Ambisonics. Part one: General system description

Peter Fellgett[*]

Ambisonics is a technology for surround-sound which aims specifically at not *making four (or any other number) of loudspeakers audible as separate sources of sound. It is designed using appropriate engineering methods and psycho-acoustic theory that has shown good predictive value to make best use of available channels of communication (two or more), and of loudspeakers (a limitation often forgotten), to give stable and uncoloured acoustic images in any position, keeping the physical means of reproducing the sound as unobtrusive as possible. It claims wide freedom of recording methods and of source material, as well as protection of recorded material from obsolescence.*

Ambisonics in the perspective of surround-sound technology

MONOPHONIC REPRODUCTION provided information about direction and distance only implicitly, through ambience labelling. Stereo added explicit directional information over a front-sector not exceeding 60° in width (ie ± 30°, although some discriminating listeners prefer to set the limit for really satisfactory stereo-blending at ±15°).

Beyond stereo, the technology can be developed in several ways:

1. By using more loudspeakers.
2. By using more channels of communication.
3. Making better use of the available number of loudspeakers and channels.
4. Extending directional information from the 60° front-sector of stereo to a full 360° surrounding the listener in the horizontal plane, or to complete spherical surround reproduction including height.

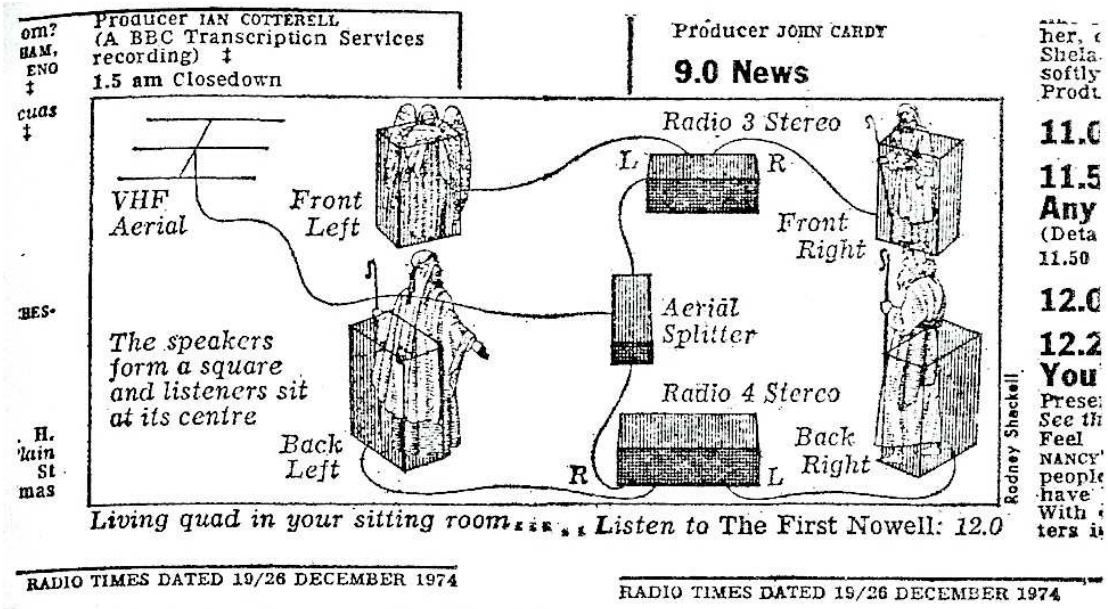
These ways are distinct, but of course the greatest opportunities for enhanced capability lie in combining them as an integrated whole. In general terms, this is the aim of the NRDC Ambisonic technology.

In ordinary life we are bathed in sound from all directions; so much do we take this for granted that it often passes unnoticed until it is cut off, eg in an anechoic chamber, when its loss is keenly felt. Except in so far as the reverberation of the listening-room can supply the deficiency, stereo reproduction subjects us to this deprivation. A major aim of developments beyond stereo has therefore rightly been

[*] Professor of Cybernetics and Instrument Physics, University of Reading

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

Ambisonics. Part one: General system description



Above: *An illustration of the restrictive limitations of the four-source 'quadrifontal' approach to surround-sound.*

the extension to surround-sound reproduction.

The first attempts at surround-sound (apart from some early experiments) used an approach generally called 'quadraphonic'. This term has not of course been precisely defined, and usage is not always consistent. We shall therefore use, as a label for this general approach, the more accurate term *quadrifontal*, meaning 'four-source'. This will be taken to mean that there are assumed to be just four signal sources which are to be connected to exactly four loudspeakers in a one-to-one manner through four respective channels.

Based on existing practice relating to four track master tape, and on the probability that most surround-sound listeners will (at least at first) be constrained by the size of their pockets and the shape of their rooms to use four loudspeakers, these assumptions have a superficial plausibility, but further consideration suggests them to be inadequate in several important ways:

1. Four track master tape is by no means the only source of signals to be considered: there is multitrack and multi-microphone material available for surround pan-potting, and of course the natural sound-field of performed music

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

including reverberant as well as direct sound.

2. There are very good reasons (see Part Two) for not being restricted for ever to some fixed number of loudspeakers, especially not to four.
3. The assumed objective can be attained only if four channels are available, and in their absence can only be imitated more or less unsatisfactorily. This requirement precludes the direct use of the many two-channel recording broadcasting media at present used for stereo, whereas satisfactory surround-reproduction is perfectly practicable in a system designed from the start to use two available channels.
4. Independent access to the signals reaching each loudspeaker appears at first sight to give the producer or recording engineer maximum freedom, but in fact denies it to him because it presents him with a problem analogous to opening a lock without having the key. *Except in the special case of an image in the direction and at the distance of one loudspeaker* it does not suffice simply to squirt independent sounds from each speaker. In general each loudspeaker should radiate a wave of amplitude and phase carefully calculated to combine in the listening-space so as to reconstruct a simulacrum of the intended surround-field fulfilling relevant psychoacoustic criteria[1]. This reconstruction is in some ways analogous to an acoustic hologram, and account must be taken of the size and shape (unknowable at the time of recording) of the individual listener's loudspeaker array.

   A competent system takes care of these precise interrelations automatically, just as a key automatically brings the levers of a lock into register so that the bolt can slide freely. Unless the system does this, there is virtually no chance of achieving clean stable images other than in a restricted set of directions. The result is the familiar practical restriction of 'quadraphonic' reproduction to corner positions, front sector, and perhaps rear centre, with side or positions virtually unusable (see illustration).

A particularly unfortunate form of the quadrifontal approach, which may perhaps be called the 'full quadrifrontal' form, assumes in addition that the four source-signals are pair-wise blended. This imposes further restrictions and disadvantages:

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

1. It makes less than full use of the information capacity of the available channels; it is possible to do as well or better with less than four channels.
2. The implied directional coding, having discontinuities of slope, cannot be realised by pick-up from any combination of ordinary directional microphones. Natural sound, including its indirect reverberant content, is therefore excluded.
3. The format gives poor results when replayed directly through four loudspeakers (see Part Two). It is therefore particularly inappropriate to take this unsatisfactory form of playback as the standard of comparison for surround-sound.
4. It places undesirable restrictions[3] on the encoding loci that can be realised subsequently, particularly in two channel format; for example the Japanese 'Regular Matrix' definitions[2] cannot be implemented by matrixing pairwise blended material. This (and other) criticisms of this widely-used four track format were at first strongly resisted (using blanket 'commercial' assertions) but are now recognised even in the so-called 'pairwise' and 'optimum' loci incorporated in the provisional 'matrix quadraphonic' standards of the USA RIAA[4].

The basic fault of pairwise blending is that each musical instrument (or other source) considered by itself activates only one pair of channels or loudspeakers. This is not beyond stereo in concept or capability, but merely extends stereo to less and less suitable speaker-angles and hearing sectors as we go from front, to back, and to the sides.

Quadrifontal assumptions underlie the regrettable practice of using 'four channel' (or even 'quad') as if it were a synonym for surround sound. Two channel surround systems are then called 'matrixed four-channel', creating the need to distinguish systems that do actually use four channels as 'discrete four-channel' (although the channels are in fact continuous, blended and frequently multiplexed!); we eschew such misleading terms. Any reference to 'the original four-track tape' lies of course entirely within quadrifontal assumptions; in reality *the original to be reproduced is nothing else than the producer's or recording engineer's intentions, and any intermediate format is to be adjudged good or bad according as it helps or hinders the realisation of this original.*

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

Against this background, Ambisonics may be seen as a basically straightforward technology for surround-sound reproduction, designed from the beginning to accept all competent source material, and to make the best use of the available resources in channels (two or more) and loudspeakers (any reasonable number), neither seeking to reproduce a derived 'original' nor attempting or pretending to communicate more channels-worth of information than there are channels in the system. Its methods conform to established principles of sound engineering, applied in new ways and making use of newly acquired knowledge of psychoacoustics. It is not of course perfect, since perfection would require many thousands of channels and a million or so loudspeakers.

Requirements for a surround-sound system

Necessary or desirable requirements in a viable surround-sound system include the following:

1. Ability to accept any competent source material. This includes at least:
   a) Natural sound-fields. This is important for all kinds of performed music, and must include reverberant as well as direct sound. It does not suffice to provide merely a vague splash of reverberation having an appropriate decay time. There should be a structured association of direction and delay of indirect sound giving specific information about the acoustic ambience of the performance, including definite impressions about the size and shape of the hall. Recent research has shown that ambience labelling according to place of origin of each sound is an important aid both to image localisation and the ability to discern inner lines in a musical texture despite differences of intensity level. Balance is thus made less critical and (particularly in pop) greater musical complexity becomes acceptable by a given audience.

   b) Multitrack and multi-microphone material for pan-potting; that is to say, mono signals on which synthetic directionality is to be imposed. There is never, of course, any

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

difficulty in principle in building pan pots to conform to any encoding standard whatsoever in any system. The more demanding requirement is for a means of adding artificial reverberation having subjectively smooth and uncoloured characteristics. (In the present state of technology artificial ambience, as distinct from mere reverberation, is best obtained naturally, if this paradoxical way of putting it can be excused.)

c) Existing pairwise-blended material. This should be seen as a rescue operation for historical material, optimising the compromises inseparable from the limitations (discussed in the opening section) inherent in this format, which should therefore not be used for new recordings where alternatives are available. This is a nevertheless an important requirement in view of the large investment of the industry in material recorded in this form.

2. Availability of a format, for studio use, robust to the inevitable small errors of intermediate recording and providing good facilities for processing, including a versatile gamut of 'effects'. This format should above all preserve explicit directional information, and thus preserve options for the future.

3. Encoding standards, for public issue, having at least the following properties:

   a) Unambiguous encoding of *every* possible sound-direction.
   b) Low sensitivity to errors of transmission and of decoding.
   c) Freedom for the listener to decode into any reasonable number of loudspeakers in any reasonable array. In particular, rectangular speaker arrays should be catered for, since few rooms are square. The desirability of not restricting the number of loudspeakers to the conventional four has already been indicated, and is further discussed in Part Two.
   d) Capable of being decoded so as to give accurate and

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

stable localisation, and freedom from coloration, according to the best available psychoacoustic criteria.

4. Mono and stereo compatibility. In one sense, this is a special case of the listener's freedom (as in 3 (c)) to use any number of loudspeakers, ie one for mono and two for stereo. The special difficulty is that mono and stereo represent methods of decoding (representable indeed by matrices) prescribed by useage, which are nonetheless definite for being trivial in the sense of requiring no explicit decoder but only suitably connected pieces of wire. Unfortunately the implied decoding matrices ('obvious' though they seem) can be shown to be incompatible in any context of two channel surround sound encoding. This is not the 'fault' of any surround sound system (not even quadrifontal) but is the result of any unhappy historical accident. Some compromise therefore has to be made between mono and stereo compatibility in *any* two channel surround system. The available means of effecting this compromise are fortunately sufficient to reduce the technological incompatibility to the level of the inevitable artistic compromises between mono and stereo; they need not affect the surround reproduction characteristics, essentially because the choice of surround decoder is still open at the design stage.

5. Capability of growth to give protection against obsolescence. Primary interest should be (and probably is) in two channel systems, because of the extensive ready-made commercial outlets available in two channel media of recording and broadcasting originally developed for stereo use. It is important however for the two channel technology to be compatible with extensions into more channels as they come into wider use (as without doubt they will, sooner or later) eg three channel fm broadcasting making full use of the triple audio bandwidth of the Zenith-GE system, multiplexed vinyl discs, multitrack magnetic

tape, and the use of video disc technology. As soon as at least three channels become available, the possibility of including height information has to be considered, even if only as a contingency for when the public may be ready for it. It is particularly important to hold master-tapes in a format that will not be prejudiced by such developments in the foreseeable future.

Ambisonic systems and characteristics

The basic NRDC Ambisonic realisation is a two-channel pantophonic system, ie giving 360° horizontal surround and needing only stereo recording or broadcasting media for dissemination. It is extendable to three channel pantophony, and to three or four channel periphony (ie with height). Particular attention has been given to the use of a third channel of reduced bandwidth (as in the Nippon Columbia TMX system). Five channel pantophony and nine channel periphony have also been studied theoretically; although these are not of current commercial interest, it is reassuring to know that compatible developments are possible well beyond presently foreseeable needs. All systems share a uniform technological design, which includes the following essential steps and signal formats:

1. Transduction or synthesis of signals representing both the desired sound waveform and its directionality. A signal format directly related to cardinal directions has been standardised as *A- format*.

2. Conversion, where a separate step is required, into studio and recording *B-format*.

3. Encoding for public dissemination; this coded form is defined as *C-format*.

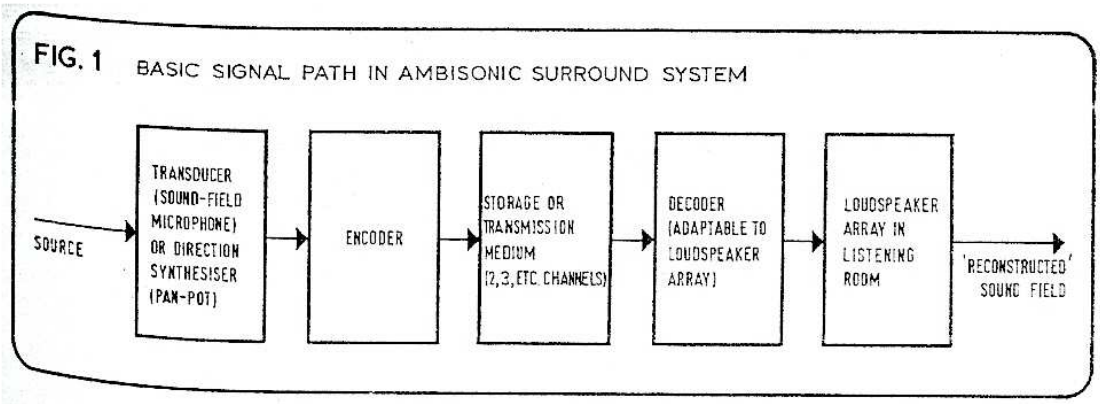
4. Decoding into signals suitable for driving loudspeakers. This *D-format* cannot be precisely standardised since it necessarily depends on the number and layout of the listener's loudspeakers; suitable options and adjustments are provided in Ambisonic decoders.

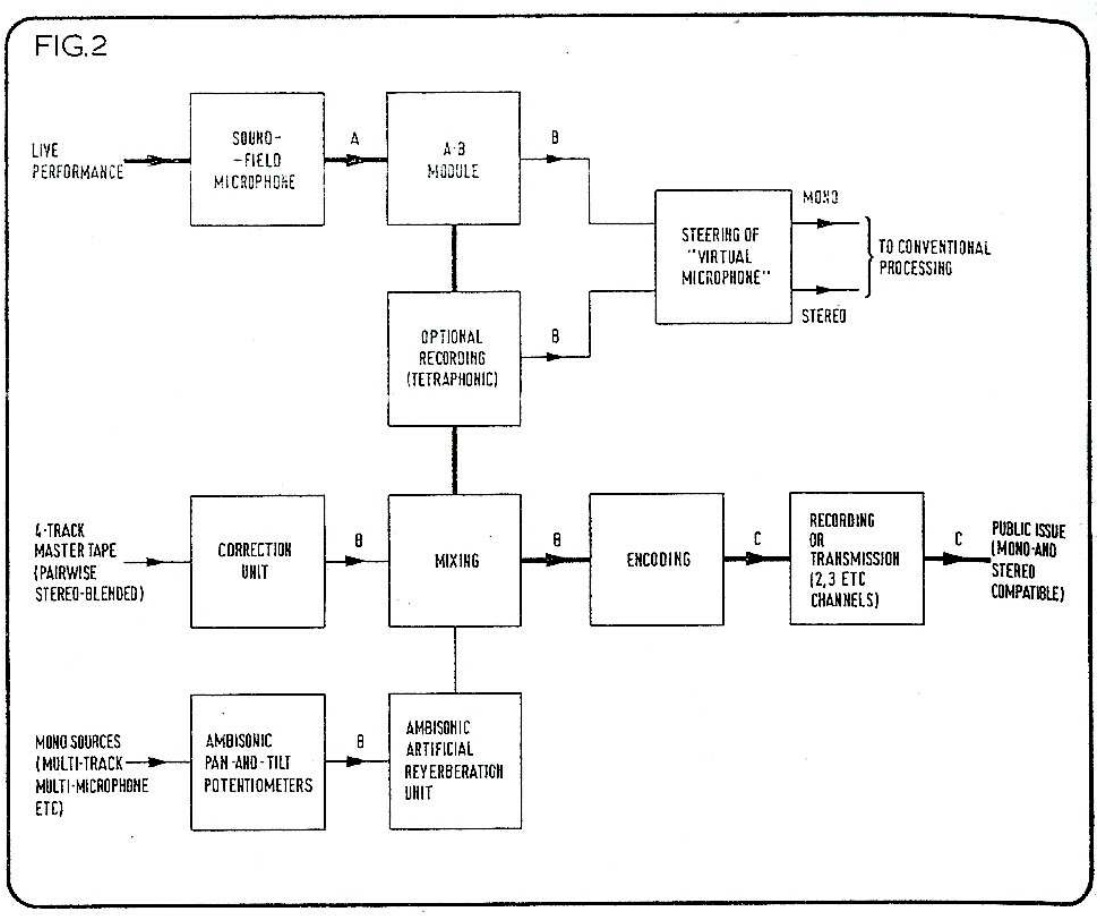
It is hoped shortly to release a set of Reports* giving details and

* Enquiries can be made Attn. P. Thompson Esq, National Research Development Corporation, PO Box 236, Kingsgate House, 66/74 Victoria Street, London SW1E 6SL

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

Ambisonics. Part one: General system description





specifications, which the present article is of course too small to contain. To the best of our knowledge the NRDC Ambisonic system alone fulfils some of the individual requirements previously set out, and is almost certainly unique in fulfilling all of them.

The basic features of an Ambisonic surround-reproduction chain are displayed in **fig. 2**, together with

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

some of the facilities that can be provided. The following aspects may be particularly noted:

1. Sound-field microphone. This is an omnidirectional microphone in the true sense, which is the *opposite* of non-directional; it characterises in a symmetrical manner the waveform and directionality of sound arriving from any direction (including vertical components). In its present first-order implementation it does so in terms of four signals corresponding to the spherical harmonic of directionality of order zero, and the three of order unity. Recordings of these signals (or their equivalent), especially in B-format, are called *tetraphonic* (there is of course no correspondence with the four signals assumed in quadrifontics). An important by-product of recording in tetraphonic mode is that the complete capture of first-order directional information enables *any combination of non-directional figure-of-eight, cardioid or hypercardioid microphones to be simulated and their directions steered* (including vertically) *after the recording session*. Although remote real-time adjustment of directionality has long been available, for example in the AKG *C24* microphone, post-session freedom of adjustment is believed to be new. In addition the virtual microphones are *truly coincident* (see Part Two), a requirement of conventional stereo hitherto unfulfilled. (It is worth noting that the sound-field microphone depends on placing capsules in accordance with Sampling Theory on a sphere, and the associated circuits are an integral part of it; the superficial resemblance of the lowest-order form to the well-known tetrahedral array of separate microphones is mainly misleading).
2. The choice of a C-format encoding standard is central to the whole design. It must be mathematically compatible at the two interfaces respectively to the source-material and to the listener's equipment so as to be capable of correct decoding. A basic tool for design and characterisation of two-channel encoding is the Poincaré-Stokes sphere, first used in this connection by Scheiber[5] and much developed by Gerzon[3]. Like a circuit diagram, it gives a geometrical picture from which system properties may be inferred (including the weaknesses of systems claimed to be 'quadraphonic').

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

It is now known that for surround-reproduction the horizontal pan-locus on the Poincaré-Stokes sphere should be nearly a great-circle. The Nippon Columbia BMX system uses a great-circle locus, and a different one is defined by the Japanese RM specifications. The Ambisonic two-channel encoding also uses a great circle but takes account of the freedom to tilt and otherwise modify it slightly to improve stereo and mono compatibility, and to implement other refinements; details will be published in due course. This work is very consonant with conclusions reached by the BBC[7].

3. Compatibility between surround and stereo playback has often been discussed without explicit realisation that original sound (including reverberation) may come form any direction, will necessarily be encoded in some manner, and therefore must be located (more or less well) in some direction in stereo playback. Stereo compatibility therefore involves making these directions as acceptable as possible while recognising that for example originally rear positions cannot be correct in stereo playback; ie it is essentially a mapping operation. Once this is realised the basic choice is seen to be largely prescribed by the need to avoid contradictions[6] and this choice is followed in Ambisonics. It appears not satisfactory to map the comparatively small stereo front-sector on to itself, letting the devil take the hindmost and sometimes encode it as front-sector.

4. Effects. Although the unique ability correctly to treat natural sound is an important feature of Ambisonics, it also provides facilities for all the usual artificial 'effects', and some not available (or which cannot even be defined) in some other systems (see Part Two).

Preserving options in the recording studio

Much of the present uncertainty concerning methods and standards for surround reproduction may be ascribed to systems having been promulgated at a stage which now appears to have been premature in relation to the formulation of adequate aims, and to the development of adequate theoretical tools for implementing these aims. The consequence is not

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

only lack of agreed standards between systems, but also a tendency to patch up deficiencies so that there is variation of encoding standards even within what is nominally one system. So-called 'logic', ie signal-controlled gain, may be seen as a response to such deficiencies, and it is not difficult to show that this stratagem is at most of only limited value.

Technological reality is however now coming to the fore, and it is not surprising that systems proposed with the benefit of later knowledge seem to have been preferred by independent critics, notably the Nippon Columbia BMX system among current two-channel commercial proposals [7,8,9]; perhaps it will not be thought too disingenuous to mention that the NRDC Ambisonic system (with which BMX is essentially compatible) has come still later. Research by various groups around the world has in fact converged on surprisingly concordant conclusions about how surround reproduction should be implemented, and there is now less doubt that the present confusion will be resolved along these lines than about exactly when it may happen.

So long as uncertainty remains, it is prudent for any studio to record in a form that is likely to remain useable no matter which viable standards are eventually adopted by the industry. Part Two consequently lays stress on preserving options in this way.

References

1. Gerzon, M.A. 'Surround–sound psycho-acoustics', Wireless World, December 1974, p 483.
2. 'Regular Matrixed Disk Records', Japan Phonograph Record Association, Engineering Committee Standard, dated Tokyo, March 25, 1972.
3. Gerzon, M.A. 'A Geometric model of 2-channel 4-speaker matrix stereo systems' (in press).
4. 'Standards for Quadraphonic disc records'. United States RIAA, Bulletin No E7, 12 September 1973.
5. Scheiber, P. AES Journal Vol **19**, p 835 (1971).
6. Fellgett, P.B. 'Directional Information in Reproduced Sound' Wireless World Vol **78**, p 413 (1972).
7. Ratliff, P.A. 'Properties of hearing related to Quadraphonic Reproduction' BBC Research Department Report 1974/38.

Crompton, T. W. J. (based on work by) 'Subjective

Reproduced from *Studio Sound*, Vol. 17, pp 20-22, 40 (August 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

performance of various quadraphonic matrix systems. BBC Research Department Report 1974/29.
8. Shorter, G. 'Matrix Decoding'. Hi-Fi News & Record Review Vol **20**, No. 3, March 1975.
9. Holmes, P. 'Elektor Vol 1' Dec 1974 p 33.