

# Getting Ambisonics Around

Richard Elen discusses how he and Peter Carbines\* are bringing old surround mixes – and new ones – to life through the medium of ‘G-Format’ – Ambisonic recordings decoded before they are issued.



## Goals

The purpose of the experiments described in this article was to develop and evaluate a process for generating widely-distributable, low cost planar surround recordings rendered from Ambisonic (2-channel UHJ and planar B-Format) source material – without the traditional challenge of requiring the listener to have access to an Ambisonic decoder – by delivering Ambisonically-decoded speaker feeds (commonly referred to as “G-Format”) using a commonly-available surround distribution format.

Our primary original involvement with Ambisonics was to develop technology for creating Ambisonic mixes from multitrack sources, and as a result we were particularly concerned that our method would work well for such material, requiring as it does a generally more precise level of localisation than may be expected of a natural recording of an acoustic event. Indeed, the work described here is part of a larger project to establish accessible and effective methods of producing and distributing Ambisonic recordings, particularly of mainstream, widely popular material, which by definition would be mixed from conventional multitrack sources.

Due to its accessibility and affordability, we chose the DTS CD as a distribution medium. Although it is lossy compared to, for example, MLP/DVD-A technology, DTS CD is a very popular medium and is playable by virtually anyone with a 5.1 system. In addition, the DTS encoding process has previously been shown to carry Ambisonic localisation information effectively. It has also been used experimentally for this purpose before, though not to our knowledge with a file-based “off-line” workflow.

It is appreciated that presenting material in this way is subject to a number of compromises, notably:

- Degradation of audio quality and possibly localisation accuracy as a result of employing lossy audio compression
- Degradation of localisation accuracy as a result of lack of congruence between the listener’s speaker array and the decode target array.

## Challenges of early Ambisonic mixing

An early challenge in the recording of Ambisonic surround music material was, interestingly enough, the availability of multitrack machines. For a ‘full blown’ recording using the UK-invented surround system, at least three channels were required to record horizontal surround, or four if it was intended to include height information (a monophonic sum signal, plus difference signals representing figure-of-eights pointing left, forward and up) in the Ambisonic ‘studio format’, properly called ‘B-Format’. For users of the Soundfield microphone, it is possible that there were multitrack machines available, but in our case – performing Ambisonic mixes from multitrack – the studio multitrack machine was already in use (replaying the source material!), and as a result we had to encode the Ambisonic mix directly to 2-channel UHJ and record on a 2-track machine. The advent of Compact Disc, and the benefits of originating masters digitally, once again meant that there were only 2-channel mastering machines available, so although we had the ability to create mixes in B-Format, we were still unable to store it: it was only used to link the outboard modules used in localising the sources. The final mix was encoded to UHJ before it ever left the processors.

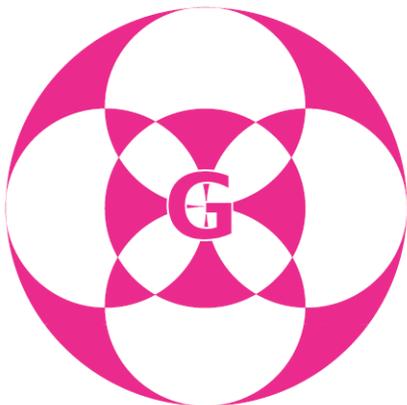
In addition, when it came to mixing multitrack to Ambisonics, the fact that the release format was 2-channel, and storing the master in any format requiring *more than* two channels was difficult if not impossible, it was easy to be lazy. A simple method of mixing to UHJ was to use the Transcoder, which takes two stereo stages, front and rear, and encodes them direct to 2-channel UHJ; any B-Format is inaccessible, inside the box. The

Transcoder also accepted a B-Format input, being based on a UHJ encoder, and both could be used simultaneously. As a result it was easy to use the Transcoder for static localisation (ie sources that did not move during the mix) by defining two pairs of groups for the front and rear stages, rather than using the more complex-to-set-up B-Format Converter, which used the console panpots to pan across quadrants and required four groups plus an aux send, but produced B-Format. Meanwhile we would use the Pan-Rotate unit (which had a B-Format output) for dynamic localisation (ie sources that *did* move during the mix, if any) and to position sounds in locations that were not covered by the Transcoder (the rear stage would only go up to about 150° width) or needed to be brought in from the periphery (using the “radius vector” controls on the Pan-Rotate). Working in this way – which was my standard method for years of mixing Ambisonically – left no possibility whatsoever of deriving a B-Format output, even if we had had something to record it on. More information on how these mixes were performed can be found in the ‘Articles’ section of [www.ambisonic.net](http://www.ambisonic.net).

Of course today, we curse this combination of factors and wish that we had been able to record master mixes in B-Format, so that later we could release them in more effective formats than 2-channel UHJ, now that they are available. 2-channel UHJ, while it did extremely well for what it was, and knocked spots off competing matrix ‘quad’ formats, did have some problems. One was a slight weirdness at the exact centre of the listening array, where your head would seem to pass through some kind of small node in which localisation shifted rapidly. Another was a tendency for the ‘South’ (centre rear) position to be a little poorly localised.

We would minimise these issues with 2-channel UHJ by monitoring via the decoder during the mix, so that at least we would know what the listener would experience. Our standard speaker layout was the minimum needed for replaying planar (horizontal-only) Ambisonics, namely four matched nearfield monitors, set in a rectangle (two in front in the usual place on the console, and two behind on mic stands) generally with approximately square aspect ratio – certainly always within the standard range of side ratios of 1:2 to 2:1.

### Decoding in the studio



Several years ago, the idea arose of what we now refer to as ‘G-Format’, namely the decoding of B-format source material into speaker feeds that could be issued on a digital multichannel

distribution medium such as DVD-A, DVD-V or DTS CD. The present authors, among others, raised this as a possible solution to the ‘chicken and egg’ situation brought about by Ambisonics requiring a special decoder that few home listeners possessed. Instead of requiring the listener to decode Ambisonic recordings, we would do the decoding *in the studio* and place the resulting multichannel loudspeaker feeds, one channel mapped to one speaker, on a multichannel disc in the usual way. The resulting disc would embody the localisation accuracy of Ambisonic technology – whether based on a Soundfield microphone or on an Ambisonic mix – including accurate side and rear imaging as well as front, without requiring the listener at home to invest in anything other than a standard surround-sound system able to decode common surround formats such as Dolby Digital, DTS or MLP.

This would allow Ambisonics adherents who were so inclined to focus on the development of studio technologies that would provide the engineer and producer with considerably more powerful and flexible surround audio tools than are currently available, without having to worry whether or not anyone would be able to experience the end results. The studio engineer could use Ambisonics in the studio, yet release the results in conventional multichannel formats. In addition, storing the surround master in B-Format would mean that the mix could be ‘rendered’ into any surround format current at the time – 5.1, 7.1, 10.2, or whatever – simply by running it through an appropriately-configured decoder, without a new mix being required.

The upside of the G-Format approach was that you could use a very high quality, studio decoder and obtain the best possible results as far as decode accuracy was concerned. The downside, however, was that one of the primary features of Ambisonics – the ability to render the Soundfield to the highest resolution (ie number and position of loudspeakers) of the user’s system, analogous to the way in which a PostScript file is rendered to resolution of a printer, was lost. In addition, the quality of Soundfield recreation was likely to be degraded as a result of any difference between the target layout for which the studio decode had been performed and the listener’s actual speaker array.

It should be noted, however, that G-Format should be regarded, at worst, as a temporary measure, due to the development of hierarchical audio rendering technologies (eg by Peter Craven) which allow a multichannel Ambisonic recording to be conveyed on a future high definition disc such that they can be recovered to the full ability of the playback system if it includes appropriate decode technology, but will offer a default level of performance (eg 5.1) on a system which does not offer additional facilities. However the current disagreement on emerging disc formats suggests that this capability may take a while to arrive in terms of mass markets, if at all, and as a result it is quite possible that G-Format may offer more than simply a temporary solution to the problem of getting Ambisonically-created recordings out to the general public.

The success or otherwise of G-Format rests, therefore, on being able to select a decode target speaker array that will be *least unlike* the average system configuration of a listener’s speaker array; and upon the robustness of an Ambisonic

decode in circumstances where a likely degree of such error exists. Anecdotal evidence – and some research, notably from Angelo Farina – seems to suggest that in fact the majority of listeners do not, in fact, have anything much like an ITU 5.1 array. In fact, they – in common with a great many recording studios mixing music surround – tend to have a rectangle with a centre front at the middle of the front side. This is in fact a much more favourable target array for a straightforward Ambisonic decoder than an ITU arrangement, which requires special configuration (via distance compensation at the least, or preferably a decode technology, generally referred to as a ‘Vienna’ decoder after its delineation by Gerzon and Barton in an AES paper presented in Vienna in the early 1990s, which can cope with asymmetrical arrays). It would also appear that the average listener tends to sit towards the rear of their essentially rectangular array.

G-Format, as has been noted, was originally described as a set of speaker feeds decoded from an Ambisonic B-Format source. In fact the concept of decoding in ‘the studio’ in this way appears to have originally been conceived by Dr Geoffrey Barton. Anecdote suggests that when he presented the idea to Michael Gerzon, the latter was not entirely in favour of the idea (no doubt due to the challenges of this approach outlined above) and dubbed it, humorously and with light disparagement, ‘Geoffrey’s Format’, which may or may not be the origin of the name.

However, with the paucity of available B-Format recordings to process to speaker feeds in this way, it was inevitable that the decoding of 2-channel UHJ to speaker feeds was also to be considered. Exactly what this process should be termed is a matter for discussion: certainly, it is reasonable to propose that the name ‘G-Format’ should be reserved for material sourced from B-Format, and that it should be evident from the name what the source material actually was, as we can expect results from decoding 2-channel UHJ to be inferior to those achieved with B-Format sources, as the former contains less information on the planar soundfield: the loss of localisation accuracy is a penalty of encoding from 3-channel planar B-Format to a 2-channel medium.

At the start of our experiments, exactly how inferior 2-channel UHJ might be was an interesting question. In early public discussions on the viability of G-Format it was suggested on a number of occasions that we should try the technique on existing 2-channel UHJ content. I was never particularly in favour of this, as I was of the opinion that the results would be less than spectacular. They could not be better than the original UHJ decoded, for example, and that would most likely not be good enough to hold its own against a ‘discrete’ surround mix of today, despite the fact that the former would still retain Ambisonic characteristics, including imaging ability between the loudspeakers beyond the front pair and a larger listening area. I reasoned that this would particularly be the case with Ambisonically-mixed recordings, where pinpoint localisation is desirable, as opposed to recordings made with a single-point Ambisonic microphone array such as the Soundfield Mic. Thus I maintained that it was not particularly worthwhile to try decoding Alan Parsons’ *Stereotomy*, for example, and putting it on a multichannel disc.

While the complexity of traditional analogue hardware Ambisonic decoders varied, the common loudspeaker configuration was what we would today call ‘4.0’ – the most basic layout that would work in a planar configuration, and thus the easiest to implement in the home. A fundamental design criterion of Ambisonics was that it should be practical in a conventional living room – often rectangular – so this should perhaps come as no surprise. And while more sophisticated decoders might offer a knob to set the ratio of sides in the rectangle, from 1:2 to 2:1, with square in the middle, cheaper units (such as those by Minim) went for a ‘least error’ approach by offering a two-position switch, selecting between one configuration that was wider than it was long, and one narrower than it was long.

## A software approach



In 2003, Nimbus Records approached Meridian Audio Ltd for assistance in developing a software-based decoder to allow them to decode their Ambisonic recordings – also predominantly 2-channel UHJ for the reasons described previously – to speaker feeds, with a view to producing a ‘proof of concept’ DVD-Audio/Video disc containing test tracks from a number of Nimbus recordings. Rhonda Wilson at Meridian provided them with a command-line decoder, ‘Ambisonics.exe’, running under DOS, which could be used to decode both 2-channel UHJ and planar B-Format .WAV files to either a single 4-channel .WAV file or to four separate .WAV files.

The initial Nimbus concept disc was circulated among a number of listeners with Ambisonic experience, and many agreed that the results were excellent. Indeed, my own view was that they were not only excellent: they were a great deal better than I had expected. Though the centre ‘node’ was evident on a well-configured system as described previously, the nature of the recordings – of actual live acoustic performances captured with a Halliday array (the Nimbus equivalent of a Soundfield Microphone, combining a physical omni and dual figure-eight mics crossed at right angles) – meant that the rear (surround) information consisted largely of ambience and

reverberation. As such, the technique did not require a high degree of localisation accuracy beyond the front stage in most cases: it required instead to offer envelopment. Nimbus went on to release a series of DVD-A/V discs (the 'Surrounded By...' series), with extensive content derived from UHJ masters using this technique: they have proved quite popular.

This encouraged me to look into the possibilities with Ambisonically-mixed material, and as a result of previous interest, I asked the people in Meridian's R&D department if they could run the Nimbus approach on the CD of *Stereotomy* at some point in their spare time. They did so, and came up with a DVD-Audio disc containing a decode of the album. The results, however, I found disappointing. Localisation was blurred and unstable, and in addition the level was unaccountably low. Switching between a 2-channel UHJ decode optimised for the current layout (using a Meridian hardware decoder) against the pre-decoded DVD version left me feeling that G-Format (or whatever we're going to call it) from Ambisonically-mixed UHJ masters was, as I had suspected, a waste of time, and the idea receded into the background.

### **New tools, new life**

A year or more later, in 2005, my interest was revived by a combination of factors. First was the increasing availability of free Ambisonic plugins for the VST and Apple Audio Units platforms, thanks to a number of contributors (listed in the 'Tools' section of my Ambisonics web site, [www.ambisonic.net](http://www.ambisonic.net)); and a resurgence of practical activity in the SurSound mailing list with the specification of an Ambisonic file format and the creation of players and other software for working with Ambisonic material. I decided to assemble a modest multitrack recording/mixing system for my small studio that could use the available VST plugins to mix to B-Format, and collaborated with former Ambisonic Mixing Project research colleague Peter Carbines to establish the parameters for an equivalent system for his own studio.

After a great deal of research it was determined that my preferred system (due to its affordability and universality), namely Digidesign's Pro Tools M-Powered version, would not support the VST plugins via a VST/RTAS 'wrapper' (allowing the VST Ambisonic plugins to function in a Real-Time Audio Suite/Pro Tools environment) not because of the wrapper but because Pro Tools – all versions of Pro Tools – would not support multichannel plugins due to a fundamentally stereo architecture. The ability to import common multitrack session file formats (to mix third-party material and import any multitrack material we wanted to use) being important, and requiring a comprehensive DAW environment in which to mix to B-Format, this left the most versatile solution as being the acquisition of Steinberg's Nuendo software. Thanks to invaluable assistance from Greg Ondo at Yamaha USA and Cam Wilder, this has now been achieved, and we have both a portable Macintosh-based recording system and a studio-based Windows mixdown system which, once we learn Nuendo and the plugins, will be capable of being used to originate new recordings as well as B-Format mixes.

Second was the arrival in my letterbox of at least two different DVDs of speaker decodes from UHJ sources, both including snippets of *Stereotomy* and one including one of the tracks

originally mixed with the prototype Audio & Design Recording Ambisonic Mastering System for the Boots Audio (the audio division of major UK drugstore chain Boots) demo-room at the Cunard Hi-Fi Show in 1981. These latter were actually some of the first Ambisonic mixes ever made, and they were performed by the original Ambisonic Mixing Project team of well-known freelance engineer George Chkiantz, with myself and Peter Carbines, then the technical advisor to the Audio Buyer at Boots and originator of the Boots Ambisonic Mini-system project. The three of us had been working on Ambisonic mixing by that time for almost a decade.

This DVD, from Eero Aro in Scandinavia, by this time a long-term collaborator with Peter Carbines on Ambisonics, was very interesting and included a decode of the entire instrumental track *Where's the Walrus?* from *Stereotomy* – an excellent test track – plus the original Ambisonic remix of the Boots demo KPM Music Library track, Keith Mansfield's *Bubble Machine*, unfortunately sourced from cassette as the original quarter-inch masters are believed lost.

Being on DVD-Video, this disc required a fairly tortuous mastering process, including sample-rate conversion from 44.1 to 48 kHz sampling, and the use of shareware software first to render the 2-channel UHJ into pseudo-B-Format (there is not enough information in a 2-channel UHJ recording to recover true B-Format, but you can extract something that can be treated as if it were B-Format as far as a decoder is concerned) and then to decode it from B-Format to 5.1 using the Gerzonic *Emigrator* plugin ([www.gerzonic.net](http://www.gerzonic.net)). However, the results were quite impressive and significantly better, in the case of *Where's the Walrus?*, than my DVD-A rendering.

We looked for a solution we could use with less complexity. The Meridian software decoder, as used by Nimbus, was an obvious answer, and Rhonda Wilson kindly supplied us with a freshly compiled version of the application she had developed for Nimbus. After some initial teething problems, both Peter and I got this working with UHJ input files and producing four output files corresponding to the speaker feeds for the default square layout.

### **Decoder parameters**

The Meridian Audio command-line decoder application is capable of rendering 2-channel UHJ and planar B-Format WAV files to four speaker feeds corresponding to the corners of a rectangle (LF, RF, LB, RB). Following experiments by Nimbus, we followed their lead and decoded to a square array: the default aspect ratio.

It has been suggested that in fact the majority of home listeners with surround replay capability do not have an ITU-style 5.1 array, with the rear speaker angle significantly wider than the front, and that instead they tend towards a rectangular array, although the listening position may generally be too far back (Farina et al).

Our own limited investigations strongly support this view and this does not come as a surprise: Ambisonics was originally designed with practicality of installation in an average living room as a significant requirement, and the original basic rectangular array reflects this. As a result we feel confident

that decoding to a rectangular array will successfully minimise errors of congruence between target decode and listener speaker positions.

However, we cannot tell whether listeners will have a rectangular array with an aspect ratio greater than 1 (longer, front-to-back, than wide) or less than 1. The square target array is thus a useful compromise – and has been shown in tests by Nimbus Records to produce the most broadly applicable result.

In fact it should be noted that in certain areas of this work we are following in the footsteps of Nimbus Records, though they have somewhat different goals and employ DVD-Audio as the release format, and we are extremely grateful for their support and in particular that of Caractacus Downes, along with that of Rhonda Wilson and others at Meridian Audio.

The decoder does not have provision for a centre (front) channel, and in our view, as Ambisonics already offers excellent front stage imaging, the only down-side of omitting a centre channel was that listeners might expect to hear a signal from it. This concern can be alleviated by noting it in the documentation accompanying a recording (eg sleeve notes).

In addition, we left the “row” setting at default: this is a parameter for selecting forward dominance, based on the idea of defining a row of seats in a concert hall. We used Row A.

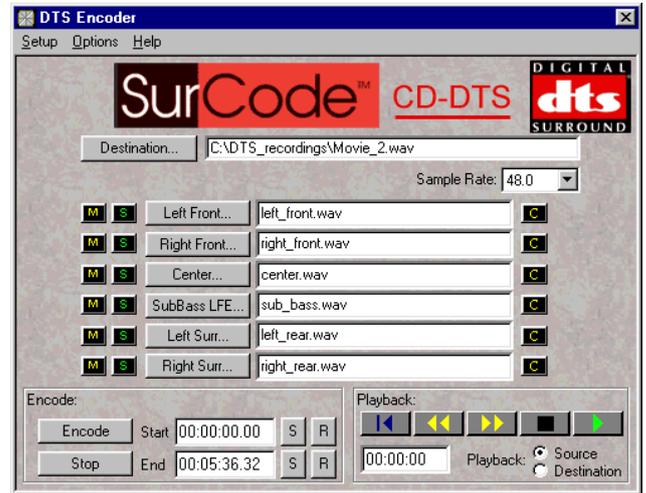
#### The DTS CD Format



We decided not to take the DVD route as Eero had done, due to the cost of authoring software and the complexities of the workflow. We decided instead to experiment with DTS CDs.

The DTS CD format has been around for a long time, and there were even, in the past, some privately produced G-Format demo discs circulated using it, based on real-time hardware decodes. The DTS CD format encapsulates a DTS 6-channel data stream in an envelope that looks, to a player, like a standard CD track at 44.1 kHz sampling, 16-bit. However, if fed from the player via a digital output and into a DTS surround decoder, the stream is extracted and decoded to 5.1. The only downside is that you must never listen to the stream as audio: it replays as solid white noise at about -10dB!

DTS CDs can be created easily and simply, using the \$99 SurCode Pro DTS CD encoder from Minnetonka Software ([www.surcode.com](http://www.surcode.com): see screen shot). The encoder has a Windows front-end and simply requires up to six input WAV files to be specified (it does not accept multichannel WAV files). It outputs a single WAV file that can then be burned on to CD using standard CD-R burning software.



With these two pieces of software in place, we had a simple and elegant workflow, initially for recovering existing 2-channel UHJ recordings, decoding them and putting them on disc, but ultimately for doing the same with any new B-Format source material that might become available.

#### Initial workflow

- First, the material has to be captured to a WAV file. This means either simply ripping a CD track to WAV, or recording analogue source material into a DAW (we are using Adobe Audition for handling UHJ) and premastering to create a WAV file. The Meridian decoder outputs the same sample rate as is input, and as the DTS CD medium requires 44.1 kHz sampling, we capture at 44.1 kHz to avoid sample rate conversion.
- Second, the required WAV tracks are processed by the Meridian decoder to produce a set of four WAV files as output, ie LF, RF, LS, RS.
- Third, the decoded WAV files are loaded in sets of four (no CF or LFE) into the SurCode encoder and processed to deliver a single WAV file.
- And finally, the DTS-encoded WAV files are ordered and pauses determined in the CD burning software and the CD written.

A simple workflow, then, but next came the real test: what did it sound like?

Perversely, both Peter and I independently chose to start with cassette-sourced material, namely the original Boots Audio demo mixes performed at Sound Developments studios in early 1981. I used a TDK metal cassette of the master mixes, which also included some later material made experimentally in 1983, one track mixed during an Ambisonic Mastering System demo at Hollywood Records in London and the other another Keith Mansfield library piece mixed with the production version of the Audio & Design mixing system. Peter used a cassette of the 1981 material sourced direct from the console output as the UHJ masters had been recorded, which included some alternate takes of some of the mixes. Both tapes were recorded at quite high level and exhibited some minor intermodulation distortion in places at the bass end.

My tape included a test tone that was evidently set at -10, and came back at about the right level, so transfer was well behaved and required only a little tweaking. I used Audition's

sample-based noise reduction on the orchestral tracks where there was a certain amount of tape noise, but didn't bother on the electronic material which had a fairly good subjective noise floor thanks to the material, the tape and Dolby B. However past experience suggests that noise can cause some problems in a UHJ decode so I recommend minimising it if it is too evident.

Following topping and tailing of each item I ran the tracks through the decoder and the result through SurCode – at some point there are some batch files that need to be written here! In addition, SurCode is a little idiosyncratic with its buttons – which a new set of source tracks requires you to press, and which you don't need to bother with. Usefully, if all the files aren't the same length (for example you loaded the wrong file into a channel) it objects; and thankfully, in a somewhat similar vein, you do not need to specify files that aren't present – the missing LFE and CF channels in our case – or prepare special files of silence or anything: you simply leave the entry blank.

Having processed all but one (substandard) number from the tape, I shuffled them into a suitable running order and burned them to CD, following the 13 items with the 13 original UHJ files for comparison. Then, off downstairs to play them back.

The centre of the array in my living room is in fact achieved by sitting on the coffee table. Thus seated and pressing Play, after a short pause I was treated to a recording of my own voice performing a synthetic walk-round test generated (with a mono mic and a hand-built switched Ambisonic panpot: much more accurate than doing it with an SFM) a quarter of a century earlier.

Starting at the North (centre front: we never used speaker locations because in Ambisonics the sound may well not be in the speakers, so compass points make more sense), my voice went round to NE, E, SE and so on, and finally back to North. And not only that: it did it quite well. East was slightly blurred, mainly due to the open door to the tiled dining room in that direction: closing the door essentially cured it. South was a little close to the back of my head, but in 2-channel UHJ, it was always thus – as was that little node in the very centre of the array, where the localisation went slightly weird. Everywhere else, the voice stayed a similar distance away from me, as it should: this panpot was very basic, with eight switched sets of resistors and no radius vector. All in all, it sounded very much as it had done when we recorded it in Peter's old South London studio 25 years before – actually rather longer ago, because this panpot went right back to our earliest interpretations of some of Michael Gerzon's earliest Ambisonic papers, quite some time before we had the prototype mixing system to play with.

Then came the first piece of music: a 'special mix' of an orchestral piece called *Saturn* by Tony Hymas, mixed from a KPM Music Library multitrack in early 1981. 'Special' because it was designed to open the Boots Audio demonstration at the Cunard Hi-Fi Show, and started with eerily sliding strings, Ligeti style, in mono; gradually widened out to stereo; then super-stereo; and at the climactic sounding of a gong, wrapped into full surround as the gong flanged and panned overhead from front to back.

And that's exactly what happened. In fact as far as I could remember across a quarter of a century, it sounded *at least* as good as it had in the studio when we originally mixed it. The passage immediately after the panned gong was particularly impressive: the entire orchestra appeared to be arrayed across the rear, until strings started to creep back in from the front as the track continued – exactly as intended. A quick comparison with the UHJ version on the same disc, decoded with the G68 Ambisonic decoder, indicated that our software-based workflow performed as well as a hardware decode with the same settings.

Next up was Keith's original *Bubble Machine*, an electronic MIDI-generated piece in the style of *Popcorn* with added acoustic drums, and again it brought back memories of the studio playback, with mix elements appearing from different locations around the room as intended. Playing the other numbers told essentially the same story: they sounded the best I had ever heard these tracks since we had mixed them, and possibly better than that.

With one success under the belt, I then proceeded to perform the same operations with *Stereotomy*, this time with the advantage of simply being able to rip the tracks from the UHJ CD. Once again, the decode was particularly impressive. Mixed with the 1983 production system, which had several added features including a radius vector on each of the panpots – and mixed by an engineer/producer far more experienced than the present writer – the tracks exhibited more sophisticated content and depth than the 1981 experiments. It was also a million miles away from my test DVD-A decode, theoretically using a similar, though not identical, workflow, suggesting that the latter had incurred some setup problems and was not representative of what the software decoder could actually do. In any event these initial playbacks vindicated the idea of decoding UHJ mixes to speaker feeds and putting the result on disc, along with identifying the chosen workflow as a cost-effective and highly practical method of distributing Ambisonic material.

### Refining the process

We encountered some initial problems, firstly with albums containing a zero-length pause between tracks, an example being the KPM/Nimbus CD *Surprise, Surprise* by Chin & Cang, one of the first two or three Compact Discs manufactured in the UK as well as being the first Ambisonically-mixed CD.

Normally, on a regular audio CD, you would simply set the pause time to zero between the tracks designed to run on (the default being two seconds), and an audio CD will simply play across the TNO boundary without interruption. Not so with a DTS CD: setting the pause to zero results in a brief dropout caused by muting and de-muting of the DTS decoder. The solution to this problem, helpfully suggested by Aldo Bazan in the SurSound list, is to burn the disc from a single, continuous DTS-encoded WAV file, using .CUE files – the CD-R-burning software equivalent of the PQ data used in traditional CD mastering – to indicate track boundaries to the burning software. Many CD burning tools, notably EAC (Exact Audio Copy) but also Nero and Roxio's Toast, can utilise these files, which are based on a proprietary standard originally developed by Golden Hawk software but now widespread.

.CUE files include a range of data required to burn a CD, including the start, end and duration of each track. It is widely believed that these files can only be used with a special type of disc image file, but this is not the case: they can equally happily reference a WAV file.

Several applications can operate on cue data, but so far we have needed to use a series of programs to do everything we need. Our current workflow is described below.

**Current workflow**

The process starts by assembling all the tracks to be burnt to CD as a single audio file rather than one for each track. In the case of a UHJ project, this can be done in a stereo editor such as Adobe Audition. Audition has the added benefit that it can be used to rip a CD to a single file, for example if the source is UHJ already on CD. In doing so, Audition automatically adds the track start and end cue points – see fig 1 below).

These points can be adjusted, but if this is done, it's necessary to make sure that Audition's frame rate is set to 75fps (CD rate) and the option to snap to frame boundaries is selected. Audition will let you display the cue list (fig. 2) but not export it. Obviously it's possible to write it down and re-enter it manually, but it is easier to use a program called CueListTool (<http://www.stefanbion.de/cueltool/>) which can read the cue data that Audition saves with a file. But note that to do this, the WAV file must be saved in Audition with the option checked to 'save non-audio data'.

At this point, the original UHJ WAV file can be processed to speaker feeds with the Ambisonic decoder app (in our case)

Label	Begin	End	Length	Type	Description
Track 01	00:00:00:00	00:00:34:15	00:00:34:15	Track	
Track 02	00:00:34:15	00:03:32:14	00:02:57:23	Track	
Track 03	00:03:32:14	00:06:47:13	00:03:14:24	Track	
Track 04	00:06:47:13	00:10:15:04	00:03:27:15	Track	
Track 05	00:10:15:04	00:13:43:14	00:03:28:10	Track	
Track 06	00:13:43:14	00:15:46:15	00:02:03:01	Track	
Track 07	00:15:46:15	00:19:36:16	00:03:50:01	Track	
Track 08	00:19:36:16	00:23:12:06	00:03:35:15	Track	
Track 09	00:23:12:06	00:25:47:22	00:02:35:15	Track	
Track 10	00:25:47:22	00:28:08:20	00:02:20:23	Track	
Track 11	00:28:08:20	00:31:17:10	00:03:08:14	Track	
Track 12	00:31:17:10	00:33:45:10	00:02:28:00	Track	
Track 13	00:33:45:10	00:37:34:04	00:03:48:19	Track	
Track 14	00:37:34:04	00:39:26:16	00:00:52:12	Track	
Track 15	00:39:26:16	00:41:13:15	00:02:46:23	Track	
Track 16	00:41:13:15	00:43:48:15	00:02:35:00	Track	
Track 17	00:43:48:15	00:46:31:00	00:02:42:10	Track	
Track 18	00:46:31:00	00:49:11:22	00:02:40:21	Track	

fig. 2: Cue listing in Audition – but you can't export it.

and the resulting files fed to SurCode DTS CD to create a single DTS-encoded WAV file. This should be virtually identical in size to the original WAV file, and, needless to say, nothing should be done, after having saved the original WAV file with cue data, that will change the duration of the file in any respect – the final DTS file must embody the same track starts and ends as the cue data. It is obviously not possible to locate cue points in the DTS file, so instead they have to be referenced in the original audio.

With the final DTS WAV file ready, the original audio WAV file is opened in CueListTool and the cue data imported. This can then be saved as a .CUE file (see figs 3 & 4 overleaf).

In doing so, the name of the audio file to which it is linked is specified. Thus the name of the final DTS file can be inserted here, along with other data, as part of the saving process (or it can be done at several other points).

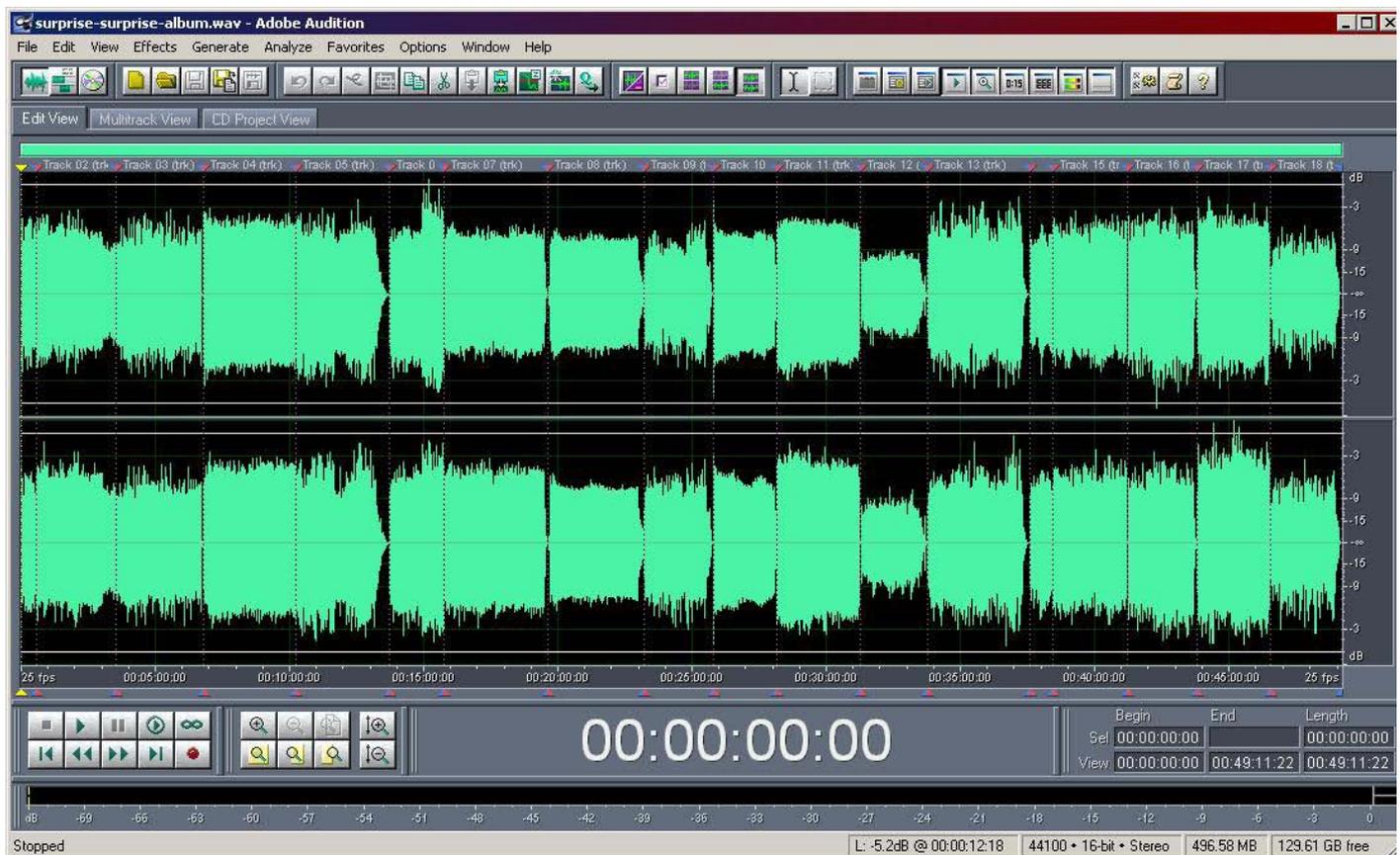


fig. 1: Ripping a CD to one file in Audition produces cue points at track boundaries automatically.

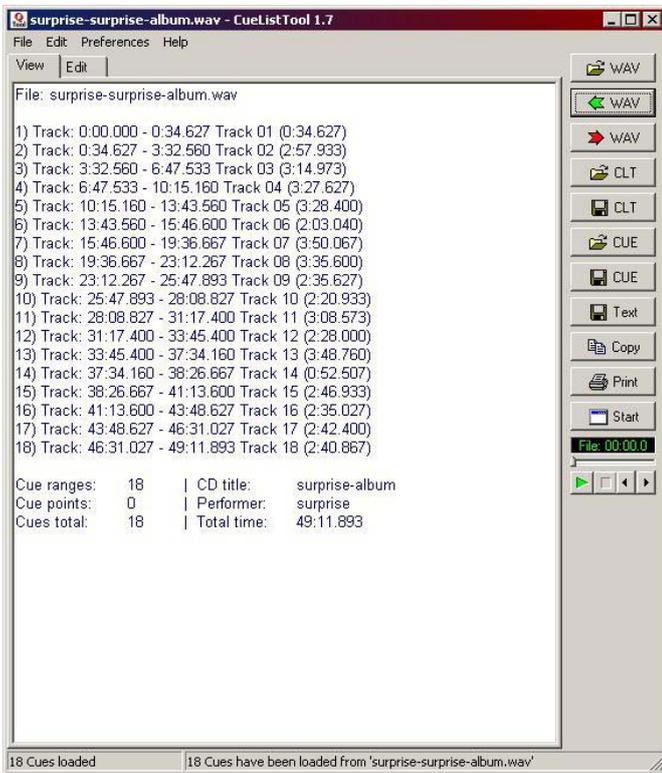


fig. 3: Opening the cue data in CuelistTool

The .CUE file can then be opened, if it is necessary to edit it, in a cue file editor such as DCSoft's CDRCue 2004 (<http://www.cdrcue.com>), as shown in fig. 5 (above right). Information entered here can be written to disc as CD-Text if the burning application and the drive support it.

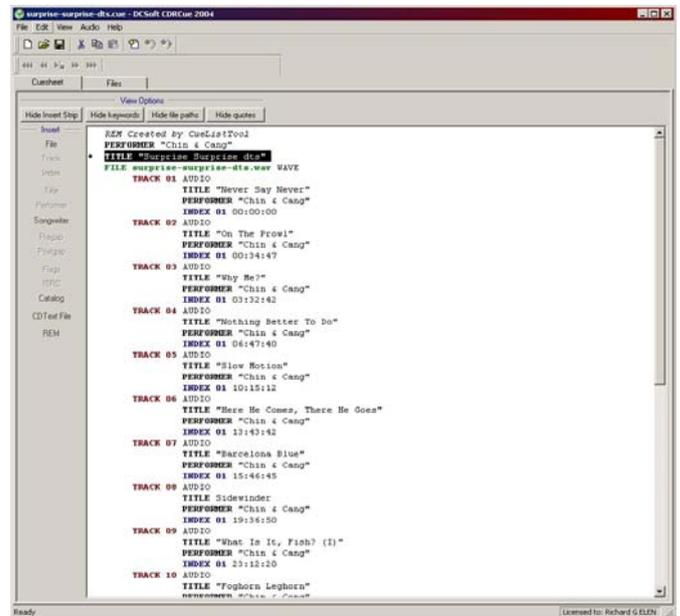


fig 5: Editing the .CUE file in CDRCue 2004

Typically, the DTSWAV file and the .CUE file should be placed in the same directory: no path information is necessary in the CUE file in this case.

The penultimate stage is to open a CD-burning app such as EAC, Nero or Toast, and set up for writing from a CUE file. EAC is used as the example here. After opening the application, the 'Write from Cue File' option is selected and the CUE file is opened. A display like that shown in fig. 6 is then visible.

All that remains then is to burn the DTS disc. Because a single file is used, there are no interruptions to the DTS stream across track boundaries of any length. It may be necessary to adjust start times to ensure that when an individual track is selected, it begins a suitable length of time before the audio: this can either be done by editing the CUE file in the cue file editor, or by going back to the original audio WAV file and changing the cue points there and then re-doing the cue data export process.

Another, related issue is the length of pause at the front of tracks required to allow the decoder to recognise the DTS stream and de-mute. Treating the encoded tracks in the same way as one would treat audio CD tracks, ie without adding an additional pause between the start of the track and actual audio, resulted in the decoder de-muting very close to the initial transient of the music – or even later.

With an audio CD, common agreement specifies how much time should elapse between the start of a track and actual audio, to allow the player to de-mute. With some older players, this time requirement could be quite long. According to DTS, no such equivalent agreement exists for DTS CDs, due to the way in which different manufacturers have implemented DTS technology. But, says Scott Esterson of DTS, "there is a safe zone: that safe zone is between 1,500 and 2,000 ms. (1.5 to 2 seconds). We vary the time depending on the music."

To do this requires the addition of digital silence (for example) at the front of a track. When ripping from a UHJ CD, the initial 2sec audio CD pause will be included in the front of the file,

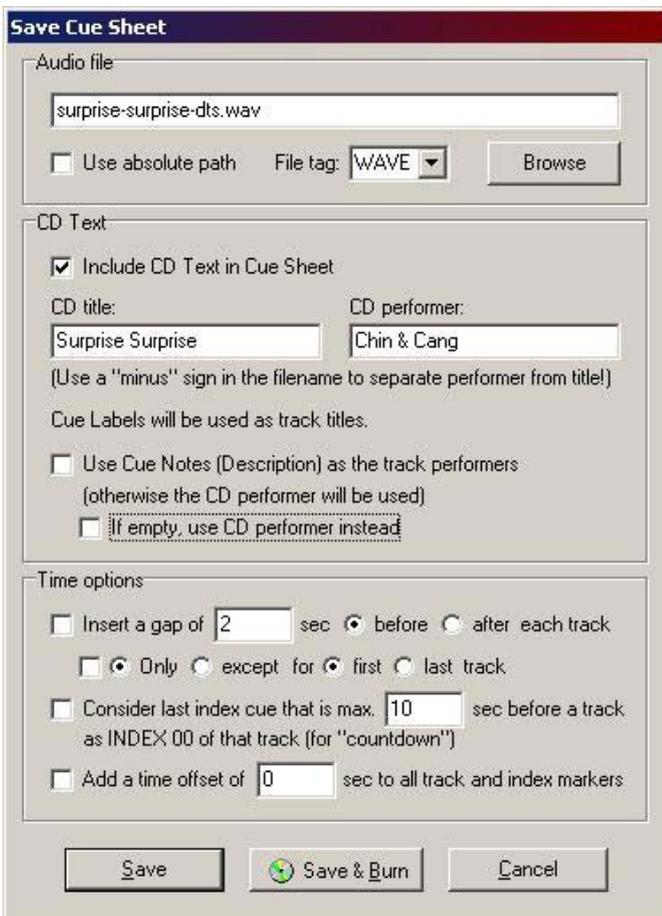


fig. 4: saving a .CUE file from CuelistTool

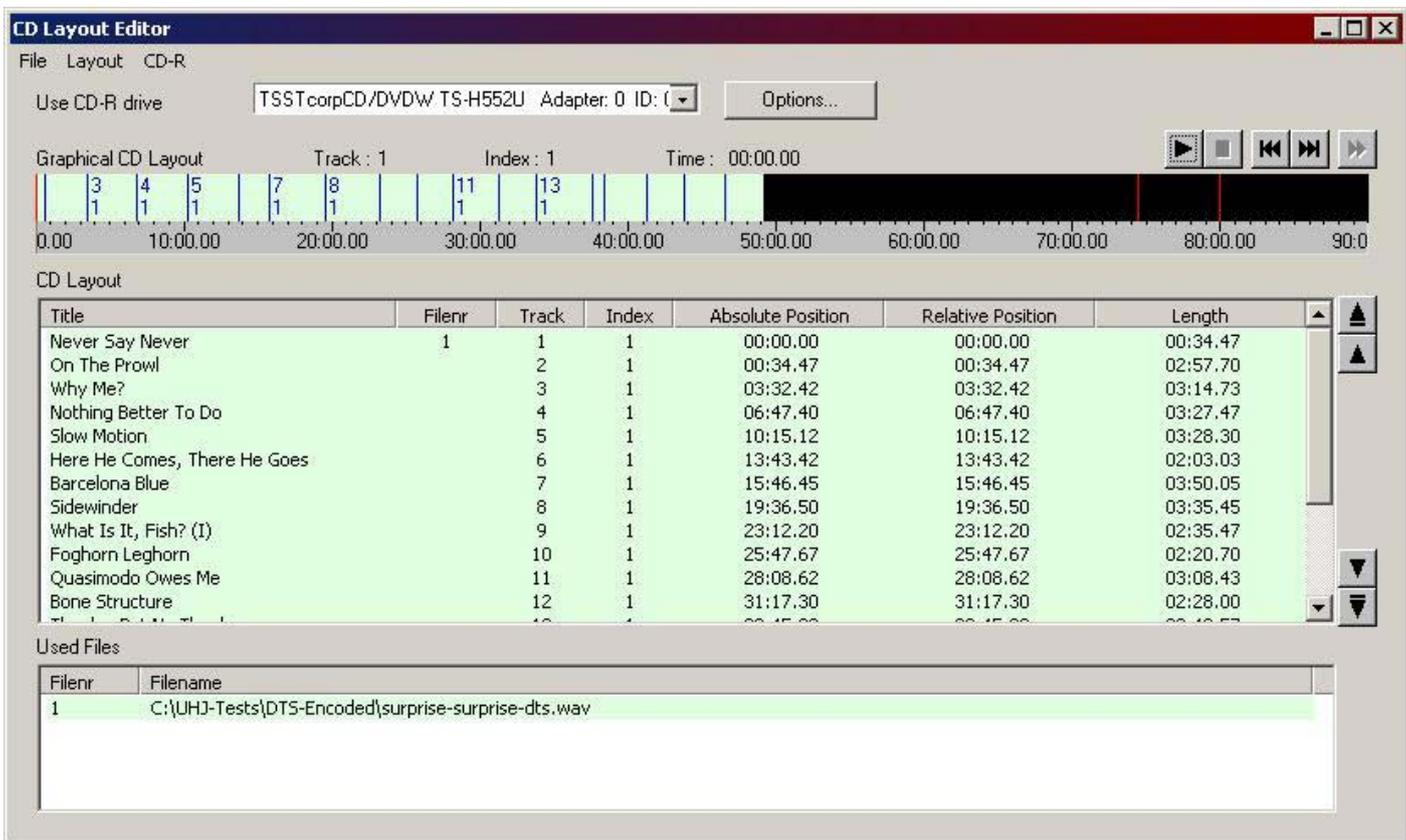
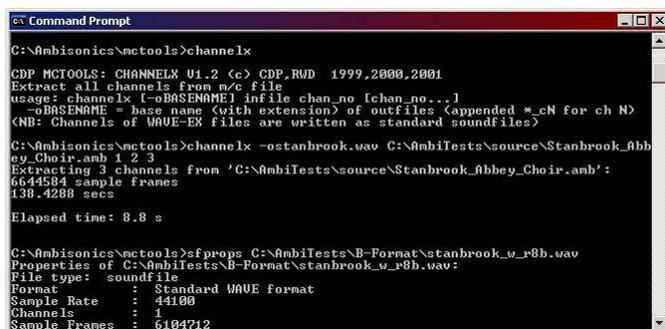


fig. 6: After choosing "Write CD from cue file" in EAC, and opening the cue file in the CD Layout Editor, a graphical representation of the CD contents is shown as seen above.

as will digital silence between tracks, so this does not have to be added to: the existing pauses will act not only as pauses between tracks but also as safety gaps to allow the DTS decoder in the replay system time to de-mute. However, in assembling new material, suitable pauses should be added at the premastering stage by inserting them *into the original audio file*.

### Working with B-Format

We now turned our attention to B-Format source material. Although when we had started there was almost nothing available, new recordings have started appearing from various sources thanks to Angelo Farina, and to Etienne Deleflie's *AmbisonicBootlegs.net* site. This made it possible for us to use our approach on B-Format material – most of these items are quite short and thus amenable to experimental applications.



In these cases, the source consisted of multichannel WAV files, including those with an ".amb" extension conforming to the recently-established WAVE-EX Ambisonic file format definition. As our processes by and large need individual mono files for each channel (the Ambisonic decoder will output multichannel files but will not accept them), we had to begin by splitting the

multichannel file into three mono files, representing W, X and Y components of B-Format. To do this we could use a number of approaches, including opening them in Adobe Audition, which automatically splits them on opening, allowing them to be saved. However, it was easier to use *channelx*, a command-line application from Richard Dobson's *Multichannel Tools* suite (see screen shot above). This simply and quickly delivers mono WAV files from a single multichannel file.

Most of the source material was originally 16-bit 44.1 kHz sampling. In these cases, the signal is treated as 16-bit throughout and leads to a fairly straightforward workflow:

- 'Disentangle' multichannel source file to three mono WAV files with *channelx*
  - Four-square B-Format decode
  - DTS CD encode
- ...and thus to a standard premastering procedure as already described.

In some cases we encountered material that was recorded at a different sample rate and word length, such as 24-bit at 48kHz sampling, and this required a slightly more complex approach. To handle the sample rate conversion we used *r8brain*, a free Windows-based sample-rate conversion application. We decided to do the sample-rate conversion as early in the chain as possible, but maintain the 24-bit path throughout.

Unfortunately, the *r8brain* sample rate converter generates an error when presented with a ".amb" file, so the SRC operation had to be carried out on the individual mono files after 'disentangling'. This gave us the following workflow prior to our standard premastering sequence:



- ‘Disentangle’ multichannel source file to three mono WAV files with *channelx*
- *r8brain* SRC to 44.1/24
- Four-square B-Format decode, setting the output word length to 24 bits
- DTS CD encode

Note that while a CD is of course 44.1/16, the DTS encoder will encode up to 24 bits of information in the source file. As a result, for 24-bit and higher source material, the output word length of the B-format decode was set to 24-bit (see below). For the other files the decode output word length was not specified, and in this case the decoder outputs the same word length as it is given, so the decoded files are 16-bit.

```
C:\Ambisonics>Ambisonics -w C:\AmbiTests\B-Format\stanbrook_v_r8b.wav -x C:\AmbiTests\B-Format\stanbrook_x_r8b.wav -y C:\AmbiTests\B-Format\stanbrook_y_r8b.wav -m C:\AmbiTests\Decoded\B-decodes\stanbrook.wav -b 24
B-Format and UHJ 4 Speaker Ambisonics Decoder
(C) Meridian Audio Limited 2003
Version 0.9 ( Jan 20 2006 14:07:29 )
Input file C:\AmbiTests\B-Format\stanbrook_v_r8b.wav
Input file C:\AmbiTests\B-Format\stanbrook_x_r8b.wav
Input file C:\AmbiTests\B-Format\stanbrook_y_r8b.wav
Output file C:\AmbiTests\Decoded\B-decodes\stanbrook_1f.wav
Output file C:\AmbiTests\Decoded\B-decodes\stanbrook_1r.wav
Output file C:\AmbiTests\Decoded\B-decodes\stanbrook_1b.wav
Output file C:\AmbiTests\Decoded\B-decodes\stanbrook_rb.wav
Row
n
Aspect ratio 1.000000
Decoding started at Sat Mar 25 11:03:57 2006
Decoding stopped at Sat Mar 25 11:04:09 2006
138.4 seconds of audio decoded
C:\Ambisonics>
```

### Redefining G Format

We’ve noted that, strictly speaking, the term “G-Format” is nominally applied to 5.1 speaker-feed decodes of B-Format source material. However, the process for rendering 2-channel UHJ content is evidently virtually identical to that used for B-Format, and is essentially a special case: 3-channel UHJ recordings, if any exist, would offer exactly the same amount of data as a B-Format recording, the difference between 2- and 3-channel UHJ being a single channel of information. As a result, it seems a little unnecessary to think up a different, special label just for this purpose, and we are tending to describe renderings of 2-channel UHJ to speaker feeds as “2G-Format”, generalising the meaning of “G-Format” in the process.

Thus we propose that the term “G-Format” be expanded to describe generally a set of speaker feeds rendered from an Ambisonic recording. If one wishes to be more specific, the source could be specified as B (for a B-Format source: “BG-Format”) or 2 (for a 2-channel UHJ source: “2G-Format”).

One could extend this further and also specify the base destination array to some extent: “2G4” would represent a 2-channel UHJ source rendered for a 4-speaker array as used here, for example.

### Conclusions

Based on early results using this workflow, we believe we have found a simple, elegant and cost-effective method of producing and distributing speaker-feed decodes of Ambisonic source material, both in the case of 2-channel UHJ (which we have tested) and planar B-Format (which so far we have not). In the former case, our experience to date indicates that the perceived results are at least as good as those originally experienced in the studio when playing back the master mixes, suggesting, perhaps unsurprisingly, that the digital decoder application we are using is superior to the original analogue hardware decoders in use a quarter of a century ago.

In addition to working with – and creating – B-Format source material, we also intend to explore different decode options, such as the *Emigrator* system available from [www.gerzonic.net](http://www.gerzonic.net), which allows a wide range of target loudspeaker arrays; however this, along with several other experiments we wish to perform, will have to wait until we have learned our Nuendo installation handling the increasing range of Ambisonic VST plugins available for a professional DAW environment.

*\*Richard Elen is a recording engineer and producer, broadcaster, writer and former Editor of Studio Sound magazine. He is the head of Creative Services at Meridian Audio Ltd in Huntingdon, UK. Peter Carbines is an audio engineer and former technical advisor to Boots Audio and is based in London. In conjunction with engineer/producer George Chkiantz, Carbines and Elen worked on Ambisonic mixing technology in the 1970s and 1980s and assisted in the development, by Dr Geoff Barton, of Audio & Design Recording’s Ambisonic Mastering System, a collection of studio outboard modules enabling B-format and UHJ mixes to be made using a conventional multitrack mixing console.*

