AMBISONIC HEMISPHERIC SOUND

The future of surround sound?

BACKGROUND

Michael Gerzon postulated in the 1970's that a full hemispheric sound was possible with 3 to 5 speakers on a horizontal plane. At the time he was a mathematics student at Oxford University. His passion was recording live sound in the conventional stereo format. He was discouraged with his ability to recreate a full perspective of what the listener heard at the live recording. Over a period of 20 years, he continued to experiment with recordings using up to 4 channels to fill in the missing ambient cues that were present in the hall used by the musicians. His observations concluded that what was missing was spatial information which is reproduced by phase relations generated in the original soundfield. This was in the analog era. His early calculations to capture this phase information started by determining microphone positions in relation to each other. Great difficulty was encountered as he was using Quad 56 bipolar electrostatic speakers rather than a single point source speaker of small dimensions which improves the arrival time and minimizes the effects of the reproduction room acoustics affecting what the listener heard. He was helped by Paul Hodges, Peter Craven, and Tony Faulkner from Guildford College.[i] By the mid 1990's, early digital technology helped evolve a better understanding of the process. Unfortunately, Gerzon died in 1996 at the age of 50 from an asthma attack. He did live long enough to write several papers that were given by the AES and Acoustical Society. They have become the core ambisonic information that has led to much interest and development of better software, quad microphones, and small digital recorders to capture information. Since 2000, many papers have been given exploring different aspects of ambisonics and interest continues to grow worldwide.

WHERE WE PRESENTLY ARE

Dolby has been the premier source of surround sound for both theater and home application. At the present, their decoding process has evolved to an infinite number of possible channels being available with the Atmos System. This allows a sound effect to be steered to any location in the hemispheric field. It is generally a mono source signal, but ambiance or delay can be added to increase realism. Most theater applications use 10 to 30 or more speakers in the surround field to give a more immersive experience. Home theater applications require 6 to 13 speakers driven by an assortment of high-quality receivers using Dolby licensing for the Atmos soundfield. Due to the speaker localization effect, the perceived sound level the listener hears is determined by his location in that soundfield. The perceived sound level quickly drops when compared to other speakers that are located further away in that soundfield. This limits the ability to best hear the soundfield as envisioned by the recording engineer to only a few seats in the middle of the theater.

As presently understood, the ambisonic format (employed to create the soundfield) is presented in first order which is 4 speakers, second order which is 9 speakers, and third order which is 13 or more speakers. The recordings are made in A format and reduced to B format for playback. Spatial information is generated by phase difference between the channels. This allows for the sensation of sound reproduced in the hemispheric field (from its source in the recording) to appear to come from a location between the speakers. Several universities and the BBC have been using large surround fields in third order configuration as the generally held opinion is it is not possible to get adequate information as to location cues in first order from only 4 speakers. The 4 speakers used are generally small home stereo type consisting of an 8 to 12 inch woofer and dome tweeter.

Presently, most soundtracks being recorded for both film presentation and gaming embed a 4 channel ambisonic track in the master recording. It would appear that the Atmos decoding uses some of this information to create their location cues in the soundfield. There has been much interest in the ambisonic format for gaming, virtual reality, simulators, and specialty theaters. Most monitoring has been in the binaural form (using headphones) as experimenters have found conventional speakers totally inadequate to recreate the phase relation cues.

A NEW APPROACH

To achieve the ambisonic format with sufficient detail to provide a satisfactory listening experience, new technologies had to be applied to fully develop the hemispheric field in the first order format of 4 speakers. This meant having the ability to hear the z axis of overhead information with the speakers employed at approximately ear level. Experimentation with a Rode Ambisonic microphone led to the conclusion that many attempts to record ambisonic information were limited to the diameter of the recorded soundfield. Information recorded in a small studio environment had a restricted soundfield limited in diameter to the recording area. Recording sound effects outdoors opened the opportunity to hear the reproduced soundfield over a much larger diameter area than what was recorded. An example is the recording of a 747 jet landing with it passing approximately 300 feet over head and fading into the distance. Upon playback, it is noticed that when the speakers are set up outdoors the jet goes over your head as recorded and you can walk outside the area quadrant of the speaker setup, and it will still fly over your head. The phase relation cues are still present. It is this observer's opinion that much needs to be learned regarding what is possible in recording with the ambisonic format when the new speakers are employed in the playback.

WHAT IS REQUIRED

It was discovered that almost all quality speakers are incapable of reproducing phase relation differences. There are several technical requirements that must be met to hear satisfactory detail in first order ambisonics using only 4 speakers in B format. The scope of this tutorial is too limited to discuss these problems. A speaker that can reproduce phase difference is a useful tool in the analysis of suitable amplifiers. The speaker is capable of revealing minute differences not observable with conventional speakers. In addition, the limitations of presently available interface units and driving software becomes painfully evident. In general, it was observed that all switching and class D amplifiers have very limited capability in reproducing phase information and most give only a partial stereo image soundfield for 2 channel reference monitoring. Many of the earlier solid-state designs, despite limitations by current standards, had no difficulty in handling the new ambisonic format and giving a good sound stage. Again, a new technology speaker was needed to reveal this.

SUMMARY

A. There are 2 basic formats that can generate an immersive hemispheric soundfield. Both have limitations and application issues. The first is a conventional speaker system that generates the surround effect by generally using a steered mono signal to recreate the movement of a specific sound within the soundfield. It reproduces the ambiance of the original soundfield by sending this information to rear speakers. The primary application of this process is presently Dolby Atmos as used in both home and commercial theaters. The second application is the use of the ambisonic format in first through third order by using phase relations. It is the opinion of the author that this format offers much greater potential in recreating original soundfields as spatial information is presented by using phase relation differences. However, it is presently limited by lack of proper hardware for reproduction and limitation in the development of recording techniques to capture this information to best advantage.

B. The development of new hardware and recording techniques will be needed to fully capture the potential of ambisonic hemispheric sound. A new speaker has been developed that when used with an appropriate amplifier and interface solves most of the problems encountered using the very simple 4 speaker first order format. It can provide the steppingstone to further development in recording the process and is becoming more available to the audio world.

Sam Saye
5-4-22
LM; AES Audio Engineering Society
LM; SMPTE Society of Motion Picture and Television Engineers

---

[i] Into the Soundfield