AMBISONICS

SURROUND SOUND IN THE AGE OF DVD

With the advent of home theater systems and DVD, there has been increased interest in surround sound. RICHARD ELEN discusses one surround system that offers features which are difficult, if not impossible, to realize with any other method, and is today poised for a rebirth — Ambisonics.

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.

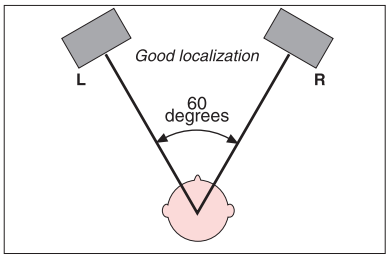

Fig.1: 'Intensity' or 'amplitude' stereo relies on level only to localize sounds between the speakers, if the ears here both speakers, (ie. they are at 60 degrees or less), and at the front, we convert the amplitude differences into phase differences: it works, as long as you sit still.

There are certain limits to stereo replayed over loudspeakers. Most important, if the speakers are further apart than 60 degrees, 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 degrees. In addition, even with the speakers at 60 degrees, there is a tendency for the stereo image to change depending on listener position: there is only a small 'sweet spot' where everything works.

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. We use phase to localize sounds between 150Hz and 1.5kHz and level between 300Hz and 5kHz. Above 2.5kHz, other directional cues are used. When two speakers in front of you are separated by no more than 60 degrees, both ears hear both speakers and LF amplitude differences are converted into phase differences between the ears. However, this effect works 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 degrees. This means that, by definition, traditional quad — and today's 5.1 systems which are their descendants — cannot work. Oops.

The fact that quad relied on recording four 'discrete' channels which used level for localization meant that good inter-channel separation was important. Here, quad

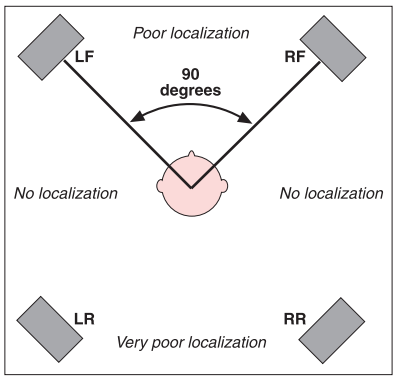

Fig.2: 'Quad' relied on speakers at 90 degrees around the listener, and level-only localization. The spacing is too great and sounds are drawn into the speakers. Localization is poor at the front, very poor at the rear, and non-existent elsewhere.

ran into a second problem: how to transmit four channels of information over then-available two-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 ➤

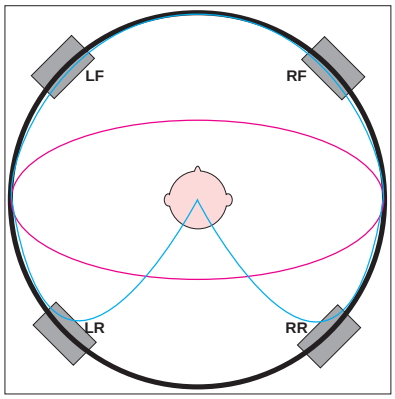

Fig.3: Matrix quad tried to get the four original channels into two and back again, which is impossible. A sound panned around the room in a circle would not play back as one: it would jump between the speakers, and traverse a flat ellipse in SQ (red) or a cardioid in QS (the foundation of Dolby Surround).

AMBISONICS

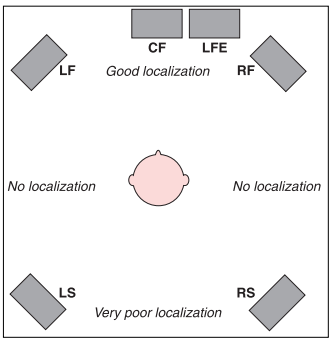

Fig.4: Today's '5.1' systems are based on 'discrete quad', with a center-front channel to fill the hole between the front speakers: fine for dialog, but difficult to integrate in music mixing. The localization elsewhere is as bad as before.

➤ jump from one speaker to another, and come in from the edge of the circle.

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. 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 Dolby Surround, adding a second surround channel (the Low Frequency Effects — 'LFE' — channel is the '0.1').

There remained the problem of carrying those — now six — 'discrete' channels from recording studio to the listener at home.

The Promise Of DVD

For the first time, we have such a medium: a multi-channel fully-digital sound carrier in the form of DVD — the Digital Versatile Disk — in its several incarnations. This technology not only offers the possibility of carrying multi-channel film soundtracks to a consumer audience: it also offers the potential for a multi-channel 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. But, if the goal of surround sound is accurate localization, 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.

However, there is another way of creating 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 over 20 years; and, to date, more original album releases have 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

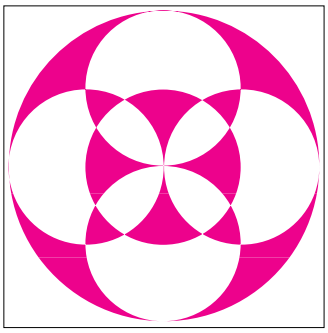

Fig.5: The Ambisonic logo indicates the components of a B-format signal: the outer circle represents the omnidirectional feed, W; the upright figure-eight is front minus back (X); the transverse figure-eight is left minus right (Y); while the center circle is the up minus down (Z) figure-eight seen from above.

manufactured in the UK by Soundfield Research) or with special console panpots and localization controls, or both. Ambisonics production equipment generates a four-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 part of surround-sound practice, despite the fact that it adds almost as much to the realism of a surround system as rear speakers do.

For replay, the B-Format signal is fed to a decoder which derives a minimum of four loudspeaker feeds. Instead of four independent signals, the speakers in an Ambisonics 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

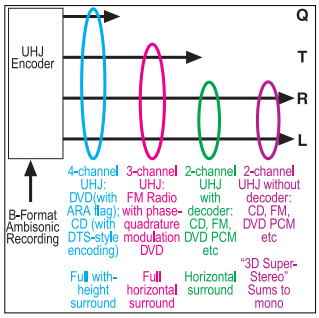

Fig.6: UHJ is a hierarchical surround encoding scheme which allows a surround signal to be experienced at the highest level according to the number of transmission channels available. A single mix satisfies all the available modes without compromise.

would be disappointing: the 'sweet spot' would, at best, be the size of a football. Ambisonics, on the other hand, uses all 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. Four-channel UHJ carries the same information as a four-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 channel can be bandwidth limited: this is referred to as a '2.5-channel' system. 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 two-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. Two-channel UHJ is a powerful '3-D stereo' tool, at least as effective as more recent two-channel surround techniques used today for computer and multimedia applications — with the added benefit that, with a decoder, true surround is available. You can even sum left and right and listen to a mono signal that's more compatible than conventional stereo. In short, to borrow a term from modern website design, Ambisonics 'degrades' almost perfectly.

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 has been recorded Ambisonically.

Ambisonics	5.1
Only requires four channels.	Needs six channels.
Includes height information.	No height information.
Only one mix required for DVD.	At least two mixes required for DVD.
Lower data density: lossless compression can be used.	Higher data density: lossy compression needed.
Imaging equally accurate in all directions.	Imaging only accurate across front stage.
Imaging possible between the speakers.	Almost non-existent inter-speaker imaging.
Sounds can be placed within the array.	Sounds appear 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 small area: image changes as you move

Fig.7: The characteristics of Ambisonics and 5.1 surround compared. It should also be noted that, while traditional Ambisonics needs a decoder, 5.1 does not. However, there is a way around this.

Ambisonics 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 two- or three-channel UHJ. The other units utilized a number of different approaches to localizing signals in an Ambisonics 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 Ambisonics panners. The 'Transcoder' turned B-Format into two-channel UHJ for release, and also allowed a quick Ambisonics 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 two-channel UHJ mix directly.

The Ambisonics 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*), Ambisonically-mixed CDs on the Collins Classics label, and Paul McCartney's *Liverpool Oratorio*, 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 of the kind only the English can manage. Despite the handling of the mass of patents by a government department designed to promote British inventions, sufficient marketing expertise and money was never available to give Ambisonics the push it needed. A discussion of the sad story appears in the present writer's previous article on this topic, 'Whatever Happened to Ambisonics?' (*Audio Media* UK, November 1991), which also covers Ambisonics principles and 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, and now Ambisonics is gearing up for the challenge of a multi-channel digital consumer medium — DVD. A proposal from a group of internationally-renowned digital audio experts called Acoustic Renaissance in Audio (ARA), for example, proposes a flag to indicate that a hierarchical surround encoding scheme — such as UHJ — has been employed, although this is just one area where Ambisonics is once again coming to the fore.

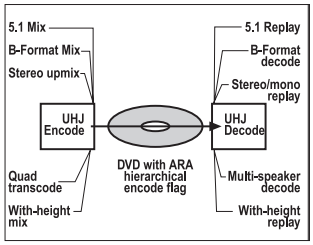

Fig.8: UHJ is an ideal medium for DVD. Any kind of surround mix can be encoded into UHJ at one end and any type of surround can be decoded at the other. However, it needs a special flag (proposed by the ARA) to tell future players to output it, and current players won't do so.

Ambisonics In The Age Of DVD

UHJ is an ideal medium for digital transmission of surround information, whatever the surround system in use. Any surround system — including 5.1 — can be encoded successfully into UHJ, and a UHJ signal can be decoded 100 percent successfully into a 5.1 speaker array. UHJ also requires only three channels to convey a high-definition horizontal surround signal while 5.1 requires six. With four channels, you can carry height too.

With a hierarchical encoding flag and gear to interpret it, you could make a recording in 5.1, transmit it via three channels of UHJ, and decode it into 5.1 at the other end. Or make an Ambisonics recording and decode it to a 5.1 array, or make a 5.1 recording and decode it with an Ambisonics decoder, with its unique and innately flexible speaker positioning capabilities. Or, of course, you could work entirely in Ambisonics. And, as B-Format can be recovered from three- or four-channel UHJ (with or without the height), you could use B-Format at either end if you so desire. UHJ would provide a seamless integration of all conceivable surround systems — and, because of the fewer channels, they would not even need lossy compression.

There would be another benefit of mixing Ambisonically. One thing that really scares today's engineers is the thought of having to do several mixes — including stereo and surround versions — of their work. If you get your sources panned into position by listening Ambisonically, and then monitor in the lowest-level format that your listeners will be likely to contend with — stereo, for example — as you carry out the final mix, you will end up not only with a good stereo mix but an excellent Ambisonic surround mix also. And, with UHJ, you simply master the one mix — and people hear a successful balance, however many channels they have access to, and whether they have a decoder or not.

The drawback here is that, to get the best out of UHJ, you need a decoder. Unfortunately, today's DVD players are equipped only to output 5.1: the only decoders they include as standard are for AC-3 and MPEG-1. Even if a UHJ flag was included in the data stream, a UHJ decoder 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 UHJ signal even to an optional external decoder.

There is, however, a way of transmitting an Ambisonics signal so that no decoder is required: simply decode the original Ambisonics recording for a standard 5.1 speaker layout and record the result.

5.1 decoded speaker feeds — now generally referred to as 'G-Format' — could be encoded with AC-3 (in 60Hz territories, generally those using the NTSC TV standard) or preferably multichannel MPEG-2 (but only in 50Hz or PAL TV

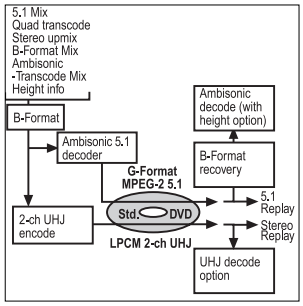

Fig.9: The author's 'G+2' proposal includes Ambisonics information pre-decoded for 5.1 ('G-Format') for direct surround replay, plus a two-channel UHJ mix that can be replayed by any player and decoded for surround if desired.

territories at this time), suffering little or no damage, and placed onto a standard 'video' DVD — with no special flags required. There would either need to be two disks (NTSC and PAL) or a double-sided disk (this is not uncommon) with PAL on one side and NTSC on the other.

G-Format feeds contain enough data in them to enable full horizontal B-Format to be recovered for those who want the added flexibility of true Ambisonics decode. Even height information could be encoded into the original mix and decoded with a Periphonic decoder. G-Format may, in the end, offer almost as many possibilities as UHJ… with one exception: the disk would still need to carry multiple mixes. But, as we have seen, the obligatory stereo mix could be developed in the studio automatically during the Ambisonics mixing process.

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. During the decoding of an Ambisonics recording to 5.1 for the multi-channel datastream on the disk, a two-channel UHJ mix could be generated automatically to be stored in linear PCM format 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 Ambisonics decoder, while the listener able to handle the multi-channel 5.1 would get excellent surround without a decoder. The sophisticated listener with a special Ambisonics decoder could even hear any height information that may be provided as well as gaining other 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 5.1 pre-decoded Ambisonics feeds recorded in AC-3 or multichannel MPEG-2, plus simultaneous two-channel UHJ in LPCM) be adopted as the standard format for carrying Ambisonics material on DVD, whatever options may finally be accepted for DVD-audio. Anyone could use Ambisonics recording and/or mixing techniques to create a G+2 disk; any DVD 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 Ambisonics Mastering System. Although many hit records still use analog mixers, and the original Audio + Design equipment, now manufactured by Cepiar Ltd in the UK, is completely compatible with such systems, today's Ambisonics outboard processor 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.

Research is already advanced on digital studio production equipment, with the features required to produce G-Format and UHJ-encoded material, and it is likely that the equipment required to produce G+2 disks will be available in the near future. It seems likely that Ambisonics may provide answers to a number of difficult questions, in addition to superior performance, as surround sound becomes the medium of choice for recorded audio. ❑

Richard Elen has worked in publishing, PR and advertising, and as a producer/engineer. He is currently VP of Marketing at Apogee Electronics. You can contact the author, Richard Elen, at relen@igc.org. A collection of his articles on Ambisonics, and links to other Ambisonics sites can be found at www.audiomedia.com/ambisonics.htm