

feature

surround sound
explained

you are surrounded

SURROUND SOUND EXPLAINED * PART 3

Though a commercial failure in the 1970s, the Ambisonics surround sound system was technically flawless, and survives to this day -- indeed, today it is more relevant than ever. Hugh Robjohns unravels the mysteries of the B-format...

Before turning my attention to modern 5.1 surround systems and how they work, I'd like to spend one more instalment of this series explaining another form of surround sound that pre-dates 5.1 -- Ambisonics. This system is finally beginning to enjoy recognition for its undoubted strengths, proving it really was way ahead of its time.

Ambisonics was conceived in the late 1960s as a complete recording and reproduction system capable of recreating accurate three-dimensional sound stages from original recordings. The format was developed using complex mathematics and psychoacoustics, all based on the original work on coincident stereo led by Alan Blumlein in the early 1930s (see the 'History' box elsewhere in this article for more on the system's development).

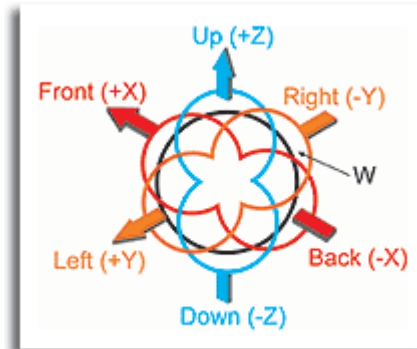


Ambisonics failed to attract the significant attention of either manufacturers or the public in the '70s, principally because too many companies had had their fingers burned by the quadraphonic fiasco (see part 1 of this series, in *SOS* August 2001). Fortunately, the numerous strengths of the system were recognised by enough professionals to keep it alive, if only on the periphery of the industry, and today it is beginning to receive the attention it truly deserves. Even though the technology is over 30 years old, it stands up to scrutiny, just like Blumlein's coincident microphone techniques.

Technology

Although some post-production equipment exists to allow Ambisonics surround images to be created artificially from multitracked or close-miked sources, most Ambisonic material is captured by using a special microphone array. The best known of these is undoubtedly the Soundfield Microphone and its associated controller (see right), but there are other incarnations, including a three-mic array designed by Dr Jon Halliday, Research Director at Nimbus Records.

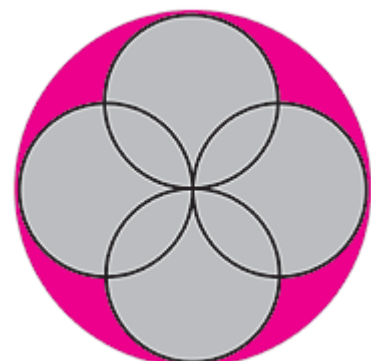
The purist single-point recording side of Ambisonics can be thought of as using an overblown middle-and-side microphone technique (for more on this subject, check out the article on the subject in SOS February 1997, or surf to www.sound-on-sound.com/sos/1997_articles/feb97/stereomiking.html). However, the cleverness of Ambisonics really manifests itself in the decoding -- converting the microphone's output signals into a form suitable to drive a number of loudspeakers. It is the complexity (and therefore the cost) of decoding which played a large part in the failure of Ambisonics to gain a hold in domestic hi-fi equipment. Very few companies have manufactured decoders for the system, whether professional or domestic, over the last 30 years, although with the increasing prevalence of digital signal processing in consumer equipment, Ambisonic decoding options are beginning to become more apparent, for example in high-end hi-fi systems from the likes of Meridian Audio and Briston.



The Ambisonic concept is based on the idea of capturing the sound impinging on a single point in space from any and every direction. This is, of course, completely impossible -- a multi-directional microphone cannot be made that small. However, if incident sound can be measured across the surface of a sphere, it is a mathematical exercise to calculate what *would* be detected at an infinitely small point at the centre of that sphere. In practical terms, the incident sound can be measured by a combination of microphone capsules mounted in coincidence -- or as near coincidence as is physically possible.

The principle of an Ambisonic microphone -- whether a dedicated single-point mic like the Soundfield, or a combination of coincident conventional mics, such as the Nimbus-Halliday array -- relies upon a combination of orthogonal bi-polar patterns with an omnidirectional, pressure sensitive capsule. The output of the omni capsule is referred to as the 'W' signal, and provides information about the overall amplitude of sound impinging on the microphone array. The bi-polar or figure-of-eight capsules provide the directional information -- that is, their outputs can be used to determine the direction from which each element of sound arrives. One of these capsules points front-back (providing the 'X' signal), another left-right ('Y') and the third up-down ('Z'). These four signals convey everything there is to know about the amplitude and direction of acoustic signals arriving at the microphone (see diagram, left).

The four signals together are known as B-format signals, and if recorded on four discrete tracks can provide a record of the original sound, captured with total three-dimensional accuracy. By replaying this B-format material through a suitable decoder, a full 'periphonic' recreation of the Ambisonic sound stage is possible (ie. one that includes both horizontal *and* vertical imaging). Although the discrete-channel surround systems currently in vogue only provide horizontal imaging over four, five, six, seven, or eight channels (ie. in one plane only), Ambisonics has always been capable of true three-dimensional imaging -- something which is only now being suggested for discrete-channel surround systems with the 10.2 system recently proposed by Tom Holman (more on this next month).



It is also possible to take B-format signals to post-production and manipulate them to obtain the most aesthetic and appropriate result for a given format. Alternatively, the signals can be combined immediately after capture to allow storage on fewer tracks, albeit with correspondingly less flexibility in

post-production -- more on this in a moment.

By combining the W, X, Y and Z signals in various ways, it is possible to recreate the effect of any conventional microphone polar pattern (from omni, through cardioid, hypercardioid and figure-of-eight), pointed in any direction. This works in exactly the same way as a conventional stereo middle-and-side microphone, only in three dimensions instead of just one (left-right). With the right combinations of W, X, Y and Z signals, it is therefore possible to replicate the signals that would have been obtained from, say, a stereo pair of crossed cardioids with a 110-degree mutual angle, dipped 25 degrees from horizontal.

To see how this works, consider combining the output of the X and Y signals in equal amplitude. The result would be a new figure-of-eight pattern with its positive-polarity lobe pointing 45 degrees to the front-left. Reducing the amount of X signal slightly will twist the pattern further to left, say to 55 degrees. Now mix in a little minus-Z (ie. phase-inverted Z) which will have the effect of tilting the figure-of-eight pattern downwards slightly. If this combination is then combined with an equal proportion of the omnidirectional W pattern, a cardioid response is created, representing the left microphone in the notional crossed pair described above. Repeating the whole thing again, substituting minus-Y for Y, will create a cardioid pattern facing 55 degrees right and tilted slightly downwards. Hey presto, a totally accurate representation of a stereo crossed cardioid array has been produced.

In Practice

That's the theory -- how are Ambisonic signals captured in practice? The Nimbus-Halliday microphone arrangement (see diagram, right) dispenses with vertical directionality, and employs just three separate microphones mounted such that their capsules are as near-coincident as is practical. A pair of figure-of-eight capsules provide the X and Y signals (looking front-back and left-right), while an omni microphone adds the W omnidirectional component. These signals are then recorded and later combined, or, more typically, combined and recorded to create the required end result on a two-channel medium (which can then be mastered onto CD). This material is fully compatible with replay over conventional stereo (or even mono) equipment, but also allows a horizontal surround image to be recreated with a suitable decoder.

The best-known Ambisonic microphone, the Soundfield, uses four sub-cardioid capsules mounted in such a way that they effectively form the faces of a regular tetrahedron (see photo below). Their outputs are collectively known as A-format signals, and, predictably, require further processing to derive B-format signals.

This physical arrangement of

Ambisonics -- A Little History

Ambisonics -- as a concept -- began over 30 years ago as the brainchild of a small group of academics including the late Michael Gerzon. Their efforts were supported and backed by the UK government's National Research and Development Council (NRDC) with assistance along the way from such august bodies as the BBC. When the NRDC eventually folded in a Government shake-up, the patents were sold to the record label (and later CD manufacturing company) Nimbus.

Back in 1978, the pro-audio manufacturer Calrec started manufacturing the Soundfield microphone to originate B-format recordings (yes, the mic was called that long before the Soundfield company was created!). When Calrec were acquired by AMS Industries in 1987, a graphical representation of the W, X and Y patterns of the Soundfield mic (also used to identify commercial UHJ recordings) became the company's new logo -- and despite no longer having any connection with Ambisonics, AMS Neve rather bizarrely retain a variation on this distinctive graphic as their logo to this day.


AMS were a resolutely digital company and the Soundfield microphone was eventually acquired from them by Ken Giles of Drawmer Distribution, who set up the current Soundfield company. The fundamental design of this unique microphone has not changed since its first incarnations, although it has benefited from incremental

capsules was chosen as the most pragmatic means of obtaining a representative spherical pickup with the minimum number of capsules. Part of the signal processing applied to the A-format signals is inter-capsule time correction to compensate for the physical

separation between the four capsules, thereby replicating the signal which would be obtained at a point at the centre of the array. This electronic correction ensures the B-format signals are accurately time-coincident for all frequencies up to about 15kHz, improving the spatial accuracy of the system way beyond that available from any other multi-microphone array.

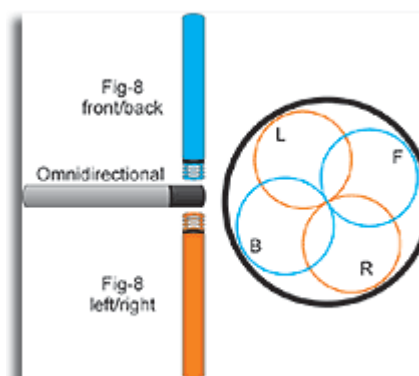
advances in electronics to become quieter and more precise and neutral than ever.

Despite a rough ride, both the Ambisonics concept and the Soundfield microphone have survived the test of time... there must be something in it after all!

 www.ambisonic.net

Another important aspect of the Soundfield design is that the B-format signals it produces are corrected to provide nearly-flat frequency responses for sounds arriving from all directions. This is simply not achievable with conventional microphones, which always exhibit some change in frequency response with direction of sound incidence. In fact, this particular characteristic of the Soundfield microphone has helped it to acquire an enviable reputation as a high-quality stereo microphone in its own right, uniquely providing flat responses both on and off axis, with none of the off-axis coloration typical of other microphones. Furthermore, imaging accuracy is greater than with any other coincident microphone, because the polar responses remain uniform with frequency, and the two-channel outputs are very accurately time-aligned.

The Soundfield microphone has to be used with a matching control unit, and there are currently three versions available. The rackmount SPS422 provides only a stereo output, but allows the effective mutual angle and polar pattern of the nominal coincident pair it is replicating to be continuously adjusted. The ST250 provides the same facilities but in a portable package, and with the ability to record the four-channel B-format signals if required. However, it is the large rackmounting Soundfield Mark V controller which provides the full range of control possibilities. Not only are the four B-format signals made available for output at fully balanced professional levels, the unit can also accept a B-format input for post-production. This signal can be manipulated to produce a stereo-compatible signal providing horizontal surround sound through a suitable decoder. As on the SPS422 and ST250, controls allow the user to manipulate the effective polar patterns and mutual angles of the nominal stereo microphone outputs. The Mark V provides facilities to steer the forward axis of this notional array in a full 360-degree horizontal circle and to elevate it over a ± 45 -degree arc, and also features a Dominance or Perspective control, which simulates moving the microphone closer to or further from the sound sources. These controls may be manipulated to produce the desired stereo-compatible output live, or in post-production from B-format signals.



One occasion which left a big impression on me was the post-production, one year after the event, of a B-format recording made of Kiri Te Kanewa singing at the wedding of Prince Charles and Princess Diana. A Soundfield mic had been suspended up in the dome of St Paul's Cathedral in London, and it was possible, through the manipulation of the decoder's controls, to literally point and zoom the 'virtual' microphone anywhere one desired. It was quite uncanny to be able to focus the sound on almost any instrument in the orchestra, or hear the audience rustling, or allow Kiri Te Kanewa's voice to dominate -- and all this long

after the original performance! Despite being rather flawed, the best comparison I can muster is of an 'audio camera' which can be zoomed, focused, panned and tilted to fine-tune the overall sound picture.

Four Channels Into Two

Thirty years ago, reliable phase-coherent four-channel consumer storage media were not particularly common, and so B-format signals were thought not to be suitable for consumers. A hierarchical collection of transmission or storage formats was therefore defined, using (rarely) four, three, 2.5, or, most commonly, two channels (the half-channel in the 2.5 format refers to the use of a possible limited-bandwidth channel). Collectively, these codings are known as UHJ, but sometimes they are also referred to as C-format signals. With modern multi-channel digital recording formats, there is no reason why the B-format signals should not be used in their native form, although some encoding would still be necessary for two-channel formats.

The four-channel version of UHJ is capable of full three-dimensional periphonic sound reproduction but, as information has to be discarded for the other formats, the latter are only capable of horizontal surround reproduction, with decreasing resolution corresponding to the number of channels you have available.



A two-channel UHJ format carries signals called, not surprisingly, 'L' and 'R' -- left and right stereo-compatible signals. These are derived from the W, X and Y signals of the B format, the three-into-two problem being solved with an elaborate, carefully designed matrix encoder, which optimises both stereo and mono compatibility, as well as minimising any ambiguity during the process of decoding into loudspeaker signals.

Of course, as with the Dolby Surround system described last month, you can't convey four channels over two without losing something. Dolby lost channel separation, and attempted to mask the problem using active steering -- with notable success most of the time. Ambisonics loses resolution in its frontal imaging accuracy -- which typically manifests itself as a slight blurring or phasiness. Most decoders have a facility which allows the listener to compensate for this, increasing the frontal imaging accuracy at the expense of side and rearward precision -- which is usually far less critical anyway.

If an additional channel is available, the matrix problem is eased, as the additional information can be conveyed separately, allowing the resolution and accuracy to be increased throughout the surround sound stage. This third channel carries the so-called 'T' signal and a full-bandwidth third channel will allow all three of the original W, X and Y signals to be recovered completely. A limited-bandwidth third channel inevitably conveys less 'T' information, and in this case, the W, X and Y signals cannot be recreated with total precision, although a worthwhile improvement in horizontal imaging accuracy can still be obtained.

A four-channel UHJ system incorporates a 'Q' signal in addition to L, R and T, which conveys height information (as with the Z element of the B-format signal). Whilst this complete family of UHJ formats have been fully specified and documented, I have only come across commercial two-channel UHJ media, and professional four-channel B-format master recordings.

Ambisonic Advantages

Although it might not seem an obvious choice at first, there are a number of advantages to using

Decoding

As with Dolby Surround, the really clever part of the Ambisonics process is in the decoding, which is a two-stage affair. The first part is to reconstruct an approximation of the original B-format signals from the UHJ format. Clearly, the more channels available to the UHJ format, the more precise this approximation will be.

The second stage involves the application of particular EQ curves to these signals, dependent on their relative directions. In effect, this filtering mimics the way our own hearing uses different localisation techniques for sounds in different frequency ranges -- level differences for high-frequency sounds, and phase or time-of-arrival differences for low-frequency sounds. Such psychoacoustic filtering techniques are often referred to as 'head-related transfer functions' or HRTFs, and it is in their application that the Ambisonics system stands or falls.

All discrete-channel surround systems use panning to locate sound sources -- in other words, a mono sound source is typically routed to more than one channel via a pan pot, and those channels are replayed directly over corresponding loudspeakers. In contrast, the Ambisonics system starts off by encoding the precise direction of the original sound source (relative to the microphone's position) in the B-format signal. The difference is that this directional information is then used to condition the signals specifically to suit their reproduction over the particular loudspeakers installed in the listening environment. In this way, the original sound stage is recreated with precision and accuracy -- and sound source directions remain stable over a surprisingly wide listening area, rather than merely in a small 'sweet spot'.

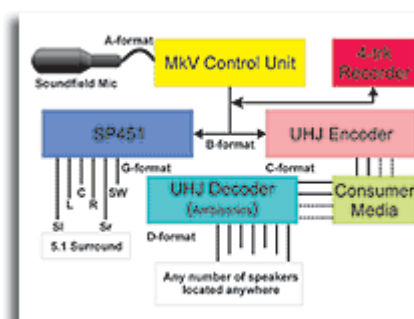
No other surround systems employ this kind of sophisticated psychoacoustic processing to optimise the signals prior to reproduction over user-determined speaker positions. Dolby Surround uses very simplistic low-pass filtering and delays to enhance the illusion of the rear-channel effects, but it pales beside the complexity of the Ambisonic decoder. The encoded and equalised signals, prepared for the specific speaker arrangement to which the decoder is connected, are referred to as 'D-format' signals. Whereas the A-, B- and C-format signals are universal, the D-format is only relevant to a specific loudspeaker and room configuration.

The way in which the reconstructed B-format signals are equalised is dependent not only on their original directionality, but also on the position and number of available loudspeakers in the listening environment, since the latter will determine where the signals can be regenerated. The decoder has therefore to be programmed in some way with information concerning the number and location of loudspeakers, as well as the approximate position of the listener(s) in relation to them.

the B format for the production of surround material. For a start, an Ambisonics system requires just four full-bandwidth channels, not five and a little bit (or seven, or more...!)

Secondly, the process of 'folding-down' multi-channel mixes to stereo and mono is much easier than with surround signals from spaced microphone arrangements, because B-format signals' inherent time-coincidence means that there are no phase problems when the signals are combined.

Thirdly, the nature of the B format is that it is output-independent -- in other words, a B-format master can be decoded to suit any present or future loudspeaker configuration. The format only stores information about the original sound localisation, not its reproduction. Consequently, if the recently proposed 12-channel 10.2 loudspeaker format becomes fashionable in a few years, a B-format master would only have to be passed through a suitable decoder to recondition the material.



Ambisonic decoders vary in their sophistication but there is normally some sort of facility to adjust the 'sharpness' of imaging in the 'preferred direction' (usually the front). Many of the more sophisticated decoders also incorporate a 'Super-Stereo' mode which uses much of the Ambisonics psychoacoustic decoding principles on stereo material which is *not* UHJ-encoded. It works well on multitrack studio material or coincident-pair stereo recordings, but can be extremely unpredictable on recordings with timing differences between channels (eg. recordings made with spaced-pair or Decca-Tree mic arrays).

Ambisonics & 5.1


Since the 5.1 surround sound format is defined as a loudspeaker arrangement, there is no reason why surround material should not be captured with a Soundfield mic (or other Ambisonics-compatible mic array) and subsequently decoded into a form suitable for 5.1 replay.

That is precisely what Soundfield have to offer with their new SP451 Surround Processor. This 1U rackmount unit is, in effect, a dedicated decoder which accepts B-format input signals from a Soundfield mic system (Mark V or ST250, for example) or a four-channel B-format recording, and effectively produces the appropriate D-format signals corresponding to a standardised 5.1 loudspeaker arrangement. Incidentally, it seems that people are beginning to describe this kind of output (ie. B format decoded to 5.1-compatible signals) as the G format.

The advantage of the G format is that Ambisonically recorded material can be played back within a 5.1 production without the need for a domestic Ambisonics decoder. However, the disadvantage over the original B format is the inability to decode the signal to accommodate any number of loudspeakers (in any position) -- the speaker positions and numbers are defined by the nature of the 5.1 system -- and there is no ability to resolve height information.

Whilst it is unlikely that Ambisonics will ever become a major release format in its own right -- there are far too many vested interests in other discrete-channel formats -- it is certainly far from outdated, and is actually showing every sign of becoming a significant and very flexible element in the origination and post-production of surround material. Not bad for a 30-something!

Next Month

Next month, I will explain discrete-channel surround, focusing on 5.1 surround and DVD systems, as these seem to offer the greatest commercial potential at present. 

Glossary

http://www.sospubs.co.uk/sos/regular_htm/glossary.htm

SOUND ON SOUND

Europe's No1 Hi-Tech Music Recording Magazine Sound On Sound

Go To SOS
Website 
Homepage 

Media House, Trafalgar Way, Bar Hill, Cambridge CB3
8SQ, UK.

Telephone: +44 (0)1954 789888 Fax: +44 (0)1954
789895

Email: info@sospubs.co.uk Website:

www.sospubs.co.uk

© 2001 Sound On Sound Limited. The contents of this article are subject to worldwide copyright protection and reproduction in whole or part, whether mechanical or electronic, is expressly forbidden without the prior written consent of the Publishers. Great care has been taken to ensure accuracy in the preparation of this article but neither Sound On Sound Limited nor the Editor can be held responsible for its contents. The views expressed are those of the contributors and not necessarily those of the Publishers or Editor.
