Ambisonic Synthesis of Complex Sources*

DYLAN MENZIES, *AES Member,* **AND MARWAN AL-AKAIDI**
(rdmg@dmu.ac.uk) (mma@dmu.ac.uk)

De Montfort University, Leicester, LE1 9BH, UK

Exterior expansions of complex sound sources are presented as flexible objects for producing Ambisonic sound-field 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.

0 INTRODUCTION

Real sound fields are frequently the result of sound from numerous sound sources, each localized to a well-defined region. Synthesizing these using loudspeaker-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 synthesize a diffuse source approximately, at varying distance [3], [4], which can be useful in a creative setting. Sound-field synthesis of an object with nonuniform directivity has been considered in the far field using spherical harmonic representation [3], [4]. With the development of high-order acoustic field construction, the simulation of near-field sources becomes feasible. 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 loudspeaker array directly [8], [9].

A localized source typically differs in two respects from a simple monopole source. The sound radiates from a region of nonzero width, and the directivity of radiation is not uniform. Near the object the sound field will be reactive, like a monopole's, but possibly will 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 among 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 represent a complex source accurately 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. Fig. 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 loudspeaker array or headphones. The wavefield approach lacks this intermediate stage, as well as suffering worse spatial aliasing artifacts [7]. Binaural rendering of high-order Ambisonics, over headphones, including the near field, has been considered [10].

This paper is organized as follows. First the source representation is 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, as a development of the first part, we consider the encoding of the reverberant field from a complex source, and how this can be modified for other listening and source positions, using variants of the transformation introduced in the first part.

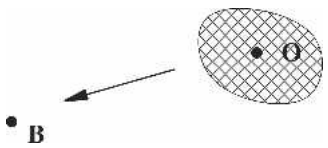

Fig. 1. Overall scheme. *O*—extended source object; *B*—listener.

*Manuscript received 2006 February 22; revised 2006 December 28 and 2007 May 17 and July 26.

1 SOURCE REPRESENTATION

We wish to use a representation that 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 modeled with several monopoles. This would be appropriate if it actually had this structure, or because a rough and fast model was 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.

1.1 Exterior Harmonic Expansion

Multipoles in their original form consist of infinitesimal arrangements of monopole sources. A multipole of sufficient order can represent 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 increases rapidly in complexity with the order. Closely related is the exterior expansion for the wave equation. This has basis functions in the frequency domain using the 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 wave number. 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 nonuniform directivity of a source, as will be discussed, 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

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

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

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

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

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

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

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

so that in the far field, where $h_m(kr)$ tends to $i^{m+1}e^{-ikr}/kr$, the field becomes

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

The $O_{mn}(k)$ coefficients then express the nonuniform directivity in this regime directly, 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 Eq. (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 calculated readily from measurements of the field on a sphere at any radius r outside the source region. Applying an integral over the sphere $\int Y_{mn}(\theta, \delta)\,d\Omega$ to Eq. (4) gives

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

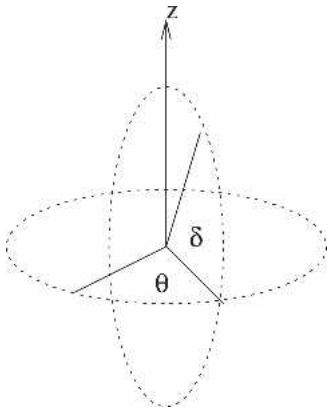

Fig. 2. Spherical coordinates used.

Note that $h_m(kr)$ 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 located on a sphere a fixed distance from the source. In the far field, 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 sound, the process can be simplified by repeated response measurements with a single microphone that is moved. Loudspeaker simulations might for instance be useful in high-end architectural simulations. When the $O_{mn}(k)$ responses are convolved with input signals for the loudspeakers, the expansion signals are generated. Spatial impulse responses may also be useful for simulating some resonant objects such as a violin body.

1.2 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 center. A monopole source at position $\boldsymbol{r}'$ has the following multipole expansion in $\boldsymbol{r}$ about the origin, valid for $r > r'$ [11],

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

where $j_m(kr')$ is the spherical Bessel function of the first kind. The arrangement is illustrated in Fig. 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, which just excludes the monopole. However, the field inside an extended source is not fully determined by its exterior field.

Figs. 4 and 5 show some cross-sectional plots of $\mathrm{Re}(p)/|p|$ 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 near field 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] agrees 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 of 1 m width can be approximated well with $m_{\max} = 14$. If we do not require good construction close to the limit $r = r'$, then the order can be reduced further. The far-field accuracy is not of so much interest, since at all orders the far field 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 far field, but not the near field.

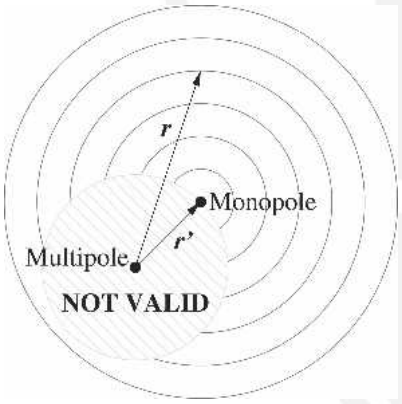

Fig. 3. Multiple representation of a monopole.

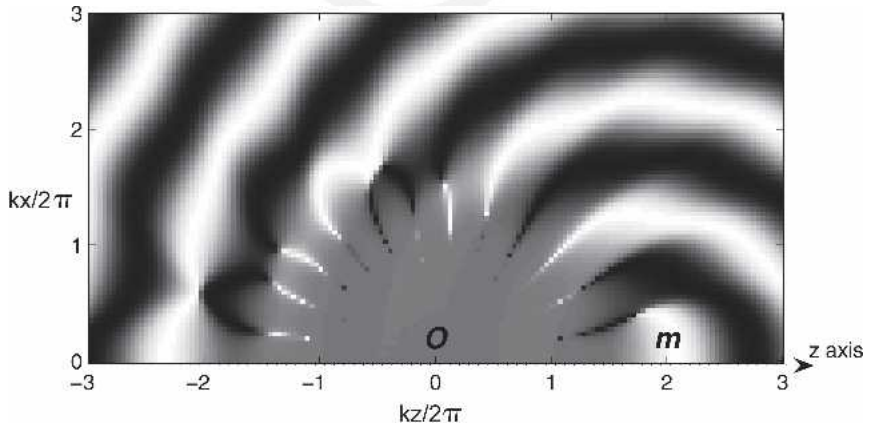

Fig. 4. Multipole approximation (center O) of displaced monopole (center m). $r' = 2\lambda$; $m_{\max} = 12$; cross section $\theta = 0$; x, z—Cartesian coordinates in units of length.

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 would not 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 near field of a general object, although this is not true across all possible sources.

1.3 Multiresolution 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 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. Fig. 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 multiresolution techniques common in computer graphics [14].

2 AMBISONIC ENCODING OF MULTIPOLES

2.1 Free-Field Expansion

High-order Ambisonics is founded on the interior expansion, which here we shall also call free-field expansion, to emphasize that it is used to describe a sourceless region around the listener, or more precisely the region at the listener if the listener were removed. Eq. (8) is the version of the expansion 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 ≈1% error for a regular free field, such as a plane wave, is $m_{\max} \approx kr$ [12], [15],

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

2.2 Monopole Encoding

For fields containing sources it is still possible to create a free-field 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 free-field expansion center. Eq. (7) can be recast as the free-field 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 $B_{mn}(k)$ can be read, 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. Fig. 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 free field, although it no longer matches the source field. Overall convergence behavior within the valid region is like any other free field, although close to the monopole, $\delta < \lambda$, the order required to achieve a given error is increased compared to a smooth free field, as we would expect [12]. The limit set to the region of free-field convergence by the source cannot be exceeded by increasing the free-field order.

Higher multipole sources have similar free-field expansions since they can be generated as composites of infini-

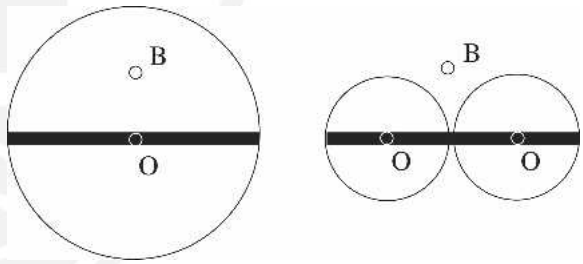

Fig. 6. When table sound comes from two sources (right), listener *B* can be closer to table while still outside any invalid region.

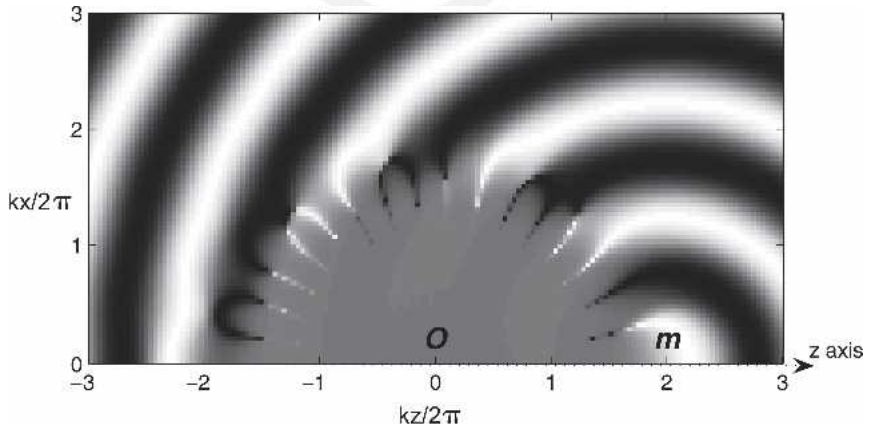

Fig. 5. Multipole approximation (center *O*) of displaced monopole (center *m*). $r' = 2\lambda$; $m_{\max} = 16$; cross section $\theta = 0$; *x, z*—Cartesian coordinates in units of length.

tesimal monopoles. The convergence condition is then that the free-field expansion be valid within a radius that does not include any of the sources.

2.3 Multipole to Free Field Coefficient Transformation

The main task in this section is to find $B_{mn}(k)$ in the presence of a multiple 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 is adopted, yielding eventually a manageable integral expression. To begin, Eqs. (8) and (4) are equated. The notation is modified according to Fig. 8,

$$\sum_m \mathrm{i}^m j_m(kr_B) \sum_n Y_{mn}(\theta_B, \delta_B) B_{mn}(k) = k \sum_m \mathrm{i}^{-m-1} h_m(kr_O) \sum_n Y_{mn}(\theta_O, \delta_O) O_{mn}(k). \quad (9)$$

To isolate $B_{mn}(k)$, the operator $\int Y_{m'n'}(\theta_B, \delta_B)\, \mathrm{d}\Omega_B$ 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

$$4\pi \mathrm{i}^{m'} j_{m'}(kr_B) B_{m'n'}(k) = k \sum_m \mathrm{i}^{-m-1} \sum_n O_{mn}(k) \times \int Y_{m'n'}(\theta_B, \delta_B) Y_{mn}(\theta_O, \delta_O) h_m(kr_O)\, \mathrm{d}\Omega_B. \quad (10)$$

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

$$B_{mn}(k) = \sum_{m',n'} M_{mnm'n'}(k, \boldsymbol{r}) O_{m'n'}(k) \quad (11)$$

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

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

The term $j_m(kr_B)$ in the denominator has zeros for kr_B at approximately regular intervals with period π. However, Eq. (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 that the symmetry axis is in the direction of $\boldsymbol{r}$. This defines a new matrix that only depends on r, and, as we shall see shortly, the symmetry gained reduces the complexity of the matrix. Fig. 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 $\boldsymbol{r}$,

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

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

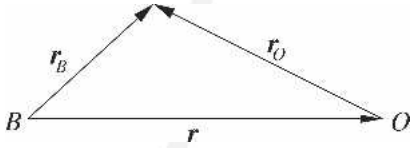

Fig. 8. Vector notation.

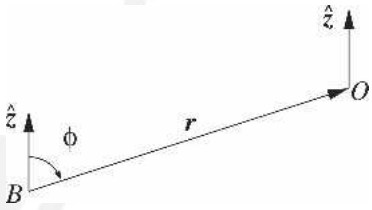

Fig. 9. Finding components relative to $\boldsymbol{r}$.

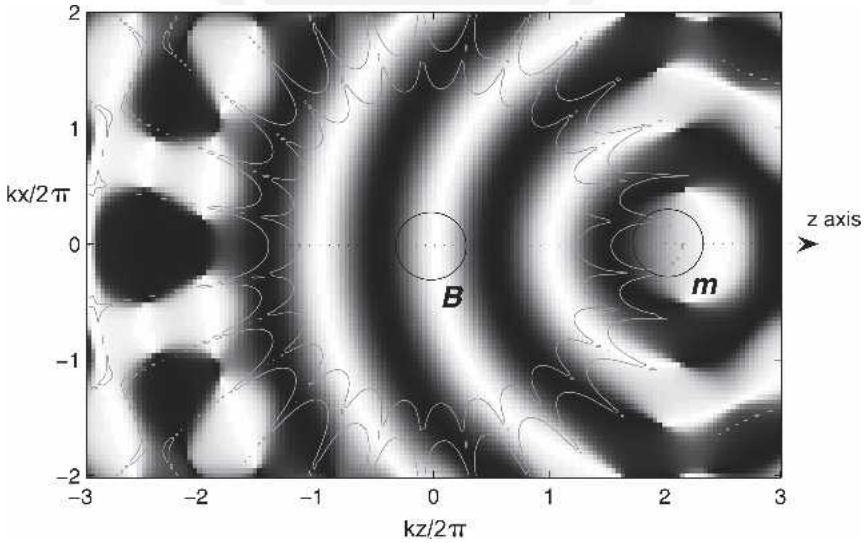

Fig. 7. Cross section of field plot for 13th-order free-field expansion (center at B) of monopole (center O). Error contours are shown at 1% and 10% levels. Cross section $\theta = 0$; x, z—Cartesian coordinates in units of length.

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

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

The coordinates in the integral are now relative to $\boldsymbol{r}$, although they have not 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

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

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

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

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

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

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 three-index coefficient matrix to be defined, which can be further simplified by factoring out a term $1/r$ and reexpressing 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,

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

where the new three-index matrix coefficient is

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

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 the B-format order required. The new filter coefficients are given in terms of one parameter, k. The actual filter acting in Eq. (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 four significant figures for coefficient magnitudes between 10 000 and 0.0001 over indices up to $m, m' = 30$ and up to $k = 10\,000$. The accuracy of each coefficient can be checked by using multiple values of α, and using the symmetry described later. Small k coefficients will also be discussed. 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 behavior in this region is quite predictable.

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

$$B_{mn}(k) = \sum_{n'} R_{mnn'}(\theta, \phi) \sum_{m'} \frac{1}{r} M_{mn'm'}(kr) \times \sum_{n''} R_{m'n'n''}(\theta, -\phi) O_{m'n''}(k). \tag{22}$$

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 $M_{mn'm'}(kr)$.

2.4 Validation and Properties

To provide an immediate confidence test that the derived formulas are correct, a random test fifth-order multipole was constructed, shown in Fig. 10, and compared with the 13th-order free-field expansion calculated using the matrix Eq. (22), shown in Fig. 11. The error contours in Fig. 11 at 10% and 1% levels are for deviations from the original multipole shown in Fig. 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 $F_m(kr) = \mathrm{i}^{-m}h_m(kr)/h_0(kr)$. 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 $S(k) = (\mathrm{e}^{\mathrm{i}kr}/r)O_{00}(k)$. In order to isolate the part matching $F_m(kr)$, we look at the adjusted value, $(1/r)M_{m00}(kr)/(\mathrm{e}^{-\mathrm{i}kr}/r)/Y_{m0}(0, \pi/2)$. With $r = 1$, this produces the plots shown in Fig. 12, matching previous results [5]. Further plots reveal how M extends to

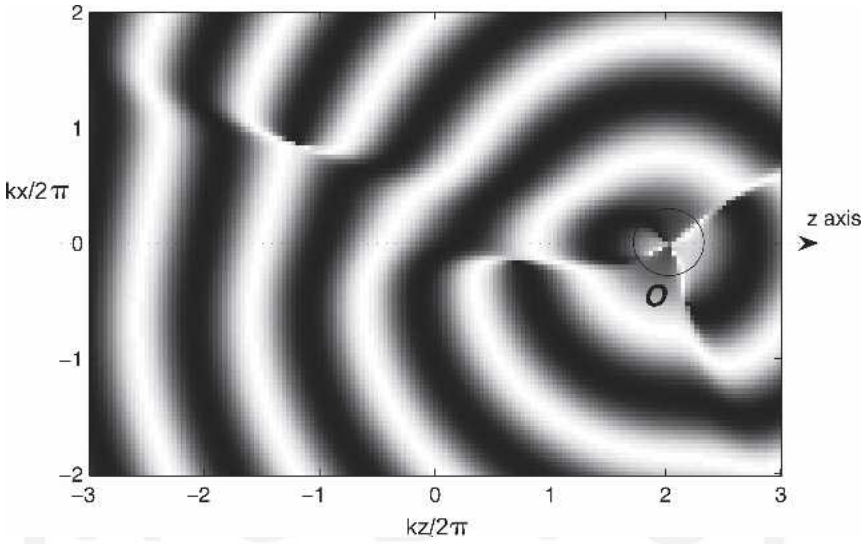

Fig. 10. Cross section of field plot for 5th-order multipole (center at *O*). Cross section $\theta = 0$; *x, z*—Cartesian coordinates in units of length.

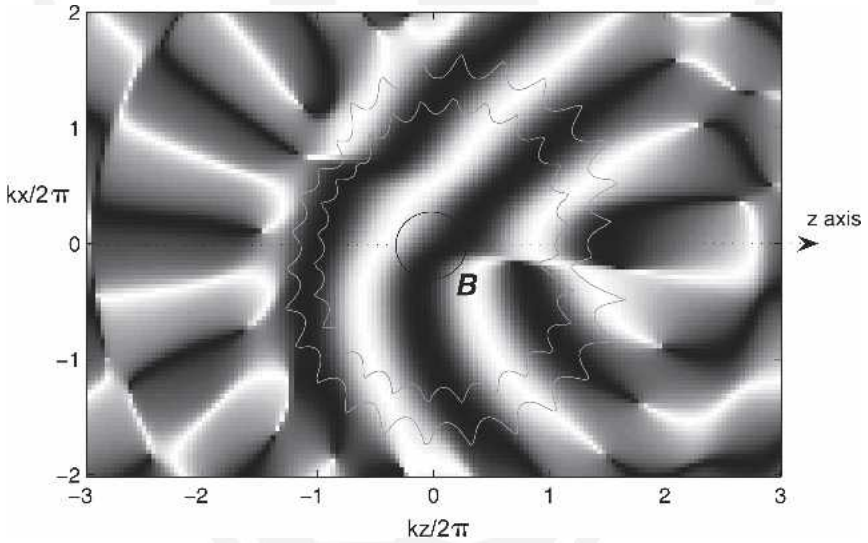

Fig. 11. Cross section of field plot for 13th-order free-field expansion (center at *B*) of multipole (center *O*). Error contours are shown at 1% and 10% levels. Cross section $\theta = 0$; *x, z*—Cartesian coordinates in units of length.

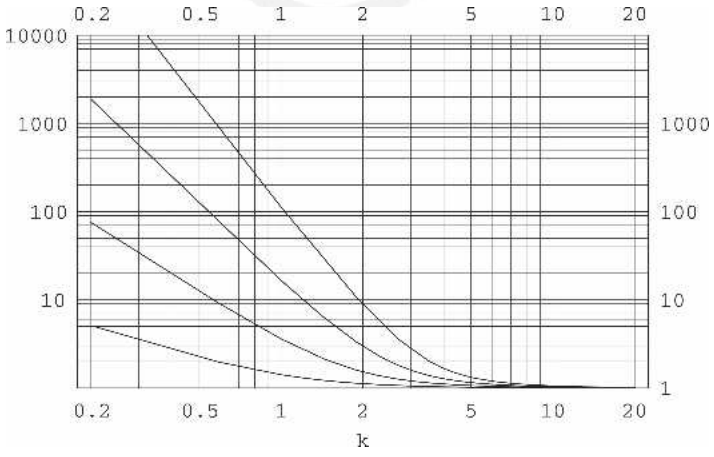

Fig. 12. Amplitude response for $M_{m00}(k)/e^{-ik}/Y_{m0}(0, \pi/2)$; $m = 1, 2, 3, 4$.

encode higher multipoles. Fig. 13 is an example showing the amplitude response for $M_{211}(k)/e^{-ik}$ along with $kM_{211}(k)/e^{-ik}$ to make the large k behavior clear. Fig. 14 shows the corresponding phase response. Another example, with $n = 2$, is shown in Figs. 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.

2.4.1 Symmetries

$M_{mnm'}(k)$ has symmetries that reduce the number of filters 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 involve distinct calculations.

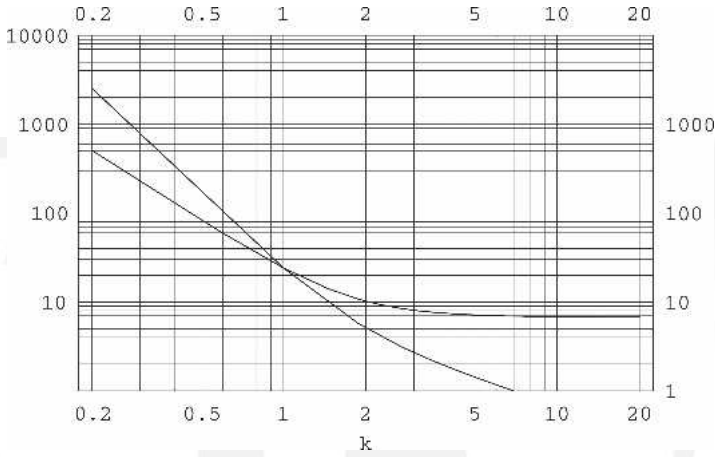

Fig. 13. Amplitude response for $M_{211}(k)/e^{-ik}$ and $kM_{211}(k)/e^{-ik}$.

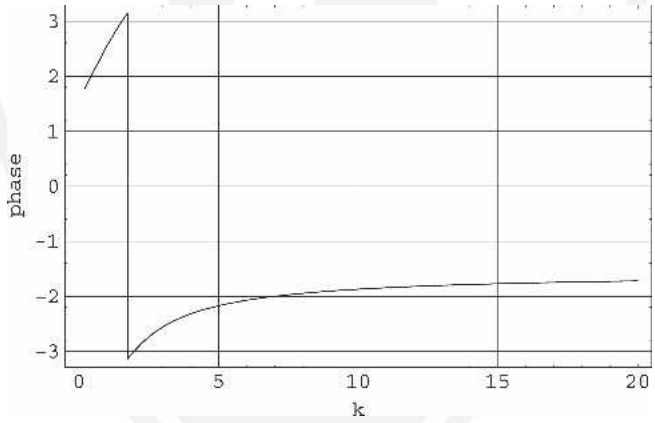

Fig. 14. Phase response for $M_{211(k)/e^{-ik}}$.

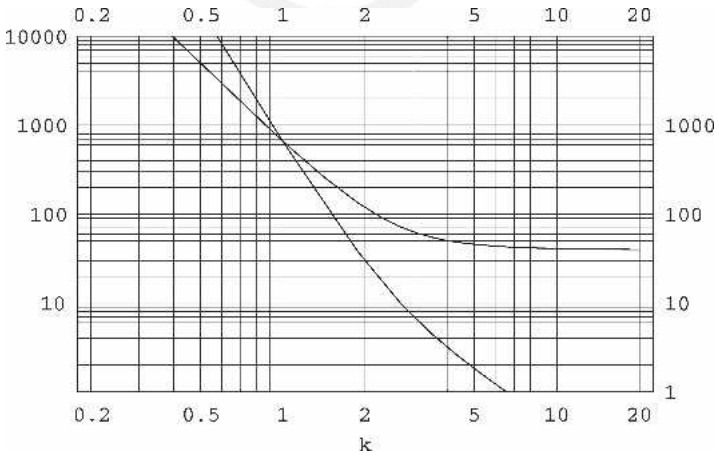

Fig. 15. Amplitude response for $M_{223}(k)/e^{-ik}$ and $k^2M_{223}(k)/e^{-ik}$.

2.4.2 Small k and Convergence

In [5] filters were commuted from loudspeaker-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 much higher order at low k, for the same order m of the B-format encoding. It has already been noted in Section 2.1 that the free-field 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 Figs. 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 is $\lesssim 10$ for all m, m' up to the maximum of 30 tested. From Section 1.2 an object that requires up to order m_O has typical radius r_O, where $r_O k \approx m_O$. This suggests that the free field 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 Fig. 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 affecting reconstruction accuracy significantly. 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 found with 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 centimeters from the source, and then this is only practical using a binaural rendering system. More investigation is needed on the convergence properties of free-field multipole reconstruction, although current results indicate that the scheme presented here is workable.

2.4.3 Digital Filters

The implementation of the digital filters is not investigated here in detail, but note that in general it will take a form similar 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 Eq. (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.

3 REVERBERATION ENCODING, FREE FIELD AND MULTIPOLE TRANSFORMATION

This paper has so far focused on the synthesis of the direct signal from complex objects in the near field. In this section we consider the encoding of reverberant sound from complex sources, and a modification of the work in the previous section, so that a free-field expansion can be transformed to a free-field expansion about another point. This is then extended further to include source translation. Although preliminary, these findings are included to show how the approach taken in the previous extension can be applied in other ways.

3.1 Encoding Source Directivity in Reverberation

Conventional room responses are mono to mono, mono to multichannel, or sometimes stereo to multichannel, where multichannel could include B format. This means that little consideration is given to how the reverberant

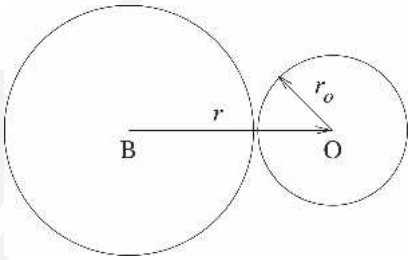

Fig. 17. Listener convergence region B limited by source object region O.

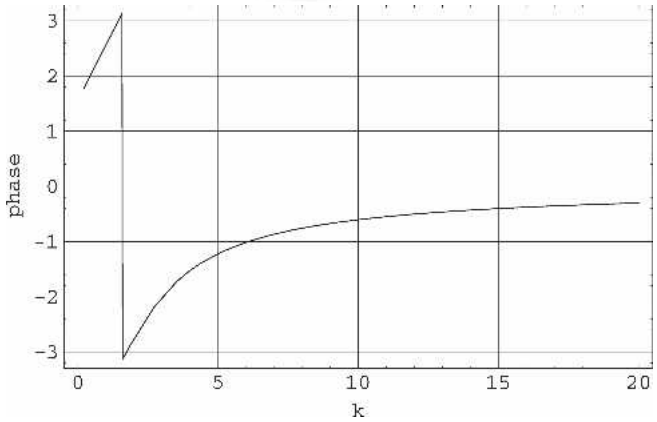

Fig. 16. Phase response for $M_{223}(k)/e^{-ik}$.

sound received from a source changes according to how the source is oriented, whereas real reverberation can vary considerably with source orientation. Fig. 18 illustrates the problem schematically.

In general the reverberant field has a linear relation to a directional source, which depends on the source position. So mathematically the general form is similar to that for the direct sound discussed already. For a source described by $O_{mn}(k)$ and a field expansion by $B_{mn}(k)$,

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

where $\boldsymbol{x}_B$ and $\boldsymbol{x}_O$ are the position vectors of the field expansion and source. This could be viewed as a set of transfer functions, one for each pair of source harmonic and field expansion harmonic. $M^{\text{rev}}_{mnm'n'}$ could be measured in a real space using directional sources and microphones, measuring the response on each microphone component for each source component. In practice the harmonics would be measured indirectly using loudspeaker arrays with directivities that collectively span the order required, for example, using a dodecahedral array such as that described in [17].

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 delivers 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.

$M^{\text{rev}}_{mnm'n'}$ could also be calculated from an image-source, ray-tracing, or other numerical simulation to much higher order than from measurement. The details of these procedures are beyond the scope of this paper.

3.2 Free-Field Transformation with Listening Position

Given an expansion about a point inside a sourceless region, such as B_{mn} from Eq. (23), we can ask what the expansion $B'_{mn}(k)$ would be about another point in the region. This would allow the listener to move around a reconstructed complex sound field, within a valid region that may include reverberant sound, and possibly direct sound from sources as well. In [18] the free-field expansion of a reverberant field from a monopole source is calculated by an image-source method to enable fast calculation of the pressure field. However, for three-dimensional listening we also need to transform the whole expansion to the current listening point.

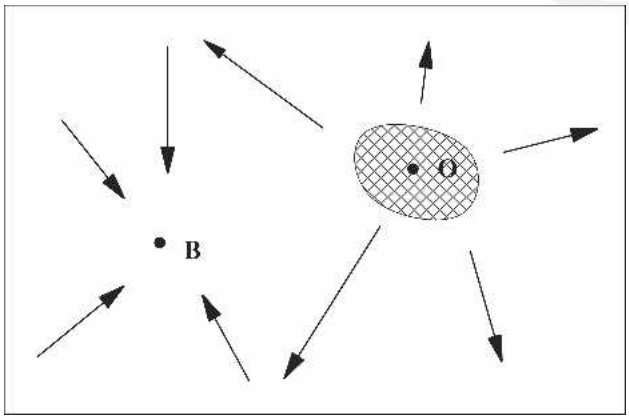

Fig. 18. Listening to reverberation at B from complex source at O.

Reworking the previous calculations leading to Eq. (21) produces the following result in terms of a new matrix $M^{BB}_{mnm'}$,

$$B'_{mn}(k) = \sum_{n'} R_{mnn'}(\theta, \phi) \sum_{m'} \frac{1}{r} M^{BB}_{mn'm'}(kr) \times \sum_{n''} R_{m'n'n''}(\theta, -\phi) B_{m'n''}(k) \tag{24}$$

where

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

with $k_{B'}$, $s_{B'}$ k_B, and s_B replacing k_B, s_B, k_O, and s_O, respectively.

The main difference between Eqs. (25) and (21) is that $h_{m'}(k_B)$ is replaced by $j_{m'}(k_B)$. For verification, a random 12th-order field B_{mn} has been synthesized, shown in Fig. 19, and transformed using Eq. (24) to an 18th-order field B'_{mn} centered at a distance 2λ from B_{mn}, as shown in Fig. 20. The error contours show clear agreement that 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 reproduces high-definition spatialized reverberation, at different listening positions, more efficiently than by direct simulation, as the computational complexity is limited by the order of $M^{\text{rev}}_{mnm'n'}$, $m_{\max} = rk$, determined by the spatial and frequency extent required, and not the complexity of the simulated environment, as noted in [18] for the monopole–source case. However, as already noted, higher orders can only be practically determined for a numerical room simulation.

3.3 Source Transformation with Source Position

The previous section showed how the expansion of a reverberant field can be transformed to a new listening position. Can a similar effect be applied to a source? Then in principle only a single reverberation matrix $M^{\text{rev}}_{mnm'n'}$ is required to encode between a wide range of source and listening positions. Fig. 21 shows the sequence of transformations to go from a generally placed source O' to a generally placed listener at B'.

This is conceptually striking, but again only of possible use for encoding high-quality simulated reverberation. The

required transformation from O' to O exists, and an example has already been described by Eq. (7), where a monopole is transformed to a multipole at another point. The general transformation follows using the same reasoning as before, except that the integration sphere radius must be larger than the separating distance to place it in the valid region of O. In other words $\alpha > 1$,

$$Q_{mn}(k) = \sum_{n'} R_{mnn'}(\theta, \phi) \sum_{m'} \frac{1}{r} M^{OO}_{mn'm'}(kr) \times \sum_{n''} R_{m'n'n''}(\theta, -\phi) O'_{m'n''}(k) \quad (26)$$

where

$$M^{OO}_{mnm'}(k) = \frac{\mathrm{i}^{m-m'}}{2h_m(k_O)} \int_{-1}^{+1} \hat{P}_{mn}(s_O) \hat{P}_{m'n}(s_{O'}) h_{m'}(k_{O'})\, \mathrm{d}s_{O'} \quad (27)$$

with k_O, s_O, $k_{O'}$, and $s_{O'}$ replacing k_B, s_B, k_O, and s_O, respectively, in Eq. (21). $h_m(k_O)$ is complex and has no zeros, so the value of α affecting k_O is not restricted further.

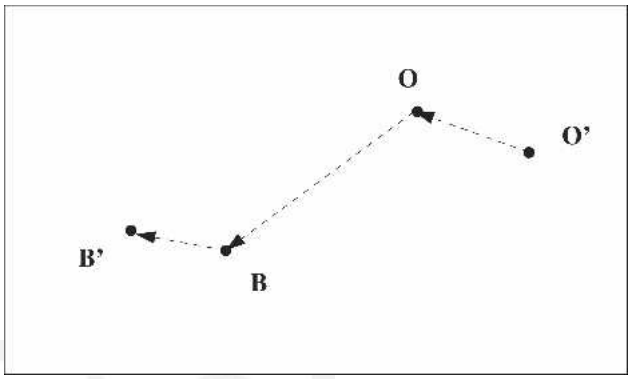

Fig. 21. Fully generalized reverberation from O' to B' using transfer between fixed source O and listener B.

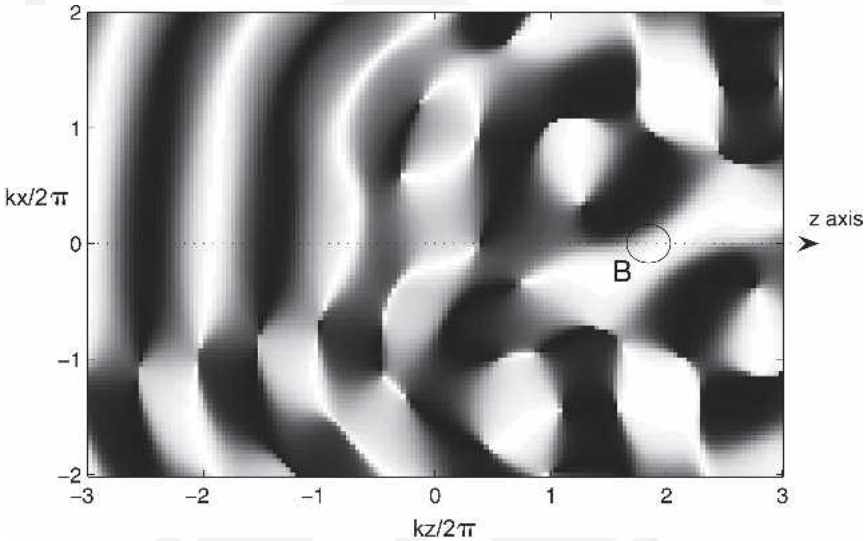

Fig. 19. Cross section of field plot for 12th-order free-field harmonic expansion (center at B). Cross section $\theta = 0$; x, z—Cartesian coordinates in units of length.

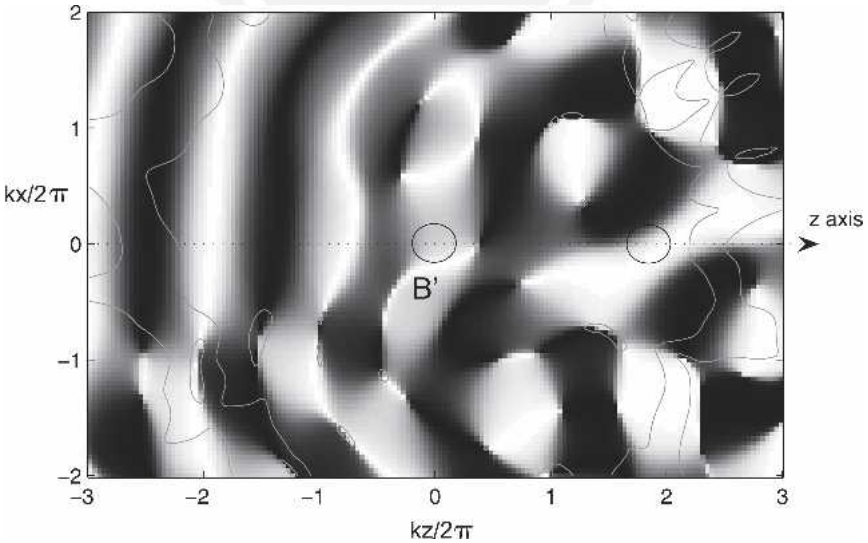

Fig. 20. Cross section of field plot for 18th-order free-field expansion (center at B') of free-field (center B) in Fig. 19. Error contours are shown at 1% and 10% levels. Cross section $\theta = 0$; x, z—Cartesian coordinates in units of length.

The invalid region centered around O that extends to O' implies the restriction that B cannot lie inside an image of this region, according to an image-source representation of the space. This will not be a problem unless O and B lie close to a wall, which can be arranged not to occur. There is no reason why O and B cannot be at the same point when encoding simulated reverberation.

In view of the listener and source transformations just described, the determination of $M^{\mathrm{rev}}_{mnm'n'}$ for two fixed points O, B can be seen as a postprocessing step for encoding a simulated acoustic in a more efficient form. To find the transfer between two new points the original simulation does not need to be run again.

4 CONCLUSION

A method has been presented for encoding a general acoustic source and transcoding it to a high-order Ambisonic signal, depending 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 [10] 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 a radius of 0.2 m, up to 1500 Hz, the required order $m = rk \approx 6$. In a loudspeaker-rendering environment the valid listening region is necessarily fixed to accommodate multiple listeners. This places constraints on how near-field 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 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 near-field rendering. A possible exception would be a small loudspeaker array designed for one person. Closely related results have also been presented for the encoding and transformation of reverberation, with the surprising conclusion that a single reverberation transfer between two points is sufficient to find the reverberation transfer between ranges of source and listener locations. In the future we hope to investigate realizations of these methods.

5 ACKNOWLEDGMENT

The authors would like to acknowledge the valuable comments made by the reviewer.

6 REFERENCES

[1] M. A. Gerzon, "Ambisonics in Multichannel Broadcasting and Video," *J. Audio Eng. Soc.,* vol. 33, pp. 859–871 (1985 Nov.).

[2] M. A. Gerzon, "General Metatheory of Auditory Localisation," presented at the 92nd Convention of the Audio Engineering Society, *J. Audio Eng. Soc.* (*Abstracts*), vol. 40, p. 447 (1992 May), preprint 3306.

[3] D. Menzies, "New Performance Instruments for Electroacoustic Music," Ph.D. thesis, University of York, York, UK (1999).

[4] D. Menzies, "W-Panning and O-Format, Tools for Object Spatialisation," in *Proc. AES 22nd Int. Conf.* (2002).

[5] J. Daniel, "Spatial Sound Encoding Including Near Field Effect," in *Proc. AES 23nd Int. Conf.* (2003).

[6] A. J. Berkhout, D. D. Vries, and P. Vogel, "Acoustic Control by Wave Field Synthesis," *J. Acoust. Soc. Am.,* vol. 93, pp. 2764–2778 (1993).

[7] J. Daniel, R. Nicol, and S. Moreau, "Further Investigations of High-Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging," presented at the 114th Convention of the Audio Engineering Society, *J. Audio Eng. Soc.* (*Abstracts*), vol. 5, p. 425 (2003 May), convention paper 5788.

[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. Int. Congr. on Acoustics* (2004).

[10] D. Menzies and M. Al-Akaidi, "Nearfield Binaural Synthesis," *J. Acoust. Soc. Am.,* vol. 121, pp. 1559–1563 (2007 Mar.).

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

[12] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions* (Elsevier Science, Amsterdam, Netherlands, 2005).

[13] J. Daniel, "Représentation de Champs Acoustiques—Application à la Transmission et à la Reproduction de Scènes Sonores Complexes dans un Contexte Multimédia," Ph.D. 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, New York, 2005).

[15] D. B. Ward and T. D. Abhayapala, "Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers," *IEEE Trans. Speech Audio Process.,* vol. 9, p. 697 (2001 Sept.).

[16] J. Daniel and S. Moreau, "Further Study of Sound Field Coding with Higher Order Ambisonics," presented at the 116th Convention of the Audio Engineering Society, *J. Audio Eng. Soc.* (*Abstracts*), vol. 52, p. 784 (2004 July/Aug.), convention paper 6017.

[17] D. Trueman and P. Cook, "Bossa: The Deconstructed Violin Reconstructed," in *Proc. Int. Computer Music Conf.* (Beijing, China, 1999).

[18] R. Duraiswami, D. N. Zotkin, and N. A. Gumerov, "Fast Evaluation of the Room Transfer Function Using Multipole Expansion," *IEEE Trans. Audio, Speech, Language Process.,* vol. 15, pp. 565–576 (2007 Feb.).

THE AUTHORS

D. Menzies

M. Al-Akaidi

Dylan Menzies received a B.A. degree in mathematics from Trinity College, Cambridge, UK, in 1994, and an M.Sc. degree in music technology and a Ph.D. degree in electronics from the University of York, York, UK, in 1995 and 1999, respectively.

Subsequently he worked for a startup company in Oxford, Mathengine, developing physical audio for simulations, and then for Sony Professional Audio working on the SACD. Following a period as a researcher in cosmology at the University of Notre Dame, he returned to Leicester, UK, as a senior lecturer at De Montfort University in the faculty of Computer Science and Engineering.

Dr. Menzies' audio research interests are broad but focus on interactive synthesis and spatialization. He is a frequent performer in acoustic and electronic music ensembles, often playing his saxophones.

●

Marwan Al-Akaidi joined the Department of Electronic Engineering at De Montfort University, Leicester, UK, in 1991. In 2000 he became professor of communication and signal processing. His main research interest is in the field of digital signal processing and digital communications. This includes speech coding, processing, recognition, and wireless and mobile communication.

Mr. Al-Akaidi is the editor and chairman of more than 10 national and international conferences, including 3G Mobile Communication Technologies and Joint Modular Languages Conferences. He is the guest editor of *Simulation*'s special issue on Telecommunication in Digital Signal/Image Processing. In the last five years he has been involved in setting SCS (Society for Computer Simulation) chapters in various countries and has chaired modeling and simulation conferences sponsored by SCS. He achieved Chartered Engineering status in 1990 and fellowship of the Institute of Analyst and Programming in 1991. He is a senior member of the Institute of Electrical and Electronics Engineers and a fellow of the Institution of Electrical Engineering. In 1999 he was appointed chairman for the IEEE UKRI Signal Processing Society and in 2000 became IEEE UKRI conferences chair. He is a member of the board of IEEE Industrial Relations and has won the award of the IEEE UKRI in recognition of outstanding leadership as a chapter chair for the years 2001 and 2002.