Ambisonic Synthesis of Complex Sources

Dylan Menzies

dylan@dmu.ac.uk

Marwan Al-Akaidi

mma@dmu.ac.uk

De Montfort University, Leicester, UK

June 2, 2007

Abstract

Exterior expansions of complex sound sources are presented as flexible objects for producing Ambisonic soundfield encodings. The sources can be synthesized or recorded directly, rotated and positioned in space. Related techniques can also be used to efficiently add high quality reverberation depending on the orientation and location of the source and listener.

1 Introduction

Real soundfields are frequently the result of sound from numerous sound sources, each localized to a well defined region. Synthesizing these using speaker-based and headphone-based approaches is a natural goal. Good results have been achieved for distant sources, which reach the listener as plane waves. Distance perception can be simulated using distance filtering and reverberation balance. It is also possible in low-order Ambisonic systems,[1, 2], to approximately synthesize a diffuse source at varying distance, [3, 4], which can be useful in a creative setting. Soundfield synthesis of an object with non-uniform directivity has been considered in the farfield using spherical harmonic representation, [3, 4]. With the development of high-order acoustic field construction, the simulation of nearfield sources becomes feasible.
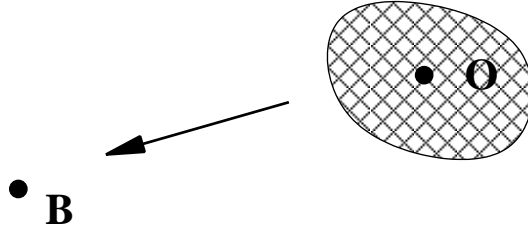
1

Figure 1: Overall scheme. O denotes the extended source object, and B the listener.

This has been developed for a monopole source in the context of high-order Ambisonics by reconstructing a monopole field about the listener, [5]. Using the wavefield approach,[6, 7], the directional properties of objects have been encoded with filters that feed the speaker array directly, [8, 9].

A localized source typically differs in two respects from simple monopole source. The sound radiates from a region of non-zero width, and the directivity of radiation is not uniform. Near to the object the soundfield will be reactive, like a monopole's, but possibly have a much more complex geometry. We should fully expect this added richness to be exploited by the auditory system for its information content, and so to have perceptual significance. Although this does not appear to have been studied in detail, informal listening provides strong evidence of spatial perceptual variety amoung complex objects. The study of directional objects using the wavefield approach also supports the hypothesis. For both practical and creative applications it would be desirable to find a way to accurately represent a complex source and encode it into Ambisonic B-format. The conversion from source encoding to Ambisonic encoding depends on the location and orientation desired of the source. From a single source encoding, that source can be rendered anywhere and in any orientation around the listener. Figure 1 illustrates this scheme. The advantage of Ambisonic modularity is apparent here, in that we seek a process that encodes into a format that is independent of the details of the rendering mechanism, whether it be a particular speaker array or headphones. The wavefield approach lacks this intermediate stage, as well as proving less accurate for given order in some studies [7]. Binaural rendering of high-order Ambisonics, over headphones, including the nearfield, has been considered, [10].

The article is organized as follows. First the source representation is

2

discussed, followed by the main part, the development of a method to transform a source encoding, with knowledge of its position and orientation, into an Ambisonic encoding. Some simulations are provided for verification and illustration. Finally we consider how the approach can be adapted for the Ambisonic encoding of reverberation depending on source and listener positions and orientations.

2   Source Representation

We wish to use a representation which can encode any source to any desired accuracy, relates well to direct measurements of the field, and can be manipulated efficiently. The following possibilities suggest themselves. A source can be modelled with several monopoles. This would be appropriate if it actually has this structure, or because a rough and fast model is required. The source can be positioned and orientated using standard cartesian transformations. For more accuracy we can attempt to use many monopoles distributed over the source volume or surfaces. It is far from obvious how this would be done for a general source. Such a representation contains considerable redundancy since it describes the structure of the object as well as the sound produced.

2.1   The exterior harmonic expansion

Multipoles in their original form consist of infinitesimal arrangements of monopole sources. A multipole of sufficient order can represent a the field around a given extended object arbitrarily well. Although they are operated on by simple cartesian operations, their infinitesimal nature does not lend itself to direct numerical manipulation. Also the relationship of multipole parameters to the directionality of the field rapidly increase in complexity with order. Closely related is the *exterior expansion* for the wave equation. This has basis functions in the frequency domain using spherical coordinates, $h_m(kr)Y_{mn}(\theta, \delta)$, where $h_m(kr)$ are the spherical hankel functions of the second kind, [11]. m is the multipole order of each function, and $k = 2\pi/\lambda$ is the wavenumber. The type of hankel function chosen gives an outward moving wave when associated with a positive frequency time piece $e^{i\omega t}$, the same convention used in [5].

An infinitesimally defined multipole of order m can always be expressed exactly using an exterior expansion with terms up to order m. For this reason

an exterior expansion is alternatively called an *exterior multipole expansion* or just a *multipole*, [12]. Another term used is *singular expansion*, since the center of the expansion has a singularity. The exterior expansion relates closely to the non-uniform directivity of a source, as discussed below, and our principal goal shall be to manipulate it to provide an Ambisonic source encoding. By multipole we shall mean an exterior expansion, unless otherwise stated.

The remainder of this section reviews the exterior expansion and introduces the conventions that will be used. In keeping with the high-order Ambisonic literature the real-valued N3D spherical harmonic set will be used throughout, [13, 5]. The components are defined for $m \geq 0$ and $m \geq n \geq 0$ by

$$Y_{mn}^{\sigma}{}^{(N3D)}(\theta, \delta) = \sqrt{2m+1} \; \tilde{P}_{mn}(\sin \delta) \times \begin{cases} \cos n\theta & \text{if } \sigma = +1 \\ \sin n\theta & \text{if } \sigma = -1 \end{cases} \tag{1}$$

$$\tilde{P}_{mn}(\sin \delta) = \sqrt{(2 - \delta_{0,n}) \frac{(m-n)!}{(m+n)!}} \; P_{mn}(\sin \delta) \tag{2}$$

For $n = 0$, σ only takes the value $+1$. θ here measures the angle around the coordinate symmetry axis. $\pi/2 - \delta$ is the angle between the axis and the coordinate direction, so that δ would normally be called the elevation, as shown in Figure 2. The symmetry axis is normally called the z axis, which is not neccessary, but aids labeling in diagrams.

We shall use a slightly simplified notation that removes σ by extending n to negative values as used in more conventional harmonic sets,

$$Y_{mn} = \begin{cases} Y_{mn}^{+1} & \text{if } n \geq 0 \\ Y_{m|n|}^{-1} & \text{if } n < 0 \; . \end{cases} \tag{3}$$

For convenience we define coefficients, $O_{mn}(k)$, by a general exterior expansion,

$$p(\boldsymbol{r}, k) = k \sum_m i^{-m-1} h_m(kr) \sum_n Y_{mn}(\theta, \delta) O_{mn}(k) \; , \tag{4}$$

so that in the farfield where $h_m(kr)$ tends to $i^{m+1} e^{-ikr}/kr$, the field becomes
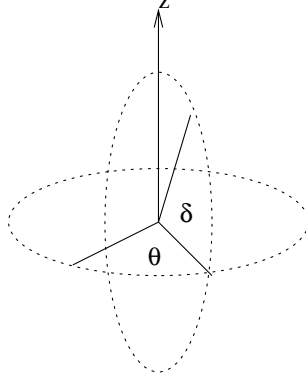
Figure 2: Spherical coordinates used.

$$p_{far} = \frac{e^{-ikr}}{r} \sum_{m,n} Y_{mn}(\theta, \delta) O_{mn}(k) \ . \tag{5}$$

The $O_{mn}(k)$ coefficients then directly express the non-uniform directivity in this regime, where locally the field tends to an outward moving plane wave. The signals $O_{mn}(k)$ coincide with the *O-format* encoding used previously for Ambisonic synthesis, [3, 4]. The same name will be used here for the more general case described by (4). We emphasize that this is just a convention, for convenience and appropriate to its context, in the same sense as B-format is defined. Nothing essentially new is added.

$O_{mn}(k)$ can be readily calculated from measurements of the field on a sphere at any radius r outside the source region. Applying an integral over the sphere, $\int d\Omega \ Y_{mn}(\theta, \delta)$ to (4) gives

$$O_{mn}(k) = \frac{i^{m+1} \int d\Omega \ Y_{mn}(\theta, \delta) p(\boldsymbol{r}, k)}{4\pi k h_m(kr)} \ . \tag{6}$$

Note $h_m(kr)$ does not have zeros for $r > 0$ so this is well defined at all r. For a real object the field could be measured approximately with pressure microphones placed located on a sphere a fixed distance from the source. In the farfield where the field becomes planar, inwardly pointing directional microphones are equally effective given the appropriate equalization including phase. For devices such as loudspeakers that convert electricity to soundly, the process can be simplified by repeated response measurements with a

single microphone that is moved. Speaker simulations might for instance be useful in high-end architectural simulations. When the $O_{mn}(k)$ responses are convolved with input signals for the speakers, the expansion signals are generated.

2.2   Source approximation order and error

We consider now the order to which a source is approximated, m_{max}. We wish to minimize this subject to reconstruction error constraints. A source can be arbitrarily small and still have power up to any multipole order, for example using the explicit definition of infinitesimal multipoles. However this is unusual in a real acoustic source because opposed component sources are not usually found very close together. To gain insight into the more usual case, we examine multipole fields for a source consisting of a monopole offset from the expansion centre. A monopole source at position r' has the following multipole expansion in r about the origin, valid for $r > r'$, [11],

$$\frac{e^{-ik|r-r'|}}{|r-r'|} = ik \sum_{m=0}^{\infty} j_m(kr')h_m(kr) \sum_{n=-m}^{m} Y_{mn}(\theta', \delta')Y_{mn}(\theta, \delta) \qquad (7)$$

$j_m(kr')$ is the spherical Bessel function of the first kind. The arrangement is illustrated in Figure 3.

Note that in this special case, fixing r and varying r' gives a valid field expansion for the spherical region centered on the origin, that just excludes the monopole. However, the field inside an extended source is not fully determined by its exterior field.

Figures 4 and 5 show some cross-sectional plots of $Re(p)/|p|$, for different orders and offset $r' = 2\lambda$, where $\lambda = 2\pi/k$ is the wavelength. Only one half the plane is shown because the field is symmetric about the line from multipole to monopole. The nearfield resolves sharply as the order is increased. For order $m_{max} \approx 2\pi r'/\lambda = kr'$, the relative error compared with a real monopole is $< 1\%$ for $r > r' + \lambda$. Detailed error analysis of the multipole approximation, [12], agree with these observations, and fast convergence is cited as one of the key attributes of the spherical multipoles.

Rewriting for m_{max} that will ensure good reconstruction of a field from a region of diameter d, up to frequency, f_{max}, we find $m_{max} = \pi d f_{max}/c$. For example, frequencies up to 1500 Hz from an object 1 m width can be approximated well with $m_{max} = 14$. If we do not require good construction
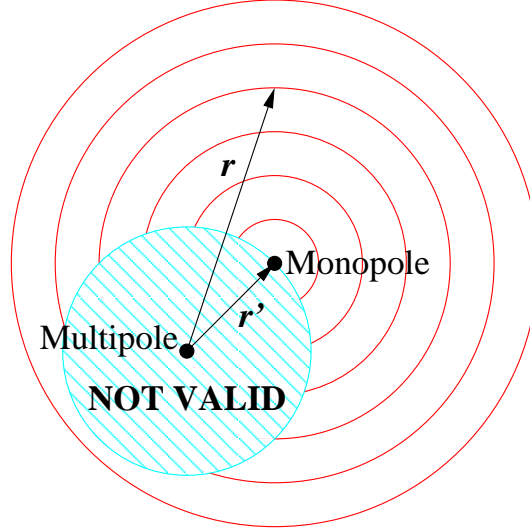
6

Figure 3: Multipole representation of a monopole.

close to the limit $r = r'$, then the order can be reduced further. The farfield accuracy is not of so much interest, since at all orders the farfield tends locally to a plane wave. This can be conventionally encoded in Ambisonics using $B_{mn} = Y_{mn}(\theta_s, \delta_s)p(k)$, where the direction is to the source and $p(k)$ is the pressure from the source measured at the listener. Increasing the multipole order in this approximation can improve reconstruction in the farfield, but not the nearfield.

For a general source with diameter d, we cannot expect to use a lower order than the displaced monopole example for similar accuracy, because it would be unusual that different parts of the object would cancel out at higher orders. Conversely we wouldn't expect higher orders to be required, because that would require even more cancellation in order to make higher orders relatively significant. In summary, the formula for m_{max} provides a first estimate for the order required to represent the nearfield of a general object, although this is not true across all possible sources.

2.3 Multi-resolution sources

So far single multipole expansions have been considered for each object. In some cases a hybrid approach may be more appropriate, in which a source
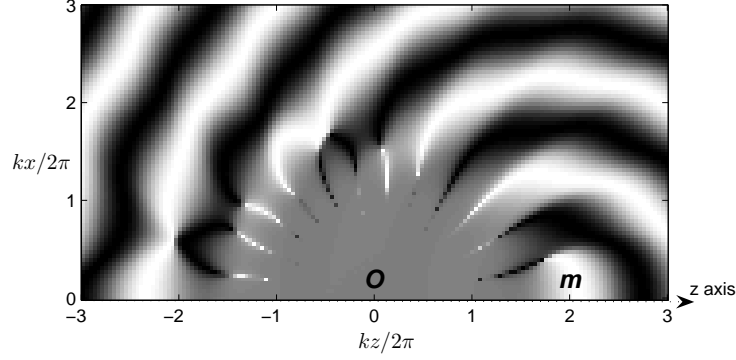
Figure 4: Multipole approximation, center **O**, of a displaced monopole, center **m**, $r' = 2\lambda$, $m_{max} = 12$. The cross-section is $\theta = 0$. x, z are cartesian coordinates in length units.

is represented using several multipoles. This is neccessary whenever we wish to find the field at a free space inside the bounding sphere of an object, for example nearer to a table surface than its length. Figure 6 illustrates this.

Outside a bounding sphere a single multipole is sufficient, by expanding around the center of the sphere. As we move closer to some part of the source more multipoles are neccessary. Far from the object compared to its size the field can be approximated as a plane wave, using the O-format coefficients to determine the direction dependence. This scheme of successive simplification resembles the multi-resolution techniques common in computer graphics, [14].
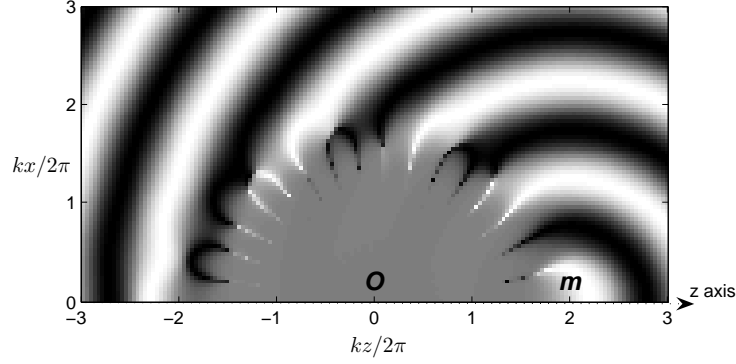
Figure 5: Multipole approximation, center **O**, of a displaced monopole, center **m**, $r' = 2\lambda$, $m_{max} = 16$. The cross-section is $\theta = 0$. x, z are cartesian coordinates in length units.
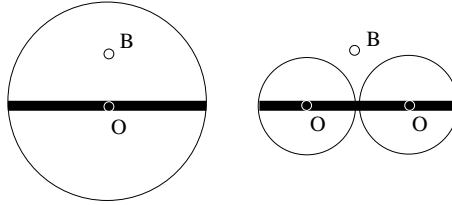

Figure 6: When the table sound comes from two sources (right), the listener, B, can be closer.

3 Ambisonic encoding of multipoles

3.1 Freefield expansion

High-order Ambisonics is founded on the *interior expansion* that we shall also call the *freefield expansion* here, to emphasize that it is used to describe a sourceless region around the listener. Eq. (8) is the version of the expan-

sion using N3D harmonics, $Y_{mn}(\theta, \delta)$, and defines the B-format coefficients, $B_{mn}(k)$, [5]. The expansion converges quickly on any source-free field, up to a given radius r. The typical order required to achieve $\approx 1\%$ error for a regular freefield, such as a planewave, is $m_{max} \approx kr$, [12, 15].

$$p(\boldsymbol{r}, k) = \sum_m i^m j_m(kr) \sum_n Y_{mn}(\theta, \delta) B_{mn}(k) \tag{8}$$

3.2 Monopole encoding

For fields containing sources it is still possible to create a freefield expansion, however it is only valid within a region that does not contain any source. Consider first a field containing a single monopole set away from the freefield expansion center. Eq. (7) can be recast as the freefield expansion for a monopole, by fixing $\boldsymbol{r}$ and instead varying $\boldsymbol{r}'$. The form of this expansion is then consistent with Eq. (8), from which the values of the $B_{mn}(k)$ can be read off, as shown in [5]. The condition of convergence $\boldsymbol{r}' < \boldsymbol{r}$ now implies that the expansion converges *within* a circle that just touches the monopole source. Figure 7 shows a field plot for such a monopole reconstructed to the 13th order, and set at a distance 2λ from the expansion center. Outside this area the expansion is a valid freefield, although no longer matches the source field. Overall convergence behavior within the valid region is like any other freefield, although close to the monopole, $\delta < \lambda$, the order required to achieve a given error is increased compared to a smooth freefield, as we would expect, [12]. The limit set to the region of freefield convergence by the source can not be exceeded by increasing the freefield order.

Higher multipole sources have similar freefield expansions, since they can be generated as composites of infinitesimal monopoles. The convergence condition is then that the freefield expansion is valid within a radius that does not include any of the sources.

3.3 Multipole to freefield coefficient transformation

The main task in this section is to find $B_{mn}(k)$ in the presence of a multipole described by $O_{mn}(k)$ at a given position. It would be desirable to find a generalized closed form expression, as for the monopole case in [5]. However, it is not very apparent how this could be done or even if it would be the most practical method of calculation, so instead a more pragmatic approach
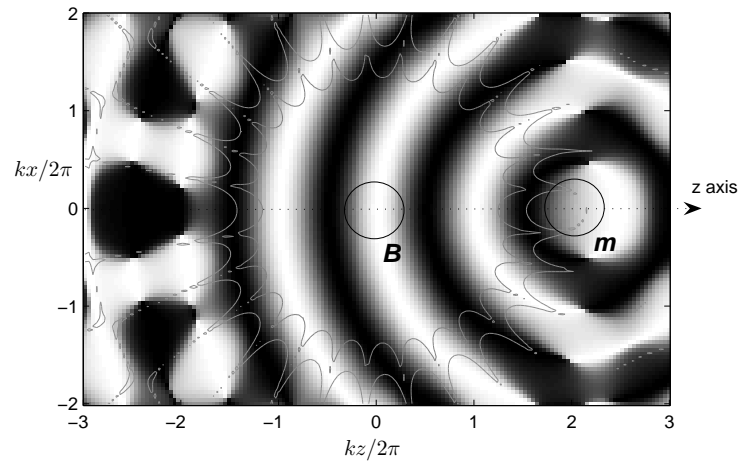
Figure 7: Cross-section of a field plot for a 13th order freefield expansion, center at **B**, of a monopole, center **O**. Error contours are shown at the 1% and 10% levels. The cross-section is $\theta = 0$. x, z are cartesian coordinates in length units.
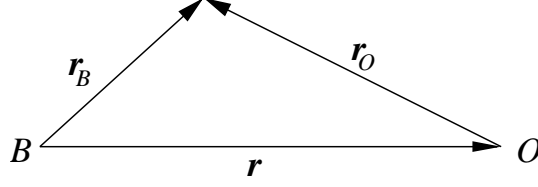
Figure 8: Vector notation

is adopted yielding eventually a manageable integral expression. To begin (8) and (4) are equated. The notation is modified according to Figure 8,

$$\sum_m i^m j_m(kr_B) \sum_n Y_{mn}(\theta_B, \delta_B) B_{mn}(k) = k \sum_m i^{-m-1} h_m(kr_O) \sum_n Y_{mn}(\theta_O, \delta_O) O_{mn}(k)$$

(9)

To isolate $B_{mn}(k)$ the operator $\int d\Omega_B \, Y_{m'n'}(\theta_B, \delta_B)$ is applied, with r_B a freely chosen constant, and θ_O, δ_O and r_O are functions of the vector $\boldsymbol{r}_B$, yielding

$$4\pi i^{m'} j_{m'}(kr_B) B_{m'n'}(k) = k \sum_m i^{-m-1} \sum_n O_{mn}(k) \int d\Omega_B \, Y_{m'n'}(\theta_B, \delta_B) Y_{mn}(\theta_O, \delta_O) h_m(kr_O).$$

(10)

Relabeling indices, $B_{mn}(k)$ can be written as

$$B_{mn}(k) = \sum_{m',n'} M_{mnm'n'}(k, \boldsymbol{r}) O_{m'n'}(k),$$

(11)

where the filter matrix $M_{mnm'n'}(k, \boldsymbol{r})$ is

$$M_{mnm'n'}(k, \boldsymbol{r}) = \frac{ki^{-m-m'-1}}{4\pi j_m(kr_B)} \int d\Omega_B \, Y_{mn}(\theta_B, \delta_B) Y_{m'n'}(\theta_O, \delta_O) h_{m'}(kr_O). \quad (12)$$

The term $j_m(kr_B)$ in the denominator has zeros for kr_B at approximately regular intervals with period π. However, (12) is well defined, because the zeros can be shifted by changing the free parameter $\alpha = r_B/r$. We return to this later. The $\boldsymbol{r}$ direction dependence in the matrix can be factored out by transforming the components $B_{mn}(k)$ and $O_{mn}(k)$ so they are relative to $\boldsymbol{r}$.
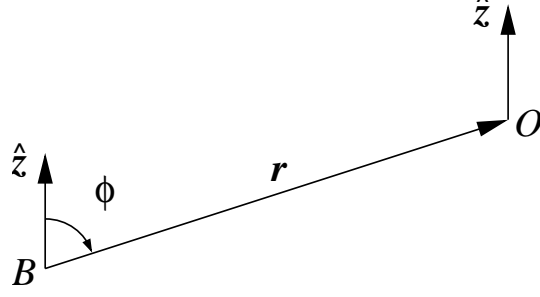
12

Figure 9: Finding components relative to $\boldsymbol{r}$

Figure 9 shows the relationship between the initial coordinate axis, $\hat{\boldsymbol{z}}$, and the vector, $\boldsymbol{r}$, connecting the centres B and O. $\phi = \pi/2 - \delta$ together with θ specify a rotation mapping $\hat{\boldsymbol{z}}$ onto $\boldsymbol{r}$, written in components as $R_{m'n'n}(\theta, \phi)$. The third degree of freedom is unspecified, although it must be consistent. Therefore $R_{m'n'n}(\theta, -\phi)$ transforms $B_{mn}(k)$ and $O_{mn}(k)$ to find their coordinates relative to $\boldsymbol{r}$,

$$B'_{m'n'}(k) = \sum_n R_{m'n'n}(\theta, -\phi) B_{m'n}(k) \tag{13}$$

$$O'_{m'n'}(k) = \sum_n R_{m'n'n}(\theta, -\phi) O_{m'n}(k) \tag{14}$$

Now Eq. (11) can be written with an $\boldsymbol{r}$-direction-independent matrix, $M_{mnm'n'}(k, r)$,

$$B'_{mn}(k) = \sum_{m',n'} M_{mnm'n'}(k, r) O'_{m'n'}(k), \tag{15}$$

where

$$M_{mnm'n'}(k, r) = \frac{ki^{-m-m'-1}}{4\pi j_m(kr_B)} \int d\Omega_B \; Y_{mn}(\theta_B, \delta_B) Y_{m'n'}(\theta_O, \delta_O) h_{m'}(kr_O). \tag{16}$$

The coordinates in the integral are now relative to $\boldsymbol{r}$, although they haven't been relabeled. The symmetry this brings, with r_O independent of $\theta_B = \theta_O$, can be used to factor the integral into a product of θ and δ integrals. To make this clear $Y_{mn}(\theta, \delta)$ is factored into

13

$$Y_{mn}(\theta, \delta) = \hat{P}_{mn}(\sin \delta) \times \begin{cases} \sqrt{2} \cos n\theta & \text{if } n > 0 \\ 1 & \text{if } n = 0 \\ \sqrt{2} \sin n\theta & \text{if } n < 0 \end{cases} \qquad (17)$$

where for convenience later, $\hat{P}_{mn}$ is defined,

$$\hat{P}_{mn}(\sin \delta) = \sqrt{(2m+1)\frac{(m-|n|)!}{(m+|n|)!}} \; P_{m|n|}(\sin \delta) \qquad (18)$$

and $P_{mn}(x)$ is the associated Legendre polynomial. (16) becomes

$$\begin{aligned} M_{mnm'n'}(k, r) &= \frac{ki^{-m-m'-1}}{4\pi j_m(kr_B)} \int d\delta_B \; \cos \delta_B \hat{P}_{mn}(\sin \delta_B) \hat{P}_{m'n'}(\sin \delta_O) h_{m'}(kr_O) \times 2\pi \delta_{nn'} \\ &= \frac{\delta_{nn'} ki^{-m-m'-1}}{2j_m(kr_B)} \int_{-1}^{+1} ds_B \hat{P}_{mn}(s_B) \hat{P}_{m'n'}(s_O) h_{m'}(kr_O), \qquad (19) \end{aligned}$$

where $s_B = \sin \delta_B$ and $s_O = \sin \delta_O$. s_O and r_O can be found from r, r_B and s_B using $r_B s_B - r_O s_O = r$. $r_O = r\sqrt{1 + \alpha^2 - 2\alpha s_B}$ and $s_O = r(\alpha s_B - 1)/r_O$, where $\alpha = r_B/r$. The term $\delta_{nn'}$ in Eq. (19) allows a simplified 3-index coefficient matrix to be defined, which can be further simplified by factoring out a term $1/r$ and re-expressing the remainder in terms of the product kr. The implication is that each digital filter derived can be varied according to distance, r, by frequency scaling by r.

Eq. (15) can now be rewritten,

$$B'_{mn}(k) = \sum_{m,n} \frac{1}{r} M_{mnm'}(kr) O'_{m'n}(k), \qquad (20)$$

where the new 3-index matrix coefficient is,

$$M_{mnm'}(k) = \frac{ki^{-m-m'-1}}{2j_m(k_B)} \int_{-1}^{+1} ds_B \hat{P}_{mn}(s_B) \hat{P}_{m'n}(s_O) h_{m'}(k_O), \qquad (21)$$

and $k_B = \alpha k$ and $k_O = k\sqrt{1 + \alpha^2 - 2\alpha s_B}$. Clearly this is defined only for $n < m$ and $n < m'$, so for a given source the number of filters increases only linearly with B-format order required. The new filter coefficients are given in terms of one parameter, k. The actual filter acting in (20) is scaled in frequency by the radius r and there is a distance factor $1/r$.

Choosing $\alpha = m/k$ keeps $j_m(k_B)$ close to its first maximum, avoiding the zeros. For small k, α is limited < 1, otherwise the integral fails. This does not reintroduce a zero. For $m = 0$, $\alpha = 1/k$ is used. With the MATLAB integrator the coefficients can be evaluated to 4 significant figures for coefficient magnitudes between 10000 and 0.0001 over indices up to $m, m' = 30$ and up to $k = 10000$. The accuracy of each coefficient can be checked by using multiple values of α, and using the symmetry described later. Small k coefficients are also discussed below. At this stage the accuracy is sufficient to prove the scheme, and create prototype filters. Alternative methods of calculation might prove to be valuable, especially for large k in some cases, although the asymptotic behaviour in this region is quite predictable.

Combining (20) and (21) with (13) gives a new expression of $B_{mn}(k)$ in terms of $O_{mn}(k)$,

$$B_{mn}(k) = \sum_{n'} R_{mnn'}(\theta, \phi) \sum_{m'} \frac{1}{r} M_{mn'm'}(kr) \sum_{n''} R_{m'n'n''}(\theta, -\phi) \; O_{m'n''}(k)$$

(22)

An orientation rotation could be included into the first rotation acting on O_{mn}. With rotations included, the number of filters required for reconstructing the field from a given source is still linear in the maximum B-format order, m_{max}, for $m_{max} > m'$, owing to zeros in $M_{mn'm'}(kr)$.

3.4 Validation and properties

To provide an immediate confidence test that the derived formulas are correct, a random test 5th order multipole was constructed, shown in Figure 10, and compared with the 13th order freefield expansion calculated using the matrix (22), shown in Figure 11. The error contours in Figure 11 at 10% and 1% levels are for deviations from the original multipole shown in Figure 10. The region of agreement extends as far as the center of the original multipole, as expected, and indicates that the calculations described in this section are correct.

Next we examine $M_{mnm'}(kr)$ by checking that it is consistent with previous results for the monopole case,[5], in which the encoded signal is given by $B_{mn} = S(k)F_m(kr)Y_{mn}(\theta, \delta)$, where $F_m(kr) = i^{-m}h_m(kr)/h_0(kr)$. To match the alignment used to define $M_{mnm'}(kr)$, $\theta = \delta = \pi/2$. We first note that the source term $S(k)$ includes the delay and distance attenuation so that
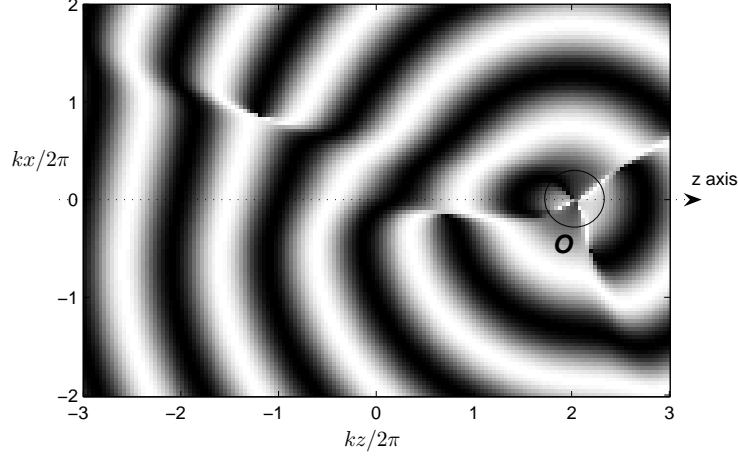
Figure 10: Cross-section of a field plot for a 5th order multipole, center at **O**. The cross-section is $\theta = 0$. x, z are cartesian coordinates in length units.

$S(k) = \frac{e^{ikr}}{r}O_{00}(k)$. In order to isolate the part matching $F_m(kr)$, we look at the adjusted value, $(1/r)M_{m00}(kr)/(e^{-ikr}/r)/Y_{m0}(0, \pi/2)$. With $r = 1$, this produces the plots shown in Figure 12, matching previous results, [5]. Further plots reveal how M extends to encode higher multipoles. Figure 13 is an example showing the amplitude response for $M_{211}(k)/e^{-ik}$ along with $kM_{211}(k)/e^{-ik}$ to make clear the large k behavior clear. Figure 14 shows the corresponding phase response. Another example, with $n = 2$, is shown in Figures 15 and 16. The general picture is that with the e^{-ikr}/r piece factored out, the response is always minimum phase. For small k the order of the filter becomes $m + m'$, while for large k it is n. The location of the transitional region increases linearly with $m + m'$ from $k \approx 2$ for $m + m' = 1$. For higher orders, the transitional region can be more complex.

3.4.1 Symmetries

$M_{mnm'}(k)$ has symmetries, that reduce the number of filters that are needed. $M_{mnm'}(k) = M_{m-nm'}(k) = M_{m'nm}(k)(-1)^{m+m'}$. The last is useful for cross checking the numerical accuracy of a value, since the two symmetric integrals
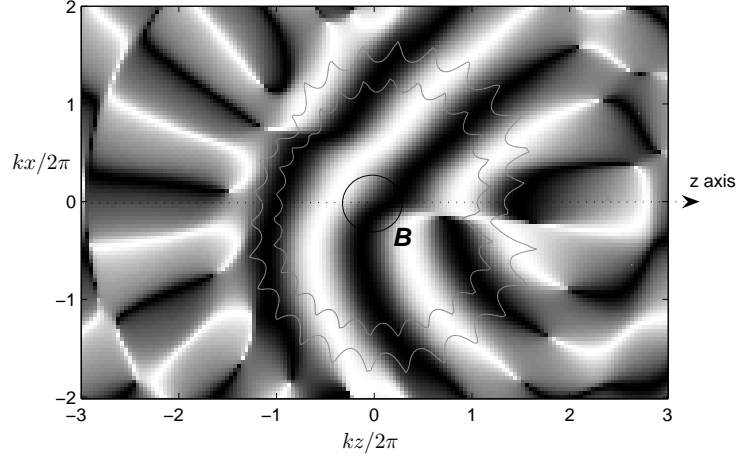
Figure 11: Cross-section of a field plot for a 13th order freefield expansion, center at **B**, of a multipole, center **O**. Error contours are shown at the 1% and 10% levels. The cross-section is $\theta = 0$. x, z are cartesian coordinates in length units.
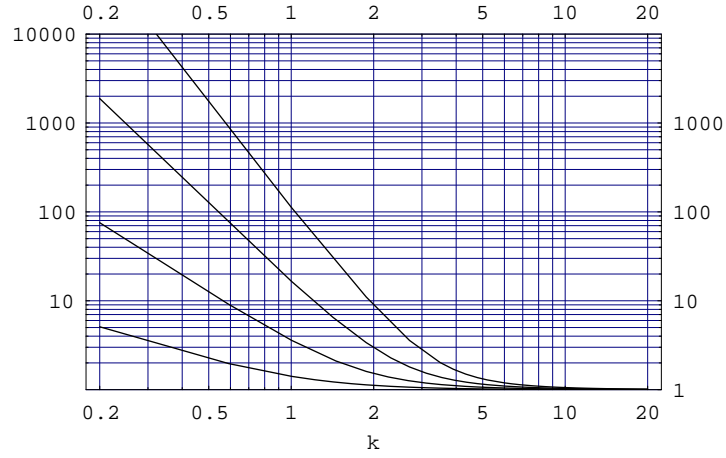

Figure 12: Amplitude response for $M_{m00}(k)/e^{-ik}/Y_{m0}(0, \pi/2)$, $m = 1, 2, 3, 4$
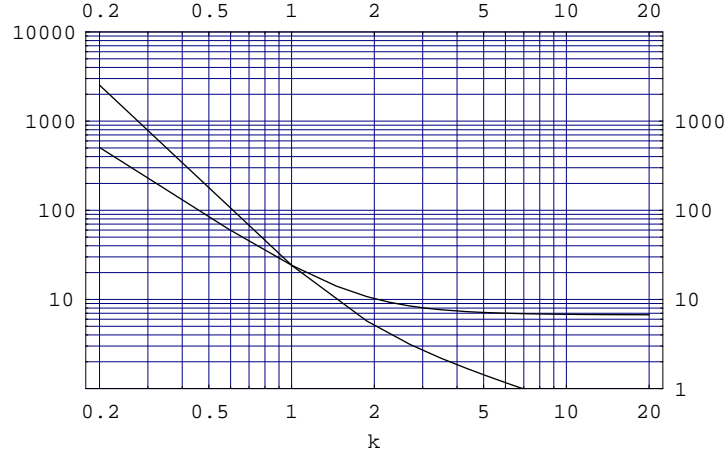
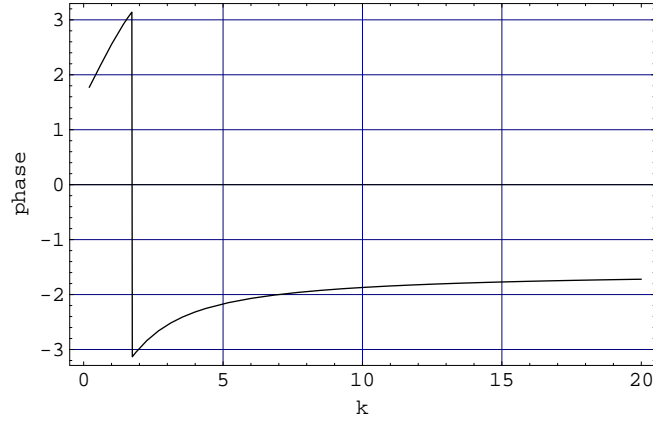Figure 13: Amplitude response for $M_{211}(k)/e^{-ik}$ and $kM_{211}(k)/e^{-ik}$

Figure 14: Phase response for $M_{211}(k)/e^{-ik}$

involve distinct calculations.

3.4.2 Small k and convergence

In [5] filters were commuted from speaker-array decoding to control the large amplitudes generated in monopole encoding at low values of k. The situation at first appears worse here because filters encoding high multipoles can have
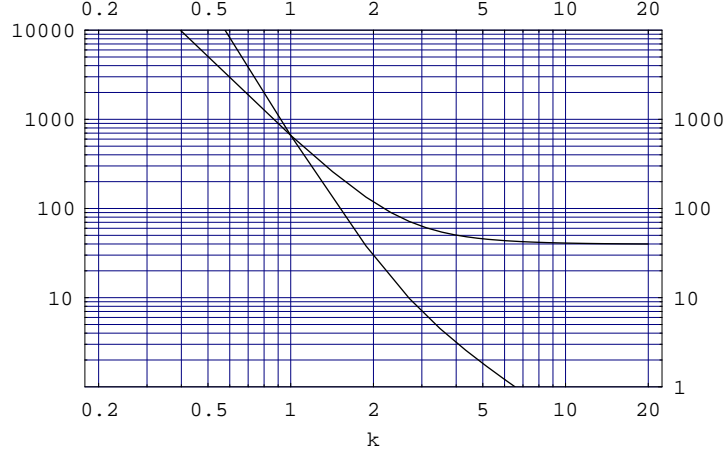
Figure 15: Amplitude response for $M_{223}(k)/e^{-ik}$ and $k^2 M_{223}(k)/e^{-ik}$
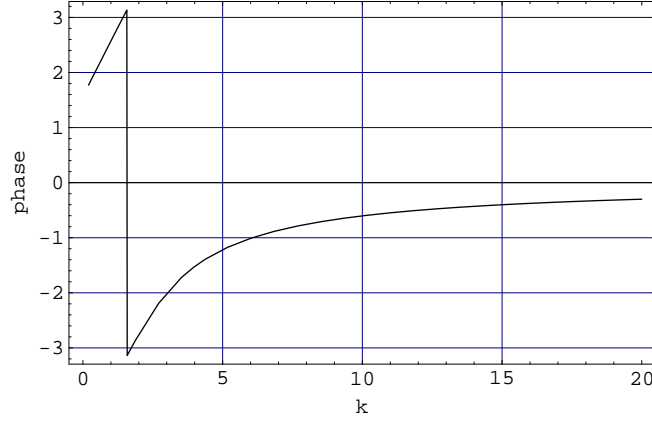

Figure 16: Phase response for $M_{223}(k)/e^{-ik}$

much higher order at low k, for the same order m of the B-format encoding. It has already been noted in section 3.1 that the freefield expansion typically converges rapidly at $m \approx rk$. Convergence to a multipole field is not as good close to the multipole, as we would expect, and depends on the complexity of the multipole. This is evident in figures 7 and 11, where $rk = 22\pi \approx 13$, and $m_{max} = 13$. For $m + m' < rk$ we find that the size of transfer coefficients are $\lesssim 10$ for all m, m' up to the maximum of 30 tested. From section 2.2 an object

19

which requires up to order m_o has typical radius r_o where $r_o k \approx m_o$. This suggests the freefield can be converged well up to the surface of the object using coefficients that have not become large, since $m_r = rk - m_o = (r - r_o)k$, see figure 17. For $rk < m + m'$ the corresponding transfer coefficient is then no longer significant in the reconstructed field. The implication for digital filter design, is that the responses for $k < m + m'$ can be limited above to make the filters stable, without significantly affecting reconstruction accuracy. In [16] a similar conclusion was reached for the monopole case. For $rk < 1$, the *proximity effect* becomes significant. This is the low frequency boost of directional microphones and hearing. The dominant contribution comes from the $m = 1, m' = 0$ component, so this filter should be extended sufficiently to cover the range of boost required. Small sources, that can be approached closely, are typically much weaker sources of low frequencies. Also, the listener's ears will not normally be less than a few cms to the source, and then this is only practical using a binaural rendering system. More investigation is needed on the convergence properties of freefield multipole reconstruction, although current results indicate that the scheme presented here is workable.
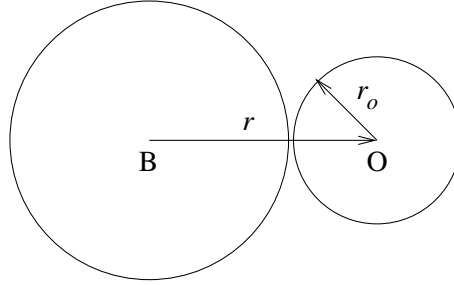

Figure 17: Listener convergence region, B, limited by source object region, O.

3.4.3 Digital filters

The implementation of the digital filters is not investigated in detail here, but note that in general it will take a similar form to that described for the monopole case, [5]. Because the filters are evaluated numerically, they must be converted to IIR form by pole/zero fitting. Variation according to object

distance can then achieved by frequency scaling the poles and zeros. From the previous section, it is not required to evaluate (21) for small k, where the numerical behavior eventually breaks down. The filters for higher values of m, n, m' in $M(m, n, m')$ will be more costly. This will increase the linear rise in overall cost with m_{max} mentioned earlier.

4 Reverberation encoding and transformation

This article has focused on the synthesis of the direct signal from complex objects in the nearfield. Related techniques can be applied to the synthesis of a reverberant signal originating from a complex source.

4.1 Encoding source directivity in reverberation

Conventional room responses are mono to mono, or mono to multichannel, where multichannel could include B-format. This means that the reverberant sound from a source is the same no matter how it is oriented, where as real reverberation can vary considerably. Using a harmonic matrix of the form Eq. (12) it is possible, because of linear supposition, to fully encode the reverberant response to source directivity, for a given pair of source and listener positions. For a source described by $O_{m'n'}(k)$ and a listening field by $B_{mn}(k)$,

$$B_{mn}(k) = \sum_{m',n'} M^{rev}_{mnm'n'}(k, \boldsymbol{x}_B, \boldsymbol{x}_O)O_{m'n'}(k) \quad , \tag{23}$$

where $\boldsymbol{x}_B$ and $\boldsymbol{x}_O$ are the position vectors of the listener and source. The $M^{rev}_{mnm'n'}$ could be measured in a real space using directional sources and directional microphones, measuring the response on each microphone component for each source component. There are many ways this could be achieved, but in any case the directional components must be able to cover the desired order of spherical harmonic resolution. The order of the source components is not so important compared to the order of the microphones, since the latter must support the spatial resolution in the listener's hearing. A small source order increase is expected to deliver a considerable advantage over the conventional zero order source. The encoding of source directivity could be particularly valuable in an interactive application where the listener controls the orientation of a sound source, and so can probe the surrounding acoustic.

21

$M^{rev}_{mnm'n'}$ could also be calculated offline by applying the source-microphone technique to simulations using ray-tracing or other models.

4.2   Transforming reverberation with listening position

A modified matrix technique can be used to transform reverberation according to a change of listening position. This allows a single $M^{rev}_{mnm'n'}$ to be used for a range of listener positions, by transforming B_{mn} from Eq. (23) to the reverberant signal at the new position, $B'_{mn}(k)$. In [17] there is a partial description of this method. The first pressure or zero-order component of the reverberant signal is found everywhere in a room using a single expansion. Here we present the full high-order encoding of reverberation about any point. Reworking the previous calculations leading to Eq. (21) produces the following result in terms of a new matrix $M^{BB}_{mnm'}$,

$$B'_{mn}(k) = \sum_{n'} R_{mnn'}(\theta, \phi) \sum_{m'} \frac{1}{r} M^{BB}_{mn'm'}(kr) \sum_{n''} R_{m'n'n''}(\theta, -\phi) \ B_{m'n''}(k) \ \ ,$$

(24)

where

$$M^{BB}_{mnm'}(k) = \frac{k i^{m'-m}}{2 j_m(k_{B'})} \int_{-1}^{+1} ds_B \hat{P}_{mn}(s_{B'}) \hat{P}_{m'n}(s_B) j_{m'}(k_B) \ \ , \qquad (25)$$

with $k_{B'}$, $s_{B'}$, k_B and s_B replacing k_B, s_B, k_O and s_O respectively. To verify this, a random 12th order field B_{mn} has been synthesized, shown in Figure 18, and transformed using Eq. (24) to an 18th order field B'_{mn} centered at a distance 2λ from B_{mn}, shown in Figure 19. The error contours show clear agreement which extends beyond the center of B_{mn} to a radius r where $kr \approx m_{max} = 18$, as expected. As noted previously, when transforming a multipole source it is not possible to extend beyond the original center in this way.

This technique can reproduce high definition spatialized reverberation more efficiently than by direct simulation with ray tracing and similar methods, since the computational complexity is limited by the reconstruction order and not the complexity of the simulation.
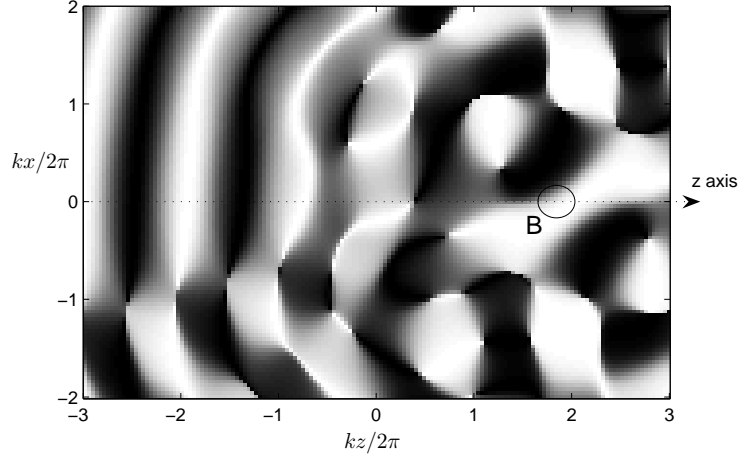
Figure 18: Cross-section of a field plot for a 12th order freefield harmonic expansion, center at **B**. The cross-section is $\theta = 0$. x, z are cartesian coordinates in length units.

5   Conclusion

A method has been presented for encoding a general acoustic source, and transcoding it to a high-order Ambisonic signal, dependent on source orientation, and position relative to the listener. The method also lends itself to the direct measurement of real sources using an array of surrounding microphones. The approach is considerably more elaborate and costly than plane wave or monopole synthesis, however it is expected that in the context of complex sources displayed with a high-quality rendering system, the efforts are worthwhile. Binaural headphone reproduction is particularly attractive, because the encoding only needs to be of sufficient order for a single listener rather than a listening area, so reducing computational costs. For instance for radius 0.2 m, up to 1500 Hz, the required order, $m = rk \approx 6$. In a speaker rendering environment the valid listening region is necessarily fixed to accommodate multiple listeners. This places constraints on how nearfield sources can be arranged relative to the listener, so for example it is impossible for a listener to experience near sources directly on the left and the
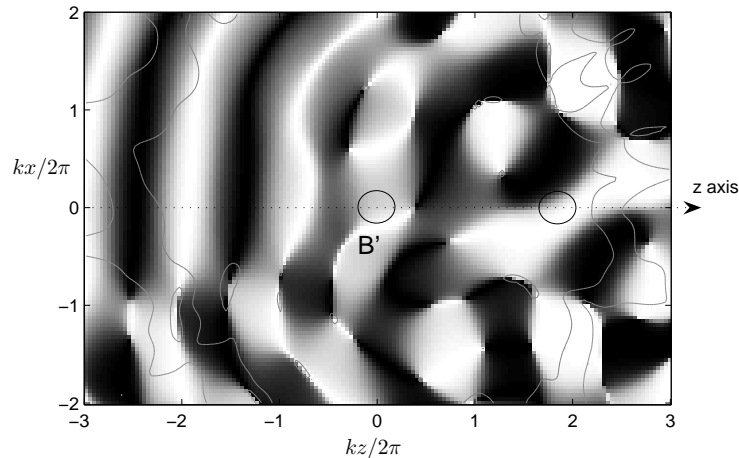
23

Figure 19: Cross-section of a field plot for a 18th order freefield expansion, center at **B'**, of a freefield, center **B** shown in Figure 18. Error contours are shown at the 1% and 10% levels. The cross-section is $\theta = 0$. x, z are cartesian coordinates in length units.

right sides while also having a large listening area that can hold multiple listeners. Binaural reproduction does not suffer this constraint, and so is the more natural method for nearfield rendering. A possible exception would be a small speaker array designed for one person. Related results have also been presented for the generation and transformation of reverberation. In the future we hope to investigate realizations of these methods.

References

[1] M.A. Gerzon. Ambisonics in multichannel broadcasting and video. *Journal of the Audio Engineering Society*, 33:859–871, 1985.

[2] M.A. Gerzon. General metatheory of auditory localisation. In *92nd AES Convention Preprints*, 1992.

[3] D. Menzies. *New Performance Instruments for Electroacoustic Music.* PhD thesis, University of York, UK, 1999.

[4] D. Menzies. W-panning and O-format, tools for object spatialisation. In *Proc. AES 22nd International Conference*, 2002.

[5] J. Daniel. Spatial sound encoding including near field effect. In *Proc. AES 23nd International Conference*, 2003.

[6] A.J. Berkhout, D.D. Vries, and P. Vogel. Acoustic control by wave field synthesis. *J. Acoust. Soc. Am.*, 93:2764–2778, 1993.

[7] J. Daniel, R. Nicol, and S. Moreau. Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging, 2003.

[8] T. Caulkins, E. Corteel, and O. Warusfel. Synthesizing realistic sound sources in wfs installations. In *Proc. DAFX04*, 2004.

[9] O. Warusfel. Reproduction of sound source directivity for future audio applications. In *Proc. International Congress on Acoustics*, 2004.

[10] D. Menzies and M. Al-Akaidi. Nearfield binaural synthesis. *J. Acoust. Soc. Am.*, 121(3):1559-63., March 2007.

[11] P.M. Morse and K.U. Ingard. *Theoretical Acoustics.* McGraw-Hill, 1968.

[12] N. A. Gumerov and R. Duraiswami. *Fast multipole methods for the Helmholtz equation in three dimensions.* Elsevier Science, 2005.

[13] J. Daniel. *Représentation de Champs Acoustiques, Application la Transmission et à la Reproduction de Scnes Sonores Complexes dans un Contexte Multimdia.* PhD thesis, University of Paris 6, Paris, France, 2000.

[14] N. A. Dodgson, M. S. Floater, and M. A. Sabin. *Advances in Multiresolution for Geometric Modelling.* Springer, 2005.

[15] D. B. Ward and T. D. Abhayapala. Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Transactions on Speech and Audio Processing*, 9(6):697, September 2001.

[16] J. Daniel and S. Moreau. Further study of sound field coding with higher order ambisonics. In *AES 116th Convention Preprints*, 2004.

[17] R. Duraiswami, N. A. Gumerov, D. N. Zotkin, and L. S. Davis. Efficient evaluation of reverberant sound fields. In *Proc. of the 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2001.