AMBISONICS SYMPOSIUM 2009
June 25-27, Graz

UNIFIED DESCRIPTION OF AMBISONICS USING REAL AND COMPLEX SPHERICAL HARMONICS

Mark Poletti[1]

[1] Industrial Research Ltd (m.poletti@irl.cri.nz)

Abstract: *The theoretical description of higher order Ambisonics uses both real and complex forms of spherical harmonic solutions to the wave equation in spherical coordinates. This paper develops a description of sound fields using consistent definitions of either form, and shows how the expansion coefficients of each are related. Both descriptions produce complex field coefficients. We show that these complex coefficients are represented in quadrature modulated form in a set of real Ambisonics signals. We also show that recording the complex form of Ambisonics signals requires the use of the Hilbert transform to obtain the analytic form of each microphone signal. We develop mode matching and simple source expressions for the decoder matrices for spherical loudspeaker arrays for both the complex and real spherical harmonic forms and show that these are equivalent. The simple source decoder for plane waves has a simple form that can be computed directly from the spherical harmonics.*

Key words: Surround sound, Ambisonics, Spherical Harmonics

1 INTRODUCTION

Ambisonics is an approach to the recording and reproduction of three-dimensional sound fields which is founded on the description of sound fields in spherical coordinates. The sound field may be expressed as a sum of orthogonal terms with polar responses which are real spherical harmonic functions [1-6]. A recording microphone must implement these polar responses to produce the Ambisonics encoded signals that describe the 3D sound field. Ambisonics also describes two-dimensional sound field recording and reproduction using a subset of the spherical harmonics (sectoral) or cylindrical harmonics, but this case is not considered here.

Spherical harmonics can be defined either in terms of real [6] or complex [7-8] functions. Much of the recent development of higher order Ambisonics has been based on the use of complex spherical harmonics [9-15]. Since the signals in Ambisonics recordings are real, this raises the question of how compatible the complex-based theory is with the real-signal Ambisonics format. Furthermore, the sound field coefficients in any spherical harmonic expansion are typically complex, which raises the question of how the complex coefficients can be contained in real Ambisonics signals.

This paper provides a unified description of surround sound recording and reproduction using both complex and real forms of the spherical harmonics. We use definitions of the real and complex spherical harmonics that are consistent with each other, and show how the real and complex forms of sound field expansions are related. We also derive the decoding matrix for a given loudspeaker reproduction array for both cases, and show how they are related.

We commence with a discussion of real and complex solutions to the wave equation. We note first that the wavenumber $k = \omega / c$, where ω is radian frequency and c is the speed of sound, and we will use k or ω in accordance with the conventions in [7].

2 REAL AND COMPLEX SOLUTIONS TO THE WAVE EQUATION

2.1. One dimensional wave equation

The general solution to the one-dimensional wave equation is [16]

$$p(x,t) = f(x-ct) + g(x+ct) \tag{1}$$

The first term describes a solution that propagates in the positive *x* direction and the second term propagates in the negative *x* direction. At a single frequency, real solutions have the general form

$$\begin{aligned} &p(x,\omega,t) \\ &= a_1 \cos(k(x-ct)+\phi_1) + a_2 \cos(k(x+ct)+\phi_2) \\ &= a_1 \cos(\omega t)\cos(kx+\phi_1) + a_1 \sin(\omega t)\sin(kx+\phi_1) \\ &\quad + a_2 \cos(\omega t)\cos(kx+\phi_2) - a_2 \sin(\omega t)\sin(kx+\phi_2) \end{aligned} \tag{2}$$

We note that the positive and negative traveling waves each consist of two terms in phase quadrature. The cosine term is the in-phase term and the sine term is the quadrature term. If $a_1 = a_2$ and $\phi_1 = \phi_2$ then the quadrature terms cancel, and the solution becomes a standing wave field, with no net propagation.

The complete solution to the 1D wave equation must include all frequencies present in the sound field. Hence from equation (2), with a_1 and a_2 functions of frequency;

$$p(x,t) = \int_0^\infty a_1(\omega)\cos\left(k(x-ct)+\phi_1\right)d\omega + \int_0^\infty a_2(\omega)\cos\left(k(x+ct)+\phi_2\right)d\omega \tag{3}$$

This may also be written in terms of the complex field

$$z(x,\omega,t) = e^{i\omega t}\left[a_1(\omega)e^{-i(kx+\phi_1)} + a_2(\omega)e^{i(kx+\phi_2)}\right] \tag{4}$$

where $p(x,\omega,t) = \mathrm{Re}\{z(x,\omega,t)\}$, as

$$p(x,t) = \mathrm{Re}\left\{\int_0^\infty \left[a_1(\omega)e^{-i(kx+\phi_1)} + a_2(\omega)e^{i(kx+\phi_2)}\right]e^{i\omega t}d\omega\right\} \tag{5}$$

where the spectrum

$$P(x,\omega) = a_1(\omega)e^{-i(kx+\phi_1)} + a_2(\omega)e^{i(kx+\phi_2)} \tag{6}$$

is defined for positive ω.

The complex field

$$z(x,t) = \int_0^\infty \left[a_1(\omega)e^{-i(kx+\phi_1)} + a_2(\omega)e^{i(kx+\phi_2)}\right]e^{i\omega t}d\omega \tag{7}$$

is termed the analytic signal [17]. It may be obtained from $p(x,t)$ by removing the negative frequency terms. For example, writing $\cos(\omega t)$ and $\sin(\omega t)$ in equation (2) in terms of complex exponentials, and removing the negative frequency terms, yields equation (4). More generally, $z(x,t)$ is obtained from $p(x,t)$ at any point $x=x_0$ by filtering $p(x_0,t)$ with a complex filter which eliminates the negative frequencies. The complex filter has ideal impulse response [18]

$$h_a(t) = \delta(t) + \frac{i}{\pi t} \tag{8}$$

The Fourier transform of this complex impulse response is the transfer function

$$H_a(\omega) = \int_{-\infty}^{\infty}\left[\delta(t) + \frac{i}{\pi t}\right]e^{-i\omega t}dt = 2U(\omega) \tag{9}$$

which is the unit step function scaled by 2. The analytic filter thus eliminates negative frequencies and scales the positive frequencies by 2 (to maintain the same total energy).

One could argue that the analytic signal can be easily generated by taking the FFT of the signal and zeroing the negative frequencies. However, this approach is not as precise as designing a Hilbert transformer using optimal filter design methods, or designing a complete complex analytic filter [19].

2.2. Three dimensional wave equation

In three dimensions solutions to the wave equation which are spherically symmetric can be expressed in spherical coordinates as [16]

$$p(r,t) = \frac{1}{r}f(r-ct) + \frac{1}{r}g(r+ct) \tag{10}$$

where the first term describes outgoing waves and the second incoming waves. Real solutions at a single frequency have the form

$$p(r,\omega,t) = \frac{a_1}{r}\cos\left(k(r-ct)+\phi_1\right) + \frac{a_2}{r}\cos\left(k(r+ct)+\phi_2\right) \tag{11}$$

Again, this can be written in quadrature form, and the quadrature component is necessary to describe propagating waves, and only disappears for a standing wave field $(a_1,\phi_1) = (a_2,\phi_2)$.

As discussed above, the complete solution including all frequencies can be written in terms of the analytic signal as

$$p(r,t) = \mathrm{Re}\{z(r,t)\} = \mathrm{Re}\left\{\int_0^\infty \left[\frac{a_1(\omega)}{r}e^{-i(kr-\phi_1)} + \frac{a_2(\omega)}{r}e^{i(kr+\phi_2)}\right]e^{i\omega t}d\omega\right\} \tag{12}$$

3 REAL AND COMPLEX SPHERICAL HARMONIC SOLUTIONS TO THE WAVE EQUATION

3.1. Complex spherical harmonic solutions to the wave equation

The general complex solution to the wave equation at frequency ω may be written in spherical coordinates (r,θ,ϕ) as [7]

$$z(r,\theta,\phi,\omega,t) = e^{i\omega t}\left[\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left(U_n^m h_n^{(1)}(kr) + V_n^m h_n^{(2)}(kr)\right)Y_n^m(\theta,\phi)\right] \tag{13}$$

where $h_n^{(1)}(kr) = j_n(kr) + y_n(kr)$ is the spherical Hankel function of the first kind, $h_n^{(2)}(kr) = j_n(kr) - iy_n(kr)$ is that of the second kind, and $j_n(kr)$ and $y_n(kr)$ are the spherical Bessel functions of the first and second kind. Since $h_n^{(1)}(kr) \propto e^{ikr}$ then, with the positive frequency convention, the U_n^m coefficients describe incoming waves and the V_n^m terms describe outgoing waves.

The complex, normalised spherical harmonics $Y_n^m(\theta,\phi)$ follow the definition of [8,9]

$$Y_n^m(\theta,\phi)=\sqrt{\frac{(2n+1)}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}}P_n^{|m|}(\cos\theta)e^{im\phi} \quad (14)$$

where ! denotes factorial. This form has the property $Y_n^{-m}(\theta,\phi)=Y_n^m(\theta,\phi)^*$ and so avoids the $(-1)^m$ term that appears in [7].

The complex spherical harmonics have the orthogonality property

$$\int Y_n^m(\theta,\phi)Y_u^v(\theta,\phi)^*\,d\Omega=\delta_{nu}^{mv} \quad (15)$$

where $\delta_{nu}^{mv}=1$ for $n=u$ and $m=v$ and is zero otherwise.

Standard spherical coordinates use the angle θ from the z axis. In Ambisonics the elevation angle β above the (x,y) plane is commonly used. Since $\cos(\theta)=\sin(\beta)$ equation (14) is easily modified to use this convention.

Since surround sound involves the recording and reproduction of sound fields that occur around a given origin, we are interested in the interior solution which contains no sound sources, and which contains radial functions that are finite at the origin [7]

$$\begin{aligned}z(r,\theta,\phi,\omega,t)&=e^{i\omega t}\sum_{n=0}^{\infty}j_n(kr)\sum_{m=-n}^{n}C_n^m(k)Y_n^m(\theta,\phi)\\&=e^{i\omega t}s(r,\theta,\phi,\omega)\end{aligned} \quad (16)$$

where $s(r,\theta,\phi,\omega)$ is the complex spatial field at frequency ω. For a truncated expansion up to and including order $n=N$, there are $(N+1)^2$ complex expansion coefficients $C_n^m(k)$.

The real part of equation (16) is the sound pressure

$$\begin{aligned}&p(r,\theta,\phi,\omega,t)=\mathrm{Re}\{z(r,\theta,\phi,\omega,t)\}\\&=\cos(\omega t)s_R(r,\theta,\phi,\omega)-\sin(\omega t)s_I(r,\theta,\phi,\omega)\end{aligned} \quad (17)$$

where s_R and s_I are the real and imaginary parts of the complex spatial field.

The real propagating field is thus, as above, the sum of two static spatial real fields oscillating in quadrature.

The sound pressure over all frequencies is – as discussed in the previous section – the real part of the inverse Fourier transform of equation (16), or equivalently the integral of equation (17), over positive frequencies.

3.2. Real spherical harmonic solutions to the wave equation

The complex spherical harmonics may be separated into real and imaginary parts. The real part has *cos(mφ)* factors and is hence even in azimuth, while the imaginary part is odd. In [6], the real and imaginary terms are defined as even and odd functions. Here, we will define the normalised real spherical harmonics as

$$\begin{aligned}E_n^m(\theta,\phi)&=\sqrt{\frac{\varepsilon_m(2n+1)}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}}P_n^{|m|}(\cos\theta)\cos(m\phi)\\&=\frac{\sqrt{\varepsilon_m}}{2}\left[Y_n^m(\theta,\phi)+Y_n^{-m}(\theta,\phi)\right]\end{aligned} \quad (18)$$

and

$$\begin{aligned}O_n^m(\theta,\phi)&=\sqrt{\frac{\varepsilon_m(2n+1)}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}}P_n^{|m|}(\cos\theta)\sin(m\phi)\\&=\frac{\sqrt{\varepsilon_m}}{2i}\left[Y_n^m(\theta,\phi)-Y_n^{-m}(\theta,\phi)\right]\end{aligned} \quad (19)$$

where $\varepsilon_m=2, |m|>0$ and $\varepsilon_0=1$. These have the orthogonality properties for integration over the sphere

$$\int E_n^m(\theta,\phi)E_u^v(\theta,\phi)\,d\Omega=\delta_{nu}^{mv} \quad (20)$$

$$\int E_n^m(\theta,\phi)O_u^v(\theta,\phi)\,d\Omega=0 \quad (21)$$

for $n,u\in[0,\infty]$ and $m,v\in[0,n]$, and

$$\int O_n^m(\theta,\phi)O_u^v(\theta,\phi)\,d\Omega=\begin{cases}\delta_{nu}^{mv}, m,v>0\\0, m=v=0\end{cases} \quad (22)$$

While negative m values will not be required in wave field expansions (see below), they are well defined and $E_n^{-m}(\theta,\phi)=E_n^m(\theta,\phi)$ and $O_n^{-m}(\theta,\phi)=-O_n^m(\theta,\phi)$, so that

$$\int E_n^m(\theta,\phi)E_n^{-m}(\theta,\phi)\,d\Omega=1 \quad (23)$$

and

$$\int O_n^m(\theta,\phi)O_n^{-m}(\theta,\phi)\,d\Omega=-1 \quad (24)$$

The complex spherical harmonic is related to the real form as

$$Y_n^m(\theta,\phi)=\frac{1}{\sqrt{\varepsilon_m}}\left[E_n^m(\theta,\phi)+iO_n^m(\theta,\phi)\right] \quad (25)$$

For interior solutions of the wave equation the complex spherical harmonic expansion of the complex spatial field is given by equation (16). Writing $Y_n^m(\theta,\phi)$ in terms of the real harmonics

$$s(r,\theta,\phi,\omega)=\sum_{n=0}^{\infty} j_n(kr)\times \sum_{m=-n}^{n} C_n^m(k)\frac{1}{\sqrt{\varepsilon_m}}\left[E_n^m(\theta,\phi)+iO_n^m(\theta,\phi)\right] \tag{26}$$

Real field in terms of complex coefficients

Writing C_n^m in equation (16) in terms of its real and imaginary parts yields the real sound pressure

$$\begin{aligned} p(r,\theta,\phi,\omega,t) &= \sum_{n=0}^{\infty} j_n(kr)\sum_{m=-n}^{n}\left[\frac{C_{nR}^m\cos(\omega t)-C_{nI}^m\sin(\omega t)}{\sqrt{\varepsilon_m}}\right]E_n^m \\ &-\sum_{n=0}^{\infty} j_n(kr)\sum_{m=-n}^{n}\left[\frac{C_{nI}^m\cos(\omega t)+C_{nR}^m\sin(\omega t)}{\sqrt{\varepsilon_m}}\right]O_n^m \end{aligned} \tag{27}$$

where the (θ,ϕ) dependency of the harmonics is implicit. Hence, the real field contains the real and imaginary parts of the complex coefficients C_n^m in quadrature modulated form.

Field expansion in real spherical harmonics, and relationship to complex coefficients

Consider equation (26). Since E_n^m is even in *m* and O_n^m is odd, the summation in *m* can be written

$$\begin{aligned} &\sum_{m=0}^{n}\frac{\varepsilon_m}{2\sqrt{\varepsilon_m}}C_n^m(k)\left(E_n^m(\theta,\phi)+iO_n^m(\theta,\phi)\right) \\ &+\sum_{m=0}^{n}\frac{\varepsilon_m}{2\sqrt{\varepsilon_m}}C_n^{-m}(k)\left(E_n^m(\theta,\phi)-iO_n^m(\theta,\phi)\right) \\ &=\sum_{m=0}^{n}\sqrt{\varepsilon_m}E_n^m(\theta,\phi)\left(\frac{C_n^m(k)+C_n^{-m}(k)}{2}\right) \\ &+i\sum_{m=0}^{n}\sqrt{\varepsilon_m}O_n^m(\theta,\phi)\left(\frac{C_n^m(k)-C_n^{-m}(k)}{2}\right) \end{aligned} \tag{28}$$

and hence the expansion of the complex spatial sound field in real harmonics is

$$s(r,\theta,\phi,\omega)=\sum_{n=0}^{\infty} j_n(kr) \times\sum_{m=0}^{n}\left[A_n^m(k)E_n^m(\theta,\phi)+B_n^m(k)O_n^m(\theta,\phi)\right] \tag{29}$$

where

$$A_n^m=\sqrt{\varepsilon_m}\left(\frac{C_n^m+C_n^{-m}}{2}\right) \tag{30}$$

and

$$B_n^m=i\sqrt{\varepsilon_m}\left(\frac{C_n^m-C_n^{-m}}{2}\right) \tag{31}$$

or in matrix form

$$\begin{bmatrix}A_n^m\\ B_n^m\end{bmatrix}=\sqrt{\varepsilon_m}\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{-1}{2i} & \frac{1}{2i}\end{bmatrix}\begin{bmatrix}C_n^m\\ C_n^{-m}\end{bmatrix}=\mathbf{H}\begin{bmatrix}C_n^m\\ C_n^{-m}\end{bmatrix} \tag{32}$$

The matrix **H** is unitary for *m*>0 and $\mathbf{H}^H\mathbf{H}=0.5\mathbf{I}$ for *m*=0. Hence

$$\begin{bmatrix}C_n^m\\ C_n^{-m}\end{bmatrix}=\frac{1}{\sqrt{\varepsilon_m}}\begin{bmatrix}1 & -i\\ 1 & i\end{bmatrix}\begin{bmatrix}A_n^m\\ B_n^m\end{bmatrix} \tag{33}$$

or

$$C_n^m=\frac{1}{\sqrt{\varepsilon_m}}\left(A_n^m-iB_n^m\right) \tag{34}$$

and

$$C_n^{-m}=\frac{1}{\sqrt{\varepsilon_m}}\left(A_n^m+iB_n^m\right) \tag{35}$$

The real sound field is obtained from equation (29)

$$\begin{aligned} p(r,\theta,\phi,\omega,t) &=\cos(\omega t)\sum_{n=0}^{\infty} j_n(kr)\sum_{m=0}^{n}\left[A_{nR}^m(k)E_n^m+B_{nR}^m(k)O_n^m\right] \\ &-\sin(\omega t)\sum_{n=0}^{\infty} j_n(kr)\sum_{m=0}^{n}\left[A_{nI}^m(k)E_n^m+B_{nI}^m(k)O_n^m\right] \end{aligned} \tag{36}$$

where A_{nR}^m and A_{nI}^m are the real and imaginary parts of A_n^m. This expresses the real sound field in terms of two static fields which oscillate in phase quadrature. It can be rearranged as

$$
\begin{aligned}
& p(r,\theta,\phi,\omega,t) \\
&= \sum_{n=0}^{\infty} j_n(kr) \sum_{m=0}^{n} \left[A_{nR}^m(k)\cos(\omega t) - A_{nI}^m(k)\sin(\omega t) \right] E_n^m \\
&+ \sum_{n=0}^{\infty} j_n(kr) \sum_{m=0}^{n} \left[B_{nR}^m(k)\cos(\omega t) - B_{nI}^m(k)\sin(\omega t) \right] O_n^m
\end{aligned} \tag{37}
$$

which expresses the real field at each frequency in terms of the quadrature modulated form of the underlying complex coefficients. This expression is the fundamental basis for Ambisonics recording and reproduction. For a truncated expansion up to and including order N, there are $N(N+1)/2+(N+1)$ even harmonic terms (which includes the $(N+1)$ zonal harmonic ($m=0$) terms) and a further $(N+1)/2$ odd harmonic terms, making a total of $(N+1)^2$ real Ambisonics signals, the same number as the complex case.

3.3. Examples

A single-frequency complex plane wave has the two equivalent expansions [6,7]

$$
\begin{aligned}
e^{i\vec{k}\bullet\vec{r}} &= 4\pi \sum_{n=0}^{\infty} j_n(kr) \times \\
& \sum_{m=0}^{n} i^n \left[E_n^m(\theta_i,\phi_i) E_n^m(\theta,\phi) + O_n^m(\theta_i,\phi_i) O_n^m(\theta,\phi) \right] \\
&= 4\pi \sum_{n=0}^{\infty} j_n(kr) \sum_{m=0}^{n} i^n Y_n^m(\theta_i,\phi_i)^* Y_n^m(\theta,\phi)
\end{aligned} \tag{38}
$$

and so the expansion coefficients are

$$A_n^m = 4\pi i^n E_n^m(\theta_i,\phi_i), n \in [0,\infty], m \in [0,n] \tag{39}$$

$$B_n^m = 4\pi i^n O_n^m(\theta_i,\phi_i), n \in [0,\infty], m \in [1,n] \tag{40}$$

and the complex expansion coefficients are

$$C_n^m = 4\pi i^n Y_n^m(\theta_i,\phi_i)^*, n \in [0,\infty], m \in [-n,n] \tag{41}$$

Note that these are independent of frequency. However, a general plane-wave field will consist of multiple plane waves with frequency-dependent amplitudes, and so the expansion coefficients of a plane-wave field will be frequency-dependent.

As a second example, a single-frequency point source at (r_s,θ_s,ϕ_s) has the two equivalent expansions [6,7]

$$
\begin{aligned}
& \frac{e^{-ik|\vec{r}-\vec{r}_s|}}{4\pi|\vec{r}-\vec{r}_s|} \\
&= -ik \sum_{n=0}^{\infty} j_n(kr) h_n^{(2)}(kr_s) \\
&\times \sum_{m=0}^{n} \left[E_n^m(\theta,\phi) E_n^m(\theta_s,\phi_s) + O_n^m(\theta,\phi) O_n^m(\theta_s,\phi_s) \right] \\
&= -ik \sum_{n=0}^{\infty} j_n(kr) h_n^{(2)}(kr_s) \sum_{m=-n}^{n} Y_n^m(\theta,\phi) Y_n^m(\theta_s,\phi_s)^*
\end{aligned} \tag{42}
$$

and so the real harmonic expansion coefficients are

$$A_n^m = -ikh_n^{(2)}(kr_s) E_n^m(\theta_s,\phi_s), n \in [0,\infty], m \in [0,n] \tag{43}$$

$$B_n^m = -ikh_n^{(2)}(kr_s) O_n^m(\theta_s,\phi_s), n \in [0,\infty], m \in [0,n] \tag{44}$$

and the complex expansion coefficients are

$$C_n^m = -ikh_n^{(2)}(kr_s) Y_n^m(\theta_s,\phi_s)^*, n \in [0,\infty], m \in [-n,n] \tag{45}$$

Finally, the real part of the point source (producing outgoing wavefronts for a positive frequency) is

$$
\begin{aligned}
& \operatorname{Re}\left\{ e^{i\omega t} \frac{e^{-ik|\vec{r}-\vec{r}_s|}}{4\pi|\vec{r}-\vec{r}_s|} \right\} \\
&= k \sum_{n=0}^{\infty} j_n(kr) [-y_n(kr_s)\cos\omega t + j_n(kr_s)\sin(\omega t)] \times \\
& \sum_{m=0}^{n} \left[E_n^m(\theta,\phi) E_n^m(\theta_s,\phi_s) + O_n^m(\theta,\phi) O_n^m(\theta_s,\phi_s) \right]
\end{aligned} \tag{46}
$$

where

$$
\begin{aligned}
& \operatorname{Re}\left\{ -ih_n^{(2)}(kr_s) e^{i\omega t} \right\} \\
&= -y_n(kr_s)\cos(\omega t) + j_n(kr_s)\sin(\omega t)
\end{aligned} \tag{47}
$$

4 SOUND FIELD DECOMPOSITION

We determine the sound field coefficients of the general real sound field in equations (27) and (37) using the orthogonality equations (15), (20), (21) and (22). We assume a simple integral of the sound field over a sphere at radius R, ignoring the practical limitations caused by the zeros of the spherical Bessel functions [11,15]. We also only consider the decomposition at a single frequency ω. This simplifies the equalization of each mode to that of a simple multiplication.

4.1. Sound Field Decomposition using complex harmonics

The complex coefficients C_n^m may be found by taking the analytic signal of each microphone signal to produce the complex form in equation (16). Then the complex decomposition at the single frequency ω is

$$
\begin{aligned}
z_{nmC}(\omega,t) &= \frac{1}{j_n(kR)}\int z(R,\theta,\phi,\omega,t)\,Y_n^m(\theta,\phi)^*\,d\Omega \\
&= C_n^m(k)\,e^{i\omega t}
\end{aligned} \tag{48}
$$

The complex coefficient is thus produced directly from the analytic signal. The complete $(n,m)th$ ambisonic signal is the inverse Fourier transform

$$
z_{nm}(t) = \int_0^{\infty} C_n^m(k)\,e^{i\omega t}\,d\omega \tag{49}
$$

A full decomposition of the sound field up to the Nth order thus requires $(N+1)^2$ complex signals, whereas the real decomposition (equation (37)) requires $(N+1)^2$ real signals. However, since taking the analytic signal reduces the signal bandwidth by a factor of two, the sample rate of the analytic signal may be halved, so in principle there is no difference in the number of samples per second required to represent the complex and real decompositions. However, the generation of the analytic signal does require additional processing (see section 2.1).

The complex coefficients cannot be determined from the real microphone signals because the negative m and positive m coefficients are not separable. Specifically, since $E_n^{-m} = E_n^m$ and $O_n^{-m} = -O_n^m$, then from equation (27) the real and imaginary parts of the complex decomposition of the real pressure

$$
p_{nmC}(t,k) = \frac{1}{j_n(kR)}\int p(R,\theta,\phi,\omega,t)\,Y_n^m(\theta,\phi)^*\,d\Omega \tag{50}
$$

are (see also equations (23) and (24))

$$
\begin{aligned}
&\operatorname{Re}\{p_{nmC}(\omega,t)\} \\
&= \frac{1}{\sqrt{\varepsilon_m}\,j_n(kR)}\int p(R,\theta,\phi,\omega,t)\,E_n^m(\theta,\phi)\,d\Omega \\
&= \frac{C_{nR}^m(k)\cos(\omega t) - C_{nI}^m(k)\sin(\omega t)}{\varepsilon_m} \\
&\quad + \frac{C_{nR}^{-m}(k)\cos(\omega t) - C_{nI}^{-m}(k)\sin(\omega t)}{\varepsilon_m} \\
&= \operatorname{Re}\left\{\frac{C_n^m(k) + C_n^{-m}(k)}{\varepsilon_m}\,e^{i\omega t}\right\}
\end{aligned} \tag{51}
$$

and

$$
\begin{aligned}
&\operatorname{Im}\{p_{nmC}(\omega,t)\} \\
&= \frac{-1}{\sqrt{\varepsilon_m}\,j_n(kR)}\int p(R,\theta,\phi,\omega,t)\,O_n^m(\theta,\phi)\,d\Omega \\
&= \frac{C_{nI}^m(k)\cos(\omega t) + C_{nR}^m(k)\sin(\omega t)}{\varepsilon_m} \\
&\quad - \frac{C_{nI}^{-m}(k)\cos(\omega t) + C_{nR}^{-m}(k)\sin(\omega t)}{\varepsilon_m} \\
&= \operatorname{Im}\left\{\frac{C_n^m(k) - C_n^{-m}(k)}{\varepsilon_m}\,e^{i\omega t}\right\}
\end{aligned} \tag{52}
$$

Hence,

$$
\begin{aligned}
p_{nmC}(\omega,t) &= \operatorname{Re}\{p_{nmC}(\omega,t)\} + i\operatorname{Im}\{p_{nmC}(\omega,t)\} \\
&= \frac{1}{\varepsilon_m}\left[C_n^m(k)\,e^{i\omega t} + \left(C_n^{-m}(k)\,e^{i\omega t}\right)^*\right]
\end{aligned} \tag{53}
$$

Hence the complex result obtained from the real sound pressures includes the conjugate of the negative m coefficient. Note that $C_n^{m*} \neq C_n^{-m}$ because they have complex terms that only depend on n, see examples above. Hence the complex spherical harmonic decomposition requires that the real microphone signals are first made analytic.

4.2. Sound Field Decomposition using real harmonics

The even harmonic decomposition of the real sound field in equation (37) is obtained as

$$
\begin{aligned}
p_{nmE}(\omega,t) &= \frac{1}{j_n(kR)}\int p(R,\theta,\phi,\omega,t)\,E_n^m(\theta,\phi)\,d\Omega \\
&= \left[A_{nR}^m(k)\cos\omega t - A_{nI}^m(k)\sin\omega t\right] \\
&= \operatorname{Re}\left\{A_n^m(k)\,e^{i\omega t}\right\}
\end{aligned} \tag{54}
$$

and the odd harmonic decomposition is

$$
\begin{aligned}
p_{nmO}(\omega,t) &= \frac{1}{j_n(kR)}\int p(R,\theta,\phi,\omega,t)\,O_n^m(\theta,\phi)\,d\Omega \\
&= \left[B_{nR}^m(k)\cos\omega t - B_{nI}^m(k)\sin\omega t\right] \\
&= \operatorname{Re}\left\{B_n^m(k)\,e^{i\omega t}\right\}
\end{aligned} \tag{55}
$$

Hence, the standard Ambisonics decomposition at each frequency produces the quadrature modulated components of the underlying sound field coefficients. If the analytic signals of $p_{nmE}(t)$ and $p_{nmO}(t)$ are taken then the complex field coefficients A_n^m and B_n^m may be determined, and the complex coefficients determined from equations (34) and (35). However, the complex forms are not required for sound reproduction.

5 SOUND FIELD REPRODUCTION

In Ambisonics, the matrixing of the Ambisonics signals to the loudspeakers is termed decoding. The goal of decoding is to produce the same sound field in the region of the centre of a loudspeaker array as the original sound field. To do this, we assume that the loudspeakers are ideal monopole sources and we require that the sound field coefficients produced by the loudspeakers equal those of the original sound field. This approach may be termed the mode matching solution.

We also present the simple source solution [7,11], which is obtained by assuming ideal monopole loudspeakers and deriving the weights required to produce the desired soundfield. This solution has an analytic form and does not require an inverse matrix calculation. We present the solution derived in [11] for complex harmonics and the simple source solution obtained using real spherical harmonics.

We assume a spherical array of loudspeakers at coordinates $(\theta_l, \phi_l), l \in [1, L]$, that produce close to an equal sampling over the sphere. A number of samplings are available at [20] and these include the solid angle weightings β_l to allow an accurate discrete approximation to a continuous integral over the sphere.

5.1. Reproduction using complex harmonics

Mode matching solution

Using the complex form of the point source (equation (42)), and the complex sound field description (equation (16)), we require the approximate field produced by L loudspeakers to equal the desired field

$$\begin{aligned}\hat{z}(r,\theta,\phi,\omega,t) &= e^{i\omega t}(-ik)\sum_{n=0}^{\infty} j_n(kr)\, h_n^{(2)}(kR) \\ &\quad\times\sum_{m=-n}^{n}\left[\sum_{l=1}^{L} w_l(k)\, Y_n^m(\theta_l,\phi_l)^*\right] Y_n^m(\theta,\phi) \\ &= e^{i\omega t}\sum_{n=0}^{\infty} j_n(kr)\sum_{m=-n}^{n} C_n^m(k)\, Y_n^m(\theta,\phi)\end{aligned} \tag{56}$$

For each (n,m) this requires

$$(-ik)\, h_n^{(2)}(kR)\sum_{l=1}^{L} w_l(k)\, Y_n^m(\theta_l,\phi_l)^* = C_n^m(k) \tag{57}$$

These mode matching equations may be written in matrix form for each (n,m),

$$\mathbf{\Psi} w(k) = C(k) \tag{58}$$

where $\mathbf{\Psi}$ is an $(N+1)^2$ by L matrix, $w(k)$ is a vector of loudspeaker weights at frequency k and $C(k)$ is the vector of complex field coefficients. For the case where $(N+1)^2 \geq L$ the least squares error solution may be found at each frequency as [9,11]

$$w(k) = \left[\mathbf{\Psi}^H\mathbf{\Psi} + \lambda\mathbf{I}\right]^{-1}\mathbf{\Psi}^H C(k) \tag{59}$$

where λ is a regularization parameter. This solution tells us that the loudspeaker weights are obtained from the complex analytic Ambisonics signals $C(k)$ by filtering them with the L by $(N+1)^2$ matrix filter

$$\mathbf{Q}_C = \left[\mathbf{\Psi}^H\mathbf{\Psi} + \lambda\mathbf{I}\right]^{-1}\mathbf{\Psi}^H \tag{60}$$

The real parts of the L complex output signals $\mathbf{Q}_C$ are sent to the loudspeakers.

For $(N+1)^2 < L$ the loudspeaker weights are obtained from the minimum energy solution [9,11]

$$w(k) = \mathbf{\Psi}^H\left[\mathbf{\Psi}\mathbf{\Psi}^H + \lambda\mathbf{I}\right]^{-1} C(k) \tag{61}$$

where λ is included to allow reduction of the loudspeaker weight energy if desired.

Simple source solution

The simple source loudspeaker weights for a spherical loudspeaker array may be derived by matching the interior and exterior field expansions on the surface of the sphere [7], or equivalently by deriving the sound field produced if the array weights were sampled versions of a single spherical harmonic and building the solution by a weighted sum of these terms [11]. The solution for positive frequencies and complex harmonics is

$$w_l(k) = \frac{i\beta_l}{k}\sum_{n=0}^{N}\frac{1}{h_n^{(2)}(kR)}\sum_{m=-n}^{n} C_n^m Y_n^m(\theta_l,\phi_l) \tag{62}$$

Each harmonic term in this equation satisfies the mode matching equation (57) and the simple source solution produces similar results to the mode matching inverse approach [11]. The L by $(N+1)^2$ simple source reproduction matrix $\mathbf{Q}_{CSS}$ thus has elements

$$\mathbf{Q}_{CSS}(l,v,k) = \frac{i\beta_l}{k h_n^{(2)}(kR)} Y_n^m(\theta_l,\phi_l) \tag{63}$$

for $l \in [1, L], v \in \left[1, (N+1)^2\right]$, where $v=n^2+n+m+1$. As for the mode matching case, the real part of the complex matrix outputs are sent to the loudspeakers.

5.2. Reproduction using real harmonics

Mode matching solution

Using the spherical harmonic description of the real sound field due to a point source (equation (46)) for each of L loudspeakers at the same radius R, and requiring the weighted sum of these to match the original sound field (equation (37)) yields the real sound pressure matching equation

$$\begin{aligned} & k\sum_{n=0}^{\infty} j_n(kr)\left[-y_n(kR)\cos(\omega t)+j_n(kR)\sin(\omega t)\right] \\ & \times\sum_{m=0}^{n}\sum_{l=1}^{L} w_l(k)\left[E_n^m(\theta,\phi)E_n^m(\theta_l,\phi_l)+O_n^m(\theta,\phi)O_n^m(\theta_l,\phi_l)\right] \\ & =\sum_{n=0}^{\infty} j_n(kr)\sum_{m=0}^{n}\left[A_{nR}^m\cos(\omega t)-A_{nI}^m\sin(\omega t)\right]E_n^m(\theta,\phi) \\ & +\sum_{n=0}^{\infty} j_n(kr)\sum_{m=0}^{n}\left[B_{nR}^m\cos(\omega t)-B_{nI}^m\sin(\omega t)\right]O_n^m(\theta,\phi) \end{aligned} \tag{64}$$

For each n and m, and since the even and odd harmonics are orthogonal, this requires

$$\begin{aligned} & k\left[-y_n(kR)\cos(\omega t)+j_n(kR)\sin(\omega t)\right]\sum_{l=1}^{L} w_l(k)E_n^m(\theta_l,\phi_l) \\ & =\left[A_{nR}^m\cos(\omega t)-A_{nI}^m\sin(\omega t)\right]=p_{nmE}(t) \end{aligned} \tag{65}$$

and

$$\begin{aligned} & k\left[-y_n(kR)\cos(\omega t)+j_n(kR)\sin(\omega t)\right]\sum_{l=1}^{L} w_l(k)O_n^m(\theta_l,\phi_l) \\ & =\left[B_{nR}^m\cos(\omega t)-B_{nI}^m\sin(\omega t)\right]=p_{nmO}(t) \end{aligned} \tag{66}$$

These two equations are the real mode matching equations, derived from the real Ambisonics signals $p_{nmE}(\omega,t)$ and $p_{nmO}(\omega,t)$ (equations (54) and (55)).

In order to solve for the speaker weights, we take the analytic signal (using equation (47), (54) & (55)) producing the complex forms

$$-ikh_n^{(2)}(kR)e^{i\omega t}\sum_{l=1}^{L} w_l(k)E_n^m(\theta_l,\phi_l)=A_n^m(k)e^{i\omega t} \tag{67}$$

and

$$-ikh_n^{(2)}(kR)e^{i\omega t}\sum_{l=1}^{L} w_l(k)O_n^m(\theta_l,\phi_l)=B_n^m(k)e^{i\omega t} \tag{68}$$

This is valid since if we match the complex forms, then the real parts will also be matched.

The time variation may now be removed and we are left with the real spherical harmonic mode matching equations

$$-ikh_n^{(2)}(kR)\sum_{l=1}^{L} w_l(k)E_n^m(\theta_l,\phi_l)=A_n^m(k) \tag{69}$$

and

$$-ikh_n^{(2)}(kR)\sum_{l=1}^{L} w_l(k)O_n^m(\theta_l,\phi_l)=B_n^m(k) \tag{70}$$

These two equations may be written for all modes up to *Nth* order as

$$\mathbf{\Phi}w(k)=\begin{bmatrix}\mathbf{\Phi}_E \\ \mathbf{\Phi}_O\end{bmatrix}w(k)=\begin{bmatrix}A(k) \\ B(k)\end{bmatrix} \tag{71}$$

where $\mathbf{\Phi}_E$ is $(N+1)(N/2+1)$ by L, and $\mathbf{\Phi}_O$ is $(N+1)N/2$ by L. (The complete matrix $\mathbf{\Phi}$ is then $(N+1)^2$ by L.) For $(N+1)^2 \geq L$ the weights may then be obtained from

$$w(k)=\left[\mathbf{\Phi}^H\mathbf{\Phi}+\lambda\mathbf{I}\right]^{-1}\mathbf{\Phi}^H\begin{bmatrix}A(k) \\ B(k)\end{bmatrix} \tag{72}$$

Hence the loudspeaker weights are obtained from the analytic form of the Ambisonics signals $A(k)$ and $B(k)$ by filtering them with the L by $(N+1)^2$ matrix filter

$$\mathbf{Q}_R=\left[\mathbf{\Phi}^H\mathbf{\Phi}+\lambda\mathbf{I}\right]^{-1}\mathbf{\Phi}^H \tag{73}$$

which must be computed for each frequency k. Since this solution produces the complex weights for positive frequency ω, the real decoder filters must be obtained by using the conjugate weights for the corresponding negative frequency, or by taking the real part of the complex time filter. In practice, the decoders are most efficiently implemented using fast convolution, ie M samples of the set of $(N+1)^2$ real Ambisonics signals are taken and the FFT of each computed. At each positive frequency the vector of bins $[A(m)\ B(m)]^T$ is multiplied by $\mathbf{Q}_R(m)$. The negative bins may be set to zero, and the real parts of the inverse FFT outputs taken to produce the loudspeaker signals.

For $(N+1)^2 < L$ the controlled minimum energy solution is

$$w(k)=\mathbf{\Phi}^H\left[\mathbf{\Phi}\mathbf{\Phi}^H+\lambda\mathbf{I}\right]^{-1}\begin{bmatrix}A(k) \\ B(k)\end{bmatrix} \tag{74}$$

Equivalence of real and complex forms

The real and complex mode matching equations are equivalent. The two real-harmonic mode matching equations ((67) and (68)) may be combined as

$$\begin{aligned} & -ikh_n^{(2)}(kR)\sum_{l=1}^{L} w_l(k)\left(E_n^m(\theta_l,\phi_l)-iO_n^m(\theta_l,\phi_l)\right) \\ & =A_n^m(k)-iB_n^m(k) \end{aligned} \tag{75}$$

Scaling by $1/\sqrt{\varepsilon_m}$ yields

$$-ikh_n^{(2)}(kR)\sum_{l=1}^{L} w_l(k)Y_n^m(\theta_l,\phi_l)^*=C_n^m(k) \tag{76}$$

which is the complex mode matching equation. Hence, we have demonstrated that the complex and real approaches have the same underlying mode matching equations. However, the real harmonic approach allows the use of real Ambisonics signals which do not require the generation of the analytic signal.

Simple Source solution

As for the complex case (equation (62)), the simple source approach may also be applied to reproduction using real spherical harmonics. Expanding equation (62) in the same manner as equation (28) yields the simple source solution

$$w_l(k) = \frac{i\beta_l}{k}\sum_{n=0}^{N}\frac{1}{h_n^{(2)}(kR)}\sum_{m=0}^{n}\left[A_n^m E_n^m(\theta_l,\phi_l)+B_n^m O_n^m(\theta_l,\phi_l)\right] \quad (77)$$

This equation satisfies the sum of equations (69) and (70) (since $E_n^m(\theta_l,\phi_l)$ and $O_n^m(\theta_l,\phi_l)$ are orthogonal). The loudspeaker weights can be written directly in terms of the vectors of ambisonics signals $A(k)$ and $B(k)$ and two decode matrices $\mathbf{Q}_{ESS}$ and $\mathbf{Q}_{OSS}$ as $w(k)=\mathbf{Q}_{ESS}A(k)+\mathbf{Q}_{OSS}B(k)$ with elements [3,9]

$$\mathbf{Q}_{ESS}(l,v,k)=\frac{i\beta_l}{kh_n^{(2)}(kR)}E_n^m(\theta_l,\phi_l) \quad (78)$$

for $l\in[1,L], v\in[1,(N+1)(N/2+1)]$

where $v=n(n-1)/2+n+m+1$, and

$$\mathbf{Q}_{OSS}(l,v,k)=\frac{i\beta_l}{kh_n^{(2)}(kR)}O_n^m(\theta_l,\phi_l) \quad (79)$$

for $l\in[1,L], v\in[1,(N+1)N/2]$, where $v=n(n-1)/2+m+1$.

These matrices may be used for the reproduction of general Ambisonics signals $A(k)$ and $B(k)$. Note that $\mathbf{Q}_{OSS}$ has no elements for m=0.

For the direct synthesis of sound fields, (which can be approximated as a sum of point sources), we derive the panning functions which describe the loudspeaker weights for a single point source. Using equations (43) and (44),

$$\begin{aligned} w_l(r_s,\theta_s,\phi_s,k) &= \beta_l\sum_{n=0}^{N}\frac{h_n^{(2)}(kr_s)}{h_n^{(2)}(kR)}\sum_{m=0}^{n}E_n^m(\theta_s,\phi_s)E_n^m(\theta_l,\phi_l) \\ &+\beta_l\sum_{n=0}^{N}\frac{h_n^{(2)}(kr_s)}{h_n^{(2)}(kR)}\sum_{m=0}^{n}O_n^m(\theta_s,\phi_s)O_n^m(\theta_l,\phi_l) \end{aligned} \quad (80)$$

For a source at the loudspeaker array radius, $r_s=R$, and

$$\begin{aligned} w_l(\theta_s,\phi_s) &= \beta_l\sum_{n=0}^{N}\sum_{m=0}^{n}E_n^m(\theta_s,\phi_s)E_n^m(\theta_l,\phi_l) \\ &+\beta_l\sum_{n=0}^{N}\sum_{m=0}^{n}O_n^m(\theta_s,\phi_s)O_n^m(\theta_l,\phi_l) \end{aligned} \quad (81)$$

which is real and frequency-independent. This equation also applies for the plane wave case, and describes basic decoding for any spherical array geometry.

6 CONCLUSION

This paper has provided a unified description of 3D Ambisonics based on complex or real spherical harmonic functions. By using consistent definitions, the coefficients for the real and complex sound field expansions are simply related. The coefficients for either case are in general complex.

The complex spherical harmonic expansion coefficients can only be obtained from real microphone signals if the analytic form of each microphone signal is first taken to produce a complex signal with no negative frequencies.

The real spherical harmonic expansion coefficients can be obtained directly from the real microphone signals and they are encoded in quadrature modulated form in the real Ambisonics signals, but this is all that is required for sound reproduction, and no analytic signal generation is required.

The design of decoders for both complex and real Ambisonics signals has been discussed. The complex harmonic decoder need only process analytic signals, and so in principle it has complex impulse responses. However, the decoder would be implemented using fast convolution, which only requires multiplication of positive frequencies, and so it is relatively simple to implement, with the real part of the outputs being sent to the loudspeakers.

The real harmonic decoder contains real filter impulse responses, but since it too is most simply implemented in the frequency domain where only positive frequencies need be considered it is similar in implementation to the complex decoder.

It has been shown that the decoders for real and complex harmonics have the same underlying mode matching equations.

The primary advantage of real harmonics is that they allow the transmission of real ambisonics signals, without requiring analytic filtering. In particular, the decoder elements for source radii equal to loudspeaker array radius (or for plane waves) are also real, making decoding particularly simple.

REFERENCES

1) M. Gerzon, “Periphony: With-height sound reproduction,” *J. Audio Eng. Soc*., vol. 21, no. 1, pp 2 – 10, 1973

2) J. Daniel, “Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia,” *PhD Thesis*, University of Paris 6, 2000

3) J. Daniel, “Spatial sound encoding including near field effect: Introducing distance coding filters and a viable new ambisonics format,” *AES 23rd Int. Conference*, 2003

4) A. Laborie, R. Bruno and S. Montoya, “A new comprehensive approach of surround sound recording,” *AES 114th Conv.*, Amsterdam, preprint 5717, 2003

5) www.muse.demon.co.uk/3daudio.html, last cited May 2009

6) P. M. Morse and H. Feshbach, *Methods of theoretical physics*, McGraw-Hill, 1953

7) E. G. Williams, *Fourier Acoustics*, Academic Press, 1999

8) D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Applied Mathematical Sciences Vol. 93, 2nd Edition, Springer, 1998

9) D. B. Ward and T. D. Abhayapala, “Reproduction of a plane-wave sound field using an array of loudspeakers,” *IEEE Trans. Speech and Audio Proc.*, vol. 9, no. 6, pp 697 – 707, 2001

10) T. Betlehem and T. Abhayapala, “Theory and design of sound field reproduction in reverberant rooms,” *J. Acoust. Soc. Am*., vol. 117, Pt. 1, pp 2100-2111, April 2005

11) M. A. Poletti, “Three dimensional surround sound systems based on spherical harmonics,” *J. Audio Eng. Soc*., Vol. 53, No. 11, pp 1004-1025, November 2005

12) F. Fazi and P. A. Nelson, J. E. N. Christensen and J. Seo, “Surround system based on three dimensional sound field reconstruction,” *AES 125th Conv.*, October 2-5, 2008, preprint 7555

13) J. Ahrens and S. Spors, “Analytical driving functions for higher order Ambisonics,” *IEEE Int. Conf. on Acoustics, Speech and Signal Proc.* (*ICASSP*), 2008, March 31-April 4, 2008 pp 373 – 376

14) J. Hannemann and K. D. Donohue, “Virtual sound source rendering using a multiple-expansion and method-of-moments,” *J. Audio Eng. Soc*., Vol. 56, No. 6, pp 473-481, June 2008

15) J. Meyer and G. W. Elko, “Spherical microphone arrays for 3D sound recording,” in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Edited by Y. A. Hunag and J. Benesty, Kluwer Academic Publishers, 2004

16) L. E. Kinsler, A. R. Frey, A. B. Coppens and J. V. Sanders, *Fundamentals of Acoustics*, 4th Edition, John Wiley and Sons

17) L. Cohen, *Time-Frequency Analysis*, Prentice Hall PTR, 1995

18) A. Papoulis, *Signal Analysis*, McGraw-Hill International Student Edition 1977

19) A. Reilly, G. Frazer and B. Boashash, “Analytic signal generation – tips and traps,” IEEE Trans. Signal Proc., vol. 42, no. 11, pp 3241-3245, 1994

20) www.personal.soton.ac.uk/jf1w07/nodes/nodes.html, last accessed March 31, 2009