THEORY AND DESIGN OF HIGH ORDER SOUND FIELD MICROPHONES USING SPHERICAL MICROPHONE ARRAY

Thushara D. Abhayapala, *

Department of Engineering, FEIT &
Department of Telecom Eng, RSISE
The Australian National University
Thushara.Abhayapala@anu.edu.au

Darren B. Ward

Department of EEE
Imperial College
London SW7 2BT, UK
d.ward@ic.ac.uk

ABSTRACT

A major problem in sound field reconstruction systems is how to record the higher order (> 1) harmonic components of a given sound field. Spherical harmonics analysis is used to establish theory and design of a higher order recording system, which comprises an array of small microphones arranged in a spherical configuration and associated signal processing. This result has implications to the advancement of future sound field reconstruction systems. An example of a third order system for operation over a 10:1 frequency range of 340 Hz to 3.4 kHz is given.

1. INTRODUCTION

In three-dimensional (3D) audio systems, the aim is to give one or more listeners the impression of being immersed in a realistic acoustic environment. This requires recording a sound field in a given environment (e.g., a musical show) and reproducing it accurately over a certain region of space. Such a recording should capture the sound field not only at a single point but over the desired region of reconstruction and over the entire audio frequency band. This is the problem to be addressed in this paper.

Sound field reconstruction methods such as ambisonics [1] are based on measuring the *spherical harmonic* composition of a given sound field. Standard omni-directional microphones can record only the zero order harmonic of a sound field, and there are commercially available first order directional microphones [2], which can be used to measure up to first order spherical harmonics of a sound field. A system for recording 2D soundfields has recently been presented in [3]. Until now, however, the technology has not been advanced enough to record higher order harmonics, which are necessary to reproduce a 3D soundfield accurately over a region of space. In this paper, we provide theory and guidelines to design higher order microphones.

*This work is partially supported by the ANU Faculties New Starters Grants Scheme.

Notation

Throughout this paper we use the following notational conventions: vectors are represented by lower case bold face, e.g., $\mathbf{x}$. A unit vector in the direction $\mathbf{x}$ is denoted by $\hat{\mathbf{x}}$, i.e., $\hat{\mathbf{x}} = \mathbf{x}/\|\|\mathbf{x}\|\|$. The symbol $i = \sqrt{-1}$ is used to denote the imaginary part of a complex number.

2. HARMONIC REPRESENTATION OF A SOUND FIELD

2.1. Background

Consider a region Ω in the space and assume all sound sources are located outside of this region. Then the sound field at a point $\mathbf{x} = x[\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta]^T$ within the region Ω and frequency f is given by [4]

$$S(\mathbf{x}; f) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \gamma_{nm}(f)\, j_n(\frac{2\pi}{c} f x)\, Y_{nm}(\hat{\mathbf{x}}), \quad (1)$$

where $x = \|\|\mathbf{x}\|\|$, c is the speed of wave propagation, $j_n(\cdot)$ is the nth order spherical Bessel function of the first kind, $Y_{nm}(\cdot)$ are *spherical harmonics*, and $\gamma_{nm}(f)$ are a set of harmonic coefficients, which do not depend on radial and angular information of the point $\mathbf{x}$.

The spherical harmonics are defined as [5, p.194]

$$Y_{nm}(\hat{\mathbf{x}}) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}}\, P_{n|m|}(\cos\theta)\, e^{im\phi}, \quad (2)$$

where θ and ϕ are the elevation and azimuth angles of $\hat{\mathbf{x}}$, respectively, and $P_{nm}(\cdot)$ is the associated Legendre function (which reduces to the Legendre function for $m = 0$). The subscript n is referred to as the *order* of the spherical harmonic, and m is referred to as the *mode*. For each order n, there are $2n + 1$ modes (corresponding to $m = -n, \ldots, n$). Spherical harmonics exhibit the following orthogonality property [5, p.191],

$$\int Y_{nm}^*(\hat{\mathbf{x}})Y_{pq}(\hat{\mathbf{x}})d\hat{\mathbf{x}} = \delta_{np}\delta_{mq}, \tag{3}$$

where δ_{np} denotes the Kronecker delta function and the integration is over the unit sphere.

The representation (1) captures any arbitrary sound field due to both plane wave (farfield) sources and spherical wave (nearfield) sources. Note that the harmonic coefficients $\gamma_{nm}(f)$ are independent of location. Thus, if we can record them, the corresponding sound field can be reconstructed accurately. In our earlier work [4], we have shown how to reconstruct a sound field given a set of harmonic coefficients using an array of loudspeakers. In the sequel, we will investigate how to record the harmonic coefficients using a spherical array of microphones.

2.2. Soundfield Decomposition

Mathematically, we can calculate the harmonic coefficients $\gamma_{nm}(f)$ for frequency f by using (1) and (3) as

$$\gamma_{nm}(f) = \frac{1}{j_n(\frac{2\pi}{c}xf)} \int S(\mathbf{x};f)Y_{nm}^*(\hat{\mathbf{x}})d\hat{\mathbf{x}} \tag{4}$$

where the integration is over the unit sphere and $S(\mathbf{x};f)$ is measured on a sphere of radius x. Equation (4) is valid only if $j_n(\frac{2\pi}{c}xf) \neq 0$ for the frequency of interest f and x. We will revisit this restriction when we consider placement of microphones to realize (4). We can use (4) to decompose a sound field in a given region into a set of harmonic coefficients which describe the sound field at every point of the given region.

A higher order sound field microphone must be able to extract harmonic coefficients $\gamma_{nm}(f),\ n > 0,\ m = -n,\cdots,n$ from the sound field surrounding it. Naturally, an omni-directional microphone records the zero order harmonic coefficient $\gamma_{00}(f)$. The challenge is how to extract higher order harmonic coefficients from the sound field. Our approach is to realize (4) using an array of omni-directional microphones in a suitable 3-dimensional configuration.

3. SPHERICAL MICROPHONE ARRAY

3.1. Approximation

To facilitate a practical realization, we may approximate the integration in (4) by a finite summation. Consider Q omni-directional microphones placed on the surface of a sphere of radius R, then we can obtain the sound field measurements $S(R\hat{\mathbf{x}}_q;f)$, $q = 1,\cdots,Q$. Thus we may approximate (4) by

$$\tilde{\gamma}_{nm}(f) = \frac{1}{j_n(\frac{2\pi}{c}Rf)} \sum_{q=1}^{Q} S(R\hat{\mathbf{x}}_q;f)Y_{nm}^*(\hat{\mathbf{x}}_q)w_q, \tag{5}$$

where w_q, $q = 1,\cdots,Q$ are a set of suitable weights.

The above approximation provides a method to extract higher order harmonic coefficients $\gamma_{nm}(f)$ of the sound field using a spherical array of microphones. However, several important theoretical and practical questions naturally arise from this approximation. These are addressed below.

3.2. Number of Microphones

We can observe from (1) that a function $S(R\hat{\mathbf{x}};f)$ of a highest nonzero order N on a surface of a sphere has $(N+1)^2$ independent harmonic components. Therefore, we should be able to sample a sound field of order N, with at least $(N+1)^2$ points on a sphere of radius R without losing information. Thus, a constraint on the number of microphones is

$$Q \geq (N+1)^2. \tag{6}$$

We can place microphones in a number of different configurations. One possibility is to place them on an equiangular grid in elevation (θ) and azimuth (ϕ) directions. However, this will result in more dense packing near poles. In such an arrangement, $(N+1)^2$ microphones are not sufficient to reproduce the sound field by its samples. For equiangular configuration, there have to be $(2N-1)$ points in both elevation and azimuth directions, hence a minimum of $(2N-1)^2$ microphones are necessary [6].

Since we could like to use as few microphones as possible, intuitively we need to place them on the sphere in an equidistance to each other. The question of how to place Q points on the sphere in number of different optimum criterions is partially answered in [7, 8]. There is no general direct mathematical formula available to find these points, however, although numerical co-ordinates are available. Since these available coordinates of points on the sphere are not accurately equidistance to each other but optimum in some way, we may have to use a higher Q than the minimum number given by (6).

3.3. Radius of Sphere and Frequency Band

We need to choose the radius of the sphere R such that

$$j_n(\frac{2\pi}{c}fR) \neq 0 \text{ for } n = 0,\cdots,N \text{ and } f \in [f_l : f_u] \tag{7}$$

where f_l and f_u are the lower and upper frequencies of the bandwidth of interest, and

$$j_n(\frac{2\pi}{c}fR) = 0 \text{ for } n > N_1 > N \text{ and } f \in [f_l : f_u]. \tag{8}$$

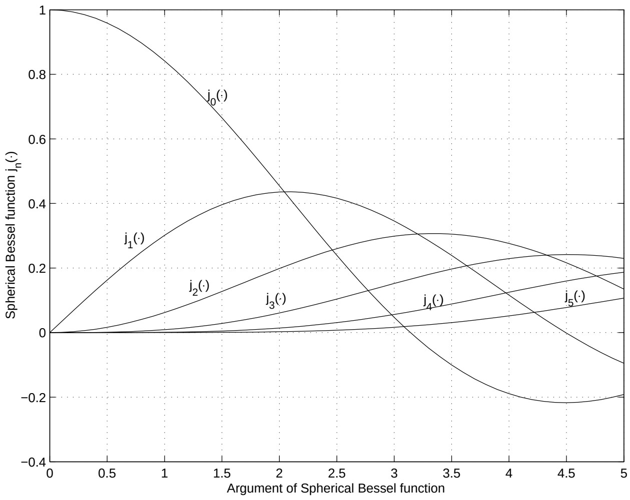

Fig. 1. Plots of spherical Bessel functions of orders $n = 0, 1, 2, 3, 4, 5$. The Bessel functions have bandpass character for orders $n > 0$.

Constraint (7) is necessary to have valid harmonic coefficients (5), and the constraint (8) guarantees that the sound field $S(R\hat{\mathbf{x}}; f)$ is order limited to N_1. The second constraint enables us to use the minimum $Q_{\text{min}} = (N_1 + 1)^2$ microphones to sample the sound field accurately.

Figure 1 depicts the spherical Bessel functions of few orders. Observe that the spherical Bessel functions have a spatial bandpass character for orders $n > 0$, and initially they are approximately zero. Also note that the higher the order n the larger the initial zero value region of $j_n(\cdot)$. This shows qualitatively why we can approximately satisfy both constraints (7) and (8) for suitable values of R, N and $N1$.

4. ERROR ANALYSIS

In general, constraint (8) cannot be satisfied exactly. One must therefore consider what error is involved in calculating the harmonic coefficients if the soundfield is not order limited on the measurement sphere of radius R.

Consider a general soundfield $S(\mathbf{x}; f)$ that can be represented by (1). Let $S_{0:N}(\mathbf{x}; f)$ denote the component of this soundfield consisting only of terms up to order N, i.e.,

$$S_{0:N}(\mathbf{x}; f) = \sum_{n=0}^{N} \sum_{m=-n}^{n} \gamma_{nm}(f)\, j_n(\frac{2\pi}{c} f x)\, Y_{nm}(\hat{\mathbf{x}}). \quad (9)$$

The true soundfield can then be written as $S(\mathbf{x}; f) = S_{0:N}(\mathbf{x}; f) + S_{N+1:\infty}(\mathbf{x}; f)$. Assume that the approximation (5) is exact[1] if the soundfield is only composed of terms up to order N, i.e.,

$$\begin{aligned}\tilde{\gamma}_{nm}(f) &= \frac{1}{j_n(\frac{2\pi}{c} R f)} \sum_{q=1}^{Q} S_{0:N}(R\hat{\mathbf{x}}_q; f) Y^*_{nm}(\hat{\mathbf{x}}_q) w_q \\ &= \frac{1}{j_n(\frac{2\pi}{c} x f)} \int S_{0:N}(\mathbf{x}; f) Y^*_{nm}(\hat{\mathbf{x}}) d\hat{\mathbf{x}} = \gamma_{nm}(f),\end{aligned}$$

where γ_{nm} are the true harmonic coefficients up to order N. However, in calculating $\tilde{\gamma}_{nm}(f)$ one does not have access to the order limited soundfield $S_{0:N}(\mathbf{x}; f)$, but only to the true soundfield $S(\mathbf{x}; f)$. Thus, the coefficients that one can actually measure are given by

$$\begin{aligned}\tilde{\gamma}_{nm}(f) &= \frac{1}{j_n(\frac{2\pi}{c} R f)} \sum_{q=1}^{Q} \Big(S_{0:N}(R\hat{\mathbf{x}}_q; f) \\ &\qquad + S_{N+1:\infty}(R\hat{\mathbf{x}}_q; f) \Big) Y^*_{nm}(\hat{\mathbf{x}}_q) w_q \\ &= \gamma_{nm} + \epsilon_{nm}, \end{aligned} \quad (10)$$

where ϵ_{nm} represents the *order-limiting error*. It is given by

$$\begin{aligned}\epsilon_{nm} &= \frac{1}{j_n(\frac{2\pi}{c} R f)} \sum_{q=1}^{Q} Y^*_{nm}(\hat{\mathbf{x}}_q) w_q \\ &\quad \times \sum_{p=N+1}^{\infty} \sum_{s=-p}^{p} \gamma_{ps}(f)\, j_p(\frac{2\pi}{c} R f) Y_{ps}(\hat{\mathbf{x}}_q) \end{aligned} \quad (11)$$

Observe that the order-limiting error in the lower order coefficients therefore depends on the true values of the higher order coefficients. Although these will in general be unknown, one can conclude that provided the higher order spherical Bessel functions $j_n(\cdot), n = N + 1, \ldots, \infty$, are of significantly smaller amplitude than the lower order spherical Bessel functions on the measurement sphere, the order-limiting error will be relatively small. However, if we can have a very large number of microphones or a 'continuous' microphone, then the order-limiting error can be made arbitrary small or eliminated. That is,

$$\begin{aligned}\lim_{Q\to\infty} \epsilon_{nm} &= \frac{1}{j_n(\frac{2\pi}{c} R f)} \sum_{p=N+1}^{\infty} \sum_{s=-p}^{p} \gamma_{ps}(f)\, j_p(\frac{2\pi}{c} R f) \\ &\quad \times \int Y_{ps}(\hat{\mathbf{x}}) Y^*_{nm}(\hat{\mathbf{x}}) d\mathbf{x} = 0,\end{aligned} \quad (12)$$

where the integration vanishes due the orthogonality property of the spherical harmonics (3). Further, quantitative error analysis will be presented in a future journal paper.

5. DESIGN EXAMPLE

To demonstrate application of the above theory, we now present a specific design example. We will consider a third

[1]The approximation error could be made arbitrarily small by choosing an appropriate number of microphones and their positions, or by using some form of 'continuous' microphone that better implements the required integration.

Order (n)	Argument
0	0.89
1	1.47
2	2.05
3	2.63
4	3.21
5	3.78
6	4.36
7	4.94
8	5.51

Table 1. Argument of spherical Bessel function for each order such that higher order terms are at least 10 dB down.

order system (i.e., it will record all orders $n = 0, \dots, 3$) for operation over a 10:1 frequency range of 340 Hz to 3.4 kHz.

Because constraint (8) cannot be satisfied exactly, in practice one must choose an appropriate radius for the desired frequency range based on the bandpass characteristic of the spherical Bessel functions. We shall choose the maximum radius as corresponding to the argument of the spherical Bessel function for which the next highest order is at least 10 dB below the maximum desired order (which should ensure that the order-limited error is small). It is straightforward to calculate these from Fig. 1, and the results are shown in Table 1 for systems up to eighth order.

Thus we see that if we restrict the argument of the spherical Bessel function to 2.63, then the fourth order terms will be down by at least 10 dB (and higher order terms are in fact down by more than 20 dB). Thus, although the system is not strictly limited to third order, higher order terms are negligible and thus will not unduly effect the approximation of the lower order terms. For the chosen frequency range, with a speed of wave propagation of $c = 342$ m/s, we therefore choose to place the microphones on a sphere of radius $R = 4$ cm. Theoretically, a third order system should only require $(N+1)^2 = 16$ microphones. More accurate results are obtained, however, if one uses extra microphones. We will therefore use a total of $(N+2)^2 = 25$ microphones placed on the surface of the 4 cm sphere at locations determined by [8]; appropriate values of the integration weights w_q to be used in (5) are also provided in [8].

As an example we considered the recording of a plane-wave soundfield using the designed third order microphone. The soundfield[2]

$$S(\mathbf{x}; f) = e^{i 2\pi f / c (\hat{\mathbf{y}}^T \mathbf{x})},$$

where $\hat{\mathbf{y}}$ is the direction of plane-wave incidence, was sampled by the spherical microphone array. Then the spherical harmonics coefficients were calculated according to (5),

[2]This is an infinite order soundfield.

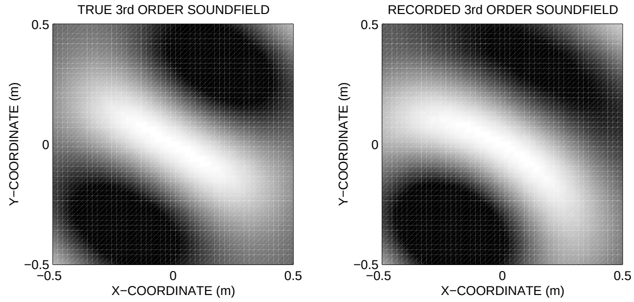

Fig. 2. True 3rd order soundfield, and recorded soundfield, for a 500 Hz plane wave.

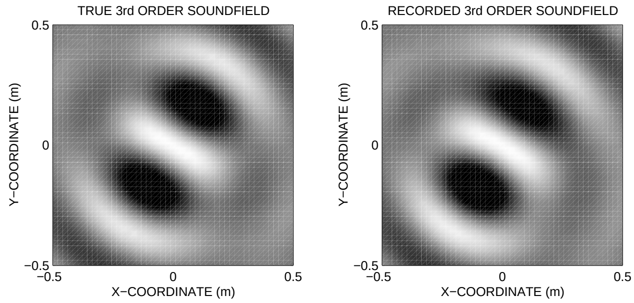

Fig. 3. True 3rd order soundfield, and recorded soundfield, for a 1 kHz plane wave.

where $S(R\hat{\mathbf{x}}_q; f)$ is the signal of frequency f sampled at microphone q. Using these approximated spherical harmonics coefficients, we then calculated the recorded soundfield according to (9), and compared it to the true 3rd order soundfield. Results of simulations at four separate frequencies are shown in Figs. 2 to 5.

We note that the spherical microphone array is able to accurately record the third order soundfield over the desired frequency range. The accuracy is worse at low frequencies (see Fig. 2), since at 500 Hz, where the argument of the spherical Bessel function on the measurement sphere is 0.37, we notice from Fig. 1 that the 3rd order components are very small. At high frequencies, although the recorded soundfield is very close to the true 3rd order soundfield, to accurately reproduce the full plane-wave soundfield[3] requires orders above $n = 3$. This is a fundamental property of soundfields, and can only be rectified by designing higher order soundfield microphones. Because of the limited frequency range over which a given spherical microphone array will be able to accurately record a soundfield of given order, we believe that higher order soundfield microphones will require arrays that are positioned on spheres of several different radii. This is currently a topic of ongoing research.

[3]This is especially true at locations far from the origin of the spherical microphone array.

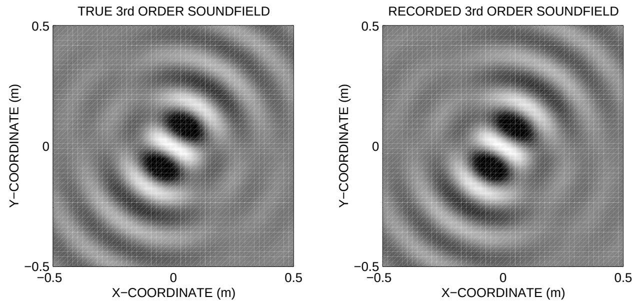

Fig. 4. True 3rd order soundfield, and recorded soundfield, for a 2 kHz plane wave.

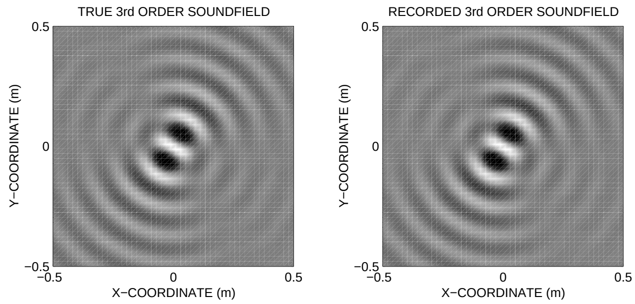

Fig. 5. True 3rd order soundfield, and recorded soundfield, for a 3 kHz plane wave.

6. CONCLUSION

We have established the theory and demonstrated the practical design of a system for recording higher order soundfields. This system is implemented using a spherical microphone array. Simulation results show that the higher order soundfield microphone is able to accurately record a soundfield over an extended region of space within a desired frequency range.

7. REFERENCES

[1] M.A. Gerzon, "Periphony: With-height sound reproduction," *J. Audio Eng. Soc.*, vol. 21, no. 1, pp. 2–10, 1973.

[2] Soundfield Research, *SoundField: An Introduction*, [Online]. Available: www.soundfield.com/feature.htm.

[3] M.A. Poletti, "A unified theory of horizontal holographic sound systems", *J. Audio Eng. Soc.*, vol. 48, no. 12, pp. 1155–1182, Dec. 2000.

[4] D.B. Ward and T.D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans. SAP*, vol. 9, no. 6, pp. 697–707, Sept. 2001.

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

[6] D. Healy, D. Rockmore, and S. Moore, "An FFT for the 2-sphere and applications," *Proc. ICASSP-96*, vol. 3, pp. 1323–1326, 1996.

[7] N. J. A. Sloane, R. H. Hardin, and W. D. Smith et al., *Tables of spherical codes*, [Online]. Available: www.research.att.com/ njas/packings/.

[8] J. Fliege, *Integration nodes for sphere*, [Online]. Available: www.mathematik.uni-dortmund.de/lsx/fliege/nodes.html.