Ambisonic Sound for the Masses

by Jeffery S. Bamford
Audio Research Group, Department of Physics
University of Waterloo, Waterloo, Ontario
N2L 3G1

July 18, 1994

In the past decade, television broadcasts have increasingly been available in stereo. The availability of stereo Video Cassette Recorders has also increased interest in home theatre. Films have used multi-speaker arrays to create better imaging and this has filtered down into videotapes of those films by using Dolby Surround encoding.

When a recording is made using Dolby Surround, four signals are encoded into two channels for storage. Upon playback, a Surround Sound decoder is required. Dolby Surround certainly sounds different from stereo but offers little imaging for rear sounds.

An Ambisonic system uses three or more channels to encode the sound for reproduction. By using more channels of information, a better reproduction of the image is obtained. Ambisonic systems also easily allow for the addition of more speakers. At the minimum, there must be as many speakers as channels. Adding more speakers will improve imaging.

The investigation is started by considering a plane wave. The plane wave is at an angle ψ with respect to the forward pointing x-axis. (In Ambisonics, it is customary for the x-axis to point forward and for the y-axis to point to the left). At a field (or listening) point, the plane wave can be described by $S_\psi = P_\psi e^{ikr\cos(\phi-\psi)}$, which is in polar coordinates (i.e. radius r at an angle ϕ). k is the wave number, ψ is the azimuthal angle the original plane wave makes with the x-axis, r is the distance from the origin of the co-ordinates (which is typically the centre of the listening area), i is $\sqrt{-1}$ and P_ψ is the peak pressure or amplitude of the wave.

The plane wave can be expanded in terms of spherical harmonics to be of the form: [1]

$$S_\psi = P_\psi(J_o(kr) + \sum_{m=1}^{\infty} 2\, i^m J_m(kr)[\cos(m\psi)\cos(m\phi) + \sin(m\psi)\sin(m\phi)]) \quad (1)$$

where J_o and J_m are cylindrical Bessel functions of the first kind.

For a system with N loudspeakers that are equidistant from the centre location and in a regular array (i.e. equal angles separating the speakers), the plane wave signal from the nth loudspeaker is given by: [2]

$$S_n = P_n(J_o(kr) + \sum_{m=1}^{\infty} 2\, i^m J_m(kr)[\cos(m\phi_n)\cos(m\phi) + \sin(m\phi_n)\sin(m\phi)]) \quad (2)$$

where P_n is the pressure from the nth speaker, and ϕ_n is the angle of the nth speaker.

If we sum over each of the N signals from the speakers and match this sum with the original plane wave, the following can be noted:

$$P_\psi = \sum_{n=1}^{N} P_n \quad (3)$$

$$P_\psi \cos(m\psi) = \sum_{n=1}^{N} P_n \cos(m\phi_n) \quad (4)$$

$$P_\psi \sin(m\psi) = \sum_{n=1}^{N} P_n \sin(m\phi_n) \quad (5)$$

The above equations represent the matching conditions for the spherical harmonics. By matching all the orders of spherical harmonics, the original plane wave can be reproduced at the centre of listening area. Clearly, it is unfeasible to reproduce *all* of the spherical harmonics as there are an infinite number of them. For the present case only the first and second order limits are considered.

Equation (3) gives the first of the Ambisonic signals which is the pressure of the wave, which is denoted by W. The next order, m=1, gives the next two Ambisonic signals, $X = P_\psi\cos(\psi)$ and $Y = P_\psi\sin(\psi)$. The second-order Ambisonic signals are $U = P_\psi\cos(2\psi)$ and $V = P_\psi\sin(2\psi)$. With those five signals, a plane wave can be described up to second-order.

One way of analyzing this system is to look at the integrated wavefront error. This is done by integrating at a constant value of kr around the origin the difference between the original plane wave and the sum of the speaker signals. [2]

$$
\begin{aligned}
D &= \frac{1}{2\,\pi\,|P_\psi|}\int_0^{2\,\pi}\left|\sum_{n=1}^{N} P_n\,J_o(kr)\right. \\
&+ \sum_{n=1}^{N} 2\,i\,J_I\,(kr)\,(\cos(\phi)\cos(\phi_n)+\sin(\phi)\sin(\phi_n)) \\
&- \sum_{n=1}^{N} 2\,J_2(kr)\,P_n\,(\cos(2\,\phi)\cos(2\,\phi_n) \\
&\qquad\qquad + \sin(2\,\phi)\sin(2\,\phi_n)) \\
&- \left. P_\psi\,e^{ikr\,\cos(\phi-\psi)}\right|\,d\phi \qquad (6)
\end{aligned}
$$

With this equation, it is then possible to model various sized systems. By graphing the results for different numbers of speakers and layouts, it is possible to gauge how effective an Ambisonic Sound System could be. The above equation is for a second-order (five channel) Ambisonic System. By dropping the terms involving $\cos(2\psi)$ and $\sin(2\psi)$, the model is correct to first-order.

In order to compare the Ambisonic results with regular stereo and Dolby Surround, equations for these systems are required. For the stereo case, with speakers at $\phi_1 = -\phi_2$ and with $L = P_1$ and $R = P_2$, the signals are: $P_n = \frac{1}{2}\left(\frac{X}{\cos\phi_n} + \frac{Y}{\sin\phi_n}\right)$. [2]

Stereo only matches the velocity of the plane wave but *not* its pressure. Dolby stereo adds three more speakers driven as follows: a centre speaker with $\frac{1}{\sqrt{2}}(L+R)$ and two rear channels as $\pm\frac{1}{2}(L \mp R)$. It should be noted that Dolby Stereo reproduction is not quite this simple. For example, Dolby Surround Systems put a delay of a few milliseconds on the rear speakers. The rear signals are also limited to 7kHz. [3]

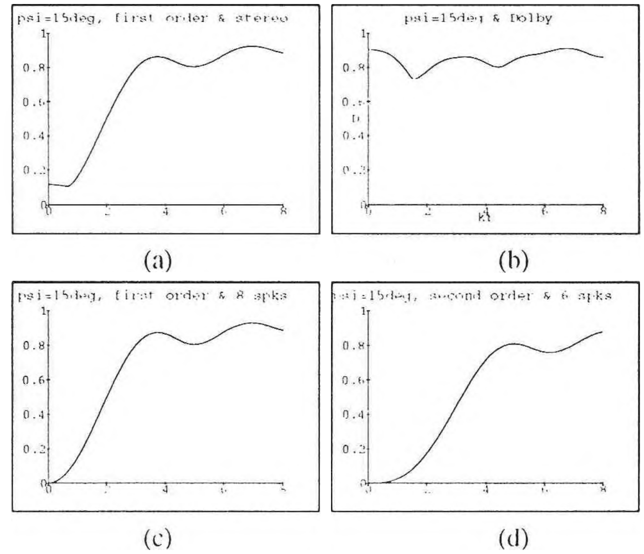
Ambisonic systems have as a general speaker signal: [4]

$$
\begin{aligned}
P_n &= (W + 2\cos(\phi_n)X + 2\sin(\phi_n)Y \\
&+ 2\cos(2\phi_n)U + 2\sin(2\phi_n)V)\,/N \qquad (7)
\end{aligned}
$$

where for a first-order system U and V are zero.

As a test of the various systems, a plane wave was modeled at an angle of 15 degrees. For reproduction, each of the systems would have to recreate a plane wave at 15 degrees. The stereo system had speakers placed at ±30 degrees. The Dolby System had the two speakers as in the stereo case, plus a centre channel and two rear channels at 150 and 210 degrees (i.e. 180±30 degrees). The Ambisonic systems had speakers placed at 30, 90, 150, 210, 270 and 330 degrees. The plots of D versus kr for the various systems are shown in Figure 1.

Both stereo and Ambisonic systems offer fairly good imaging for low values of kr. This would correspond to low frequencies or for locations close to the centre of the listening area. The Ambisonic systems are clearly much better at localizations near the centre of the listening area than is the stereo



(a) (b) (c) (d)

Figure 1: Plots of the D error versus kr for $\psi = 15$ for (a) Stereo (b) Dolby (c) Ambisonic (first-order) with six speakers and (d) Ambisonic (second-order) with six speakers

system. Dolby offers very little imaging anywhere. This is understandable as the Dolby System is mainly for ambient effects. The farther one gets from the centre of the listening area, the worse all four systems get. At kr=8 there is very little difference between the four systems.

The Ambisonic systems offer two important improvements over stereo and Dolby methods. The first is that they offer better imaging. An ambisonic system is able to image over 360 degrees, for example. The second benefit is that the effective area of listening is increased.

In conclusion, the second-order Ambisonic system offers improved imaging over a wider area than the first-order system and is suitable for larger rooms. A first-order system is better suited for the smaller home enviornment.

References

1. Morse and Ingard, *Theoretical Acoustics*. (McGraw-Hill, N.Y., 1968)

2. Vanderkooy and Lipshitz, "Anomalies of Wavefront Reconstruction in Stereo and Surround-Sound Reproduction", Audio Engineering Society Preprint 2554

3. S. Julstrom, "A High-Performance Surround Sound Process for Home Video", *Journal of the Audio Engineering Society*, vol. 35, pp. 536-549 (July/August 1987)

4. M. A. Gerzon, "Design of Ambisonic Decoders for Multispeaker Surround Sound", presented at the 58th Convention of the Audio Engineering Society, *Journal of the Audio Engineering Society (Abstracts)*, vol. 25 p. 1064