Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers

Darren B. Ward, *Member, IEEE*, and Thushara D. Abhayapala, *Student Member, IEEE*

***Abstract*—Reproduction of a sound field is a fundamental problem in acoustic signal processing. In this paper, we use a spherical harmonics analysis to derive performance bounds on how well an array of loudspeakers can recreate a three-dimensional (3-D) plane-wave sound field within a spherical region of space. Specifically, we develop a relationship between the number of loudspeakers, the size of the reproduction sphere, the frequency range, and the desired accuracy. We also provide analogous results for the special case of reproduction of a two-dimensional (2-D) sound field. Results are verified through computer simulations.**

***Index Terms*—Acoustic signal processing, loudspeaker arrays, sound field reproduction, spherical harmonics, 3-D audio.**

I. INTRODUCTION

A FUNDAMENTAL signal processing problem in acoustics is to control the sound field within a given region of space. Reproducing a particular sound field using an array of loudspeakers has application in three-dimensional (3-D) audio systems [1], where the aim is to give one or more listeners the impression of being immersed in a realistic, yet virtual, sound environment. In this paper, we derive performance bounds on how well an array of loudspeakers can recreate a 3-D sound field in free space.

There have been various studies of sound field reproduction. One of the first was by Gerzon [3], in which he proposes using spherical harmonics as a means to represent and reproduce a sound field. The outcome of Gerzon's work was the *ambisonics* system, probably the best known of sound field reproduction systems. Ambisonics effectively recreates a first-order spherical harmonics representation of a sound field at a single point in space. Higher order ambisonics representations have also been considered [3]. More recently, ambisonics has been related to the Kirchhoff–Helmholtz theorem [4] and shown to be equivalent to a Taylor series expansion of the sound field [5]. In another approach to sound field reproduction, known as *wavefield synthesis*, holographic techniques are used to reproduce a desired sound field over a relatively large area using a large number of loudspeakers [6]. Least squares techniques were used in [7] and [8] to reproduce a sound field locally using only a few loudspeakers. The two-dimensional (2-D) Fourier transform has recently been used to formulate a system for both the recording and reproduction of 2-D sound fields [9].

Here we seek to develop some fundamental performance limits for the specific problem of reproducing a plane-wave sound field in free space.[1] Specifically, we develop a relationship between the number of loudspeakers, the size of the reproduction area, the frequency range, and the desired reproduction accuracy. Since we only deal with the free field case (i.e., we ignore the effect of reverberation) the relationships we derive effectively provide an upper bound on the performance that one could expect to achieve in a real room.

The paper is organized as follows. In Section II we outline the problem addressed. Spherical harmonics analysis is described in Section III and used to derive the two main theoretical results of the paper: 1) conditions for exact reproduction of a sound field within a sphere (which would require an infinite number of loudspeakers to achieve); and 2) consideration of the error involved with using an approximation to the ideal reproduction. Practical issues relating to design of the loudspeaker array are considered in Section IV. Analogous results for the more specific case of 2-D sound field reproduction are presented in Section V. Finally, we present simulations to verify the theoretical results derived.

Notation

Throughout this paper, we use the following notational conventions: matrices and vectors are represented by upper and lower case bold face, respectively, e.g., $\mathbf{X}$ and $\mathbf{x}$. 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.

II. PROBLEM FORMULATION

Consider a plane-wave incident from the arbitrary direction $\hat{\mathbf{y}} = [\sin\vartheta\cos\varphi,\, \sin\vartheta\sin\varphi,\, \cos\vartheta]^T$, as shown in Fig. 1. The sound field produced at an arbitrary observation point $\mathbf{x} = x[\sin\theta\cos\phi,\, \sin\theta\sin\phi,\, \cos\theta]^T$ is

$$S(\mathbf{x};k) = e^{ikx(\hat{\mathbf{y}}^T\hat{\mathbf{x}})} \tag{1}$$

where $k = 2\pi f c^{-1}$ is the wavenumber (with c the speed of wave propagation and f the frequency), and $x = |\mathbf{x}|$. We assume that

Manuscript received August 8, 2000; revised April 19, 2001. This work was supported by the Australian Research Council. The associate editor coordinating the review of this paper and approving it for publication was Dr. Michael S. Brandstein.

D. B. Ward was with the Department of Telecommunications Engineering, RSISE, The Australian National University, Canberra, ACT 0200, Australia. He is now with the Department of Electrical and Electronic Engineering, Imperial College of Science, Technology, and Medicine, London SW7 2BT, U.K. (e-mail: d.ward@ic.ac.uk).

T. D. Abhayapala is with the Department of Telecommunications Engineering, RSISE, The Australian National University, Canberra, ACT 0200, Australia.

Publisher Item Identifier S 1063-6676(01)07436-3.

[1] The use of a plane-wave field is justified by noting that any general sound field can be represented as a superposition of plane waves.

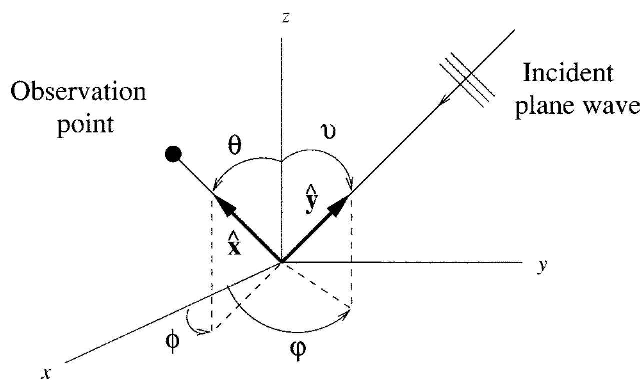

Fig. 1. Three-dimensional geometry.

c is independent of frequency, implying that the wavenumber is a constant multiple of frequency.

Assume that we want to reproduce this incident plane-wave field within some specified region χ using an array of L loudspeakers. Let the lth loudspeaker be a point source that produces a sound field

$$T_l(\mathbf{x};k) = |\mathbf{y}_l| e^{ik|\mathbf{y}_l|} \frac{e^{-ik|\mathbf{y}_l - \mathbf{x}|}}{|\mathbf{y}_l - \mathbf{x}|}, \qquad l = 1, \ldots, L \tag{2}$$

at the observation point $\mathbf{x}$, where $\mathbf{y}_l$ is the loudspeaker location. Observe that in order to relate the point-source field to the plane-wave field, we have included a normalization term $|\mathbf{y}_l| \exp(ik|\mathbf{y}|)$, chosen such that $T_l(\mathbf{x};k) \rightarrow \exp(ikx(\hat{\mathbf{y}}_l^T \hat{\mathbf{x}}))$ as $|\mathbf{y}_l| \rightarrow \infty$, i.e., the point-source field becomes the plane-wave field as the source distance goes to infinity. In the ambisonics formulation, it is assumed that the loudspeakers are far enough away that they may be considered to be plane-wave sources. Here we keep the more general point-source form.

Applying a complex-valued frequency-dependent weighting function $a_l(k)$ to the lth loudspeaker,[2] the total field at the observation point $\mathbf{x}$ due to the loudspeaker array is given by

$$\begin{aligned} T(\mathbf{x};k) &= \sum_{l=1}^{L} a_l(k) T_l(\mathbf{x};k) \\ &= \sum_{l=1}^{L} a_l(k) |\mathbf{y}_l| e^{ik|\mathbf{y}_l|} \frac{e^{-ik|\mathbf{y}_l - \mathbf{x}|}}{|\mathbf{y}_l - \mathbf{x}|}. \end{aligned} \tag{3}$$

One way to find the array weights is to minimize the least-squares error between the incident field and the reproduced field within the region χ, i.e.,

$$\min_{\mathbf{x} \in \chi} |S(\mathbf{x};k) - T(\mathbf{x};k)|^2.$$

This can be formulated as a standard numerical optimization problem by choosing a set of sampling points within the region χ, and then minimizing the least-squares error at these points. Such an approach was considered in [8]. Although this approach is a useful design technique in specific cases, as with any numerical method it provides little insight into fundamental issues of feasibility. To address such issues, in this paper we will tackle the reproduction problem by using spherical harmonics analysis.

We will define the reproduction region χ as being bounded by a sphere of some specified radius, centered on the origin. Through our analysis, we primarily seek to answer the following fundamental question: *What is the minimum number of loudspeakers that can reproduce, up to a given accuracy, the 3-D sound field generated within a sphere of given radius due to a plane-wave source of given frequency?*

III. Spherical Harmonics Analysis of Sound Fields

A. Background

At the physical level, the sound field within a given region of space is characterized by the classical wave equation. The general solution to the wave equation in the spherical coordinate system can be decomposed into *spherical harmonics*, which form an orthogonal basis set for the representation of an arbitrary wave field. Specifically, any arbitrary sound field at a point $\mathbf{x}$ and wavenumber k can be represented as

$$g(\mathbf{x};k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} G_{nm}(x;\, k) Y_{nm}(\hat{\mathbf{x}}) \tag{4}$$

where $G_{nm}(x;\, k)$ are a set of harmonic coefficients, which do not depend on angular information of the point $\mathbf{x}$. Notice that the representation (4) is similar in spirit to the Fourier series expansion.

The spherical harmonics are defined as [10, p. 194]

$$Y_{nm}(\hat{\mathbf{x}}) = A_{nm} P_{n|m|}(\cos\theta) e^{im\phi} \tag{5}$$

where

$$A_{nm} = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} \tag{6}$$

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 [10, p. 191]:

$$\int Y_{nm}^*(\hat{\mathbf{x}}) Y_{pq}(\hat{\mathbf{x}})\, d\hat{\mathbf{x}} = \delta_{np}\delta_{mq} \tag{7}$$

where δ_{np} denotes the Kronecker delta function, and integration is over the unit sphere. Hence, the harmonic coefficients in (4) can be found by

$$G_{nm}(x;\, k) = \int Y_{nm}^*(\hat{\mathbf{x}}) g(\mathbf{x};k)\, d\hat{\mathbf{x}}. \tag{8}$$

In the sequel, we will investigate the sound field reproduction problem by representing each of the plane-wave field (1) and the loudspeaker array field (3) using a spherical harmonics expansion of the form (4). Such a decomposition of sound fields provides insight into the sound reproduction problem.

[2]In the audio engineering literature, $a_l(k)$ is often referred to as a *panning* function.

B. Spherical Harmonics Expansion

The sound field (1) at an arbitrary observation point $\mathbf{x} = x\hat{\mathbf{x}}$, produced by a plane-wave source can be represented in the form of (4) using [10, p. 227] as

$$S(\mathbf{x}; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} 4\pi X_n(kx) Y_{nm}^*(\hat{\mathbf{y}}) Y_{nm}(\hat{\mathbf{x}}) \quad (9)$$

in which we define

$$X_n(kx) \triangleq i^n j_n(kx) \quad (10)$$

where $j_n(\cdot)$ is the nth order spherical Bessel function of the first kind. The spherical Bessel function is related to the ordinary Bessel function by [10, p. 194]

$$j_n(x) = \sqrt{\frac{\pi}{2x}} J_{n+1/2}(x).$$

To obtain a valid spherical harmonics expansion for the sound field produced by the loudspeakers requires the following assumptions.

Assumption 1: The observation point $\mathbf{x} = x\hat{\mathbf{x}}$ satisfies $x < |\mathbf{y}_l|, \forall\, l$, where $\mathbf{y}_l$ is the location of the lth loudspeaker.

Assumption 2: All loudspeakers are located on a sphere of radius r, i.e., $|\mathbf{y}_l| = r, \forall\, l$.

Under these mild restrictions, and using [11, p. 30], the sound field (3) due to the loudspeaker array can be written in the form of (4) as

$$T(\mathbf{x}; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} 4\pi X_n(kx) R_n(kr) \times \sum_{l=1}^{L} a_l(k) Y_{nm}^*(\hat{\mathbf{y}}_l) Y_{nm}(\hat{\mathbf{x}}) \quad (11)$$

where

$$R_n(kr) \triangleq -ikre^{ikr} i^{-n} h_n(kr) \quad (12)$$

and [10, p. 194]

$$h_n(x) = \sqrt{\frac{\pi}{2x}} \left[J_{n+1/2}(x) - iN_{n+1/2}(x) \right]$$

is the nth order spherical Hankel function of the second kind, $J_n(\cdot)$ is the nth order Bessel function of the first kind, and $N_n(\cdot)$ is the nth order Bessel function of the second kind (also known as the Neumann function). The practical significance of Assumptions 1 and 2 is that the representation we derive is valid for points located within the convex hull spanned by the loudspeaker array; this is a situation which is almost always assured in practice.

Thus, the spherical harmonics expansion shows that to exactly reproduce a plane-wave field by a loudspeaker array requires one to find the loudspeaker weights $a_l(k)$ that equate (9) with (11). It would appear that we have gained little in using a spherical harmonics expansion, since we have only succeeded in transforming (1) and (3) into the more formidable expressions (9) and (11), respectively. In the following sections, however, we will show that this spherical harmonics expansion allows us to derive bounds on the performance of sound field reproduction.

C. Exact Reproduction

Theorem 1: Consider a plane-wave with wavenumber k, incident from an arbitrary direction $\hat{\mathbf{y}}$. An array of L point-source loudspeakers, located on a sphere of radius r, can exactly reproduce this plane-wave field at all points $x < r$ if the loudspeaker weights $a_l(k)$ satisfy

$$P_{n|m|}(\cos\vartheta)e^{-im\varphi} = R_n(kr) \sum_{l=1}^{L} a_l(k) P_{n|m|}(\cos\theta_l) e^{-im\phi_l}$$
$$n = 0, \ldots \infty, \quad m = -n, \ldots, n \quad (13)$$

where θ_l and ϕ_l, respectively, are the elevation and azimuth angles of the loudspeaker direction $\hat{\mathbf{y}}_l$, and ϑ and φ, respectively, are the elevation and azimuth angles of the source direction $\hat{\mathbf{y}}$.

Proof: Equating S in (9) with T in (11), multiplying each side by $Y_{pq}^*(\hat{\mathbf{x}})$, and integrating over the unit sphere with respect to $\hat{\mathbf{x}}$, gives

$$4\pi X_n(kx) Y_{nm}^*(\hat{\mathbf{y}}) = 4\pi X_n(kx) R_n(kr) \sum_{l=1}^{L} a_l(k) Y_{nm}^*(\hat{\mathbf{y}}_l) \quad (14)$$

where we have used the orthogonality property (7). Noting that the $4\pi X_n(kx)$ term is common to both sides of (14), substitution of (5) completes the proof. ■

We will refer to (13) as the condition for exact reproduction. Notice that the observation point $\mathbf{x}$ does not appear in this reproduction equation; in other words, if (13) is satisfied for all n and m, then the plane-wave field will be reproduced exactly by the loudspeaker array at *all* observation points $x < r$.

To satisfy (13) exactly for every term in the spherical harmonics expansion, however, would require an infinite number of loudspeakers (one for each n and m term). We therefore consider the following approximate reproduction problem.

D. Approximate Reproduction

Exact reproduction of the plane-wave sound field requires that (13) is satisfied for all orders n and all modes m. If, however, most of the power of the sound field within the chosen reproduction region is contained in the first N orders, then the plane-wave field (9) could be accurately reproduced by equating only the terms for $n = 0, \ldots, N$ in (13). We will refer to this as an Nth order reproduction system.[3]

An important practical question naturally arises as to the order of expansion, N, required to sufficiently characterize the plane-wave field. Assume the plane-wave field $S(\mathbf{x}; k)$ is approximated by a field $\hat{S}(\mathbf{x}; k)$, obtained by truncating the infinite series (9) at order N, i.e., the outer summation in

[3]In effect, the ambisonics system uses only first-order reproduction.

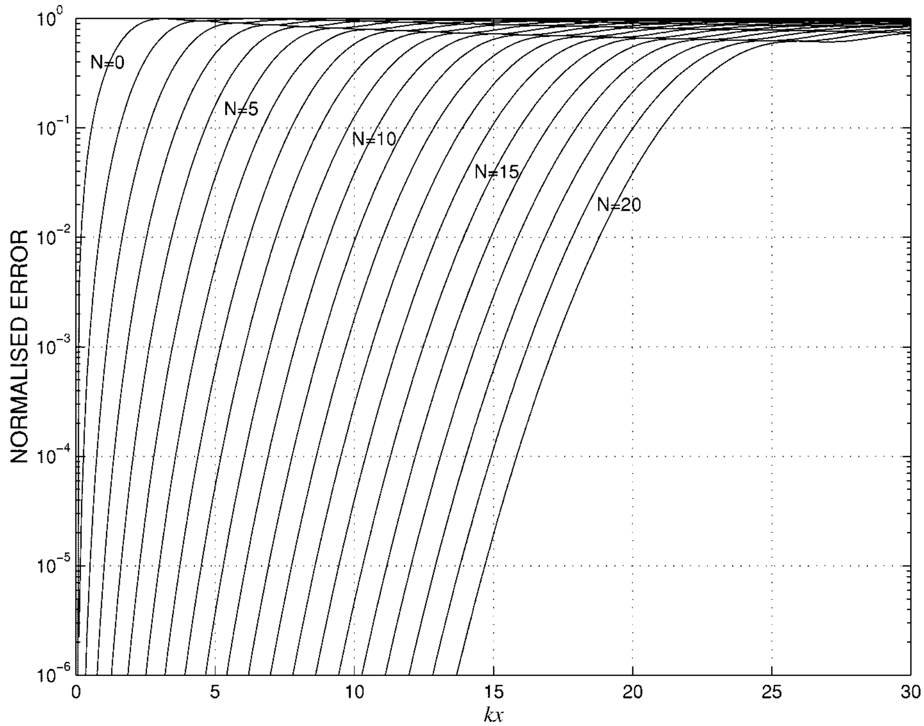

Fig. 2. Normalized truncation error (16) as a function of kx for various reproduction orders N.

(9) is only taken over $n = 0, \ldots, N$. Define the *normalized truncation error* as

$$\epsilon_N(kx) = \frac{\int |S(\mathbf{x};k) - \hat{S}(\mathbf{x};k)|^2 \, d\hat{\mathbf{x}}}{\int |S(\mathbf{x};k)|^2 \, d\hat{\mathbf{x}}} \tag{15}$$

where integration is taken over the unit sphere. This is the normalized error associated with using a finite order truncation of the infinite series (9). We then have the following result.

Theorem 2: Let $S(\mathbf{x};k)$ be the field produced by a plane-wave source of wavenumber k as measured on a sphere of radius x, and let $\hat{S}(\mathbf{x};k)$ be the corresponding field produced by truncating the spherical harmonics expansion of $S(\mathbf{x};k)$ at order N. The normalized truncation error (15) is given by

$$\epsilon_N(kx) = 1 - \sum_{n=0}^{N} (2n+1)(j_n(kx))^2. \tag{16}$$

Proof: The proof is given in the Appendix. ■

We make the following comments regarding this result.

1) Normalized truncation error is independent of the source direction $\hat{\mathbf{y}}$.
2) Normalized truncation error depends only on the product of the wavenumber, k, and the radius of the sphere, x. Thus, for a given order N, a higher operating frequency will result in a smaller reproduction sphere.

Observe that the normalized truncation error is specified on a sphere of particular radius x. Since the problem we address is to reproduce the field within a sphere of given radius x_0 (i.e., not just on the surface of the sphere) one must ask what the error is for all spheres of smaller radius. More specifically, one would like to know whether $\epsilon_N(kx) \leq \epsilon_N(kx_0) \; \forall \, x < x_0$. In other words, if the field is accurately produced on the surface of a sphere, is it accurately produced for all points within the sphere?

To answer this question qualitatively, we show in Fig. 2 the normalized truncation error (16) as a function of kx for various order expansions. We note that for any given N, the error decreases monotonically below a certain kx. In all cases, it is only for very high errors (above about 50%) that the error is not monotonically decreasing. We therefore assert that in cases of practical interest, $\epsilon_N(kx) \leq \epsilon_N(kx_0) \; \forall \, x < x_0$. Thus, if an Nth order expansion is sufficient to accurately represent the desired plane-wave field on a sphere of radius x_0, then the field is also accurately reproduced at all points within the sphere.

This assertion also follows from the Kirchhoff–Helmholtz theorem, which states that the sound field at any point within a source-free volume is fully defined by the sound pressure and pressure gradient on the continuous surface enclosing the volume [12]. The sound pressure is S, and the pressure gradient is given by taking the derivative of S with respect to x at $x = x_0$, where x_0 is the radius of the reproduction sphere. Observe from (9) that the only term in S that depends on x is $X_n(kx) = i^n j_n(kx)$, and note that [10, p. 197]

$$\frac{d}{d\zeta} j_n(\zeta) = j_{n-1}(\zeta) - \frac{n+1}{\zeta} j_n(\zeta).$$

It follows that satisfying (13) up to the Nth order equates both the sound pressure and the pressure gradient on the sphere for $n = 0, \ldots, N$. Thus, accurately reproducing the pressure field on the surface of the sphere using spherical harmonics also accurately reproduces the pressure gradient. From the Kirchhoff–Helmholtz theorem, this thereby ensures that the pressure field within the sphere is also accurately reproduced.

In practice, the required expansion order N for a given situation can be found from Fig. 2. For example, assume we wish to reproduce a 1 kHz plane-wave field within a sphere of radius 0.5 m with a 1% error. In this case, $kx_0 = 9.24$ (assuming a wave propagation speed of 340 m/s), and the curve in Fig. 2 indicates that an expansion order of at least $N = 10$ is required for an error of 0.01. A straightforward method for determining the required order N is as follows.

Rule of Thumb: From Fig. 2, we observe that using a reproduction order N equal to the product kx gives an error of around 4% for all values of kx. Such an error should be sufficient for most practical applications. Thus, given the wavenumber k (or equivalently the frequency) and the radius of the reproduction sphere x_0, the following simple rule of thumb can be used to determine the reproduction order:

$$N = \lceil kx_0 \rceil \tag{17}$$

where $\lceil \cdot \rceil$ denotes rounding up to the nearest integer.

IV. Loudspeaker Array Design

A. Loudspeaker Weights

In the previous section we showed that a loudspeaker array can reproduce a plane-wave sound field with a normalized error of around 4%, if the array weights are chosen to satisfy (13) for $n = 0, \ldots, N$, where N is chosen according to (17).

Let

$$\mathbf{p}_{nm} = [\, P_{n|m|}(\cos\theta_1) \quad \ldots \quad P_{n|m|}(\cos\theta_L) \,]$$

and

$$\mathbf{e}_m = [\, e^{-im\phi_1} \quad \ldots \quad e^{-im\phi_L} \,]$$

be L-vectors.

To satisfy (13) for the first N orders therefore requires

$$\mathbf{Pa} = \mathbf{b} \tag{18}$$

where

$$\mathbf{P} = \begin{bmatrix} R_0(kr)[\mathbf{p}_{00} \odot \mathbf{e}_0] \\ R_1(kr) \begin{bmatrix} \mathbf{p}_{11} \odot \mathbf{e}_{-1} \\ \mathbf{p}_{10} \odot \mathbf{e}_0 \\ \mathbf{p}_{11} \odot \mathbf{e}_1 \end{bmatrix} \\ \vdots \\ R_N(kr) \begin{bmatrix} \mathbf{p}_{NN} \odot \mathbf{e}_{-N} \\ \vdots \\ \mathbf{p}_{NN} \odot \mathbf{e}_N \end{bmatrix} \end{bmatrix} \tag{19}$$

is a $K \times L$ matrix (where $\odot$ denotes an element-by-element product)

$$\mathbf{a} = [a_1(k), \ldots, a_L(k)]^T \tag{20}$$

is the L vector of array weights, and

$$\mathbf{b} = \begin{bmatrix} P_{00}(\cos\vartheta)e^{-i0} \\ P_{11}(\cos\vartheta)e^{i\varphi} \\ P_{10}(\cos\vartheta)e^{-i0} \\ P_{11}(\cos\vartheta)e^{-i\varphi} \\ \vdots \\ P_{NN}(\cos\vartheta)e^{iN\varphi} \\ \vdots \\ P_{NN}(\cos\vartheta)e^{-iN\varphi} \end{bmatrix} \tag{21}$$

is a K vector. Note that for each order n, there are $2n+1$ modes (corresponding to $m = -n, \ldots, n$), giving a total of

$$K = \sum_{n=0}^{N} (2n+1) = (N+1)^2$$

rows in each of $\mathbf{P}$ and $\mathbf{b}$. We thus have a system of K equations involving the L unknown loudspeaker weights.

B. Determining the Number of Loudspeakers

The number of loudspeakers L specifies whether the linear system (18) can be solved exactly or not. There are three cases of interest.

For an over-determined system (i.e., $K > L$), in general there will be no exact solution to (18), and the array weights would typically be found to solve the least squares problem

$$\min_{\mathbf{a}} \|\mathbf{Pa} - \mathbf{b}\|^2$$

where $\|\cdot\|$ represents the vector 2-norm. This is a well studied problem [13, p. 236]. Note that this least squares approach attempts to find the set of loudspeaker weights that can best reproduce all of the spherical harmonics for $n = 0, \ldots, N$. As we saw in the previous section, however, the lower order harmonics carry the most energy for small reproduction spheres, with higher order modes contributing more energy to larger reproduction spheres. This suggests that for an over-determined system it may be preferable to exactly reproduce as many low-order harmonics as possible, and then use a least squares approach for the higher orders (that cannot be exactly reproduced anyway). This would ensure that reproduction was accurate for the largest reproduction sphere possible using the given number of loudspeakers.

If $\mathbf{P}$ is a square nonsingular matrix, then a unique solution to (18) exists, given by $\mathbf{a} = \mathbf{P}^{-1}\mathbf{b}$. Although this solution will satisfy (18) exactly, it is somewhat of a moot point if $\mathbf{P}$ is poorly conditioned. The conditioning of $\mathbf{P}$ is determined primarily by the loudspeaker geometry, which we consider below.

Finally, when the linear system is under-determined (i.e., $K < L$), there may either be no solution or an infinite number of solutions. In the latter case it would be most appropriate to find the array weights to satisfy

$$\min_{\mathbf{a}} \|\mathbf{a}\|^2 \text{ subject to } \mathbf{Pa} = \mathbf{b}.$$

Again, this least squares problem has a well-known solution [13, p. 271].

In general the system (18) can only be satisfied exactly if $L \geq K$. Therefore, the number of loudspeakers required for exact reproduction of the plane-wave field is

$$L \geq (N+1)^2. \tag{22}$$

We note that this is also the number of independent modes in an Nth order ambisonics representation [2], [5]. In the ambisonics formulation, however, no relationship has been established between the number of modes, the accuracy of the reproduction, or the frequency.

C. Loudspeaker Geometry

Although (22) provides an expression for the minimum number of loudspeakers required, it says nothing about where the loudspeakers should be placed. We know, however, that the applicability of the solution obtained for (18) will depend on the condition number of the $\mathbf{P}$ matrix, i.e., the ratio of its largest and smallest eigenvalues. Because of the structure of $\mathbf{P}$, finding the optimum loudspeaker positions to minimize the condition number is nontrivial. We do not attempt to find such an optimum geometry here, rather we appeal to heuristics. Specifically, a well-conditioned $\mathbf{P}$ matrix should result from a geometry in which the loudspeaker locations are maximally distributed in some sense. One way to approach this is through a problem in mathematics know as the *sphere packing problem*, defined as [14]: "Place n points on a sphere in d dimensions so as to maximize the minimal distance (or equivalently the minimal angle) between them." The solution to this problem in three-dimensions can be found in the library of 3-D packings at [14], which contains the coordinates of packings for up to 130 points. In all cases that we have tried, we have found that using these coordinates results in a well-conditioned $\mathbf{P}$ matrix.

D. Summary of Results

The design procedure for the loudspeaker array can be summarized as follows.

1) The number of loudspeakers required to accurately reproduce a plane-wave with wavenumber k within a sphere of radius x_0 is

$$L \geq \left(\lceil kx_0 \rceil + 1\right)^2.$$

This number of loudspeakers guarantees that the normalized reproduction error within the sphere is approximately 4%.

2) Each loudspeaker must be placed at a distance greater than x_0 from the center of the sphere. If the loudspeakers are placed on a sphere of radius $r > x_0$, their locations can be found from [14].

3) The loudspeaker weights are chosen to satisfy (18).

V. Application to 2-D Reproduction

The spherical harmonics expansion used in the previous section is appropriate for the reproduction of a 3-D sound field within a sphere. In many practical situations, however, it is of interest to reproduce a sound field in two dimensions only, typically using an array of loudspeakers placed on a ring around a listener. Hence, we now consider the special case of reproducing a 2-D sound field in a plane. Without loss of generality we define the reproduction region as being bounded by a circle in the plane $\theta = \pi/2$, centered on the origin and with radius x_0.

In the 2-D case, the spherical harmonics expansion for the plane-wave field becomes

$$S = 4\pi \sum_{m=-\infty}^{\infty} \sum_{n=|m|}^{\infty} X_n(kx) A_{nm}^2 P_{n|m|}^2(0) \times e^{im\phi} e^{-im\varphi} \tag{23}$$

where we have exchanged the order of the summations, substituted (5), and noted that $\cos(\pi/2) = 0$ in the plane. Similarly, the field produced by the array is

$$T = 4\pi \sum_{m=-\infty}^{\infty} \sum_{n=|m|}^{\infty} X_n(kx) R_n(kr) A_{nm}^2 P_{n|m|}^2(0) \times e^{im\phi} \sum_{l=1}^{L} a_l(k) e^{-im\phi_l}. \tag{24}$$

Equating (23) with (24), multiplying both sides by $e^{iq\phi}$, and integrating over the unit circle with respect to ϕ gives

$$Q_m(kx) e^{-im\varphi} = \tilde{Q}_m(kx,\, kr) \sum_{l=1}^{L} a_l(k) e^{-im\phi_l}, \quad m = -\infty, \ldots, \infty \tag{25}$$

where we have used the orthogonality property of complex exponentials, i.e.,

$$\int_0^{2\pi} e^{-im\phi} e^{iq\phi}\, d\phi = 2\pi \delta_{mq} \tag{26}$$

and defined

$$\tilde{Q}_m(kx,\, kr) \triangleq 4\pi \sum_{n=|m|}^{\infty} X_n(kx) R_n(kr) A_{nm}^2 P_{n|m|}^2(0) \tag{27}$$

and

$$Q_m(kx) \triangleq 4\pi \sum_{n=|m|}^{\infty} X_n(kx) A_{nm}^2 P_{n|m|}^2(0) = i^m J_m(kx) \tag{28}$$

where the proof of the final equality is given in the Appendix.

Thus, for the special case of a 2-D plane-wave field,[4] the requirement for exact reproduction is given by (25). It is important to note that, unlike the 3-D case, the resulting linear system is dependent on the radial distance to the observation point x. This suggests that there is no single set of array weights that can give exact reproduction for all points within the reproduction circle. However, the Kirchhoff–Helmholtz theorem again tells us that,

[4]Here we have adapted our general result to the 2-D case. One would obtain a similar expression, however, by starting with an orthogonal expansion that is better suited to the specific 2-D case, e.g., a cylindrical expansion.

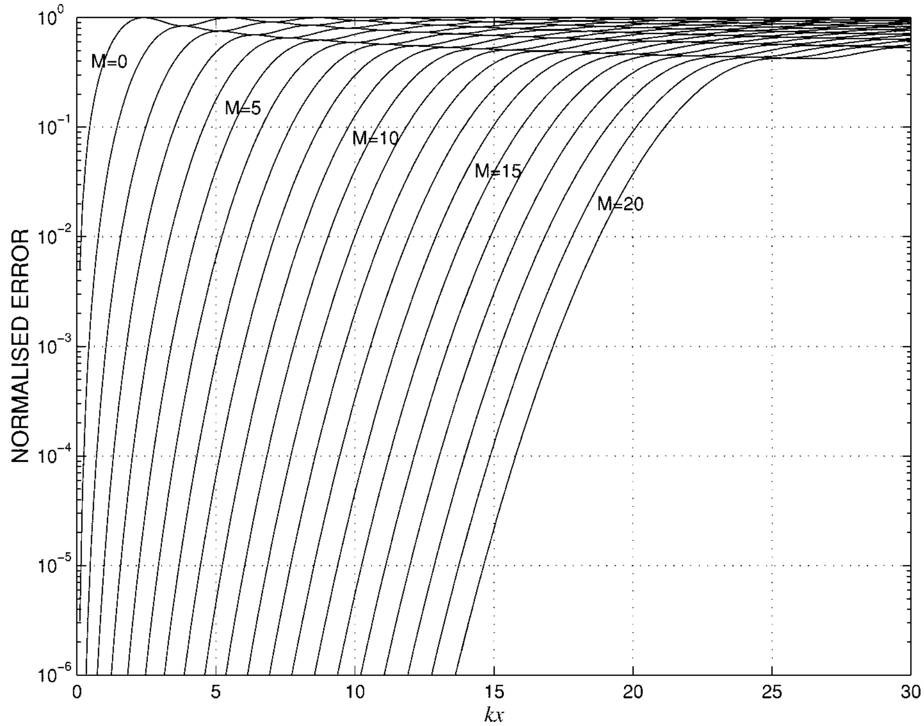

Fig. 3. Normalized truncation error (29) for a 2-D field as a function of kx for various reproduction orders M.

as long as we can satisfy (25) on the circle bounding the reproduction region, then the reproduction will also be accurate for points within this region. Thus, in solving for the array weights it is only necessary to solve (25) at $x = x_0$.

As in the 3-D case, we must consider reproduction by a finite series expansion, and we have the following results.

Proposition 1: Let S be the field produced by a 2-D plane-wave source, and let $\hat{S}$ be the corresponding field produced by truncating the series expansion of S at mode M, i.e., using modes $m = -M, \ldots, M$. Then the normalized truncation error is

$$\epsilon_M(kx) \triangleq \frac{\int |S - \hat{S}|^2 \, d\phi}{\int |S|^2 d\phi} = 1 - \sum_{m=-M}^{M} J_m(kx)^2. \tag{29}$$

Proposition 2: For exact reproduction of a 2-D plane-wave field up to the Mth mode, i.e., using modes $m = -M, \ldots, M$, the number of loudspeakers required is

$$L \geq 2M + 1. \tag{30}$$

The proofs parallel those given in Section III and are not repeated here.

The normalized truncation error (29) is shown in Fig. 3 as a function of kx for various reproduction orders M. Comparing this with the truncation error for the general 3-D case, we note that for an error less than around 10%, the curve for a given M in Fig. 3 is essentially identical to that for $N = M$ in Fig. 2. Thus, the rule of thumb for determining the required reproduction order for 2-D is the same as for 3-D, i.e.,

$$M = \lceil kx_0 \rceil \tag{31}$$

where k is the wavenumber and x_0 is the radius of the reproduction circle.

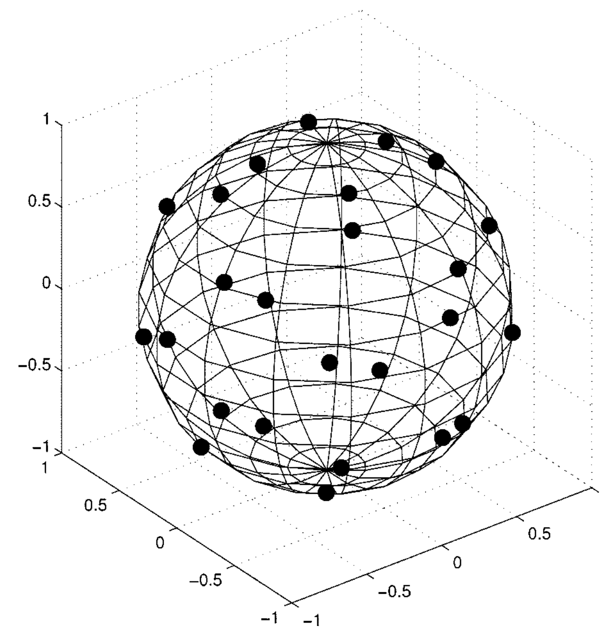

Fig. 4. Loudspeaker locations used for the examples shown in Figs. 5–7.

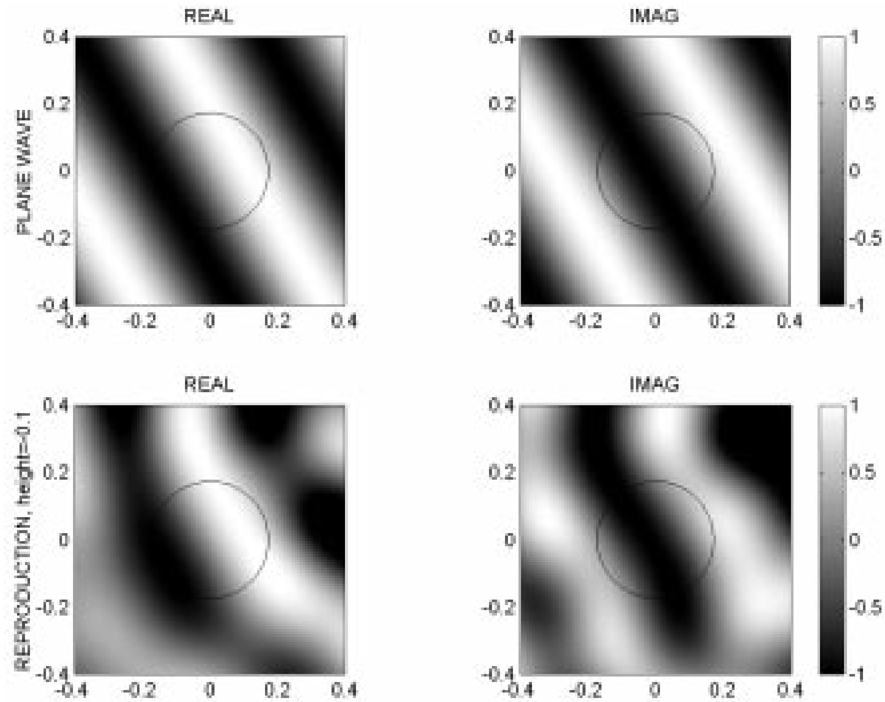

Fig. 5. Reproduction at a height of $z = -0.1$ m. Normalized reproduction error within the circle is 0.0322.

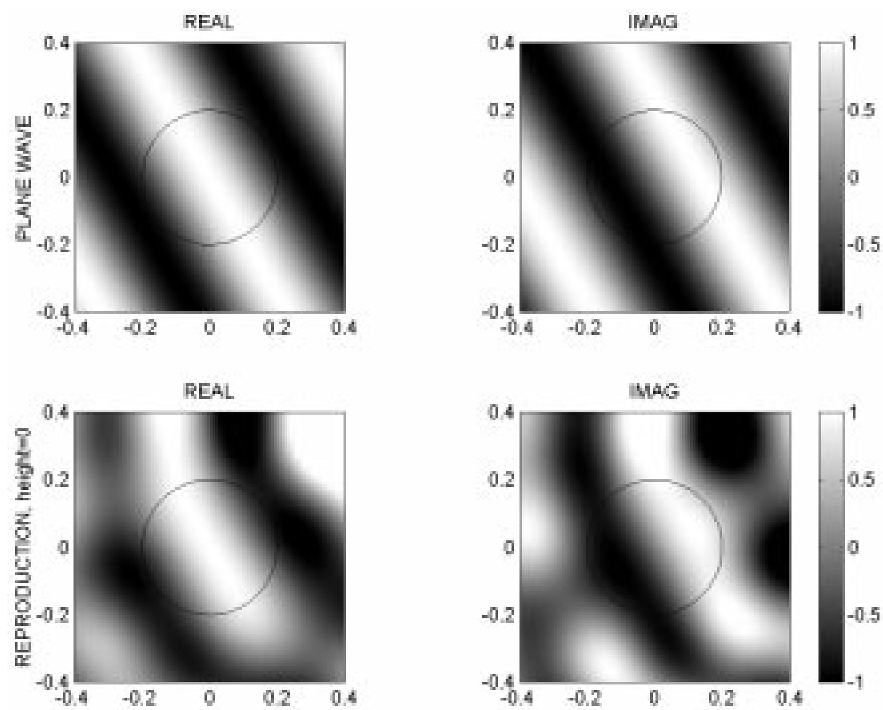

Fig. 6. Reproduction at a height of $z = 0$ m. Normalized reproduction error within the circle is 0.0359.

VI. Simulation Examples

A. Three-Dimensional Example

In the first example, we considered a monochromatic plane wave of frequency 1 kHz, incident from $[\vartheta, \varphi] = [45°, 30°]$ within a reproduction sphere of radius $x_0 = 0.2$ m. This gave $kx_0 = 3.7$, and the rule of thumb (17) suggested using $N = 4$, thus requiring $L = 25$ loudspeakers. The loudspeakers were placed on a sphere of radius $r = 1$ m, at points specified by the set "pack.3.25" in [14]. Loudspeaker locations are shown in Fig. 4. The loudspeaker weights were found from (18), and the resulting reproduced pressure fields within the sphere are shown in Figs. 5–7, displayed as slices through the sphere at heights of -0.1, 0, and 0.1 m, respectively. These figures are displayed as density plots, where acoustic pressures greater than 1 are white, pressures less than -1 are black, and pressures in

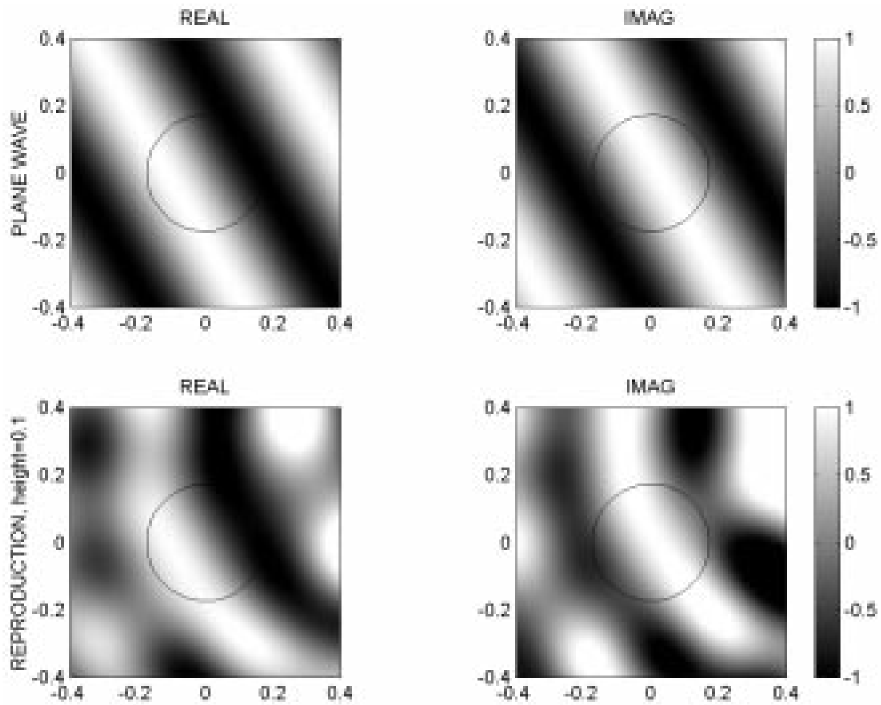

Fig. 7. Reproduction at a height of $z = 0.1$ m. Normalized reproduction error within the circle is 0.0468.

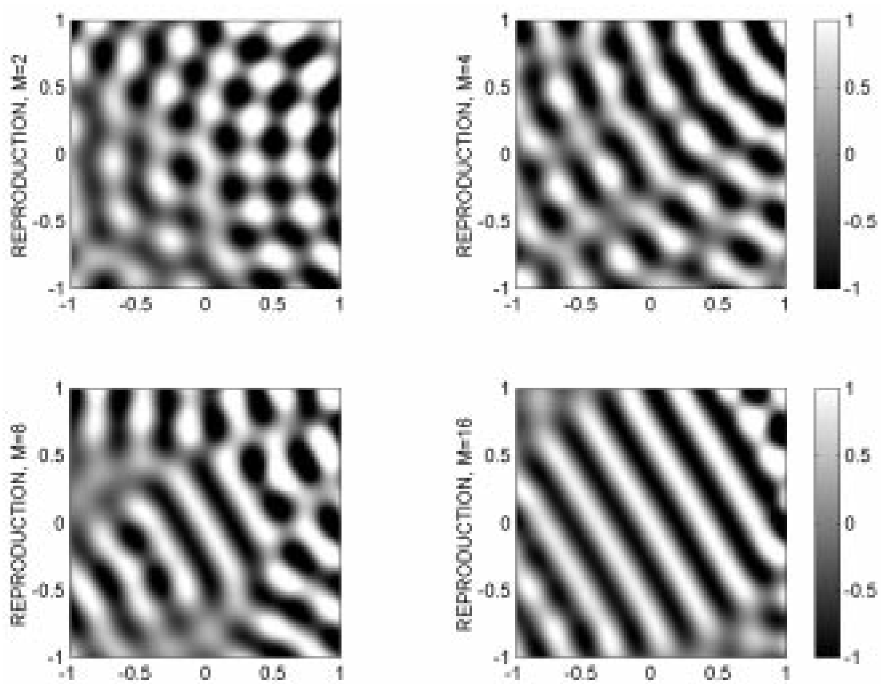

Fig. 8. Reproduction of a 2-D plane wave for various maximum modes.

between are appropriately shaded. In each figure, the top two plots show the real and imaginary parts of the ideal plane-wave field, and the bottom two plots show the field produced by the loudspeaker array. The circle indicates the boundary of the reproduction sphere at the particular height shown. The normalized reproduction error (within the sphere) in each slice ranged from 3% to 5%. These values agree very well with the expected error of 4% and thus validate our rule of thumb for choosing the required expansion order.

B. Two-Dimensional Example

In the second example we considered a monochromatic 2-D plane wave of frequency 1 kHz, incident from $\varphi = 30°$. Here we examined the effect of changing the expansion order M. Fig. 8 shows the real part of the reproduced field for $M = 2$ (top left), $M = 4$ (top right), $M = 8$ (bottom left), and $M = 16$ (bottom right). In each case we used an array of $L = (2M + 1)$ loudspeakers equally spaced on a circle of radius $r = 2$ m. Observe that as M increases, the size of the reproduction region (in which the plane wave is accurately reproduced) increases proportionally.

In closing, we briefly consider the expected performance of a five-loudspeaker array, a system that is commonly found in practice. Our results indicate that within a region of $x_0 = 0.1$ m (which is only slightly larger than the average adult human head), accurate reproduction of a 2-D plane-wave sound field will only be accurate up to around 1100 Hz.

VII. Conclusions

Spherical harmonics are a powerful tool that can be used to analyze the propagation of wave fields. Here we have used a spherical harmonics expansion of the acoustic field produced by a plane-wave source to derive performance bounds on the reproduction of this sound field within a sphere. We have shown that for a wavenumber k, the field within a sphere of radius x_0 can be accurately reproduced using a $N = \lceil kx_0 \rceil$ order expansion, requiring $L \geq (N+1)^2$ loudspeakers for a 3-D field, or $L \geq (2N+1)$ loudspeakers for a 2-D field.

Appendix

Proof of Theorem 2

Let $\hat{S}$ denote the field produced by truncating the series expansion (9) of S at order N. By definition,

$$\hat{S} = 4\pi \sum_{n=0}^{N} \sum_{m=-n}^{n} X_n(kx) Y_{nm}^*(\hat{\mathbf{y}}) Y_{nm}(\hat{\mathbf{x}}),$$

and

$$S - \hat{S} = 4\pi \sum_{n=N+1}^{\infty} \sum_{m=-n}^{n} X_n(kx) Y_{nm}^*(\hat{\mathbf{y}}) Y_{nm}(\hat{\mathbf{x}}).$$

The squared error over the unit sphere is

$$\begin{aligned} &\int |S - \hat{S}|^2 \, d\hat{\mathbf{x}} \\ &= (4\pi)^2 \sum_{n=N+1}^{\infty} \sum_{m=-n}^{n} \sum_{p=N+1}^{\infty} \sum_{q=-p}^{p} X_n(kx) X_p^*(kx) \\ &\quad \times Y_{nm}^*(\hat{\mathbf{y}}) Y_{pq}(\hat{\mathbf{y}}) \int Y_{nm}(\hat{\mathbf{x}}) Y_{pq}^*(\hat{\mathbf{x}}) \, d\hat{\mathbf{x}} \\ &= (4\pi)^2 \sum_{n=N+1}^{\infty} |X_n(kx)|^2 \sum_{m=-n}^{n} |Y_{nm}(\hat{\mathbf{y}})|^2 \end{aligned}$$

which follows from the orthogonality property (7) of the spherical harmonics.

The addition theorem of Legendre functions states that [11, p. 27]

$$\sum_{m=-n}^{n} Y_{nm}^*(\hat{\mathbf{y}}) Y_{nm}(\hat{\mathbf{x}}) = \frac{2n+1}{4\pi} P_n(\cos\gamma) \qquad (32)$$

where γ denotes the angle between $\hat{\mathbf{y}}$ and $\hat{\mathbf{x}}$. Using this addition theorem with $\hat{\mathbf{x}} = \hat{\mathbf{y}}$ [noting that $P_n(\cos 0) = 1 \; \forall n$], gives

$$\int |S - \hat{S}|^2 \, d\hat{\mathbf{x}} = 4\pi \sum_{n=N+1}^{\infty} (2n+1)(j_n(kx))^2$$

after substituting (10).

It can similarly be shown that

$$\int |S|^2 \, d\hat{\mathbf{x}} = 4\pi \sum_{n=0}^{\infty} (2n+1)(j_n(kx))^2.$$

Hence,

$$\int |S - \hat{S}|^2 \, d\hat{\mathbf{x}} = \int |S|^2 d\hat{\mathbf{x}} - 4\pi \sum_{n=0}^{N} (2n+1)(j_n(kx))^2.$$

Substitution of (1) for S shows that

$$\begin{aligned} \int |S|^2 \, d\hat{\mathbf{x}} &= \int_0^{2\pi} \int_0^{\pi} e^{ikx(\hat{\mathbf{y}}^T\hat{\mathbf{x}})} e^{-ikx(\hat{\mathbf{y}}^T\hat{\mathbf{x}})} \sin\theta \, d\theta \, d\phi \\ &= 4\pi \end{aligned}$$

thus completing the proof.

Proof of (28)

Here we show that $Q_m(kx) = i^m J_m(kx)$, as in (28). The Bessel function can be written as [15, p. 210]

$$J_m(kx) = \frac{1}{2\pi i^m} \int_0^{2\pi} e^{ikx\cos\phi} e^{im\phi} \, d\phi.$$

Using a spherical harmonics expansion, we can write [10, p. 227]

$$e^{ikx\cos\phi} = \sum_{n=0}^{\infty} i^n (2n+1) j_n(kx) P_n(\cos\phi)$$

and using (32) and (5), $P_n(\cos\phi)$ can be expressed in a series as

$$P_n(\cos\phi) = \sum_{p=-n}^{n} \frac{4\pi}{(2n+1)} A_{np}^2 P_{n|p|}^2(0) e^{ip(-\phi+0)}.$$

Substitution gives

$$\begin{aligned} J_m(kx) &= \frac{4\pi}{2\pi i^m} \sum_{n=0}^{\infty} i^n j_n(kx) \\ &\quad \times \sum_{p=-n}^{n} A_{np}^2 P_{n|p|}^2(0) \int_0^{2\pi} e^{im\phi} e^{-ip\phi} \, d\phi. \end{aligned}$$

From the orthogonality property of complex exponentials (26) this becomes

$$J_m(kx) = \frac{4\pi}{i^m} \sum_{n=|m|}^{\infty} i^n j_n(kx) A_{nm}^2 P_{n|m|}^2(0)$$

thus completing the proof.

REFERENCES

[1] C. Kyriakakis, P. Tsakalides, and T. Holman, "Surrounded by sound," *IEEE Signal Processing Mag.*, vol. 16, pp. 55–66, Jan. 1999.

[2] M. A. Gerzon, "Periphony: With-height sound reproduction," *J. Audio Eng. Soc.*, vol. 21, pp. 2–10, Jan. 1973.

[3] J. S. Bamford, "An analysis of ambisonic sound systems of first and second order," M.S. thesis, Univ. Waterloo, Waterloo, ON, Canada, 1995.

[4] R. Nicol and M. Emeritt, "3-D-sound reproduction over an extensive listening area: A hybrid method derived from holophony and ambisonic," in *Proc. AES 16th Conf. Spatial Sound Reproduction*, Rovaniemi, Finland, Apr. 1999.

[5] G. Dickins and R. A. Kennedy, "Toward optimal sound field representation," in *Proc. 106th Conv. AES*, Munich, Germany, 1999.

[6] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Amer.*, vol. 93, no. 5, pp. 2764–2778, May 1993.

[7] O. Kirkeby and P. A. Nelson, "Reproduction of plane wave sound fields," *J. Acoust. Soc. Amer.*, vol. 94, no. 5, pp. 2992–3000, Nov. 1993.

[8] O. Kirkeby, P. A. Nelson, F. Orduña-Bustamante, and H. Hamada, "Local sound field reproduction using digital signal processing," *J. Acoust. Soc. Amer.*, vol. 100, no. 3, pp. 1584–1593, Sept. 1996.

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

[10] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. New York: Academic, 1999.

[11] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*. New York: Springer-Verlag, 1997.

[12] P. A. Nelson and S. J. Elliot, *Active Control of Sound*. New York: Academic, 1992.

[13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.

[14] 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/.

[15] P. M. Morse and K. U. Ingard, *Theoretical Acoustics*. New York: McGraw-Hill, 1968.

Darren B. Ward (S'93–M'96) was born in Mareeba, Australia, in 1970. He received the B.E. (hons) degree in electronic engineering and the B.A.S. degree in computer science from the Queensland University of Technology, Brisbane, Australia, in 1993, and the Ph.D. degree in engineering from the Australian National University (ANU), Canberra, in 1997.

From 1996 to 1998, he was a Post-doctoral Member of Technical Staff at Bell Labs, Lucent Technologies, Murray Hill, NJ. He has held academic positions at the University of New South Wales and ANU, and is currently a Lecturer at Imperial College, London., U.K. His research interests are in the fields of signal processing and digital communications.

Thushara D. Abhayapala (S'96) was born in Colombo, Sri Lanka, in 1967. He received the B.E. (hons) degree in interdisciplinary systems engineering in 1994, and the Ph.D. degree in engineering in 1999, both from the Australian National University, Canberra, ACT.

From 1995 to 1997, he was a Research Engineer at the Arthur C. Clarke Center for Modern Technologies, Moratuwa, Sri Lanka, and is currently a Research Fellow at the Department of Telecommunication Engineering, Research School of Information Sciences and Engineering, Australian National University. His research interests are in the fields of digital communications and signal processing.