This paper considers the recording and reproduction of three dimensional (3D) sound fields, based on spherical harmonic expansions of the field. It is shown that the plane wave description is insufficient for the description of fields with point sources which have wavefront curvature. The recording of a sound field requires the measurement of the coefficients of the spherical harmonic expansion. The use of spherical and general arrays for recording the coefficients is discussed. The reproduction of the sound field requires the resynthesis of the field using the spherical harmonic coefficients. It will be shown that there are two approaches to the determination of the speaker weights. The mode matching approach leads to a pseudo-inverse solution. The simple source approach is formally introduced, and it is shown that its application yields a matrix transpose approach. Computer simulations of soundfield synthesis are given to illustrate the two approaches.

1 Sound Field Description

1.1 Spherical harmonic expansion

For the interior case – where all sources lie outside the region of interest – the sound pressure at a single frequency can be expressed as [3]

\[
p(r, \theta, \phi, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} A_n^m(k) j_n(kr) Y_n^m(\theta, \phi)
\]

where \(k\) is the wavenumber, \(j_n(.)\) is the spherical Bessel function of the first kind and the spherical harmonics are defined as [6]

\[
Y_n^m(\theta, \phi) = \frac{(2n+1)(n-|m|)!}{4\pi (n+|m|)!} P_n^m(\cos \theta) e^{im\phi}
\]

Each harmonic is the product of an elevation term and the azimuthal variation \(\exp(im\phi)\) which appears in 2D ambisonics theory. In practice the sum is truncated to a maximum \(n=\bar{N}\), and there are \((\bar{N}+1)^2\) terms in the expansion with \(2n+1\) terms in \(m\) for each \(n\).

For the exterior case the field outside a region containing sources can be expressed as

\[
p(r, \theta, \phi, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} B_n^m(k) h_n(kr) Y_n^m(\theta, \phi)
\]

where \(h_n(kr)\) is the spherical Hankel function [3].

1.2 Plane and spherical wave expansions

The spherical harmonic expansion of the spatial variation of a plane wave arriving from angle of incidence \((\theta_i, \phi_i)\) is

\[
e^{\vec{r} \cdot \vec{r}_i} = 4\pi \sum_{n=0}^{\infty} \vec{r} \cdot \vec{r}_i \sum_{m=-n}^{n} Y_n^m(\theta, \phi) Y_n^m(\theta_i, \phi_i)
\]

and that of a spherical source at \(\vec{r}_s = (r_s, \theta_s, \phi_s)\) is...
\[ G(\vec{r} | \vec{r}^* ) = \frac{e^{-i (\vec{k} \cdot \vec{r}^*)}}{4\pi |\vec{r}^*|} \]

\[ = k \sum q_n^m(k) Y_n^m(\theta, \phi) \]

We assume a harmonic dependency of \( e^{i (\vec{k} \cdot \vec{r})} \) so that equation 5 represents propagation outwards from \( r_s \).

### 1.3 Fourier description

For a distribution of plane waves with complex amplitudes \( Q(\theta, \phi) \), the Fourier transform of the field is only non-zero at a single spatial frequency \( k_0 \). \( Q \) can be written in terms of its spherical harmonic expansion

\[ Q(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} q_n^m(k_0) Y_n^m(\theta, \phi) \]

and the inverse Fourier transform in spherical coordinates becomes

\[ p(r, \theta, \phi, k) = \frac{1}{(2\pi)^2} \int \int Q(\theta, \phi) e^{i (\vec{k} \cdot \vec{r})} \sin(\theta) d\theta d\phi \]

Substituting equations 4 and 6 into 7 yields the same form as equation 1 with \( A_n^m(k_0) = 4\pi i^n q_n^m(k_0) \), i.e. the coefficients of the plane wave field are simply related to those of the far field distribution. This is the 3D form of the expansion in [2].

If the soundfield includes point sources, the Fourier transform is nonzero for all \( k \) and the plane wave description is invalid. For example, the sound field produced by a single point source at \( \vec{r} \) generating a frequency \( f_0 \) is the solution to the Helmholtz equation

\[ \nabla^2 \hat{p}(\vec{r}) + k_0^2 \hat{p}(\vec{r}) = -\delta(\vec{r} - \vec{r}_s) \]

yielding

\[ \hat{p}(\vec{k}) = \frac{e^{-i \vec{k} \cdot \vec{r}_s}}{k^2 - k_0^2} \]

and the inverse Fourier transform yields equation 5. A general sound field requires an integral over all sources and all spatial frequencies are required to produce the divergence due to localized sources [3].

### 2 Soundfield Recording

A soundfield may be recorded by determining its spherical harmonic coefficients. This produces a set of spherical harmonic responses which include the first order ambisonics responses. Two approaches are reviewed here.

#### 2.1 Sphere Decompositions

Multiplying equation 1 by \( Y_n^m(\theta, \phi)^* \) and integrating over the sphere yields the \( l.g. jth \) response [7]

\[ A_l^j(k) = \frac{1}{l_{l,j}(k)} \int \int p(r, \theta, \phi) Y_l^j(\theta, \phi)^* \sin(\theta) d\theta d\phi \]

In a similar manner to the 2D case [2], this produces zeros where \( j_l(k) = 0 \). One solution is to use outward-facing first order microphones (a weighted sum of pressure and acoustic-impedance-scaled radial velocity responses) which has the ideal form

\[ s_i(r, \theta, \phi, k) = \alpha p(r, \theta, \phi, k) - (1-\alpha) \rho c v_i(r, \theta, \phi, k) \]

producing the first order decomposition

\[ A_k^l(k) = \sum_{n=0}^{\infty} A_n^m(k) Y_n^m(\theta, \phi) \]

in which case the coefficients are

\[ A_k^l(k) = \frac{1}{D_l^0(k)} \int \int s_i(r, \theta, \phi, k) Y_l^j(\theta, \phi)^* \sin(\theta) d\theta d\phi \]

where the "mode" amplitude is

\[ D_l^0(k) = \alpha j_l(k) - i(1-\alpha) j_l(k) \]

which produces no zeros.

An alternative approach is to use pressure microphones mounted in a solid sphere [8]. The sound pressure produced by a solid sphere of radius \( a \) is the sum of the incident field (equation 1) and a scattered field described by equation 3. The sum of the two fields has the form [3]

\[ p_i(r, \theta, \phi, k) = \sum_{n=0}^{\infty} j_n(k) - j_n(k) \]

yielding

\[ j_n(k) = \frac{1}{k_n^2} \left[ j_n(k) - j_n(k) \right] \]

and the inverse Fourier transform yields equation 5. A general sound field requires an integral over all sources and all spatial frequencies are required to produce the divergence due to localized sources [3].

312
2.2 General sampling

A more general approach to the above is to use a sampling of the sound pressure field at \( L \) arbitrary positions. The field can be recreated at the same relative positions in a room using \( L \) speakers if the matrix of transfer functions from the speakers to those positions is known [9]. Alternatively, the method can be applied to the measurement of spherical harmonics. Expressing the pressure at \( L \) positions \( p(r, \theta, \phi) \) as in equation 1, the vector of pressures, \( \mathbf{p} \), can be written

\[
\mathbf{p}(k) = \Lambda(k) \mathbf{A}(k)
\]

where \( \mathbf{A} \) is the vector of \((N+1)^2\) spherical harmonic coefficients to be determined and \( \Lambda \) is the matrix with elements \( \frac{j_n(kr_i)}{r_i} Y_{l}^{m*}(\theta, \phi) \) with row indexed by \( l \) and column indexed by \( n \) and \( m \). Typically, \( L<K \), the system is overdetermined, and the vector of coefficients can then be obtained as the regularized least squares inverse of equation 17 [10]

\[
\mathbf{A} = [\Lambda^\dagger \Lambda + \mu I]^{-1} \Lambda^\dagger \mathbf{p}
\]

When the sample points are at the same radius, \( r_i \), then equation 17 can be written

\[
\mathbf{p} = \mathbf{Y} \mathbf{J} \mathbf{A}
\]

where \( \mathbf{Y} \) is the matrix of spherical harmonic terms and \( \mathbf{J} \) a diagonal matrix with elements \( j_n(kr_i) \). If the positions are those of a regular polyhedron and \( L=K \), then \( \mathbf{Y} \) is a unitary matrix, and

\[
\mathbf{A} = \mathbf{J}^{-1} \mathbf{Y}^\dagger \mathbf{p}
\]

which is the discrete matrix form of equation 10. Hence the general approach includes the open sphere approach as a special case. It can also accommodate directional element responses [10].

3 Soundfield Synthesis and Reproduction

To reconstruct a sound field, a spherical array of \( L \) loudspeakers at positions \((R, \theta, \phi)\) in a free field is assumed. Each speaker produces a field which has a spherical harmonic expansion, and the sum of these must equal the expansion of the desired field in a region of space near the origin. We describe two approaches.

3.1 Mode matching approach

Consider first the generation of a single plane wave [11,12]. If it is assumed that \( R \) is large then each speaker produces a plane wave near the origin. If the amplitude of each speaker signal is \( w_i \), the synthesised field is a sum of \( L \) plane waves with expansions given by equation 4. This must equal the expansion of the desired plane wave arriving from \((\theta, \phi)\), yielding

\[
\sum_{l=1}^{L} w_i (\theta, \phi) Y_{l}^{m*}(\theta, \phi)^* = Y_{l}^{m*}(\theta, \phi)^*
\]

This must be solved for the required weights \( w_i \). At low frequencies the weights are the panning functions for plane wave synthesis for the particular speaker layout [1].

If the plane wave approximation for the speaker fields is replaced by a spherical point source approximation then matching the \( L \) spherical wave expansions (equation 5) to that of the plane wave yields [12]

\[
\sum_{l=1}^{L} w_i (\theta, \phi) Y_{l}^{m*}(\theta, \phi)^* = \frac{4\pi i}{i k h_n (kR)} Y_{l}^{m*}(\theta, \phi)^*
\]

The panning function solutions include the Hankel function for the speaker sources.

If the desired field is due to a point source at \((r, \theta, \phi)\), then the mode matching equation is [13]

\[
\sum_{l=1}^{L} w_i (r, \theta, \phi) Y_{l}^{m*}(\theta, \phi)^* = \frac{h_l(kr)}{h_n(kR)} Y_{l}^{m*}(\theta, \phi)^*
\]

In this case the panning functions are different for each source radius. Finally, for an arbitrary field with coefficients \( A_{l}^{m}(k) \) the mode matching equation is

\[
\sum_{l=1}^{L} w_i Y_{l}^{m}(\theta, \phi) = A_{l}^{m}(k)
\]

For a finite \( N \)th order expansion this can be written

\[
\Psi \mathbf{w} = \mathbf{d}
\]

where \( \Psi \) is a \((N+1)^2\) by \( L \) matrix of spherical harmonics and \( \mathbf{d} \) is the vector containing the terms on the right hand side of equation 24

\[
\mathbf{d} = \mathbf{H}^{-1} \mathbf{A}
\]

where \( \mathbf{H} \) is a diagonal matrix with entries \( i k h_n (kR) \) appearing \( 2n+1 \) times. \( \Psi \) is the reproduction equivalent of the general sampling matrix \( \Lambda \) in equation 17.

If \( K<L \) the minimum energy solution is [12]

\[
\mathbf{w} = \Psi^* \left[ \Psi \Psi^* \right]^{-1} \mathbf{H}^{-1} \mathbf{A}
\]

If \( K=L \) then

\[
\mathbf{w} = \Psi^* \mathbf{H}^{-1} \mathbf{A}
\]
and if \( K \geq L \) the regularised least squared error solution is
\[
w = \left[ \Psi^\dagger \Psi + \lambda I \right]^{-1} \Psi^\dagger \mathbf{H}^\dagger \mathbf{A}
\]
In practice the lowest errors were found when \( K=L \). Equation 29 was used as it allowed control of the error. (In practical systems, however, there may be less recorded modes than speakers.)

3.2 Simple Source Approach

In [3] it is shown that there are alternatives to the Kirchoff-Helmholtz integral which do not require both pressure and normal velocity on the surface to be known. One of these is the simple source approach, for which the field is obtained from a distribution of monopole sources over a surface \( S \) as
\[
p(r, \theta, \phi, k) = \int_0^{2\pi} \int_0^n \mu(\tilde{r}) \frac{e^{ik r - i k r}}{4\pi r - r} \sin(\theta) d\theta d\phi
\]
The required simple source distribution is given by [3]
\[
\mu(\tilde{r}) = \frac{\delta p_r(\tilde{r})}{\delta n} - \frac{\delta p_\theta(\tilde{r})}{\delta n}
\]
where \( n \) is the inward facing normal, \( p_r(\tilde{r}) \) is the exterior field produced by a source distribution confined in \( S \) and \( p_\theta(\tilde{r}) \) is the interior field produced by a source distribution outside \( S \), with the condition that the two fields are equal on the surface. Equating the general interior and exterior expansions on a sphere of radius \( R \) yields
\[
B_n^m(k) = \frac{j_m(kR)}{h_n(kR)} A_n^m(k)
\]
Substituting the two expansions into 31, and employing the Wronskian relation [3] yields the simple source solution
\[
\mu(R, \theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{A_n^m(k)}{ik h_n(kR)} Y_n^m(\theta, \phi)
\]
In practice \( L \) speakers and a finite order expansion is used, and the \( lth \) speaker weighting is
\[
w_l = \frac{g_l}{ik} \sum_{n=0}^{N} \sum_{m=-n}^{n} \frac{A_n^m(k)}{h_n(kR)} Y_n^m(\theta_l, \phi_l)
\]
where \( g_l \) is a weighting term associated with the discrete approximation to the integral (which depends on the speaker geometry – see for example [14]). The solution can be written in matrix form as
\[
w = \mathbf{G} \Psi \mathbf{H}^\dagger \mathbf{A}
\]
where \( \mathbf{G} \) is the diagonal matrix of weights \( g_l \). Equation 35 contains the transpose of the mode matrix, as opposed to the inverse forms in equations 27-29 and is therefore robust to small singular values. Equations 28 and 35 are identical when \( K=L \) and the loudspeaker geometry regular. The mode matrix is then unitary and \( \mathbf{G} = \mathbf{I} \) [4]. However the solution in equation 35 is also valid for non-regular geometries.

Errors in the simple source solution can be controlled by applying a windowing function \( \Omega_n^m(k) \) to the spherical harmonics, as in the 2D case [2]. Letting \( \Omega \) be the diagonal matrix of weights the matrix solution is
\[
w = \mathbf{G} \Psi \Omega \mathbf{H}^\dagger \mathbf{A}
\]
A simple window is \( \Omega_n^m = W_1(n)W_2(m) \) where \( W_1 \) is one-sided and \( W_2 \) is a standard window. An exponential window \( W_i = \exp(-\gamma \pi n/N) \) and Kaiser window in \( m \) will be used in the example below.

3.3 Reproduction error

A useful measure of the error in the reproduced field is the normalised radial error, which is the magnitude of the difference between the actual and ideal fields at radius \( r \) integrated over all angles [1,12];
\[
\pi(r) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi |p(r, \theta, \phi) - \tilde{p}(r, \theta, \phi)|^2 \sin(\theta) d\theta d\phi
\]
A useful lower bound to the error is the truncation error due to limiting the spherical harmonic expansion of the field to a maximum of \( n=N \). The plane wave truncation error is given in [12]. For a spherical point source the truncation error is
\[
\pi_T(r) = 1 - \frac{\sum_{n=0}^{N} (2n+1) j_n^2(kr) h_n^2(kr)}{\sum_{n=0}^{N} (2n+1) j_n^2(kr) h_n^2(kr)}
\]
This assumes that only spherical harmonics up to order \( N \) are created by the speaker array and ignores alias terms.

4 Examples

We present a simulation example for both mode matched and simple source cases. The number of speakers required for plane wave synthesis at a given frequency and radius is approximately
\[
L \geq ([k r] + 1)^2
\]
where \([.]\) denotes the next highest integer [12]. For example for a reproduction area of radius 0.1m and a frequency of 4.5 kHz \( L \) must exceed 100 speakers. We assume \( L=100 \) speakers at a radius of 2m, which allows the use of spherical harmonics up
to order 9. The speaker angles and weights were obtained from [14].

The radial errors are shown in figure 1 for a point source at radius 2.5 m on the x axis.

For small $kr$, the mode match error reduces to zero, since the number of speakers is sufficient to represent all significant spherical harmonics. For $kr$ above 10 the mode match approach cannot match the higher modes which are required. Regularisation reduces this error at the expense of increased error for small $kr$.

The simple source error is higher than the mode match error at small $kr$, but is lower at large $kr$. However, the windowed simple source solution produces the lowest error at large $kr$ and is smaller than the mode match error at small $kr$. Reducing the mode match regularization to equal the simple source performance at low $kr$ increases the error further at large $kr$. Typically, the simple source solution is better performing than the regularized mode match solution.

The field reproduction errors for the regularized mode match and windowed simple source cases are shown in figures 2 and 3. The upper plots show the speaker weight magnitudes, sorted by azimuthal angle. The lower plots show the real part of the field in $x$.

Finally, as has been demonstrated in the 2D case [4], sources can be synthesized within the array of speakers. Figure 4 shows the field produced in the $(x,y)$ plane using the simple source approach for a spherical source on the x axis at 1.2 m, using an array with 256 speakers. The pressure for radii less than 1.2 m is correctly reproduced, but the error is large for radii between 1.2 and the speaker radius. Figure 5 shows that a greater number of speakers must be activated to generate the field, and the error is larger for radii greater than 1.2 m. Since the field amplitudes can become very large near the speakers high power amplifier and speaker dissipations would be required to generate internal sources.

![Figure 2: Mode match solution with regularisation, Upper: speaker weight magnitudes, lower: Actual (--) and synthesised (-) field on x axis](image2)

![Figure 3: Simple source solution with windowing Upper: speaker weight magnitudes, lower: Actual (--) and synthesised (-) field on x axis](image3)
5 Summary

The spherical harmonic theory of sound recording and reproduction has been examined. Two methods for sound recording have been reviewed and two approaches to sound synthesis and reproduction considered in more detail, with synthesis simulations given. It has been shown that the spherical harmonic approach provides an alternative general foundation to sound reproduction to the Kirchoff-Helmholtz approach. The simple source method is a useful alternative to mode matching, and windowing of the spherical harmonic components allows for control of interference in the field while maintaining accuracy at the center of the array.

Both recording and reproduction of 3D sound fields require large numbers of transducers, making practical feasibility an issue, although simplifications are possible for particular reproduction systems [15]. Furthermore, the systems described here do not compensate for the semi-reverberant fields encountered in most listening rooms, which requires sounding and calibration for the desired listening positions.

References


