Ambisonics: The Surround Alternative

by Richard Elen

Surround-sound has returned, and this time, thanks to home theater and DVD, it's here to stay. But is 5.1, the system designed for movie sound, the most appropriate for music? Perhaps there is a better way.* Richard Elen *discusses Ambisonics – a surround sound system that offers features which are difficult, if not impossible, to realize with other methods – and how it can coexist with 5.1-based techniques.

The Quadraphonic Quandary

QUADRAPHONY – the ancestor of today's "5.1" surround-sound systems and sharing many of the same shortcomings – is simply an extension of "intensity", or "amplitude" stereo, where sounds are positioned between two speakers in front of the listener by using solely the relative level of the two channels to provide the localization information. Quad sought to extend this into two dimensions by panning sounds between four speakers – two speakers at the front and two at the rear.

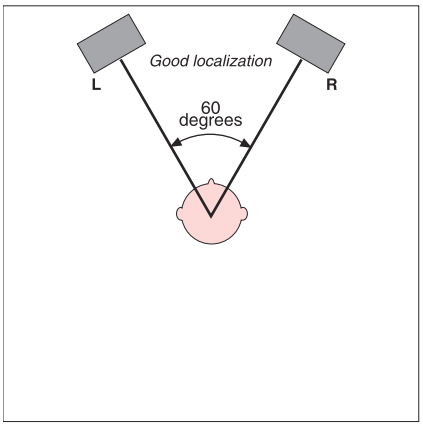

Fig 1. "Intensity" or "amplitude" stereo relies on level only to localize sounds between the speakers. If the ears hear both speakers (ie they are at 60 degrees or less) and at the front, we convert the amplitude differences into phase differences: it works.

There are certain limits to stereo replayed over loudspeakers. Most important, if the speakers are further apart than 60°, a "hole in the middle" tends to be experienced and sounds are drawn into the speakers, leaving virtually no inter-speaker imaging. In a quad system, the speakers are at 90°.

The problem is the use of level as the sole means of localization. The human ear/brain combination relies on a number of different localization techniques at different, overlapping frequencies.

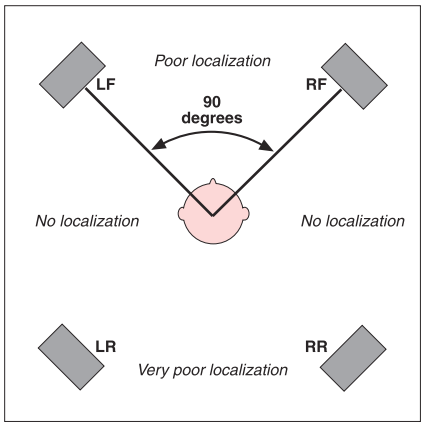

Fig 2. "Quad" relied on speakers at 90 degrees, around the listener. The spacing is too great and sounds are drawn into the speakers: localization is poor at the front, very poor at the rear and nonexistent elsewhere.

We use phase to localize sounds between 150 Hz and 1.5 kHz and level between 300 Hz and 5 kHz. Above 2.5 kHz, other directional cues are used. When two speakers in front of you are separated by no more than 60°, both ears hear both speakers and LF amplitude differences are converted into phase differences between the ears. However, this effect works only poorly if the speakers are behind you, and not at all if the speakers are to the side – or if they are separated by more than about 60°. This means that, by definition, traditional quad – and today's 5.1 systems which are their descendants – cannot offer optimum localization.

The fact that quad relied on recording four "discrete" channels which used relative levels between pairs of channels for localization (so-called "pairwise mixing") meant that good interchannel separation was important. Here, quad ran into a second problem: how to transmit four channels of information over then-available 2-channel media. One approach was the use of subcarriers on vinyl disks. More common, and less successful, was "matrixing" the four original channels into two for normal vinyl disks or FM broadcast, using a decoder to recover the original channels on replay (a so-called "4-2-4" system). The big problem: it is mathematically impossible. There's always a loss of information, resulting in significant localization errors.

Imagine a sound panned around the control room in a circle at constant speed. In the case of almost all quad systems (including all but one of the subcarrier systems), the sound on replay would not follow the same path. It would jump from one speaker to another, and come in from the edge of the circle.

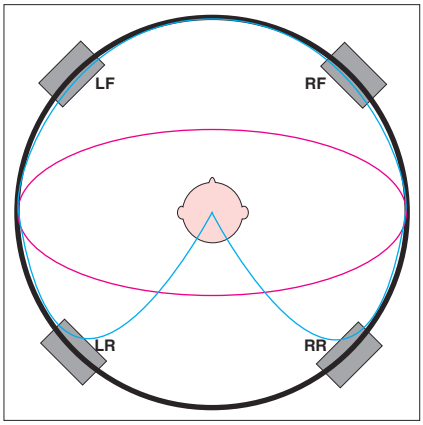

Fig 3. Matrix quad tried to get the four original channels into two and back again, which is impossible. A sound panned around the control room in a circle (black) would be replayed as a flat ellipse by SQ (red) and a cardioid shape (blue) by QS (the foundation for Dolby Surround).

Quad as a home system failed at the time, but it survived in the movie industry, where a descendant of the Sansui QS matrix system became the standard for movie theater surround, with a single surround channel, an LF effects channel, and a stereo front stage.

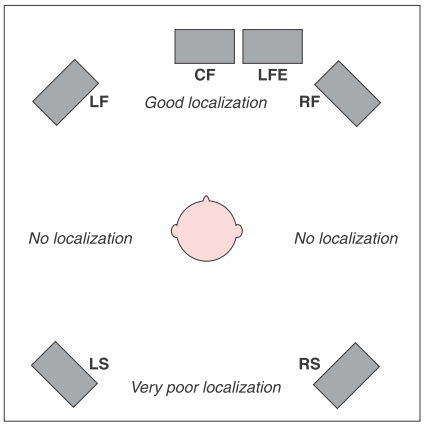

Fig 4. Today's "5.1" systems are based on "discrete quad", with a center-front channel to fill the hole between the front speakers - fine for dialog, but difficult to integrate in music mixing. And the localization everywhere else is as bad as before.

The separation of the front speakers resulted, inevitably, in a "hole in the middle", so a center channel was added to carry dialog.

Today's "5.1" systems build on the original model by adding a second surround channel. The Low Frequency Effects (or "Enhancement") – "LFE" – channel is the ".1".

The Promise of DVD

Today, for the first time, we have digital media that can provide multi-channel fully-digital sound carriers: in the form of DVD – the Digital Versatile Disk, in its several incarnations, and the Super-Audio CD (SACD). These technologies not only offer the possibility of carrying multichannel film soundtracks to a consumer audience: they also offer the a multichannel high-quality audio medium.

Today, it is being generally assumed that a system designed to make sound effects sound impressive in a movie theater will be ideal for high quality music in surround. Unfortunately, if the goal of surround sound is to reproduce in the home the creativity and subtlety of studio work, or accurately capture the acoustic environment of an orchestral performance, to name but two examples, 5.1 – even with perfect transmission from end to end – cannot, by definition, do it.

But there is another way of doing surround sound that can achieve this goal in an all-digital environment. The technology has been around for almost as long as quad; it has been in continuous use for music recording for almost 30 years; and, until the advent of multi-channel digital carriers, more original album releases had been created using it than all other surround-sound systems put together. That technology is *Ambisonics.*

The Arrival of Ambisonics

Ambisonics was the brain child of a group of British researchers, notably the late Michael A. Gerzon at the Mathematical Institute in Oxford, and Professor Peter Fellgett of the Cybernetics department at Reading University. They and their colleagues worked to develop a surround sound system that would enable a musical performance to be captured for replay in a conventional living room in which, as far as possible, the original sound and acoustic environment of the original performance (real concert hall or multitrack mix) would be recreated. The system was christened Ambisonics – meaning, simply, "surround sound".

The system was designed from the beginning to enable recordings to be made with a special surround microphone (the Soundfield Microphone, now manufactured in the UK by Soundfield Research) or with special console panpots and localization controls, or both. Ambisonic production equipment generates a 4-channel signal, called "B-Format", that embodies all the information in the soundfield, resolved into left-right, front-back and up-down information plus a mono reference signal. Interestingly, as early as the mid-Seventies, Ambisonics included the capability to record and reproduce height information, which even now is not a common part of surround-sound practice, despite the fact that it adds a lot to the realism of a surround system.

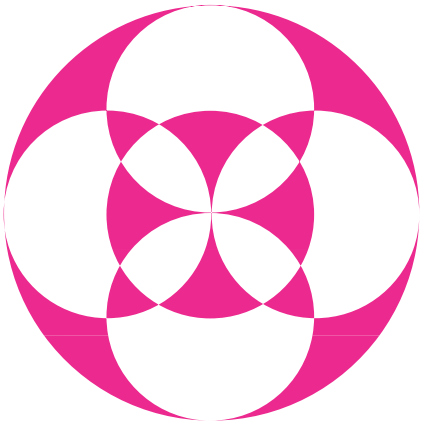

Fig 5. The Ambisonic logo indicates the components of a B-Format signal: the outer circle represents the omnidirectional feed, W; the vertical figure-8 front minus back (X); the horizontal figure-8 left minus right; and the center circle is the up minus down (Z) figure-8 seen from above.

For replay, the B-Format signal is fed to a decoder which derives a minimum of four loudspeaker feeds (for horizontal, or "planar" surround). Instead of four independent signals, the speakers in an Ambisonic replay system are fed with signals, each of which contain virtually all the elements of the recording, but with different relationships. The speakers work together to recreate the acoustic and ambience of the original recording.

If all that was involved was to capture the wavefronts impinging upon a soundfield mic and recreate them during replay, the results would be disappointing: the "sweet spot" would at best be the size of a football. Ambisonics, on the other hand, uses many of the methods of localization employed by the ear/brain combination to localize sound sources. There is wavefront reconstruction to a degree, but it is overlaid with level, phase and other directional cues to mimic human hearing.

The result is a number of noticeable benefits. First of all, you can put the speakers more or less where you want them – you tell the decoder where they are. Second, the surround effect is pronounced and stable over a very wide listening area. You can even stand outside the speaker array and experience a kind of "sonic image" emanating from within the array.

And finally, Ambisonics offered the promise of something that was beyond the capability of any other practical consumer surround system available then or now: the reproduction of height information. This "full-sphere" surround was christened *"Periphony"*, from Greek roots meaning "sound around the edge".

With the exception of discrete quad, all the extant surround systems were compatible, to a greater or lesser extent, with conventional stereo and mono. Yet B-Format consisted of sum-and-difference signals, like Blumlein M-S recording, which could not be listened to directly. Ambisonics needed a stereo/mono-compatible matrix as well. The UHJ hierarchy was developed to satisfy this need.

UHJ is an example of what is referred to as a "hierarchical" surround encoding scheme, offering an increasing gamut of capabilities depending on the number of transmission channels available and on the decoder. 4-channel UHJ carries the same information as a 4-channel B-Format signal, including height information. If three channels are available, the fourth channel can be dropped, leaving a high-resolution horizontal surround signal. If necessary, the third T channel can be bandwidth limited: this is referred to as a "2.5 channel". If only two channels are available, they can be used with a decoder to provide a very effective horizontal surround capability, although the accuracy of localization is not quite as high as in a 2.5- or three-channel version.

If no decoder is available, a 2-channel UHJ recording can be treated as a stereo signal. In this case, the listener experiences a "super stereo" effect that goes way beyond the speakers. 2-channel UHJ is a powerful "3D stereo" tool, at least as effective as more recent 2-channel surround techniques used today for computer and multimedia applications – with the added benefit that with a decoder, true surround is available. To borrow a term from modern Web site design, Ambisonics "degrades" almost perfectly.

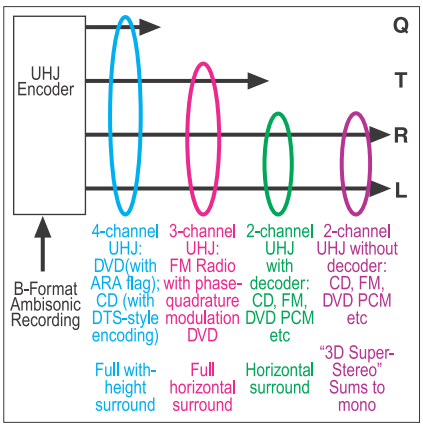

Fig 6. UHJ is a hierarchical surround encoding scheme which allows a surround signal to be experienced at the highest level according to the number of transmission channels available. A single mix satisfies all the available modes without compromise.

Ambisonics quickly attracted the attention of audiophile label Nimbus Records, and apart from a few initial stereo and QS releases, *every album ever released by Nimbus was recorded Ambisonically.*

Ambisonic mixing from multitrack became a reality in the early Eighties with the launch of a series of rack-mounted outboard processors by Audio+Design Recording in the UK, designed by Dr Geoff Barton and including a professional studio decoder, capable of decoding B-Format and 2- or 3-channel UHJ. The other units utilized a number of different approaches to localizing signals in an Ambisonic soundfield, and were all designed to integrate with a conventional console as much as possible.

The "**Pan-Rotate**" unit offered eight mono inputs that could be localized anywhere in two-dimensional space (including, for the first time, the ability to bring sounds in from the edge of the speaker array); the "**B-Format Converter**" was fed from four console groups and an aux send, and allowed console panpots to be used as Ambisonic panners; and the "**Transcoder**" turned B-Format into 2-channel UHJ for release, and also allowed a quick Ambisonic mix to be created by feeding a front and a rear stereo stage (of variable width) into the unit from a conventional console, to produce a 2-channel UHJ mix directly.

Members of the Ambisonic fraternity approached artists, engineers and producers and encouraged them to use the system, primarily by demonstrating its superiority. Although there were few people engaged in this kind of promotion, they had quite a good deal of success, and within a few years Ambisonics had been used by a varied range of artists, including Alan Parsons (*Stereotomy*), Tina Turner (*Break Every Rule*), Steve Hackett (*Till We Have Faces*), Adrian Legg (*Lost For Words*), and Paul McCartney's *Liverpool Oratorio* along with Ambisonically-mixed CDs on the Collins Classics label, to name but a few. In each case, the albums received rave reviews for their sound quality and the extent of the sound-stage, and in each case, the use of Ambisonics was the engineer or producer's decision, just as the use of any other kind of outboard equipment might be.

Barriers to Acceptance

It is entirely reasonable at this point to ask the question, "If Ambisonics is so amazing, why aren't we all using it?" The simple answer is money; the longer answer involves bureaucratic bungling

Ambisonics	5.1
Only requires four channels	Needs six channels
Includes height information	No height information
Only one mix required	At least two mixes required
Lower data density: lossless compression can be used	Higher data density: lossy compression required
Imaging equally accurate in all directions	Imaging only accurate across front stage
Imaging between the speakers	Poor inter-speaker imaging
Sounds can be placed within the array	Sounds must be placed around the edge of the array
Speakers can be placed almost anywhere	Speakers must be in special positions
"Sweet spot" large enough to enjoy image outside the array	"Sweet spot" only covers a relatively small area within the speakers.

Fig 7. The characteristics of Ambisonics and 5.1 compared. There are disadvantages too: Ambisonics, if transmitted via UHJ, needs a decoder, while 5.1 does not.

of the kind only the British can manage. Despite the handling of the mass of patents by a government department designed to promote British inventions, Ambisonics, like many of Britain's inhabitants, suffered dearly under the hand of Margaret Thatcher. A discussion of the sad story appears in the present writer's previous article on this topic, *Whatever Happened to Ambisonics?* (AudioMedia UK, November 1991), which also covers Ambisonic principles in techniques in more detail.

Despite the disappearance of quad, and the emergence of Dolby surround and home theater, Ambisonics has never gone away: it has simply taken a back seat. Recordings are still being made and issued, including recordings on DVD-Audio and SACD.

A group of internationally-renowned digital audio experts called *Acoustic Renaissance in Audio* (ARA) proposed a flag to indicate that an Ambisonically-based *hierarchical surround encoding scheme* has been employed. The capability came into being in principle with the acceptance of Meridian Lossless Packing (MLP) as the mandated lossless compression scheme for DVD-Audio. As a result, it is theoretically possible for a DVD-Audio disk to include Ambisonic material encoded using a scheme presented by Michael Gerzon and Dr. Geoffrey Barton at the 1992 AES in Vienna. This technology also includes decoding systems that can deal with irregular loudspeaker arrays such as "standard" 5.1 configurations. In fact, the first public demonstration of MLP was from a Compact Disc encoded from Ambisonic B-Format.

Ambisonics in the Age of DVD

Such a scheme would, however, still require an Ambisonic decoder at the receiving end to recover the information and decode it for a 5.1 speaker configuration: a so-called "Vienna" decoder.

Unfortunately, today's DVD players are equipped only to output 5.1: the only decoders they include as standard are for basic MLP (DVD-A players only), AC-3, DTS and (in Europe) MPEG-2. Even if an Ambisonic flag was included in an MLP data stream, an Ambisonic decoder to decode the data would add to the cost of the player, which would rule it out for many manufacturers – and in any event existing players would not be able to provide the signal even to an optional external decoder.

There is, however, a way of transmitting an Ambisonic signal so that no decoder is required at home. Instead, the "Vienna" decoder is installed *in the studio*, and simply decodes original Ambisonic B-Format material to a standard ITU 5.1 speaker layout.

The resulting 5.1 speaker feeds – now generally referred to as "*G-Format*" – can be encoded with Dolby AC-3, DTS, or MPEG-2, and mastered on to a standard multi-channel SACD, DVD-Video disc, or with MLP on DVD-A: no special flags are required.

The Ambisonic decoding parameters are slightly different from those in consumer decoders, so that the resulting G-Format is "reversible" – Ambisonic B-Format can be recovered from G-Format with a suitable device. If the home listener wants more than a conventional 5.1 array, they can recover the B-Format and decode it with an Ambisonic decoder – which can drive a wealth of different loudspeaker arrays and configurations.

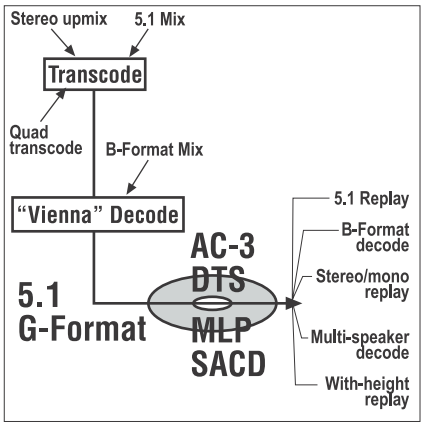

Fig 8. G-Format can be produced directly from a B-Format mix or mic source via a "Vienna" decoder that can decode to ITU 5.1, or indirectly by transcoding other sources or mixes into an intermediate B-Format and then decoding them. Height information may be included. The result can be authored on to any 'discrete" digital transmission medium with 6 (or so) channels, including DVD-V, DVD-A and SACD. No special decoder is required to listen to the results. G-Format is reversible, so audiophiles can recover B-Format and decode it Ambisonically to any speaker array.

Even height information can be encoded into the original mix. The LFE channel is not required in modern music systems: it was designed for low frequency effects such as crashing asteroids and dinosaur footfalls. These do not normally occur in music – and even if they did, modern digital distribution systems offer six channels of full bandwidth. The LFE is thus redundant, and may be used to carry height information.

In fact for some engineers the Center Front channel is *also* unnecessary for music, and one record label, Chesky Records, has released G-Format albums which provide height information for two elevated side speakers, driven from the LFE and CF channels (for example, the DVD-A release of *Swing Live* by Bucky Pizzarelli).

Every DVD disk has to include a stereo mix so that the listener with only a basic system will hear something when they play it. Many SACD discs have a stereo high-density area and sometimes a Red Book CD-compatible layer. During the decoding of an Ambisonic recording to 5.1 for inclusion on the multichannel part of the disk, a 2-channel UHJ mix could be generated to be stored in the disk's stereo space. This would give the listener with a basic system excellent "super stereo" as well as full mono compatibility, and even reasonable horizontal surround with an external Ambisonic decoder, while the listener able to handle the multichannel 5.1 would get excellent surround without a decoder. The sophisticated listener with a special home Ambisonic decoder could enjoy additional benefits of the system such as freedom of speaker placement.

I would like to propose that this system, which I am calling "G+2" (G-Format – ie an Ambisonic signal decoded for ITU 5.1 loudspeaker positions – plus simultaneous 2-channel UHJ) be adopted as the standard format for

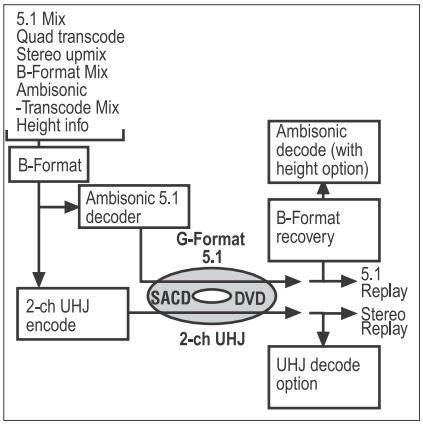

Fig 9. The author's "G+2" proposal includes Ambisonic information pre-decoded for 5.1 ("G-Format"), for direct surround replay, yet providing for B-Format recovery, plus a 2-channel UHJ mix that can be replayed by any player and decoded for surround if desired.

carrying Ambisonic material on DVD and SACD. The multichannel part of the disk would contain the G-Format 5.1 material, while the stereo part (and the Red Book compatible layer on a hybrid multichannel SACD) would contain the 2-channel UHJ mix, generated automatically alongside the 5.1.

Anyone could use Ambisonic recording and/or mixing techniques to create a G+2 disk; any DVD or SACD player, present or future, could play it; and any listener would gain an audio experience beyond the capabilities of other systems.

There is now a need for a new version of the old Ambisonic Mastering System. Although many hit records still use analog mixers, and the original Audio+Design equipment is completely compatible with such systems, today's Ambisonic outboard processors could benefit from additional facilities to take advantage of the possibilities of DVD. Digital processing would be an obvious choice, and in fact a new system could operate entirely in the digital domain. They could equally take the form of software applications or plug-ins to operate in existing computer-based DAW environments. Some software tools and prototype hardware already exist, and it is to be hoped that a digital Ambisonic G+2 production system will be available in the near future.

Richard Elen has worked in publishing, PR and advertising, and as a producer/engineer. He is currently VP of Marketing at Apogee Electronics. 20 years of his Ambisonic writing is available at www.ambisonic.net.