PLANE-WAVE DECOMPOSITION ANALYSIS FOR SPHERICAL MICROPHONE ARRAYS

Ramani Duraiswami, Zhiyun Li, Dmitry N. Zotkin, Elena Grassi, Nail A. Gumerov

Perceptual Interfaces and Reality Laboratory, UMIACS, University of Maryland, College Park,
{ramani,zli,dz,gumerov}@umiacs.umd.edu, egrassi@isr.umd.edu

ABSTRACT

Spherical microphone arrays have attracted attention for analyzing the sound field in a region and beamforming. The analysis of the recorded sound has been performed in terms of spherical wave-functions, and recently the use of plane-wave expansions has been suggested. We show that the plane-wave basis is intimately related to the spherical wave-functions. Reproduction in terms of both representations satisfies certain band-limit criteria. We provide an error bound that shows that to reproduce the spatial characteristics of a sound of a certain frequency we need to be able to accurately represent sounds of up to a particular order, which establishes a Nyquist-like criterion. The order of the sound field in turn is related to the number of microphones in the array necessary to achieve accurate quadrature on the sphere. These results are illustrated with simulations.

1. INTRODUCTION

Many authors have proposed the use of spherical microphone arrays for beamforming and for capturing the spatial structure of the sound field in the region of space occupied by the array and in the nearby area. In [1] the sound field was captured using an open spherical microphone array. Microphones can also be positioned on the surface of a rigid sphere to make use of the scattering [2]. The design of the spherical arrays was considered in [3], and the analysis of the recorded sound field using plane-wave expansion was described in [4]. We present further details of the mathematical theory here, adding the effects of the finite number of microphones, the influence of the sound frequency, and error bounds.

While plane-waves are often used to represent sound sources in the far-field, they also constitute a remarkable basis for the wave equation that can represent the sound field in both the near and far-fields. This representation of the wave field is being exploited in developing the fast multipole method for the Helmholtz equation [5]. Using error bounds presented in [6], we build on all the papers mentioned above, show how analysis in terms of both wavefunctions and plane-waves can be used interchangeably, and provide explicit error bounds and frequency dependence.

2. ANALYSIS OF THE SOUND FIELD

Sound propagation and sound scattering off the spherical array are governed by the wave equation (1, left). Using the Fourier transform (2, left), we can convert the wave equation to the Helmholtz equation in the frequency domain (1, right)

\[
\frac{1}{c^2} \frac{\partial^2 \psi(r,t)}{\partial t^2} = \nabla^2 \psi(r,t) + k^2 \psi(r) = 0, \quad \text{equation in the frequency domain (1, right)}
\]

\[
\psi(r,\omega) = \int_{-\infty}^{\infty} \exp[i\omega t] \psi(r,t) dt, \quad \psi(r,\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp[-i\omega \omega] \psi(r,\omega) d\omega.
\]

where \(k\) is the wavenumber and \(\omega = 2\pi f\) is the circular frequency. The wave propagation and scattering problem can be reduced then to solving the Helmholtz equation for a number of frequencies. To obtain the time domain solution, we take the inverse Fourier Transform (2, right) of \(\psi\). For scattering problems the total (measured) field can be decomposed as

\[
\psi(r) = \psi_{in}(r) + \psi_{scat}(r).
\]

In this paper, we consider two problems: sound field analysis and beamforming. The goal of sound field analysis is to build an approximation to \(\psi_{in}\) at the array location (center) using the sound recorded on the surface at the microphones, \(\psi\), which is a combination of \(\psi_{in}\) and \(\psi_{scat}\). When such an approximation is constructed, the sound field in a nearby area can be computed by evaluating it (because the approximation consists of eigenfunctions of the wave propagation equation, it describes accurately the wave propagation process and therefore can be extended spatially). The goal of beamforming is to pick out the portion of the sound coming from a particular direction \(s_0\), and if the incoming sound field is represented as a sum of plane waves from various directions, then a solution to the beamforming problem is provided simply by determining the intensity of the plane wave that comes from the direction of interest, \(s_0\).

2.1. Representations of the incident field

We consider two representations of the sound field: using the spherical wave functions and the plane-wave expansions. We want to analyze the sound field using the spherical array and (later) to reconstruct the sound in the region occupied by the array and in the nearby region, as if the array is not present. We assume that the region in which we are modeling the sound field does not contain any scattering objects or sound sources, except for the spherical array itself. This means that in the modeling area the incident field is a regular function of \(r\).

Spherical wave functions: Because the sound field is a regular solution of the Helmholtz equation, it can be expanded into an infinite series over the basis of its elementary spherical regular solutions, \(\{R_{nm}(k;r)\}\):

\[
\psi_{in}(k;r) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} A_{nm} R_{nm}(k;r), \quad R_{nm}(k;r) = j_n(kr) Y_{nm}(\theta, \phi),
\]

where \(A_{nm}\) are expansion coefficients. Here \((r, \theta, \phi)\) are spherical coordinates of \(r\), \(j_n(kr)\) are spherical Bessel functions of the first kind, and \(Y_{nm}(\theta, \phi)\) are spherical harmonics

\[
y_{nm}(\theta, \phi) = (-1)^n \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} j_n(kr) \cos m \phi, 
\]

where \(n, m\) are non-negative integers.
where \( n = 0, 1, 2, \ldots \) and \( m = -n, \ldots, n \). For future reference we note the addition theorem for spherical harmonics: if \( s_i = (\theta_i, \phi_i) \) and \( s_j = (\theta_j, \phi_j) \) are two points on the unit sphere, then the spherical harmonics and the Legendre polynomial of order \( n \) of the angle between \( s_i \) and \( s_j \) satisfy

\[
P_n(s_i \cdot s_j) = \frac{4\pi}{2n+1} \sum_{m=-n}^{n} Y_n^m(s_i) Y_n^m(s_j).
\]

**Band-limited plane-wave expansions:** This plane-wave basis, also called the Herglotz wave function basis, represents solutions of the Helmholtz equation \( \psi \) in integral form. It can be interpreted as representing the incident field as a superposition of plane waves \( e^{ik \mathbf{r} \cdot \mathbf{s}} \) propagating in all possible directions \( \mathbf{s} \), with the magnitude and phase of the plane wave in direction \( \mathbf{s} \) characterized by the complex amplitude \( \mu_{in}(k; \mathbf{s}) \). (Note we take the convention that the plane wave is propagating in the direction \( \mathbf{s} \). A few authors have \( \mathbf{s} \) as the direction the wave is coming from.) Thus we write

\[
\psi_{in}(k; \mathbf{r}) = \frac{1}{4\pi} \int_{S_u} e^{ik \mathbf{r} \cdot \mathbf{s}} \mu_{in}(k; \mathbf{s}) dS(\mathbf{s}),
\]

where integration is taken over all directions (notionally over a unit sphere surface), \( \mathbf{s} \) is the unit vector on the sphere with Cartesian coordinates \( \mathbf{s} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \), and \( \mu_{in}(k; \mathbf{s}) \) is a surface function, also known as the signature function. The dependence of \( \psi_{in}(k; \mathbf{r}), R_n^m(k; \mathbf{r}) \), and \( \mu_{in}(k; \mathbf{s}) \) on \( \mathbf{k} \) will be suppressed from now on for clarity.

In fact, both representations of \( \psi_{in} \), (4) and (7), are related due to the Gegenbauer expansion of the plane wave [7]

\[
e^{ik \mathbf{r} \cdot \mathbf{s}} = 4\pi \sum_{n=0}^{\infty} \sum_{m=-n}^{n} i^n Y_n^m(\mathbf{s}) R_n^m(\mathbf{r}),
\]

where \( Y_n^m(\mathbf{s}) = Y_n^m(\theta, \phi) \). Thus the expansion coefficients in (4) and the signature function \( \mu_{in} \) in (7) are related as

\[
\mu_{in}(s) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} i^n A^m_n Y_n^m(s), \quad A^m_n = \int_{S_u} \mu_{in}(s) Y_n^m(s) dS(s).
\]

In practice we work with samples of a function; therefore, the integral over the sphere needs to be performed via some quadrature rule, which replaces integration over the sphere by a summation of function values at \( L \) quadrature points multiplied by appropriate weights. The microphones on the surface of the spherical array represent these points. The locations and weights of the quadrature points are chosen to compute the surface integral exactly for the functions that can be represented by a “band-limited” spherical harmonic expansion of a particular degree up to \( p \) (Eq. (10), right)

\[
\int_{S_u} F(s) dS = \sum_{j=0}^{L-1} F(s_j) w_j; \quad F(s) = \sum_{n=0}^{p-1} \sum_{m=-n}^{n} C^m_n Y_n^m(s),
\]

where \( C^m_n \) is the set of arbitrary coefficients. Thus the concept of band-limitedness is applicable to both the plane-wave and the truncated spherical wave-function expansions. From now on, we will assume that we work in a band-limited representation in either case. The goal is to choose a band-limit that is sufficiently accurate for our purposes. To control the accuracy, we need an error bound.

### 2.2. Choice of band-limit via an error bound

**Truncated spherical wavefunction expansions of a plane wave:** As the series (4) converges absolutely and uniformly in the domain of interest, the function \( \psi_{in}(r) \) can be approximated there by band-limited functions \( \psi_{in}^{p^*}(r) \). The value of \( p^* \) chosen depends on the frequency and on the size of the domain. The truncation error in representing a single plane wave input field by regular spherical basis functions using (8) is

\[
\epsilon_p(s, r) = e^{ik\mathbf{r} \cdot \mathbf{s}} - 4\pi \sum_{n=0}^{p-1} \sum_{m=-n}^{n} i^n Y_n^m(s) R_n^m(r),
\]

\[
\sum_{n=p}^{\infty} (2n+1)i^n j_n(kr) P_n \left( \frac{r}{r} \right)
\]

Assume that the domain of interest can be enclosed inside a sphere of radius \( R \). Then the following general error bound can be found using the inequality [7]

\[
|j_n(kR)| \leq (kR)^n / (2n+1)!
\]

(where ! is a double factorial) and the error bound for the Taylor expansion of the exponent:

\[
|\epsilon_p(s, r)| \leq \sum_{n=p}^{\infty} \sum_{m=-n}^{n} (kR)^n j_n(kR) P_n \left( \frac{r}{r} \right) \leq \frac{2}{p!} \left( \frac{kR}{2} \right)^{p+1} \exp \left( \frac{kR}{2} \right) = \delta_p,
\]

\( p \gg 1 \).

For relatively low \( kR < 1 \) or moderate \( kR \approx 1 \) frequencies, Eq. (12) provides relatively low \( p \) (e.g., for \( kR = 2 \) we have \( |\epsilon_p| < 2e/p! \)). For higher frequencies \( kR \gg 1 \) asymptotic analysis (e.g., [6]) shows that \( p \) should be always larger than \( kR \) and

\[
|\epsilon_p(s, r)| \leq \frac{1}{3} \left[ \frac{p-kR}{(kR)^{1/3}} \right]^{3/2} = \delta_p, \quad kR \gg 1.
\]

As \( \epsilon_p(s, r) \) can be uniformly bounded and the incident field can be represented as a superposition of plane waves, we can obtain from Eq. (7) the overall error of approximation of the incident field by the band-limited function \( \psi_{in}^{p^*}(r) \) inside a sphere of radius \( R \):

\[
\left| \psi_{in}(r) - \psi_{in}^{p^*}(r) \right| \leq \frac{1}{4\pi} \int_{S_u} |\epsilon_p(s, r)| |\mu_{in}(s)| dS(s)
\]

\[
|\mu_{in}(s)| \leq \max_{s} \left( |\mu_{in}(s)| \right) \leq \delta_p \max_{s} |\mu_{in}(s)| = \epsilon_s,
\]

where \( \delta_p \) can be selected according to Eq. (12) or Eq. (13). The latter formula can be inverted to determine \( p \) based on the specified accuracy \( \epsilon_s \):

\[
p \approx kR + \frac{1}{2} \left( 3 \ln \max_{s} |\mu_{in}(s)| \right)^{2/3} (kR)^{1/3}, \quad kR \gg 1.
\]

In multifrequency analysis, it is best in practice to increase \( p \) along with the frequency as guided by Eq. (15) to avoid numerical errors. In Section 3, we verify the obtained error bounds and the rules for setting the truncation number \( p \) in a synthetic data experiment.
2.3. Scattering off a spherical microphone array

Solution of the problem of a plane wave scattering off a sound-hard sphere of radius \(a\) is classical and can be found, for example, recently in [2]. For a general band-limited incident field with coefficients of expansion \(A_m^m\) and sound-hard boundary conditions, the solution at a point with angular coordinates \(\theta, \phi\) on the sphere is

\[
\psi_S(\theta, \phi) = \frac{i}{(ka)^2} \sum_{n=0}^{\infty} \frac{1}{h_n(ka)} \sum_{m=-n}^{n} A_m^m Y_n^m(\theta, \phi),
\]

where \(\psi_S\) is the measured field on the sphere surface. Particularly, when the incident field is a band-limited plane wave propagating in the direction \(\theta'\), it follows from Eq. (8) that \(A_0^m = 4\pi i^m Y_n^m(\theta')\). Using Eq. (6), we can write that the measured sound field at \(\theta, \phi\) due to a plane wave propagating towards \(\theta'\) is

\[
\psi_S(\theta, \phi) = K(\theta, \phi') = \frac{i}{(ka)^2} \sum_{n=0}^{\infty} \frac{i^n (2n+1) P_n(\cos \theta) \cos \phi}{h_n(ka)}. \quad (17)
\]

Let the sound received at the sphere be represented as a superposition of plane-waves (Eq. (7)). Then the measured sound field on the sphere can also be represented as a superposition of the scattered plane waves as

\[
\psi_S(\theta, \phi) = \frac{1}{4\pi} \int_{S_h} K(\theta, \phi') \mu_{in}(\theta', \phi') dS' = \sum_{j=0}^{L-1} w_j K(\theta, \phi'; \theta_j, \phi_j) \mu_{in}(\theta_j, \phi_j). \quad (18)
\]

We must determine the input coefficients \(\mu_{in}(\theta_j, \phi_j)\) at the \(L\) quadrature points to determine the incident field \(\psi_{in}\). One way to proceed is to choose particular measurement (microphone) locations \(\theta_j, \phi_j\) and particular quadrature points \(\theta_j, \phi_j\) for \(\mu_{in}(\theta_j, \phi_j)\) and to solve the following (possibly over-determined) linear system of equations in \(\mu_{in}(\theta_j, \phi_j)\):

\[
\psi_S(\theta_j, \phi_j) = \sum_{j=0}^{L-1} w_j K(\theta_j, \phi_j) \mu_{in}(\theta_j, \phi_j), \quad j = 1, \ldots, L_M, \quad (19)
\]

where \(L_M\) is the total number of microphones. Using the fact that we work in a band-limited representation, we can also write an explicit expression (below) for the incoming plane-wave coefficients \(\mu_{in}(\theta_j, \phi_j)\) in terms of the measured \(\psi_S(\theta_j, \phi_j)\).

**Explicit computation of plane-wave coefficients:** The coefficients \(A_m^m\) can be found from Eq. (16) as

\[
A_m^m = -i (ka)^2 h_n(ka) \int_{S_h} \psi_S(\theta, \phi) Y_n^m(\theta, \phi) dS(\theta, \phi). \quad (20)
\]

Substituting this into Eq. (9) and using Eq. (6)

\[
\mu_{in}(\theta_j, \phi_j) = \sum_{n=0}^{L-1} \sum_{m=-n}^{n} i^n A_m^m Y_n^m(\theta_j, \phi_j) = -i (ka)^2 h_n(ka) \int_{S_h} \psi_S(\theta, \phi) Y_n^m(\theta, \phi) dS(\theta, \phi) \quad (21)
\]

\[
\begin{align*}
\mu_{in}(\theta_j, \phi_j) &= \frac{i}{4\pi} \sum_{n=0}^{L-1} \sum_{m=-n}^{n} i^n h_n(ka) \int_{S_h} \psi_S(\theta, \phi) Y_n^m(\theta, \phi) dS(\theta, \phi) \\
&= \frac{i}{4\pi} \sum_{n=0}^{L-1} \sum_{m=-n}^{n} (2n+1) h_n(ka) \int_{S_h} \psi_S(\theta, \phi) P_n(\theta, \phi') dS(\theta, \phi).
\end{align*}
\]

Using quadrature (Eq. (10)) for the sphere, we obtain

\[
\mu_{in}(\theta_j, \phi_j) = \frac{-i (ka)^2}{4\pi} \sum_{n=0}^{L-1} \sum_{m=-n}^{n} \frac{(2n+1)}{i^n} h_n(ka) \sum_{j=0}^{L-1} w_j P_n(\theta_j, \phi_j) \psi_S(\theta_j, \phi_j) \quad (22)
\]

where the symmetric kernel \(M\) is

\[
M(\theta_j, \phi_j) = \frac{-i (ka)^2}{4\pi} \sum_{n=0}^{L-1} \sum_{m=-n}^{n} (2n+1) i^n h_n(ka) P_n(\theta_j, \phi_j).
\]

2.4. Quadrature for discrete integration

As we work with band-limited functions, the quadrature must be able to accurately reproduce functions of the specified bandwidth \(b\). Furthermore, as quadrature points are to be used for placement of microphones, intuitively we want “uniform” distribution on the sphere surface. A surprising result on order \(p\) quadrature over the sphere [8] is that any quadrature formula must have more than \(p^2\) points (i.e., microphones) and that a quadrature rule can be obtained with \(4p^2\) points in more or less arbitrary point distribution.

It follows from Eq. (15) and \(L = 4p^2\) condition that the number of microphones \(L\) needed for the head-sized \((a = 0.09m)\) and the tennis-ball-sized \((a = 0.03m)\) microphone array is as follows:

<table>
<thead>
<tr>
<th>a</th>
<th>f</th>
<th>(ka)</th>
<th>(p)</th>
<th>(L)</th>
<th>f</th>
<th>(ka)</th>
<th>(p)</th>
<th>(L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3cm</td>
<td>4kHz</td>
<td>2.1</td>
<td>3</td>
<td>36</td>
<td>12kHz</td>
<td>6.3</td>
<td>17</td>
<td>196</td>
</tr>
<tr>
<td>9cm</td>
<td>4kHz</td>
<td>6.3</td>
<td>7</td>
<td>196</td>
<td>12kHz</td>
<td>18.9</td>
<td>20</td>
<td>1600</td>
</tr>
</tbody>
</table>

Clearly a lot of microphones are needed. For a special grid, which is equispaced in \(\varphi\) and arranged in \(\theta\) as zeros of the Legendre polynomial of degree \(p\), we have \(L = 2p^2\), reducing the number of points in half. However, these points are inconveniently distributed, and it would be hard to manufacture such an array. Furthermore, if the bandwidth of all functions \(\psi_S(\theta, \phi)\), \(K(\theta, \phi'; \theta_j, \phi_j)\), and \(\mu_{in}(\theta_j, \phi_j)\) is \(p\), then the quadrature for the bandwidth of \(2p\) is actually necessary as we need to integrate products of functions, which further increases the required number of microphones.

If we are willing to use approximate quadrature formula, then it may be possible both to reduce the number of points and to have uniform point distribution. In [9] quadrature nodes and weights of various number of nodes are computed using an optimization procedure. These weights and nodes are not related to the spherical harmonics, and no analytical guarantees on their integration are available. On the other hand, the point distributions in [9] are relatively uniform, which simplifies the manufacturing of a spherical microphone array. Because of that, we verified whether these nodes provide good quadrature for our purposes.

We performed a numeric integration of spherical harmonic pairs of various orders (up to order \(p\)) over the sphere using the Fliege quadrature nodes and weights. We used two Fliege grids, one with 64 points and another with 324 points, setting \(p = 7\) and \(p = 17\), respectively. It was found that the Fliege nodes are in fact well-suited for integration of band-limited functions and that the orthogonality condition is satisfied with low error. In Figure 1 the 64 node quadrature of [9] is used, and the difference between the theoretical value of the integral (1 when order and degree are equal and 0 in all other cases) and its computer value is shown as a pixel.
gray level. We note that the product of $Y_p^n$ and $Y_q^m$ is a spherical harmonic of order $p+q$, and while it may not be justified to expect reasonable performance beyond $p=3$, these nodes perform remarkably well. For higher order products, a maximum error of 0.03 is introduced, and most errors are much smaller. In practice we expect these errors to be small compared to other errors introduced (finite-bit representations, noise, etc.).

3. VERIFICATION AND TESTING

We performed several sets of experiments recreating the simulated sound fields from synthetic capture data and evaluated the accuracy of the reconstruction. When the theory of near-field source sound field reconstruction using the plane-wave basis is presented to people unfamiliar with it, usually some discussion arises. In the following, we show that it is indeed possible to perform such reconstruction. It must be recognized first that our algorithm recreates only the projection of the sound field (up to a particular bandwidth) using the regular spherical eigenfunctions of the Helmholtz equations. We thus must compare the reconstructed field not with the original field but rather with the bandlimited original field. The comparison below reveals that with Fliege node sampling we are able to achieve the required reconstruction with low error.

In Figure 2, we show the reconstruction of the incident sound field created by a point source located 0.2 m from the array center. The first simulation (left column) has a source of frequency 1 kHz and an array radius of 0.1m. This corresponds to $ka = 1.83$ and a predicted $p$ of 3 for reconstruction in a region of the array size. As it is seen from the plot, reconstruction errors are low in the region where the reconstruction should hold (the neighborhood of the sphere shown as a circular area in this cross-section through the $z = 0$ plane). The 64-node Fliege grid was used. The right column shows the same plot for a case where the source frequency is 8 kHz, an array radius is 0.08m, $ka = 1.173$, $p = 13$, and Fliege grid has 324 nodes. Again, good reconstruction is observed in the region where it should be valid. In other work we have used the expressions for $\mu_{\ell m}$ for plane-waves from particular directions to perform beamforming and to recreate spatial audio, and will be presented elsewhere.

4. REFERENCES