG and Directivity Index, while others provided better source separation for sources placed near the array surface. The results are motivating enough to be considered in medium sized spherical arrays with upto fifty microphones and in speech and audio applications where source and interference are directionally co-incident. Future work will explore constrained optimisation formulations for improving the gain of the radial response in the look radius.

7. REFERENCES

[1] Jens Meyer and Gary Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*. IEEE, 2002, vol. 2, pp. II–1781.

[2] Yahong Rosa Zheng, Rafik Goubran, Mohamed El-Tanany, et al., "Robust near-field adaptive beamforming with distance discrimination," *Speech and Audio Processing, IEEE Transactions on*, vol. 12, no. 5, pp. 478–488, 2004.

[3] Lalan Kumar, Kushagra Singhal, and Rajesh M Hegde, "Near-field source localization using spherical microphone array," in *Hands-free Speech Communication and Microphone Arrays (HSCMA), 2014 4th Joint Workshop on*. IEEE, 2014, pp. 82–86.

[4] Thushara D Abhayapala, Rodney Kennedy, Robert C Williamson, Darren B Ward, et al., "Nearfield broadband adaptive beamforming," in *Signal Processing and Its Applications, 1999. ISSPA'99. Proceedings of the Fifth International Symposium on*. IEEE, 1999, vol. 2, pp. 839–842.

[5] Rodney Kennedy, Thushara D Abhayapala, Darren B Ward, et al., "Broadband nearfield beamforming using a radial beampattern transformation," *Signal Processing, IEEE Transactions on*, vol. 46, no. 8, pp. 2147–2156, 1998.

[6] Thushara D Abhayapala, Rodney A Kennedy, and Robert C Williamson, "Nearfield broadband array design using a radially invariant modal expansion," *The Journal of the Acoustical Society of America*, vol. 107, no. 1, pp. 392–403, 2000.

[7] Etan Fisher and Boaz Rafaely, "The near-field spherical microphone array," in *Proceedings of Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*. IEEE, 2008, pp. 5272–5275.

[8] Etan Fisher and Boaz Rafaely, "Dolph-chebyshev radial filter for the near-field spherical microphone array," in *Applications of Signal Processing to Audio and Acoustics, 2009. WASPAA'09. IEEE Workshop on*. IEEE, 2009, pp. 169–172.

[9] Etan Fisher and Boaz Rafaely, "Near-field spherical microphone array for speech," *Transformation*, vol. 10, pp. 17, 2010.

[10] Etan Fisher and Boaz Rafaely, "Near-field spherical microphone array processing with radial filtering," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 19, no. 2, pp. 256–265, 2011.

[11] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, Eds., *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, NY, 2010, Print companion to [22].

[12] I. Cohen, J. Benesty, and S. Gannot, "Microphone arrays: Fundamental concepts," *Speech Processing in Modern Communication: Challenges and Perspectives*, vol. 3, 2010.

[13] Earl G Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*, Academic press, 1999.

[14] Boaz Rafaely, "Analysis and design of spherical microphone arrays," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 1, pp. 135–143, 2005.

[15] Thushara D Abhayapala and Darren B Ward, "Theory and design of high order sound field microphones using spherical microphone array," in *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*. IEEE, 2002, vol. 2, pp. II–1949.

[16] Zhiyun Li and Ramani Duraiswami, "Flexible and optimal design of spherical microphone arrays for beamforming," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 15, no. 2, pp. 702–714, 2007.

[17] Boaz Rafaely, Barak Weiss, and Eitan Bachmat, "Spatial aliasing in spherical microphone arrays," *Signal Processing, IEEE Transactions on*, vol. 55, no. 3, pp. 1003–1010, 2007.

[18] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, Oxford, U.K.: Elsevier Academic Press, 2005.

[19] Harry L Van Trees, *Detection, estimation, and modulation theory, optimum array processing*, John Wiley & Sons, 2004.

[20] Israel Cohen, Jacob Benesty, and Sharon Gannot, *Speech processing in modern communication: challenges and perspectives*, vol. 3, Springer Science & Business Media, 2009.

[21] Jens Meyer and Gary W Elko, "Analysis of the high-frequency extension for spherical eigenbeamforming microphone arrays.," *The Journal of the Acoustical Society of America*, vol. 127, no. 3, pp. 1979–1979, 2010.

[22] "NIST Digital Library of Mathematical Functions," http://dlmf.nist.gov/, Release 1.0.10 of 2015-08-07, Online companion to [11].