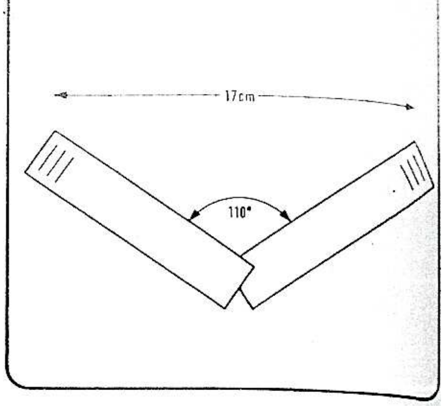
Dummy Head Recording

Michael Gerzon

Recently, there has been a lot of interest in dummy-head or binaural recording for reproduction via headphones. This has been presented by some as 'the answer to quadraphony', and some ill-informed comment has thoroughly confused people as to the advantages and disadvantages of binaural techniques. Issues and available information are summarised, together with an indication of areas of doubt.

FIG.1 ORTF TECHNIQUE



FIRST, IT SHOULD BE emphasised that binaural recordings, i.e. recordings made for reproduction via headphones, contain three main types of sound-localisation cue that is absent from conventional stereo, and that there has been some conflict of opinion as to which cue is most important. The first cue is that of time delays between the ears. It is clear that a sound from, say, the left will arrive at the left ear before the right ear. For a sound on the extreme left, this time delay at the right ear is about 0.62 ms, and for a sound arriving from (say) 30° from the left of front (or back, or up, or down), the time delay is about 0.24 ms. Clearly, the time delay cue cannot distinguish front from behind from above from below. All it indicates is the angle of arrival of the sound from the axis of symmetry of the ears. One technique of binaural recording makes use mainly or only of this cue, and that is the ORTF technique of using a pair of microphones spaced apart by about 17 cm. In order to improve compatibility with stereo loudspeaker reproduction, the microphones are directional ones pointing to the left and right respectively, angled about 110° one from the other (see **fig. 1**). For this application, cardioids are to be preferred, as the anti-phase lobes of hypercardioids

Reproduced from *Studio Sound*, Vol. 17, pp 42-44 (May 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk

tend to give exaggerated width requiring a smaller spacing.

It is important that the spacing not generally exceed 17 cm, as otherwise the time delays are too large and everything is concentrated at the extreme left or right. Certainly, ear-spaced binaural recording gives much sharper defined images via headphones than ordinary stereo, even when recorded with a pair of omni-directional microphones. An indication of the importance of time delays between the ears is given if one makes one channel of a binaural recording 10 dB louder than the other. Surprisingly, the image shift thus caused amounts at most to a few degrees, since the time delays are unaltered. Time delays around 0.5 ms are much less important in speaker reproduction, and this leads to the possibility of an amusing paradoxical recording in which a sound appearing on the left via headphones appears on the right via speakers. Simply record a sound on the left of a pair of omni mics spaced 17 cm apart, and turn up the gain of the right channel about 10 dB.

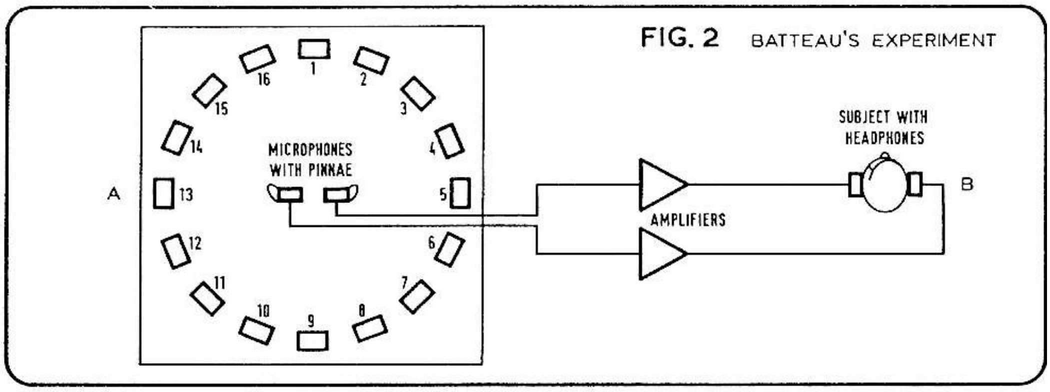
The second cue for headphone reproduction is the fact that the head casts an acoustical shadow across each ear for sounds from the opposite side. This effect is significant only in the treble (above about 500 Hz). Many early workers in binaural sound, such as de Boer and Blumlein, considered that this head obstruction effect was important in sound localisation, and some experimenters still do. However, as will be explained, the evidence seems to suggest that this is by far the least important cue. Indeed, experiments in which a mono sound is fed to both ears but with differing gains show that a relative gain of as much as 15-23 dB is required to create an illusion of a sound coming from 45° from the front, which is much more than the difference in the level at the two ears caused by a live sound from this direction. Moreover, if the ears have just previously been exposed to sounds with natural time delays between the ears, such a panned mono sound can seem to come from only about 15° off front. In other words, not only do the ears make poor use of differences in level, but when they are provided with other cues, they almost entirely disregard these level differences.

The third cue is the effect of the pinnae, which is the name for the flaps on the ears. The various ridges on the pinnae reflect and refract the sound waveform before it enters the ears, and the coloration thus produced varies according to the direction of arrival of the sound. This coloration, which mainly affects frequencies above 5 kHz, is now known to be of vital importance to the ears in localising and positioning sounds, although the way in which this coloration is used by the ear to provide information is still not understood.

An intriguing experiment of Batteau (described in *ref. 1*) demonstrates this in no uncertain fashion. In one room, he set up 16 loudspeakers (see **fig. 2a**) in a circle around a pair of microphones spaced apart by ear distance. The outputs of these were fed to a subject sitting in another room via headphones. The various speakers were then fed with sound and the subject was asked which of the 16 directions the sound appeared to be coming from. This test was performed both using ordinary omni microphones, and with microphones fitted with accurate replicas of human pinnae, but with no dummy head used in either case. When no pinnae were used, the subjects found it difficult to localise the sounds, assigning them to more-or-less random positions. However, with the pinnae fitted to the microphones, localisation was correct with no confusion between front and rear.

Other experiments have also demonstrated that pinnae are of vital importance for correct localisation. Roffler and Butler (*ref. 2*) describe experiments in which a subject's head was fixed, so that he could derive no clues from head movements, and in which a sound source was moved in the plane of symmetry of the subject's head, so that it could be above, below, in front or behind. Since the sounds reaching the two ears is then identical, conventional theories of stereo hearing would suggest that height effect cannot be heard under these conditions. However, Roffler and Butler found that a change of the sound source elevation as little as 5° could be clearly heard. On the other hand, if the subject wore a 'pinna mask' which covered up the pinnae but allowed sound to enter the ears, then no change of elevation could be heard.

So we see that the pinnae play an essential role in locating sounds, and that they should therefore be accurately modelled (preferably by taking moulds from human pinnae) if used with a dummy head. It seems that most of the recent commercial dummy head recordings have used inadequately-accurate pinnae for optimal effect. We also see that the complication of an actual dummy head between the microphones-with-pinnae may be omitted, thereby improving the visual appearance of the microphones as well as reducing some of the



coloration if the sound is played via speakers. Alternatively, a very idealised 'dummy head', such as a simple baffle to separate the microphones (as suggested by Blumlein, *ref 3*), may be used.

Dr Edmund Rolls, of the Department of Psychology, University of Oxford, has recently been conducting experiments in dummy head recording, using small microphones placed in the ears of actual people (although one conjectures that they may resent being termed "dummies"). This microphone technique, so purist that advocates of Blumlein technique must blush with shame, is capable of giving very superior binaural results, as would be expected with such accurate dummy heads. Recognising the importance of the pinnae described above, Dr Rolls has suggested a simple, ingenious and effective method of reproducing dummy head recordings via loudspeakers.

The trick is to reproduce the dummy head recording via stereo loudspeakers placed either to each side of the listener (A in **fig. 3**), or at least angled widely apart (B in **fig. 3**). I have found that angles θ (see **fig. 3**) of more than 110° work well. The listener listens wearing a pinna mask. (For listening tests, it is adequate to use the hands to cover the back part of the pinnae.) Since the sound has been past pinnae once during the recording, and since it is prevented from going over them again by the pinna masks, the ears hear just the pinna coloration inherent in the recording, and hence hear a correct directional effect, including sounds form behind or above. I have used this to demonstrate dummy head recordings to an audience of about a dozen via loudspeakers.

Correct localisation is not the only benefit produced by pinna coloration. It is well-known that headphone reproduction always gives the effect of in the head localisation (ihl). This has been

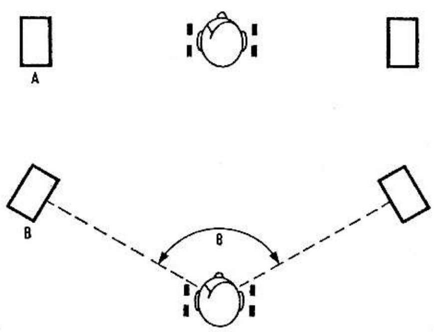


FIG. 3 LOUDSPEAKER LAYOUTS FOR LISTENING TO BINAURAL RECORDINGS WHILE WEARING A PINNA MASK.

explained as being due to the fact that a dummy head cannot move in the original sound field in the same way as the listener's head is moving, and it has been supposed that it is the information produced by such movements that prevents ihl and allows front and back to be distinguished. While head-movement information is undoubtedly of some importance in these regards (see *ref.* 4), the pinnae also are capable both of localising sounds (as we have seen) and of externalising them outside the head, without any help from head movements. Thus the conventional explanation of ihl is wrong (see also *ref.* 5, if you can read German).

A dramatic illustration of the ability to place sounds outside the head is obtained if one takes one channel only of a good binaural recording, and feeds it to *both* earpieces. Despite the fact that both ears are now hearing the same thing, the pinna coloration still allows front and back to be distinguished to some extent. Even more intriguing is what happens if such one-eared recording is made to pick up a sound to the side of the dummy head. Since the information reaching the two ears of the listener is identical, it is impossible for him to place the sound at either side, and it is difficult to say precisely where it is. Yet despite this, the sound is heard as being definitely external and not in the head at all! The experiments of Batteau mentioned above showed that this externalisation occurred if no dummy head was used so long as pinnae were affixed to the microphones.

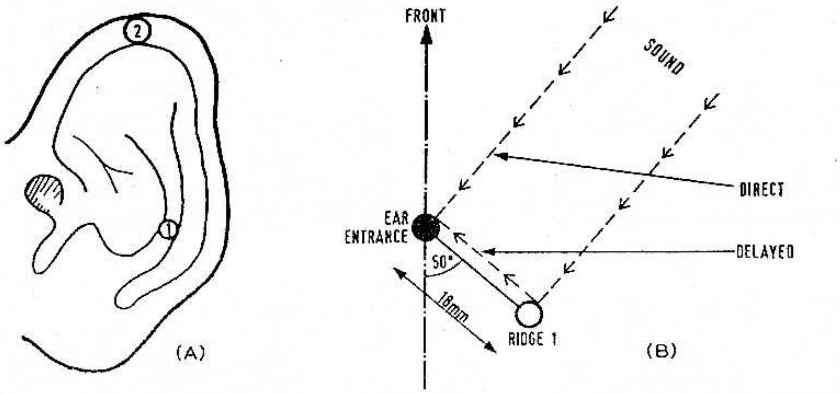
However, dummy head recording is not without its serious problems, both in its imperfections and in the technical and commercial problems. The worst problem, assuming that an accurate dummy head is used (or at least accurate pinnae), is that the least accurately defined positions tend to be at the front, just where accuracy of location is most required. In the absence both of the visual cue present live or the cues given by the effects of small head movements, frontal sounds given half a chance tend to appear to be either in the head or even slightly behind the listener. This pulling in of the frontal sound stage is disliked by most listeners, who find difficulty in being sure that frontal sounds are front or back, although back sounds are quite unambiguously at the back. If the recording contains strong clues as to when a sound is at the front (eg marked differences in room acoustics, or a commentator telling you where he is), then the ambiguity disappears. This is why on the Sennheiser Dummy Head recording No. 1 (see *ref. 6*) it is better that you listen first to the German side (assuming *now* that you *don't* understand German!) before listening to the English.

This tendency of front sounds to be localised behind is not unique to dummy head; recording anyone with practical experience of quadraphony will have experienced similar difficulties. However, with a well-made quadraphonic or ambisonic recording, suitably reproduced, the ability to move one's head often provides sufficient extra information to lock sounds at the front without ambiguity.

We can obtain some understanding of why front sounds are so unstable, and of why headphone reproduction tends to pull sounds behind the listener, if we study the effects of the pinnae in more detail. If we examine the pinna (**fig. 4**), we see that there are two main ridges from which incident sound is reflected or refracted before entering the ear, marked **1** and **2** in **fig. 4**. The effect of these ridges is for a sound impulse to arrive at the ear followed a few tens of microseconds later by delayed impulses reflected off the ridges. The delays, of course, depend on which direction the sound arrived from in the first place. These delays have been measured by Batteau and others (see *ref. 1*) for various sound directions and the results are shown in **fig. 5.** This shows the delay of the reflected impulse after the arrival of the original impulse both for sounds (**fig. 5a**) in the horizontal plane, and (**fig. 5b)** in the side to side vertical plane. It will be noted that the vertical displacement of sounds causes much larger delays (of the order of 200 μs) than horizontal displacements, which cause delays only of the order of 50 μs.

Because 50 μs is the duration of only half a cycle at a frequency of 10 kHz, the ear gets rather little information about horizontal position from the pinna effect, and so we would expect ambiguities to be worst in this plane. Moreover, sounds from the back involve no delayed impulse reflected from ridge *1* in **fig. 4**. Thus, if a sound is not perceived as having a delayed impulse delayed by around

FIG. 4 (A) THE PINNA, SHOWING REFLECTING RIDGES 1 AND 2 THAT RESPECTIVELY DETERMINE HORIZONTAL AND VERTICAL SOUND DIRECTION THROUGH THE DELAY BETWEEN THE DIRECT SOUND AND THAT REFLECTED FROM THE RIDGES SHOWN IN (B)



15-100 µs, then it will be heard as coming from behind.

This explains why normal stereo reproduced via headphones tends to seem to be slightly behind the listener in many cases, because such sound lacks any delayed impulses. However, ordinary stereo via headphones is not very convincingly right behind the listener, but rather in his head, which presumably is a result of such a sound not having the second vertical information reflection from ridge 2 (**fig. 5b**) either. One presumes that if suitable delayed impulses according to **fig. 5** were supplied in such cases, then the headphone reproduction would tend to be externalised.

However, since the short delays of the first delayed impulse for horizontal frontal sounds are difficult to disentangle from the complexities of the sound waveform, one expects the ear to miss the presence of the first delayed impulse altogether in many cases. When the delayed impulse is not detected by the ear, one would expect the ear to assume that the sound is behind the listener, since back sounds lack such a first delayed impulse together (see **fig. 5a**). This explains why back sounds are always heard at the back, but front sounds tend to be heard at the back with some degree of uncertainty. The much larger delays involved in vertical discrimination (**fig. 5b**) are much easier to detect and thus give more reliable localisation.

We thus see that the poor localisation of front sounds is inherent in headphone reproduction. For live sounds, the extra clues derived by moving one's head seem to be vital in confirming that a sound is in front.

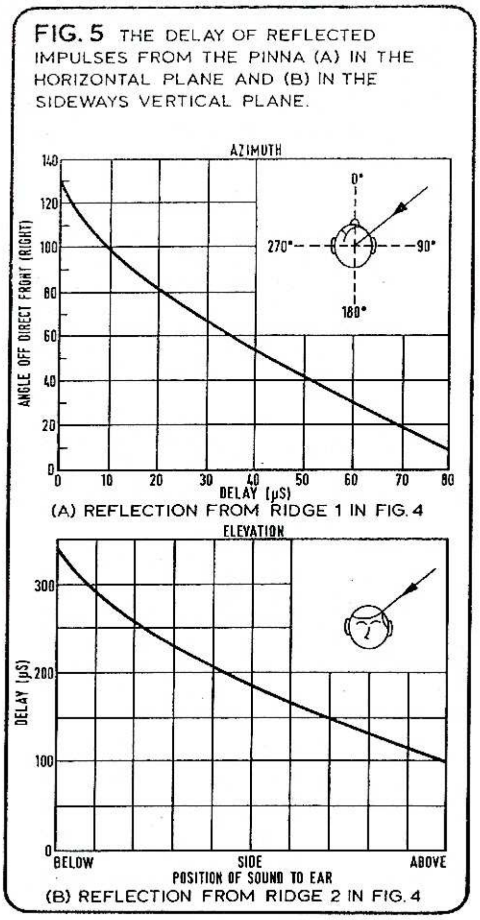


FIG. 5 THE DELAY OF REFLECTED IMPULSES FROM THE PINNA (A) IN THE HORIZONTAL PLANE AND (B) IN THE SIDEWAYS VERTICAL PLANE.

(A) REFLECTION FROM RIDGE 1 IN FIG. 4

(B) REFLECTION FROM RIDGE 2 IN FIG. 4

We do not have a complete understanding of how the delays caused by reflection from the pinna are actually pulled out of the sound waveform information by the ear. The processing involved is still something of a mystery, so that the above explanation must still be regarded as incomplete. In effect, we are saying that *if* the ear can use the delay information, then we can explain the behaviour of binaural recordings, but we don't know *how* the ear can use this information.

It is a matter of experience that if one adds the two channels of a good binaural recording together to get mono, then the overall quality of the mono obtained is very poor, certainly poorer than the mono consisting of one ear channel only. As explained earlier, the one–eared mono fed to both ears during reproduction still retains all the pinna reflection information required to externalise sounds correctly. The sum-signal mono, however, combines two separate signals, one from each ear, each with its own time delays. The extra time delays thus introduced not only cause unpleasant signal colorations, but also so confuse the listener that no sense of externalisation is obtained. Thus we see that binaural recordings inherently have very poor mono compatibility, which virtually rules out the use of binaural recording techniques for most public broadcasting applications, unless the majority of mono listeners are to be sacrificed. (Indeed, some say that the ORTF are doing just that with their preferred classical microphone technique.)

When reproduced via loudspeakers, binaural recordings also tend to give a poor stereo effect, which is unstable in the bass and rather unsharp and colored in the treble. This is partly caused by the very frequency-dependent polar diagram of dummy heads in the treble. Since we have seen that the precise form of the dummy head is unimportant providing that the pinnae and the intermicrophone spacing is correct, one could presumably choose the form of the microphone baffling very carefully so as to optimise the sharpness and quality of stereo speaker reproduction in the treble. Clearly, the design of a suitable intermicrophone baffle is very complex, and is probably as much an art as conventional loudspeaker design. For this reason, we have to leave to the interested reader the problem of designing a dummy head baffle with good stereo compatibility and retaining good binaural reproduction.

One might consider getting a good stereo image by fitting an ear-spaced pair of *directional* microphones (such as those of **fig. 1**) with replica pinnae, but there is a serious problem with this proposal. For correct effect, all the sound should enter the microphones after first having passed over the surface of the pinna. But directional microphones obtain their directionality by having more than one entrance through which the sound gains access, and they lose this directivity if some of the sound entry points are covered up. Since the pinna only has one point at which sound is allowed to gain access, we can only use replica pinnae effectively in conjunction with omni-directional microphones.

Despite their overall effectiveness, binaural recordings are seen to pose severe problems as regards mono and stereo compatibility. Added to these problems is the poor localisation of frontal sounds binaurally (unless additional clues are given to the listener), and the difficulty of achieving a binaural mixdown of multimic recordings.

References

1. William Mills, A. *Auditory Localization*, which forms chapter 8 (pages 303-348) of *Foundations of Modern Auditory Theory* (ed. J. V. Tobies) Volume II, Academic Press, New York, 1972. Mills' article is the best survey that I know of concerning how stereo hearing works.
2. Roffler, S.K. & Butler, R.A. *Factors that Influence the Localization of Sound in the Vertical Plane*, Journal of the Acoustical Society of America, Volume 43, pp. 255-9 (1968).
3. Blumlein, A.D. UK Patent 394, 325 (1931).
4. Bernfeld, B. Psychoacoustics of sound Localization – A New Approach, Audio Engineering Society Preprint, 47th Convention, Copenhagen, March 1974.

Reproduced from *Studio Sound*, Vol. 17, pp 42-44 (May 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk

5. Plenge, G. Uber das Problem der Im-Kopf-Lokalisation (in German), Acustica, Volume 26, pp. 241-252 (May 1972).
6. Sennheiser, Kunstkopf-Stereofonie, 45 rpm record.

Further references

The above Sennheiser record and the Sennheiser Dummy Head Stereo Record 2 are both available from Hayden Laboratories Ltd, Hayden House, 17 Chesham Road, Amersham, Bucks HP6 5AG. An article on dummy head recording has appeared in the September 1974 'Wireless World'.

Two more recent articles of interest on dummy head recording have been published in French and German:

Wilkens, H. *Kopfbezugliche Stereophonie–Ein Hilgsmittel fur Verglach und Beurteilung verschiedener Raumeindrucke,* 'Acustica' Volume 26, pp.213-221 (April 1972).
Céoen, Carl. *La Perception Periphonique–Casque ou Haut-parleurs?,* Conférences des Journées D'Etudes, Paris Festival International du Son Haute Fidelité Stéréophonie Editions Radio, Paris, 1974.

The Principles of Quadraphonic Recording
Part One: Are Four Channels Really Necessary?

By Michael Gerzon

THE recent growth of interest in quadraphony (i.e. sound reproduction via four loudspeakers) has encouraged the belief that four recording channels are necessary for a full quadraphonic effect. The author has recently published an account[1] of a method of four-speaker reproduction of ordinary two-channel stereo that can give a convincing all-round-sound effect, despite certain theoretical and practical limitations. This poses the problem of whether one really needs to record or transmit four channels of audio for four speaker quadraphonic reproduction.

The aim of this article is to give an elementary theoretical analysis which indicates that *three* recorded channels should be quite adequate for quadraphonic reproduction. The uses, advantages and limitations of this are discussed, and formulae are given which indicate how three-channel recordings can be reproduced via four loudspeakers, and how four-channel recordings can be reduced to three channels. The second part of this article is devoted to the use of these considerations in obtaining a system of *Periphonic* (Greek: *peri-*, around) sound reproduction, i.e. the reproduction of sound in all spatial directions, from in front, each side, behind, above and below.

While the author has used certain advanced mathematical techniques in deriving the material in this article, all the results are here stated only in terms of very elementary mathematics, and physical reasons are given for most of the phenomena. It is hoped that the information here will prove useful in designing multi-channel recording and reproducing systems, quadraphonic pan-pot circuits, and in other applications.

It is generally accepted that three speakers are not adequate for good surround sound, due to the limited listening area and the wide angle between the loudspeakers. This has led many people to assume that, because four loudspeakers are necessary for surround sound, therefore one needs to record four channels. The author has shown[1] that even two-channel recordings can be made to give a genuine surround sound (albeit with some defects), and this suggests that three channels might be quite sufficient to convey all the information required for quadraphonic reproduction.

There are strong arguments in support of deriving the sound for four loudspeakers from only three channels of recorded sound. The primary

Reproduced from *Studio Sound*, Vol. 12, pp. 338-342 (August 1970) by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

purpose of four speaker reproduction is to reproduce music realistically. It is generally recognised that the most natural recordings are obtained by coincident microphones, rather that spaced or multi-mike techniques. (It is true that the latter techniques may produce a more spectacular, 'pleasing' or analytic sound, but it is not the purpose of the present article to argue matters of taste.)

It may be thought that placing four coincident microphones with, say, cardioid directional characteristics pointing in different directions will give a reasonable four channel sound. For conventional quadraphony, which only conveys horizontal directional information, these microphones will normally have their axes pointing horizontally. However, it may not be generally known that, for microphones whose axes point horizontally, there are only three linearly independent microphone directional characteristics. Put another way, given four coincident microphones whose axes point horizontally, it is always possible to derive the audio output of at least one of the microphones by matrixing the outputs of the other three microphones together in suitable proportions. (Technically, this is expressed by saying that 'the space of horizontal microphone directional characteristics is three-dimensional'. This arises from the fact that all conventional microphone directional characteristics are linear combinations of zero and first order spherical harmonics. In future, high quality microphones may be developed whose directional characteristics involve second order spherical harmonics. Four such microphones would be capable of recording four independent channels for four speaker reproduction.)

This means that whenever a coincident microphone technique is used for sound to be reproduced over four loudspeakers arranged horizontally around the listener, as in **fig. 1** for example, only three microphones are actually needed to obtain all the audio information. The sound fed to each of the four

FIG. 1 QUADRAPHONIC SPEAKER LAYOUT

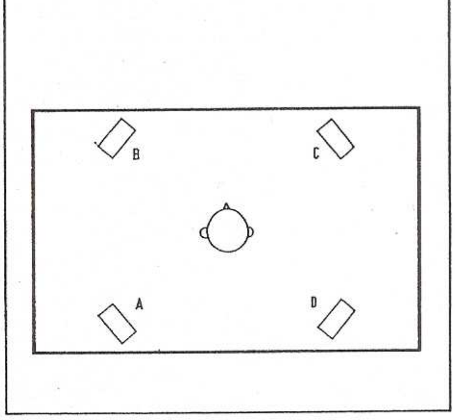
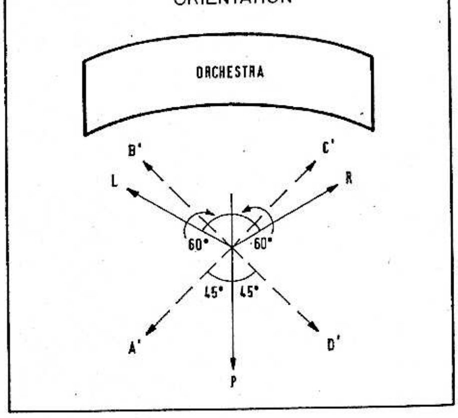


FIG. 2 COINCIDENT MICROPHONE ORIENTATION



speakers can be derived by suitable matrixing of the three microphone signals. Thus, for many purposes, only three recorded channels are needed to convey all the information reproduced by the four loudspeakers.

In order to give this assertion concrete form, it is first necessary to describe the layout of the reproducing loudspeakers. In the standard quadraphonic system, the four loudspeakers are to the rear left, front left, front right, and rear right of the listener, in a square as illustrated. It is convenient to label these loudspeakers A, B, C and D respectively, and to use these letters to indicate the four audio signals which must be transmitted to the four loudspeakers.

Consider three identical coincident microphones with, say, cardioid directional characteristics pointing in the three directions indicated by solid arrows in **fig. 2**. Thus, one microphone points 60^{o} to the left (giving an output L), one points 60^{o} to the right (giving an output R), and one points backward (giving an output P). According to what was said above, it is possible to derive all other horizontally-pointing microphone outputs from L, R and P by matrixing. Also note that the signals L and R form a good stereo signal. The signals A′, B′, C′, and D′ fed to the four loudspeakers A, B, C and D could well be the outputs that would be given by cardioid microphones pointing in the four directions (broken arrows) labelled A′, B′, C′ and D′ in fig. 2, i.e. 135^{o} to the left, 45^{o} to the left, 45^{o} to the right, and 135^{o} to the right. Rather messy trigonometric computations show that A′, B′, C′ and D′ may be obtained from L, R and P by the matrixing described in **Table 1**.

If the microphones pointing in the directions L, R and P of fig. 2, have a given identical hypercardioid directional characteristic, then the signals A′, B′, C′ and D′ obtained by the matrixing of Table 1 will be the signals which would be obtained from identical hypercardioid microphones pointing in the directions A′, B′, C′ and D′ of fig. 2.

Thus, for coincident microphone recordings, we need only record the signals L, R and P given out by identical cardioid or hypercardioid microphones pointing in the directions of the three solid arrows in fig. 2. The signals fed to the four loudspeakers can be obtained by the matrixing of Table 1. This illustrates the principle that good quadraphony can be recorded using only three channels, although four loudspeakers are needed to reproduce it.

Having shown that coincident microphone recordings need only three channels, the question naturally arises whether other types of four-speaker audio can be recorded or transmitted using only three channels. Most of the four-channel material recorded at the moment consists of more or less independent sound on each channel, due to the use of widely spaced microphones. There is one obvious way of reducing genuine four-channel recordings to three channels. This is to derive the three signals, L, R, and P that would be picked up by imagined microphones pointing along the solid arrows in fig. 2 if the sound A, B, C, and D of the four channels were played through loudspeakers in the four directions A', B', C', and D' in fig.2. While we are only imagining 'make-believe' microphones picking up imaginary loudspeakers, the computation of the signals L, R, and P picked up by these microphones does give us a prescription for reducing four channels to three. Unfortunately, the signals L, R, and P thus obtained depend on the choice of directional characteristic of these imagined microphones. This illustrates the fact that there is no unique way of reducing four-channel material to three channels.

Table 2 gives the matrixing that converts four channels to three assuming that our fictional microphones are hypercardioids with a null response 135° off axis, i.e. with a 15.31 dB front-to-back ratio.

Table 1
Converting three channels L, R, P to four channels (see fig. 2)

$$A' = 0.506\ L - 0.311\ R + 0.805\ P$$
$$B' = 0.977\ L + 0.161\ R - 0.138\ P$$
$$C' = 0.161\ L + 0.977\ R - 0.138\ P$$
$$D' = -0.311\ L + 0.506\ R + 0.805\ P$$

Table 2
Converting four channels A, B, C, D to three channels (in the manner of 135°-null hypercardioid microphones).

$$L = 0.418\ A + 0.724\ B + 0.194\ C - 0.112 D$$
$$R = -\ 0.112\ A + 0.194\ B + 0.724\ C + 0.418\ D$$
$$P = 0.612 A + 0.612\ D$$

Table 3

Reproduction of four channels transmitted via three channels (as in Tables 2 and 1)

A' = 0.739 A	+ 0.306 B - 0.127 C + 0.306 D
B' = 0.306 A	+ 0.739 B + 0.306 C - 0.127 D
C' = -0.127 A	+ 0.306B + 0.739 C + 0.306 D
D' = 0.306 A	- 0.127 B + 0.306 C + 0.739 D

Table 4

Converting four channels A, B, C, D, to three channels (in the manner of cardioid microphones).

L = 0.445 A + 0.695 B + 0.262 C + 0.012 D
R = 0.012 A + 0.262 B + 0.695 C + 0.445 D
P = 0.604 A + 0.104 B + 0.104 C + 0.604 D

The four-channel material, with signals A, B, C, and D, can be transmitted or recorded via three channels, L, R, and P by the recipe of Table 2. The signals A', B', C' and D' for the four loudspeakers can be rederived by the recipe of Table 1. Of course, something is lost in the process of reducing four channels to three. **Table 3** give the signals A', B', C', and D' emerging from the loudspeakers in terms of the original signals A, B, C, and D, after these have been reduced to three channels L, R, P and been reconstituted according to Tables 2 and 1. It will be seen that the signal B' (say) emerging from loudspeaker B consists mainly of the signal B, plus the signals A and C each attenuated by 7.66 dB, plus an out-of-phase crosstalk of the signal D attenuated by 15.31 dB. However, this crosstalk should not seriously affect the directional characteristics of the reproduced sound, as satisfactory results are obtained with the much higher degree of crosstalk obtained when two-channel stereo is reproduced via four loudspeakers.[1]

Thus, while the three-channel transmission or recording of four-channel material causes little loss of directional effect, the sound does tend to spread out among the loudspeakers. The reconstituted sound may be thought of as a spatially blurred version of the original. This is illustrated by what happens to coincident microphone recordings. Suppose that A, B, C, and D were picked up by coincident hypercardioid microphones with nulls 135° off axis pointing along the broken arrows of fig. 2. Then the reproduced signals A', B', C' and D' obtained by reducing to three channels and reconstituting as in tables 2 and 1 will be the sound that would be picked up by *cardioid* microphones pointing along the broken arrows of fig. 2.

Thus a certain amount of information is lost even with coincident microphones in the process of converting from four channels to three, and back to four again. For both spaced and coincident microphone recordings, the degree of loss depends on the chosen imaginary microphone characteristic used to convert four channels to three. **Table 4** gives the reduction from four channels to three when the fictional microphones are cardioids. The reproduced channels are then as in **Table 5**, in which the degree of sound spreading on to adjacent channels is greater that in **Table 3**. However, crosstalk on to the opposite channel is eliminated, which is a desirable requirement with spaced microphone recordings.

Table 6 gives the reduction from four channels to three when the fictional microphones are hypercardioids with a null 120^{o} off axis (i.e. with a 9.54 dB front-to-back ratio.) Such a three channel signal will not be very suitable for reconversion to four channels by the recipe of **Table 1** in many cases, as the reproduced signals will be as in **Table 7**, in which the crosstalk on each channel from the opposite channel is a rather excessive -9.54 dB. However, if the signals A, B, C, and D originate from a coincident microphone system, then the reproduced signals A', B', C' and D' will be the same as A, B, C, and D.

This shows that, if four channels are reduced to three, conversion using an imaginary 120^{o} -null hypercardioid (Table 6) works best with coincident microphone recordings, conversion using a fictitious cardioid (Table 4) has desirable properties for spaced microphone recording in which no pan-potting is used, and conversion using an imaginary 135^{o}-null hypercardioid (Table 2) is a good intermediate compromise between these conflicting requirements.

Sounds which are pan-potted exactly half-way between two adjacent speakers can be conveyed without loss via three channels, as long as the conversion to three channels uses a fictitious 120^{o}-null hypercardioid (**Table 6**). For example, consider an audio signal X pan-potted halfway between speakers A and B. Then the four-channel signal is A = 0. 707X, B = 0.707X, C = 0, D = 0. After conversion to three channels via **Table 6**, one has L = 0.787X, R = -0.079X, P = 0.354X. The matrixing of **Table 1** reconstitutes the signals

A' = 0.707X, B' = 0.707X, C' = D' = 0. Similarly, a signal pan-potted half-way between speakers B and C, C and D, or D and A can be transmitted via three channels by putting, respectively, L= 0.604X, R = 0.604X, P = -0.146X or L = -0.079X, R = 0.787X, P = 0.354X or L = 0.104X, R = 0.104X, P = 0.854X.

All the above indicates that reasonable quadraphonic sound can be conveyed via three channels. It is therefore worthwhile to examine the various domestic recording and transmission media to see what advantages three- channel recording might have over four-channel recording.

Take the problem of transmitting four-speaker sound via FM radio. The author has recently proposed[2] a system of broadcasting three channels which involves the use of no subcarrier frequencies not already use for stereo. This system inherently has a much better noise performance, and causes less adjacent-station interference than any four-channel FM multiplex system. In the case of FM broadcasting, significant improvements in technical quality can thus be obtained if quadraphonic sound is conveyed via only three channels.

The use of only three channels also has significant advantages in domestic tape recording. Current proposals for quadraphony involve recording four channels side-by-side on 6.25 mm (quarter inch) tape. The width of each tape track on the best four channel tape heads is only about 1 mm. This means that the outer tracks are badly affected by dropout, and the hiss level is rather high.

Table 5

Reproduction of four channels transmitted via three channels (as in tables 4 and 1)

A' = 0.707 A + 0.354 B + 0.000 C + 0.354 D
B' = 0.354 A + 0.707 B + 0.354 C + 0.000 D
C' = 0.000 A + 0.354 B + 0.707 C + 0.354 D
D' = 0.354 A + 0.000 B + 0.354 C + 0.707 D

Table 6

Converting four channels A, B, C, D, to three channels (in the manner of 120°-null hypercardioid microphones).

L = 0.379 A + 0.733 B + 0.121 C – 0.233 D
R = -0.233 A + 0.121 B + 0.733 C + 0.379 D
P = 0.604 A - 0.104 B – 0.104 C + 0.604 D

Table 7

Reproduction of four channels transmitted via three channels (as in tables 6 and 1)

A' = 0.750 A + 0.250 B - 0.250 C + 0.250 D
B' = 0.250 A + 0.750 B + 0.250 C - 0.250 D
C' = -0.250 A + 0.250B + 0.750 C + 0.250 D
D' = 0.250 A - 0.250 B + 0.250 C + 0.750 D

If only three tracks had to be recorded, the track width would increase to 1.6 mm, which would improve the signal-to-noise ratio by at least 2 dB, and would dramatically reduce drop-out. The three-track format would also be compatible with half-track stereo recordings. Furthermore, the cost of reasonable quality multitrack heads is rather high, and three track heads would only cost about half as much as four-track heads of comparable quality.

Alternatively, it would be economic to record quadraphonic tapes using three tracks in each direction. The quality loss involved in this would be substantially less that when four tracks each way are used. Indeed, experience with 8-track cartridges indicates that eight tracks on 6.25 mm tape cause many severe problems, due to the difficulty of accurate track alignment. The wider track widths and guard bands possible with tapes using three channels each way should reduce these problems.

It is more difficult to see whether the use of only three channels gives any advantages with gramophone records, as one first has to consider how multi-channel gramophone records might be manufactured. To preserve the low cost of gramophone records, it is essential that any multichannel disc should be manufactured by the simple process of pressing a blob of vinyl, and this rules out adding channels by modulating the colour, the dielectric constant, the magnetisation, or other esoteric properties of the disc. Conceivably, two channels could be added to ordinary stereo by modulating the slope of each of the two groove walls, but there are numerous difficulties in designing a pickup to recover this information. In practice, it seems certain that additional channels will be added by modulating an ultrasonic subcarrier with frequency between 30 and 40 kHz.

Several companies are known to be working on multichannel discs using subcarriers, with reasonably promising results. Despite the very low amplitudes of such subcarriers (about the wavelength of light…) it appears that a fairly low noise level can be achieved, thanks to the ability of heated record cutting styli to reduce noise at high frequencies. By recording two modulated subcarriers, each causing a direction of stylus motion 90° from that caused by the other, four channels can be carried on one disc. The crosstalk between the subcarriers will be poor, because of the poor channel separation of pickups at high frequencies, but even this can be minimised by recording the two subcarriers 90° out-of-phase with respect to one another.

Back to seventy-eights?

However, at the current 33⅓ RPM rotation rate there is one severe problem with the subcarrier method. When the record is tracked, the pickup will produce fairly large amounts of harmonic distortion above 5 kHz. As the usual stereo tracks will be recorded at a much higher level than the subcarriers can be, distortion products of the audio will interfere badly with the subcarrier modulations. This problem can be partly overcome by using higher rotation rates, e.g. 45 or 78 RPM, but it is known that the distortion level in the vertical component of the stylus motion is larger than in the horizontal component. Thus the easiest way of reducing the distortion's interference with the subcarriers is to record the subcarriers horizontally only. But if this is done only one subcarrier can be recorded.

Thus the technical problems associated with using modulated carriers may mean that only three channels may be available on gramophone records. In such a case, it would again be desirable to convey quadraphony via three channels. Care will be needed to ensure that the polarity, phase and frequency response of the subcarrier channel matches that of the stereo channels, at least in the mid-frequency audio range, so that the matrixing of **Table 1** can be performed accurately.

While quadraphony can be conveyed over three channels, we have seen that four channels are capable of conveying sound with less spatial spreading of the sound onto adjacent channels. What, then, is the precise nature of the additional information conveyed by four channels?

In a four channel recording, there is one audio signal essentially independent of the three audio signals L, R, and P of **Tables 2, 4 or 6**, which conveys no directional information whatsoever. This is the 'focus' signal F defined by $F = ½A - ½B + ½C - ½D$.

This is the only combination of the four signals A, B, C, and D which is always zero for coincident microphone recordings. For any four channel recording, given the signals L, R, P, and F, it is always possible to rederive the original signals A, B, C, and D by means of matrixing. When the signal F is suppressed (i.e. when only L, R, and P are transmitted), a reasonable facsimile of the original directional effect can still be obtained by means of the matrixing of **Table 1**. Thus the essential difference between three and four channel quadraphony is the addition in the latter of the 'focus' signal, which conveys no directional information, but only information about how widely a sound appearing to come from a given direction is spread out among the four loudspeakers.

The question of when a four channel recording with signals A, B, C, and D is capable of being passed through three channels without alteration has a simple answer; this can be done if and only if the focus signal is zero, and the matrixing that achieves this is that of **Tables 6** and **1**.

Despite the fact that three channels are sufficient for quadraphony, commercial pressures make it likely that practical quadraphonic media will in fact convey four channels. However, we have seen that 'focus' information can be dispensed with without excessive losses of directional information. This prompts the thought that, in four-channel recordings, perhaps the focus information can be discarded, and other information smuggled into its place. The next part of this article will describe how, by this means, 'conventional' (!) quadraphonic recordings can be used to reproduce height information via suitable reproducing equipment.

REFERENCES

1) M.A. Gerzon: 'Surround Sound–From 2-Channel Stereo' Hi-Fi News, August.
2) M.A. Gerzon: 'QUART- A System of Multichannel FM Multiplex'. Hi-Fi News, May 1970.

The Principles of Quadraphonic Recording
Part Two: The Vertical Element

By Michael Gerzon

CURRENT quadraphonic systems are designed to reproduce sound from all horizontal directions around the listener, but still fail to reproduce height information. In Part One of this article, by means of considering the types of sound pick-up associated with the use of a coincident microphone technique, it was shown that only three channels were necessary for 'horizontal quadraphony'. In the following, these arguments will be extended to the reproduction of sound from all spatial directions about the listener, both horizontally and vertically. The author has christened reproduction techniques which reproduce all spatial directions 'periphonic' from the Greek prefix *peri-* meaning about, or around.

While Granville Cooper has recently described[3] a system of periphony called 'tetrahedral ambiophony', this is only one of many possible periphonic techniques. It is the purpose of this article to establish that four channels should usually be adequate to convey periphonic sound. It will further be shown that it is possible to convey a periphonic recording via four channels such that, when it is reproduced through four loudspeakers placed in a horizontal square around the listener (as in current American "horizontal quadraphony" proposals), a good conventional quadraphonic sound reproduction will be obtained. Thus the method of conveying periphony described in the following has the advantage of complete compatibility with conventional quadraphonic reproduction. A consequence of this is that the listener has a wide choice as to how complex his reproduction system is, and he may choose to reproduce the four channels over anything between three and eight speakers, according to his pocket and preferences.

First question

The first question to be resolved is why reproduce height information at all? The case against periphony has been wittily stated by Alec Nisbett, and it is worth quoting him:[4]

"I am not being totally facetious when I suggest that if God had meant us to take an interest in the vertical separation of sounds, we would have an ear on the top of our heads. Lacking such a rainwater collector, I don't see much need to feed directional information in this sense, even though it is present in the concert hall: the horizontal component is enough. Anyway, I am going to cut short this argument by saying that if you don't like the horizontal box format it's just too bad, because that's how it's going to be – there's no turning back, unless you want another big battle over standards, which would be exhausting, expensive and, I suspect, unwinnable. So everybody please agree with me."

Reproduced from *Studio Sound,* Vol. 12, pp. 380-384 (September 1970) by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

The Principles of Quadraphonic Recording
Part Two: The Vertical Element

The fallacy with the argument is the assumption that human hearing is insensitive to height information. It is well known that it is possible to perceive the elevation of a sound quite accurately by means of small unconscious head movements.[5]

In the author's experience, height information can be of great musical importance. In orchestral and choral music, a strong impression of depth is often gained due to the fact that the orchestra frequently subtends a vertical angle of a few degrees at the listener's ears; this height information is then clearly audible with one's eyes shut. Of even greater musical importance is the existence of religious and secular music in which a large organ accompanies a choir or orchestra. In this case, composers have often (perhaps not consciously?) used the fact that the organ will be placed high up above the other performers to obtain a remote, all pervading, or ethereal effect from the organ. This effect is totally destroyed by restricting the sound to the horizontal plane. It should also be mentioned that, even using the realistic coincident microphone technique, reverberation sounds curiously 'cramped' via horizontal quadraphony, due to the lack of height dimension.

Alec Nisbett's other objection, that it is undesirable and impractical to introduce more than one system of quadraphony, ceases to hold if the periphonic recording is capable of being reproduced via a conventional quadraphonic set-up. This compatibility requirement can be completely fulfilled, as will be shown in the following.

Consider a conventional four-channel quadraphonic recording with signals A, B, C, and D corresponding to the four loudspeakers placed in a horizontal square about the listener, as in **fig.1.** It was shown in Part One that a good quadraphonic sound could be obtained even if ½A – ½B + ½C – ½D was equal to zero, and methods were described to convert arbitrary four-channel recordings into a form where this was

FIG. 1 CONVENTIONAL FOUR-SPEAKER LAYOUT

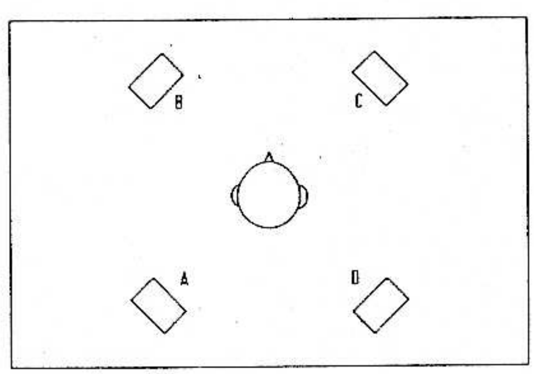
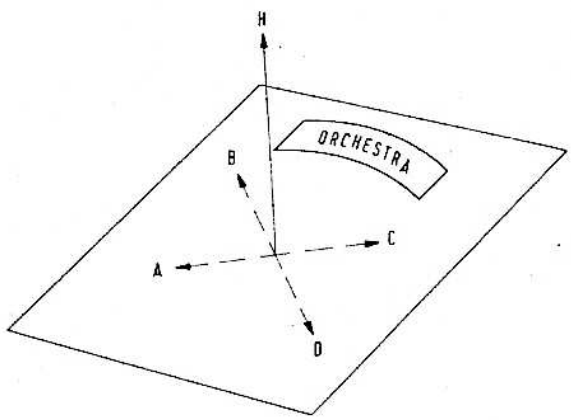


FIG. 2 PERIPHONIC COINCIDENT MICROPHONE DIRECTIONS (VERTICAL AND HORIZONTAL ASPECTS)



Reproduced from *Studio Sound,* Vol. 12, pp. 380-384 (September 1970) by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

true. Thus, in the rest of this article, we consider quadraphonic signals A, B, C, and D such that

$$\tfrac{1}{2}A - \tfrac{1}{2}B + \tfrac{1}{2}C - \tfrac{1}{2}D = 0 \qquad (1)$$

By imposing the condition (1) on our quadraphonic signals, we have produced signals that can be conveyed via only three channels, as explained in Part One.

Thus, if a four-channel recording medium is used, there is room to convey height information. Let H be a "height" audio signal, whose precise nature will be considered later. Then one can make a four-channel recording conveying the four signals

$$A^- = A - \tfrac{1}{2}H,\ B^+ = B + \tfrac{1}{2}H,$$
$$C^- = C - \tfrac{1}{2}H,\ D^+ = D + \tfrac{1}{2}H. \qquad (2)$$

The four signals A^-, B^+, C^-, and D^+ may be reproduced via the horizontal four-speaker set-up of **fig.1** without any alteration of the directional effect that would have been reproduced if A, B, C and D had been fed to the four speakers instead. The reason for this is that the 'focus' signal F′ for the four signals A^-, B^+, C^-, and D^+ is given by

$$F' = \tfrac{1}{2}A^- - \tfrac{1}{2}B^+ + \tfrac{1}{2}C^- - \tfrac{1}{2}D^+ = -H. \qquad (3)$$

As was shown in Part One, altering the focus signal in a four-channel quadraphonic recording does not alter the reproduced directional effect, but only affects the degree of crosstalk of a sound on to the other channels.

Thus, if we start off with a conventional quadraphonic recording whose signals A, B, C, D obey condition (1) and if we smuggle in a "height" signal H as in formulae (2), then we have four signals which reproduce well via a conventional four-speaker set-up, but which contain height information as well as horizontal information.

Of course, this has not yet shown either how to record the height information, nor how to reproduce it. As in Part One, examining coincident microphone recording techniques is very revealing. Only microphones with horizontally-pointed axes were then considered but we must now consider coincident microphones with axes pointing in any direction. Standard mathematical theory reveals that there are only four linearly independent microphone directional characteristics. Put another way, given five coincident microphones, it is always possible to derive the audio output of at least one of the microphones by matrixing the outputs of the other four microphones together in suitable proportions.

This means that no matter how many loudspeakers are used to reproduce the sound, only four microphones are needed to pick up all the periphonic audio information that can be obtained from coincident microphones. (Of course, this no longer holds if the microphones are not precisely coincident. Neither will it hold if new types of microphone directional characteristics are developed.) The sound fed to each of the reproducing loudspeakers may be obtained by a suitable matrixing of the signals from these four microphones.

The Principles of Quadraphonic Recording
Part Two: The Vertical Element

The nature of the four signals A^-, B^+, C^-, and D^+ will now be investigated for coincident microphone periphonic recordings. This investigation yields useful information on how to reduce spaced microphone periphonic recordings down to four channels.

Clearly there are two ways of looking at the recording of the signals A^-, B^+, C^-, and D^+ via coincident microphones. We can either consider the microphone directional characteristics required to pick up the signals A, B, C, D, and H, or we can consider the microphone directional characteristics required to pick up the signals A^-, B^+, C^-, and D^+ . The first way of considering coincident microphone periphonic recording has the advantage that it lays special emphasis on the 'horizontal quadraphony' component A, B, C, D of the periphonic signal, while the second approach reveals the essentially three-dimensional geometric nature of periphonic recording.

As the signals A, B, C, and D correspond to horizontal quadraphonic sound they must be the signals obtained by four coincident identical cardioid or hypercardioid microphones whose axes point in a horizontal direction along four directions at right angles to each other as illustrated in **fig. 2**. Four such signals obey condition (1), as observed in Part One. Sounds originating from horizontal directions around the microphones clearly contain no height information, and so require that the signal H equals zero for such sounds. The only microphone directional characteristics which gives no output for sounds from all horizontal directions is a vertically oriented 'figure-of-eight' microphone. It is convenient to assume that the 'positive' lobe of the figure of eight points upwards, rather than downwards.

Thus, for coincident microphone recordings, the four signals conveying the periphonic sound are $A^- = A - ½H$, $B^+ = B + ½H$, $C^- = C - ½H$, $D^+ = D + ½H$, where A, B, C, and D are the conventional quadraphonic signals obtained by identical coincident horizontal cardioid or hypercardioid microphones, and where H is the output of an upward pointing figure-of-eight microphone. However, this has not completely specified the nature of the periphonic signal A^-, B^+, C^-, and D^+ for coincident microphones, as we do not yet know the correct relative gains of the H signal and the A, B, C, and D signals. Put crudely, how loud should the height signal be compared to horizontal quadraphonic signal in formulae (2)?

To answer this, we look at the microphone directional characteristic required to pick up the signals A^-, B^+, C^-, and D^+. Each of these signals is obtained by adding the audio output of a horizontal cardioid or hypercardioid microphone to that of a vertical figure-of-eight microphone. Thus one may consider that the signals A^-, B^+, C^-, and D^+ are picked up by hypercardioid microphones pointing at an angle to the horizontal (see **fig.3**). Thus the signals B^+ and D^+ are the signals that would be picked up by hypercardioid microphones pointing at an angle *above* the horizontal B and D

Reproduced from *Studio Sound,* Vol. 12, pp. 380-384 (September 1970) by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

The Principles of Quadraphonic Recording
Part Two: The Vertical Element

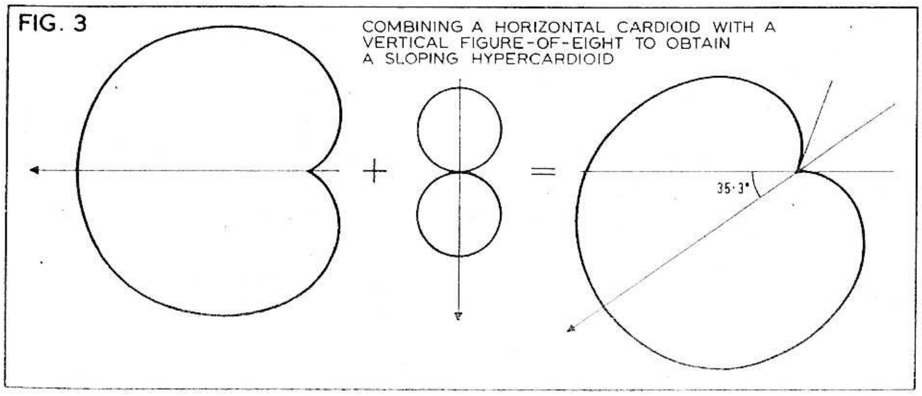


FIG. 3 COMBINING A HORIZONTAL CARDIOID WITH A VERTICAL FIGURE-OF-EIGHT TO OBTAIN A SLOPING HYPERCARDIOID

directions, and the signals A^- and C^- are the signals that would be obtained by hypercardioid microphones pointing at an angle *below* the A and C directions, as illustrated in **fig. 4**.

It is desirable that the four directions in which the A^-, B^+, C^-, and D^+ microphones point should be disposed as symmetrically as possible, and the most symmetrical arrangement possible is obtained if the four microphones point along the axes of a regular tetrahedron. (A tetrahedron is said to be *regular* if all its sides are equal.) This requirement is fulfilled if the axes of these four microphones are inclined at an angle of 35.3° to the horizontal. (This occurs if the sensitivity of the vertical figure-of-eight picking up H is √2 times the sensitivity of the figure-of-eight component of the horizontal directional characteristics used to pick up A, B, C, or D, ignoring the omnidirectional component.

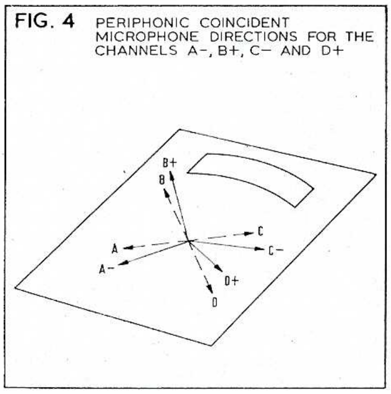


FIG. 4 PERIPHONIC COINCIDENT MICROPHONE DIRECTIONS FOR THE CHANNELS A−, B+, C− AND D+

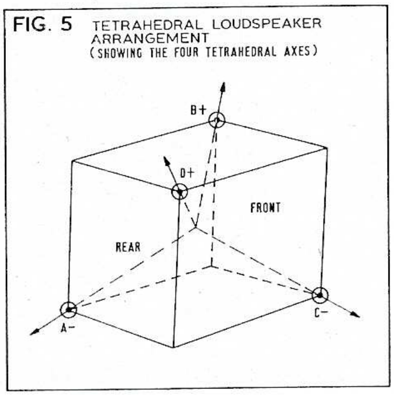


FIG. 5 TETRAHEDRAL LOUDSPEAKER ARRANGEMENT (SHOWING THE FOUR TETRAHEDRAL AXES)

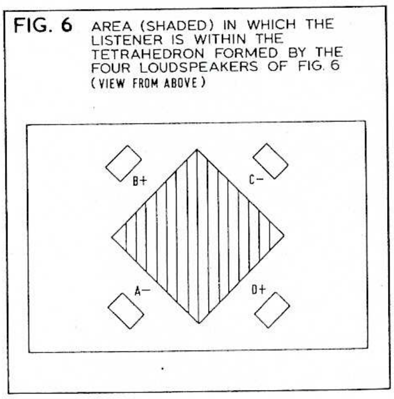


FIG. 6 AREA (SHADED) IN WHICH THE LISTENER IS WITHIN THE TETRAHEDRON FORMED BY THE FOUR LOUDSPEAKERS OF FIG. 6 (VIEW FROM ABOVE)

Thus the four periphonic signals A^-, B^+, C^-, and D^+ are required to be the signals picked up by four identical coincident hypercardioid microphones directed along tetrahedral axes pointing in a direction 35.3° above (in the case of B^+ and D^+) or below (in the case of A^- and C^-) the horizontal directions A, B, C, and D illustrated in figs. 2 and 4. The requirements of simplicity and compatibility with conventional quadraphony have thus led to a periphonic recording system in which four identical hypercardioid microphones point along four tetrahedral axes.

Reproduced from *Studio Sound*, Vol. 12, pp. 380-384 (September 1970) by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

The Principles of Quadraphonic Recording
Part Two: The Vertical Element

The proposal for periphonic recording is very similar to Granville Cooper's,[3] except that in our case the axes of the tetrahedron point in different directions. The axes of the four microphones picking up A^-, B^+, C^-, and D^+ point along the lines connecting the centre of a cube to four of its eight corners, as illustrated in **fig. 5**.

The simplest method of reproducing the original directional effect of the periphonic signals A^-, B^+, C^-, and D^+ is to feed them to four loudspeakers placed at four corners of a cube, as illustrated in **fig. 5**. A^- is fed to a floor-level rear left speaker, B^+ is fed to a ceiling-level front left speaker, C^- is fed to a floor-level front right speaker, and D^+ is fed to a ceiling level rear right speaker. For a given room height this tetrahedral speaker layout (and its mirror image) encloses a larger volume than any other possible arrangement using four loudspeakers placed at the corners of a regular tetrahedron. For this reason, the listening area in which reasonable periphonic reproduction can be obtained is likely to be larger than with any other tetrahedral arrangement of loudspeakers, including that of Granville Cooper.[3] For a listener whose ears are half way between the floor and ceiling, the portion of the room in which his head lies within the tetrahedron is indicated by the shaded area of **fig. 6**. Within this area, a reasonable periphonic effect should be obtained, although this has not been tried experimentally.

Reproduced sounds

Reproduced sounds will appear to come from a horizontal direction via the loudspeaker layout of **fig. 5** only if the signals A^-, B^+, C^-, and D^+ contain no height information. This occurs when the focus signal $\frac{1}{2}A^- - \frac{1}{2}B^+ + \frac{1}{2}C^- - \frac{1}{2}D^+$ is equal to zero. Thus a coincident-microphone horizontal quadraphonic recording will reproduce well over the fig. 5 tetrahedral loudspeaker layout. However, conventional two-channel stereo or spaced-mike horizontal quadraphonic recordings will reproduce properly via this speaker layout only if their focus information is suppressed. As explained last month, several different matrixings are capable of suppressing the focus. Tables 8 and 9 give a typical matrixing that allows conventional stereo and horizontal quadraphony to be reproduced via tetrahedral loudspeakers. In the case of ordinary stereo, it will be seen that the tilt of the sound from the front speakers is compensated for by the opposing tilt of the rear speakers.

TABLE 8 Playing a stereo signal L and R through the tetrahedral speaker layout of fig. 6

$$A^- = 0.354\ L - 0.146\ R$$
$$B^+ = 0.854\ L + 0.354\ R$$
$$C^- = 0.354\ L + 0.854\ R$$
$$D^+ = -\ 0.146\ L + 0.354\ R$$

TABLE 9 Playing a horizontal quadraphonic signal A^1, B^1, C^1, and D^1 through the tetrahedral speaker layout of fig. 6.

$$A^- = 0.854\ A^1 + 0.354\ B^1 - 0.146\ C^1 + 0.354\ D^1$$
$$B^+ = 0.354\ A^1 + 0.854\ B^1 + 0.354\ C^1 - 0.146\ D^1$$
$$C^- = {}^-0.146\ A^1 + 0.354\ B^1 + 0.854\ C^1 + 0.354\ D^1$$
$$D^+ = 0.354\ A^1 - 0.146\ B^1 + 0.354\ C^1 + 0.854\ D^1$$

TABLE 10 Corresponding horizontal and tetrahedral microphone directional pick-up characteristics (see text).

HORIZONTAL MICROPHONES		**TETRAHEDRAL MICROPHONES**	
Angle of null off axis	Front-to-back ratio	Angle of null off axis	Front-to-back ratio
Not hypercardioid	19.91 dB	180° *	∞ dB *
180° *	∞ dB *	144.7°	19.91 dB
150°	22.88 dB	135°	15.31 dB
135°	15.31 dB	125.3°	11.44 dB

*N.B. – cardioid directional characteristic

The identical cardioid or hypercardioid directional characteristics used to pick up the horizontal signals A, B, C, and D are not the same as the identical hypercardioid directional characteristics used to pick up the signals A^-, B^+, C^-, and D^+, due to the fact that the latter contain a proportion of the vertical figure-of-eight H signal. A hypercardioid directional characteristic may be specified either by its front-to-back ratio or by the angle from its axis at which its null response lies. Table 10 indicates the hypercardioid characteristics ('tetrahedral microphones') used to obtain A^-, B^+, C^-, and D^+ corresponding to each of a range of possible hypercardioid characteristics ('horizontal microphones') used to pick up the signals A, B, C, and D. It will be seen that the microphone characteristics used to pick up A^-, B^+, C^-, and D^+ are more hypercardioid, and less cardioid, than the corresponding microphone characteristics used to pick up A, B, C, and D.

Ideally, the microphone directional characteristics used to pick up A^-, B^+, C^-, and D^+ should have a good front-to-back ratio so as to prevent sounds being reproduced loudly from loudspeakers in the direction opposite to that from which the sound should appear to come. This requirement would imply that the microphone characteristics used to pick up the signals A^-, B^+, C^-, and D^+ should be cardioids, as in Granville Cooper's experimental recordings. However, it will be seen from Table 10 that the corresponding horizontal microphone pick-up is something between cardioid and omnidirectional, which means that horizontal sounds would be picked up with rather a lot of inter-speaker cross-talk. Also, for sounds originating on one of the tetrahedral axes, each of the three speakers corresponding to the other three axes in **fig. 5** would reproduce such sounds only 9.54 dB more quietly than the main speaker, and a quarter of the total audio energy would be reproduced from directions on the opposite side of the listener to the desired direction.

To reduce these effects, some degree of compromise between inter-speaker cross-talk and good front-to-back ratio has to be adopted, and it may be that a 135°-null hypercardioid characteristic for the A^-, B^+, C^-, and D^+ microphones (or, equivalently, a 150°-null hypercardioid for the horizontal pick-up) will be a good compromise. In this case, the cross-talk of a sound appearing to come from one of the tetrahedral loudspeakers on to each of the other three loudspeakers is –13.19 dB. With 135°-null hypercardioids, only 0.026 of the energy of a sound being reproduced from a direction *opposite* to that of one of the tetrahedral axes will be reproduced from the speaker on that tetrahedral axis.

An alternative loudspeaker layout for periphonic reproduction might include eight loudspeakers arranged in a cube around the listener, as illustrated in **fig. 7.** When reproducing coincident microphone recordings, each of the eight speakers should be fed with the output that would be given by a hypercardioid microphone pointing in that speaker's direction. Labelling the speakers A^-, A^+, B^-, B^+, C^-, C^+, D^-, and D^+ in the obvious way, the signals fed to the eight speakers will be:

A^-, $A^+ = \frac{1}{2}A^- + \frac{1}{2}B^+ - \frac{1}{2}C^- + \frac{1}{2}D^+$

$B^- = \frac{1}{2}A^- + \frac{1}{2}B^+ + \frac{1}{2}C^- - \frac{1}{2}D^+$

B^+, C^- , $C^+ = - \frac{1}{2}A^- + \frac{1}{2}B^+ + \frac{1}{2}C^- + \frac{1}{2}D^+$

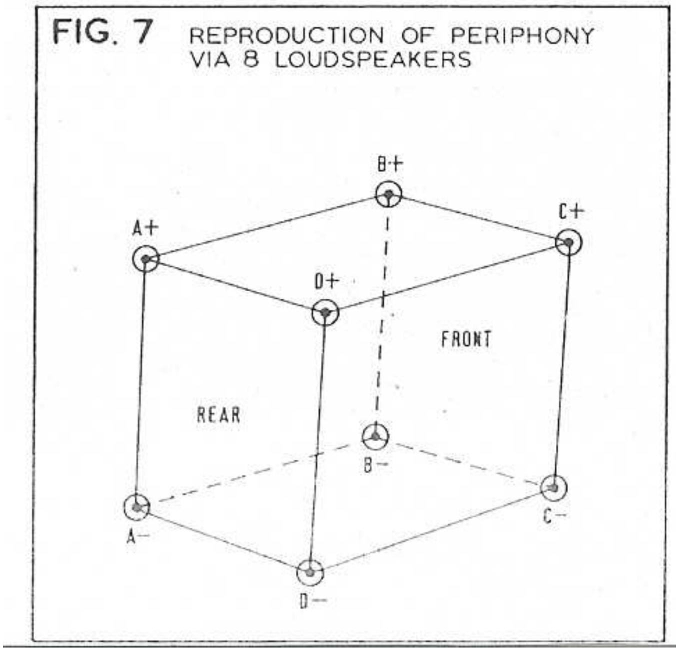
$D^- = \frac{1}{2} A^- - \frac{1}{2}B^+ + \frac{1}{2}C^- + \frac{1}{2}D^+$

and D+ respectively. (This may be seen by using formulae (2) and (3) along with the obvious fact that

$A^+ = A + \frac{1}{2}H$

$B^- = B - \frac{1}{2}H$

$C^+ = C + \frac{1}{2}H$
and $D^- = D - \frac{1}{2}H$.)



One advantage of eight-speaker periphonic reproduction (apart from the profits for speaker and amplifier manufacturers!) is the fact that conventional horizontal quadraphony can be reproduced without loss. The four-speaker tetrahedral layout of **fig. 5** can only reproduce horizontal quadraphony by suppressing its 'focus' information, as in the matrixing of Table 9. Another advantage of eight speakers is that the angle subtended between adjacent speakers at the listeners' ears is only 70.5°, as compared with 109.5° for the tetrahedral layout, and this should help to make stereo images more precise.

Clearly, much ingenuity could be expended devising various advantageous loudspeaker layouts using five, six or seven loudspeakers. For this reason, it is not proposed to investigate further loudspeaker layouts here.

The above has only discussed coincident microphone recordings, and if results are to be good with various different loudspeaker arrangements, the microphones have to be pretty coincident, with spacings of under 5 cm to avoid time-delay interference effects. The best results would thus probably be obtained by using four microphone capsules placed in close proximity, and a tetrahedral arrangement of four hypercardioid capsules placed back-to-back should prove satisfactory. In any case, separate bulky microphones would prevent the desired small spacing from being achieved. However, if only reproduction over the tetrahedral loudspeakers of fig. 5 is required, then the tetrahedral microphones need not be so precisely coincident.

The Principles of Quadraphonic Recording
Part Two: The Vertical Element

It is possible to convey and reproduce spaced microphone periphonic recordings via the four channels A^-, B^+, C^-, and D^+ by pan-potting the outputs of the spaced microphones, so that these outputs appear to come from the desired directions. An audio signal X can be pan-potted to appear to come from any desired direction in space by feeding into each of the four channels the signal that would be picked up by four imaginary coincident tetrahedral hypercardioid microphones were one to imagine the sound of X to be reproduced from a loudspeaker in that direction. One can choose the imaginary tetrahedral microphones' directional characteristic to give the best results in any particular case.

It is particularly easy to pan-pot sounds which are required to appear to come from straight ahead, straight behind, straight above, straight below, or from the left or right side. For instance, a sound X can be made to appear to come from straight ahead by putting $A^- = D^+ = 0$ and $B^+ = C^- = 0.707$ X. (This particular pan-potting simulates the sound that would be picked up by tetrahedral coincident hypercardioids with 125.3° nulls if the sound source were straight ahead). Similarly, a sound X can be made to appear to come from above by putting $A^- = C^- = 0$ and $B^+ = D^+ = 0.707$ X. This simple pan-potting is particularly useful if the recording is made with six microphones at the vertices of an octahedron, pointing forward, backward, to each side, above and below. The sounds from the six microphones can then be pan-potted into position to give a four channel periphonic recording.

It is less simple to pan-pot sounds to appear to come from other directions. If one wishes to make a sound X appear to come from some chosen horizontal direction, the four channels A^-, B^+, C^-, and D^+ must be the four signals that would be picked up by four imaginary identical coincident cardioids or hypercardioids pointing along *horizontal* directions at 90° to one another, as in fig. 2, if the sound X were to be reproduced through a loudspeaker in the desired direction. Thus, for sounds to be pan-potted in the horizontal plane, one can ignore all three-dimensional considerations. As an example, a sound X may be made to appear to come from 45° to the left by putting $A^- = 0.408$ X, $B^+ = 0.816$ X, $C^- = 0.408$ X, and $D^+ = 0$ (which simulates the sound pick-up of four coincident horizontal cardioid microphones). The tetrahedral symmetry of the channels may be used to derive similar pan-pottings for sounds in the vertical plane pointing forward and backward, or for sounds in the vertical plane pointing sideways. For example, a sound X will appear to come from 45° above straight behind if $A^- = 0.408$ X, $B^+ = 0.408$ X, $C^- = 0$, and $D^+ = 0.816$ X.

It is not all that difficult to pan-pot sounds to come from slightly above or below horizontal. The procedure is first to pan-pot the sound X in the desired direction on the horizontal plane, obtaining four signals A, B, C, and D. One then derives the signals $A^- = A - kX$, $B^+ = B + kX$, $C^- = C - kX$, and $D^+ = D + kX$, where k is a small

Reproduced from *Studio Sound,* Vol. 12, pp. 380-384 (September 1970) by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk)

number which is chosen to be positive if the sound is to come from above horizontal, and negative if from below horizontal.

The pan-potting required for a sound to appear to come from the corners of the cube of figs. 5 and 7 may be illustrated by typical examples. A sound X may be made to seem to come from the corner B^+ by putting $B^+ = 0.935$ X and $A^- = C^- = D^+ = 0.205$ X, which simulates the sound pick-up of 135°-null hypercardioids for a sound source along the B^+ axis. (The signals $B^+ = X$, $A^- = C^- = D^+ = 0$ are not really suitable, as they simulate the sound pick-up of 109.5°-null hypercardioids, and will not reproduce well over a cube of loudspeakers.) A sound X may be made to appear to come from the corner B^- of the cube of figs. 5 and 7 by putting $A^- = B^+ = C^- = 0.570$ X and $D^+ = -0.160$ X, again simulating the pick-up of 135°-null hypercardioids.

By using means of pan-potting such as described above, the sounds from any number of spaced microphones may be fed into the four periphonic channels A^-, B^+, C^-, and D^+.

Conclusions

In the two parts of this article, it has been shown that three channels are sufficient to convey horizontal quadraphonic sound, and four channels sufficient to convey periphonic sound in three dimensions. It has also been shown that it is possible to convey periphonic sound via channels A^-, B^+, C^-, and D^+ that can be reproduced via the horizontal quadraphonic 'box' speaker layout as in current American proposals, or via a tetrahedral loudspeaker layout giving three-dimensional sound reproduction over an exceptionally large listening area.

In the light of this, it would be wise for recording organisations to include height information on current quadraphonic master-tapes, to allow for the possibility that periphonic systems may become commercial. It would be feasible for companies to start issuing commercial ¼-track quadraphonic tapes conveying periphonic information almost immediately, due to the compatibility of the system described above. In order to ensure standardisation, it is recommended that the front left channel represents the output of an upward inclined microphone, rather than a downward inclined microphone.

It is further recommended that any three or four-channel system adopted for disc, radio or tape should not permit any ambiguity in the polarity of some of the channels with respect to the others, so that it will be possible to matrix signals for the various different loudspeaker layouts.

The author would like to emphasise that the above work is mainly the result of a theoretical analysis. Much remains to be done determining how well the various proposals work with different microphone techniques and different loudspeaker types.

References

[3] Granville Cooper: Studio Sound, June 1970.

[4] Alec Nisbett: Happy Birthday Ludwig! Studio Sound, March 1970.

[5] N. V. Franssen: Stereophony, Philips Technical Library, 1964 (pages 25-27).

Michael setting up the tetrahedral Calrec arrangement. (April 1971)

Close up of Calrecs in the tetrahedral arrangement.

Ambisonics. Part two: Studio techniques

By Michael Gerzon

*Conventional methods of 'quadraphonic' reproduction may not convey the original sound field to best advantage. The studio techniques for ambisonics and its applications are discussed, in the light of both accepted 'quadraphonic' techniques, and the requirements of further accurate reproduction of the sound field. Areas of compatibility and of disparity are discussed, to be read in the light of Peter Fellgett's system description (*Studio Sound, *Vol. 17, pp. 20-22, 40 (August 1975).*

The main aim in the development of NRDC ambisonics technology has been to record, to convey to the consumer, and to reproduce an accurate and repeatable surround sound directional effect. It is now well known, both from controlled experiments and from the experience of recording engineers and producers, that existing surround sound approaches (including the four channel 'discrete' approach) give extremely poor image stability for all positions except the four corners, even under ideal conditions. With 'discrete' techniques, the front stage suffers from the 'hole in the middle' effect, and the sides are virtually unusable in any less-than-ideal situation (eg when the listener is not at the centre of the speaker layout, or when the layout is non-square).

The impracticality of existing approaches has led to a careful study of each stage of the multi-channel recording and reproduction of a sound field. [1 2 3 4]. The aim of a surround sound system is to reproduce at the listener's ear accurately, reliably and repeatably, the directional sound field created in the studio either by a sound field encoding microphone array,[1] or by artificial directionality encoding devices (pan-pots) or artificial surround-reverberation devices. This aim contrasts with the aim[2] of quadraphonic systems, which is to duplicate in the home the defects of a pair-wise mixed mastertape. Without accuracy and repeatability of directional effect, the recording producer's work will not be heard correctly by the domestic listener, and artistic communication will be compromised.

In order to maintain accuracy of effect, the process of encoding the sound field on to a mastertape must be accurately specified, and the specification accurately followed in either the microphone arrays or the pan pots. Existing 'pairwise' pan pots do not give a satisfactory encoding specification[3]. Similarly, most microphone clusters use microphones with poorly defined polar diagrams that are frequency-dependent and spaced apart, and so fail to satisfy a reasonable encoding specification. Conventional mastertape encoding assigns the four speaker feed signals to four tracks of a tape, but it is obvious that this is a sub-optimal procedure, even with correctly designed microphones and pan-pots. In order to recreate the illusion of a given directional field, it is obvious that the speaker feed signals will not be the same for a rectangular speaker layout as for a square one. Thus if we are to accommodate a variety of shapes of speaker layout appropriate to differently shaped listening rooms, we need to use a decoder to derive the speaker feed signals appropriate to the layout used.

As we shall illustrate later in this article, he optimal decoder even for a square speaker layout is a rather complex frequency-dependent matrix, and so we must face the fact that the mastertape merely encodes the sound field information, and not speaker feed signals.

Studies[3] in the design of decoders for four-speaker rectangle layouts show that three channels of information are optimal for accurate image localisation. The addition of a fourth channel always degrades image localisation quality (eg by giving a 'hole in the middle' effect). Thus, the basic horizontal encoding specification has to be three-channel. Most studio equipment used for surround sound mastering has four available channels, and so the problem arises of how to use this fourth channel. Since no useful extra horizontal information may be encoded in the forth channel, it may most fruitfully be used for encoding height or elevation information.

It is not suggested that periphonic (ie with-height) reproduction will be a serious commercial proposition at the present time, but periphonic technology is now understood, and not much more complex than horizontal-only technology. The adoption of a periphonic standard at this stage will prevent the premature obsolescence of valuable mastertapes in ten or fifteen years when periphony becomes commercial, and meanwhile will permit producers to gain periphonic experience without premature commercial pressures. It is also wise to ensure that existing media can change over smoothly to a periphonic standard in the future, so as to prevent a repetition of the chaos caused by quadraphonics. At the present time, the height information gives useful additional mixdown flexibility for stereo or horizontal surround effect. If height is not required, it may be omitted, as in most of the equipment described subsequently.

For studio use, there are two four channel signal formats, termed A-format and B-format. A-format consists of four channels L_B, L_F, R_F, R_B compatible with existing 'discrete' practice for the four corner positions. Technically, the A-format signals may be described for horizontal only sounds as the outputs of four hypercardioids each having nulls 120° off-axis (or their panpot equivalent signals) pointing in the four corner directions. When height is included the A-format signals are the outputs of four hypercardioids with nulls 114.1° off axis pointing respectively 35.3° below, above, below and above the four corner directions (ie towards regular tetrahedral axes[5]).

The second format is B-format, consisting of the four signals X, W, Y, Z, where X is a forward facing figure of eight signal with frontal gain √2, W is an omnidirectional signal of gain 1, Y is a sideways-facing figure-of-eight signal with leftward gain √2, and Z is an upward figure-of-eight signal with upward gain √2.

The circuit to convert A-format to B-format, known as an 'AB module' is shown in **fig. 1**, and performs the following matrixing:

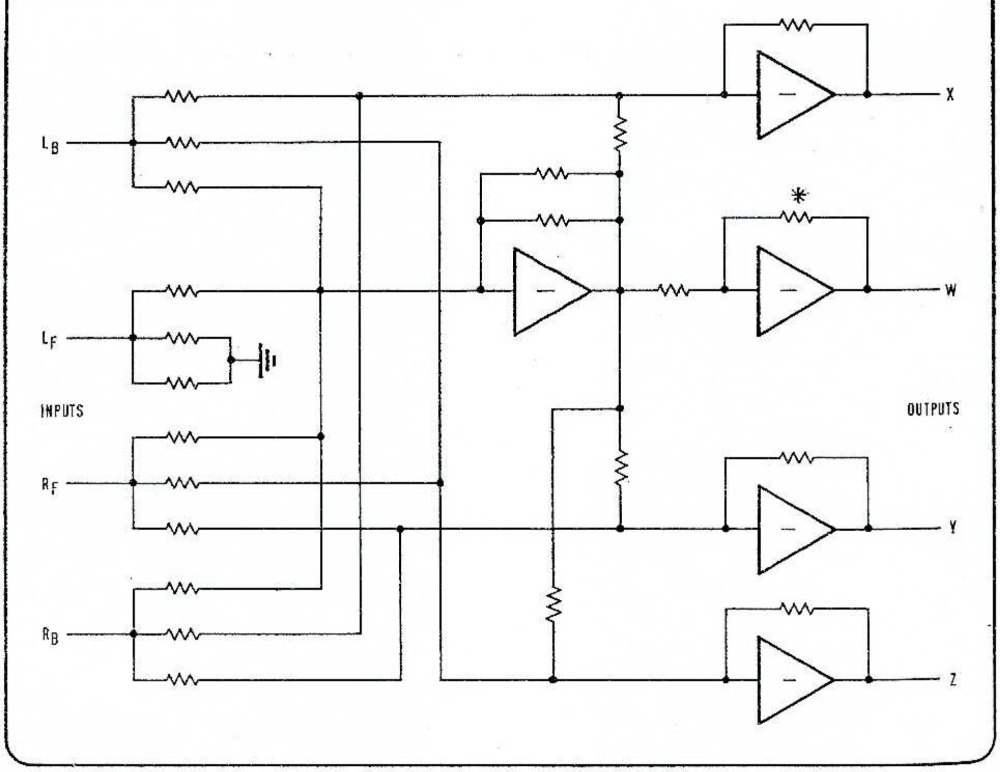
$$X = ½(-Lb+Lf+Rf-Rb)$$
$$W=½(Lb+Lf+Rf+Rb)$$
$$Y=½(Lb+Lf-Rf-Rb)$$
$$Z=½(-Lb+Lf-Rf+Rb)$$

It is important to note that for horizontal only signals, the signal Z is zero, and so may be omitted, giving a three channel B-format signal in the horizontal case.

The important thing about AB modules is that exactly the same circuit converts back from B-format to A-format. Thus, if one puts X, W, Y, Z into their respective inputs, L_B, L_F, R_F, R_B comes out. If one feeds a conventional 'discrete' or pairwise mixed signal into an AB module and then discards the Z output, then the X,W,Y signals are correctly encoded B-format signals for the four corner positions, with slight deviations from the correct encoding elsewhere. (This deviation is one of the causes of conventional discrete reproduction giving poor non-corner images – the B-format signal will not make this defect worse).

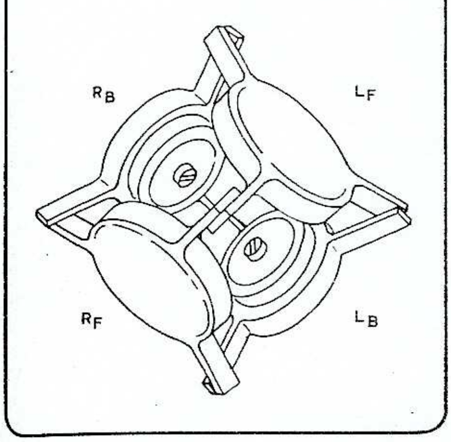
There are two main ways of producing correct B-format signals. The first is the Calrec sound field microphone (which is still

FIG. 1 AB MODULE. ALL RESISTORS SHOULD HAVE AN IDENTICAL VALUE, SAY 33KΩ.



undergoing evaluation and development). This has been developed by the present writer for the NRDC and uses (see **fig. 2**) a tetrahedral array of cardioids to

FIG. 2 ARRANGEMENT OF CAPSULES IN SOUND FIELD MICROPHONE (VIEW FROM FRONT-NOT TO SCALE)

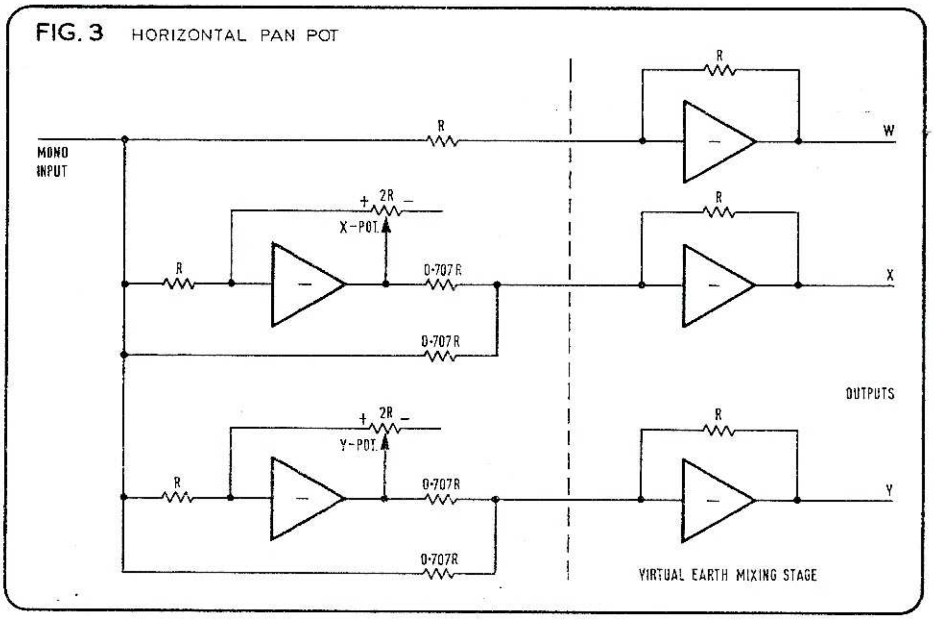


feed a frequency-dependent matrix circuit. The matrix circuit fulfils the dual function of converting to B-format and of providing electronic compensation for the spacing of the capsules, so as to give outputs that are effectively coincident and satisfy B-format encoding accurately[1]. The result is to give B-format outputs that are characterised by accurate and precisely coincident polar diagrams up to around 7.5 kHz (as compared to less than 1.5 kHz for the best existing microphone arrays). The behaviour above 7.5 kHz is arranged to be subjectively smooth, although deviating from the ideal.

In this way, the directional properties of a sound field may be captured with the minimum possible departure from the objective ideal. The microphone system should not be regarded as four microphones on a tetrahedron, but as a complete sound field transducing system. The tetrahedral configuration is purely a matter of design convenience[1] and has nothing to do with the desired form of B-format sound field encoding.

A second method of producing B-format signals it to use a panpot. **Fig. 3** shows a circuit of a panpot (feeding a virtual earth mixing stage) that uses a joystick control to meet accurately the encoding specification for B-format for horizontal sounds (so that the Z signal is zero). The 'X-pot and Y-pot' of **fig. 3** are the potentiometers that respond to the 'up-down' and 'left-right' motions respectively of the joystick (see **fig.4**). For correct results it is vital that the travel of the joystick be restricted by a mask or cut-out to that range of X-pot and Y-pot resistances in **fig. 4** such that $x^2 + y^2 \leq 1$. In other words, the corner travel must be restricted to ±0.707 of the way from the centre to the end of the pot tracks, although the full range of each pot may be covered when the other is centred.

It is also possible to convert existing pairwise pan pots to give optimal ambisonic encoding. This is clearly a worthwhile option for use with existing equipment. The modifications involve fitting a mask to the joystick control to limit its travel, and a matrix circuit positioned after the four output channels of the mixer. The matrix circuit is exactly the same as an AB module (**fig 1**), except that the gain of the W output channel is reduced to 0.707 (by reducing the resistor marked * in **fig. 1** to one equal to 0.707 of the value of the other resistors), and the Z output is not used. This gives a B format output. The masking of the joystick controls is necessary to prevent undue exaggeration of directionality of the corners. The mask for a joystick control normally travelling in a square aperture would be as illustrated in **fig 5**. The side positions remain unaffected by the mask, but the corner positions are masked so that the corner-most joystick positions in the mask give equal outputs on the X, W and Y outputs. Because most joystick pairwise pan pots are not well designed as regards constancy of sound level with direction, no guarantee can be given that the ambisonic modification will be good in this respect either.



The pan pots described give full 'interior' effects as well as a 360º azimuth coverage with accurate encoding according to B-format specifications. A similar design of about twice the complexity prior to the mixing stage, using an additional slider pot for elevation gives full-sphere periphonic encoding. Other devices of a similar nature allow the full rotation of a whole encoded sound field ('waltz' control), a facility not possible with discrete approach. When sounds are 'circled' with a B-format pan pot, the motion is smooth rather than the jerky jumping from speaker to speaker given by existing pan pots. A design is also available for a 'width' control that alters the width of the front of a sound field relative to the back without destroying the correct encoding specification.

Having obtained a horizontal or periphonic B-format signal, either it can be converted to A-format (by the AB-module of **fig. 1**) to go through existing quadraphonic equipment, or it can be recorded in

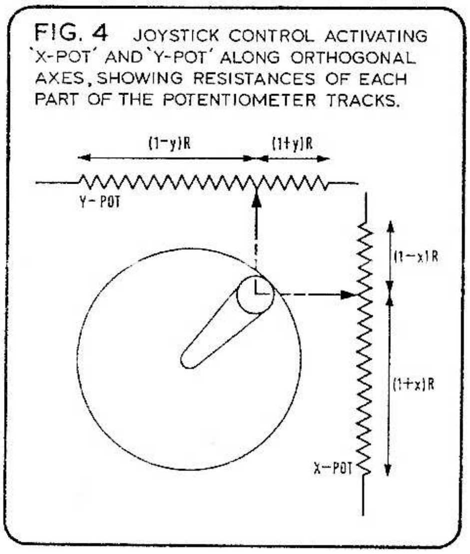
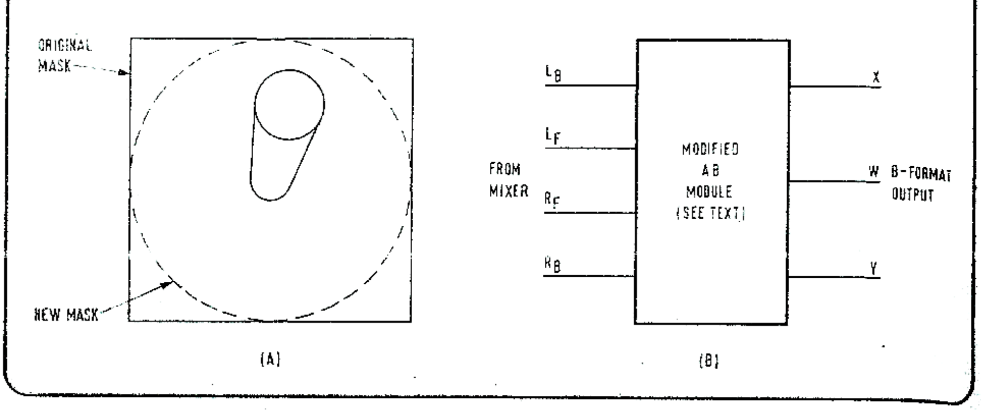


FIG. 5 TYPICAL MASKING OF PAIRWISE PAN POT (A) AND THE PROCESSING AT THE MIXER OUTPUT (B)



B-format. There are considerable advantages gained by staying in B-format in the tape recording stage. If one records in A-format then the effects of noise reduction systems are quite audible when the signal is played back. Existing stereo covers only a 60° stage and the small 'pumping' effects inevitable even with a well adjusted noise reduction system do not cause noticeable movements in the stereo image. Similarly, while conventional discrete material covers 360° its very poor image quality means that pumping effects are not noticed. However, once the image sharpness is improved by ambisonic encoding, the varying image shifts (which subjectively seem to be around 15°) start becoming comparatively objectionable. It is found that recording in B-format renders the signal much less susceptible to all forms of image degradation whether due to channel imbalances, phase errors, or noise reduction pumping. In fact, it seems likely that B-format recording should reduce signal degradation significantly even for conventional 'discrete' pan pot recordings.

Ambisonic B-format signals may be converted to A-format by an AB module, and may be used, if necessary, for feeding existing 'quadraphonic' systems via existing commercial encoding equipment (with the exception of the SQ system*), provided that the Z signal of the B-format is omitted. While the results should generally be better than with existing mastertape encoding methods, they will not be as good as they might be, except for the UD-4 system which will give correct TMX encoding.

* However, an SQ encoder may be used if it is set to 'interior' encoding mode. The results with SQ cannot be optimal.

Table 1 Gains of shelf filters in three channel decoder of **fig. 6**.

	Shelf filter 1	Shelf filter 2
low frequencies	0 dB	0 dB
high frequencies	+1.76 dB	-1.25 dB

For each of the following systems: CD4, UD4, RM and the BBC matrix systems, there is an optimal encoder design available working straight from B-format. It is understood that the BBC has evolved an encoding technique similar to B-format. The use of an optimal encoder ensures that the consumer-encoded format (C-format) accurately follows the correct specification. At the time of writing, no existing system is considered by broadcasting organisations and independent record companies to fulfil the necessary compatibility requirements adequately along with good four speaker results, and it is expected that industry discussions will be held regarding the choice of a generally acceptable system. However, ambisonic studio technology is compatible with all systems capable of good localisation for all directions from four speakers.

It is, of course, necessary to monitor the results of an encoded, A-, B- or C-format signals, and no existing decoder design is capable of optimal results. It must be remembered that it is no more desirable that the four speakers should be heard as direct sound sources than one would wish to see the individual phosphor dots on a colour tv screen. The speakers are purely a means of feeding information into the room to create a convincing (or otherwise) illusion of sounds from all directions around the listener. If the shape of the speaker layout deviates from a perfect square, the image stability and directional effect from the loudspeakers would alter unless the decoder is modified to compensate for a non-square speaker layout.

Other problems in designing decoders arise from the fact that the ears localise sounds by different mechanisms at low frequencies (< 700 Hz) and high (> 700 Hz). This means that the optimum design for a decoder is different at low and high frequencies. Moreover, when speakers are fed with signals from a matrix with no frequency variation, it is found that the tone quality of reproduction is very coloured and 'thumpy' in the bass (below 350 Hz). This is because at low frequencies the intensity at the listener is the sum of the pressures due to the four speakers, while at high frequencies it is the sum of the energies.

A design of three channel horizontal studio decoder, working off a B-format input is shown in **fig. 6**. This is intended to feed a rectangular speaker layout, and the 'layout control' adjusts the decoder to compensate for the actual rectangular shape used. The two types of 'shelf filter' used have gains as in Table 1 at low and high frequencies, and are 'phase compensated' to have identical phase response. The transition between low and high frequency gains in the step filters is gradual (using simple RC type circuits) to avoid coloration, and is centred on 350 Hz approximately.

It is found that a correctly encoded B-format signal fed to the three channel studio monitor gives stable and sharp images even at the sides of the listener and for listeners well away from the centre of the listening area. The four loudspeakers used should be matched both in frequency and phase response, and should be reasonably 'omnidirectional' in their polar diagrams over ±45^{o} off their respective axes. The layout control requires careful adjustment, but once set need not be changed.

The three channel/four speaker studio decoder fed by B-format gives what we believe is the most accurate reproduction of sounds from any desired direction around the listener possible with existing technology via four speakers. It is not perfect. In particular, it was predicted (in advance of construction) by a new 'bispectral' model for human hearing that sound waveforms with a very high degree of asymmetry would still tend to be 'pulled' to the nearest speaker position, and this is quite noticeable on clapping. Experiment and theory agree here, and theory shows that there would be no ways of overcoming this fault other than going to a five speaker decoder. Many critical listeners would possibly regard this as 'hair-splitting' by comparison with many existing four speaker systems, such as 'discrete' encoding.

It is not claimed that there is any decoder capable of giving really accurate decoding in a very large room or auditorium via four speakers. Ambisonics is essentially designed for domestic or studio listening conditions, although results in a large auditorium can be quite reasonable.

FIG. 6 THREE-CHANNEL, FOUR LOUDSPEAKER DECODER FOR HORIZONTAL STUDIO MONITORING. THE SHELF FILTERS HAVE MATCHED PHASE RESPONSES AND GAINS AT LOW AND HIGH FREQUENCIES AS IN TABLE ONE. RESISTOR VALUES GIVEN ASSUME AN 11KΩ INPUT IMPEDANCE FOR THE AB MODULE.

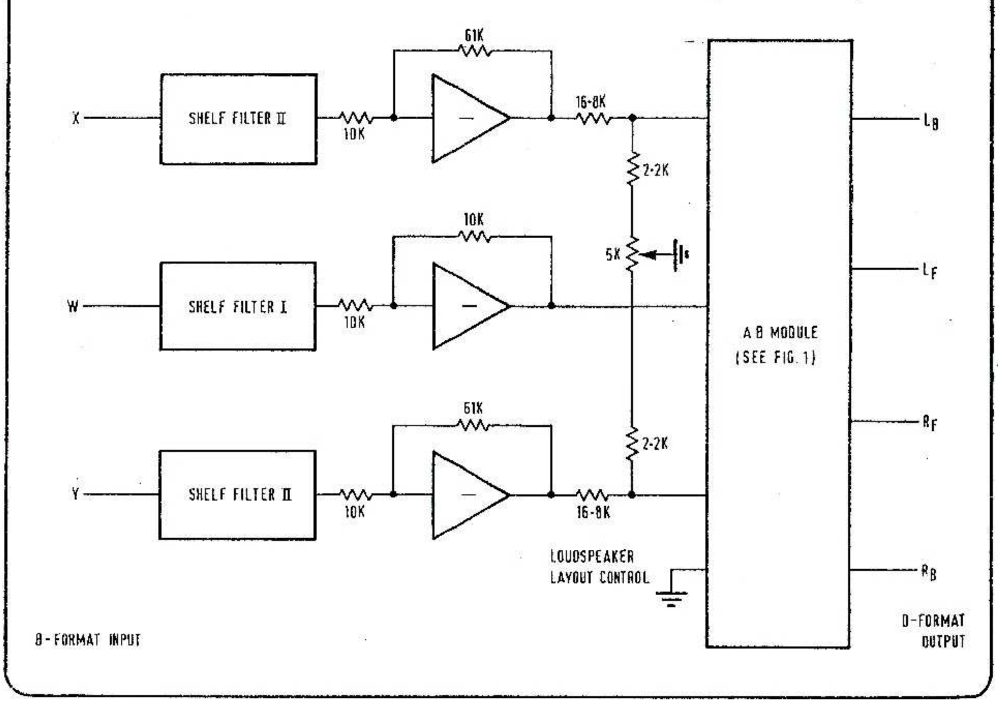


Table 2 Gains of shelf filters in two channel decoder of fig. 7.

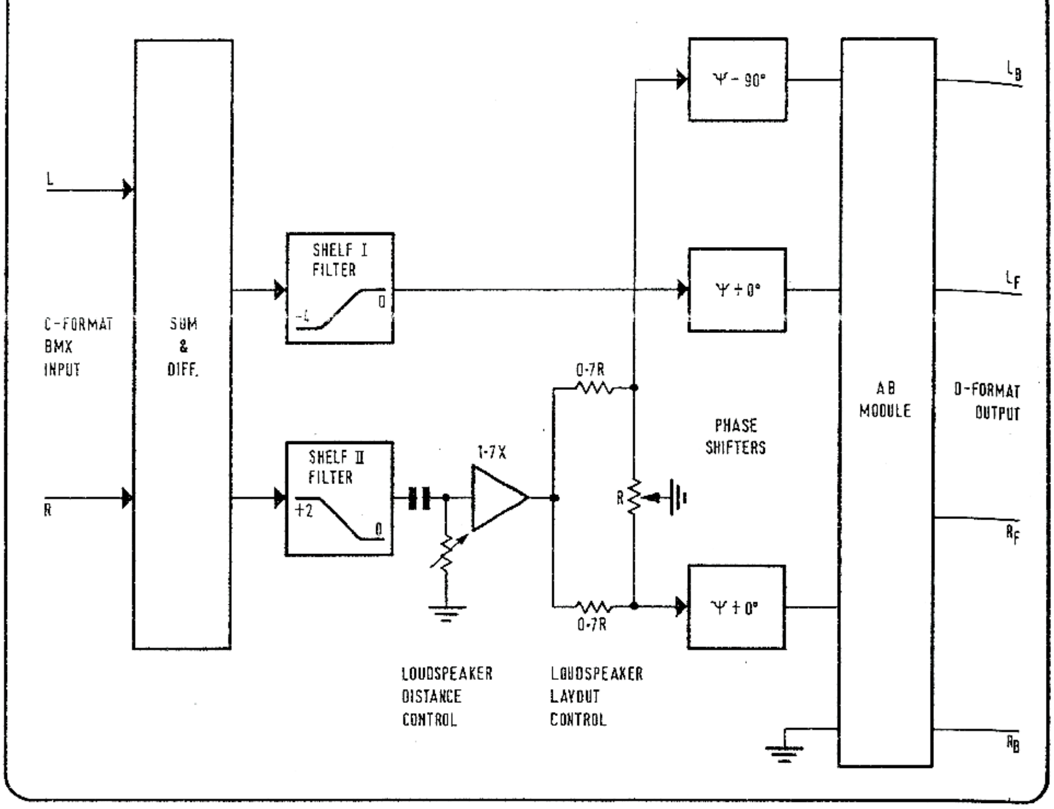
	Shelf filter 1	Shelf filter 2
low frequencies	-3.98 dB	+2.04 dB
high frequencies	0 dB	0 dB

Domestic decoders have been designed according to a range of psychoacoustic theories[3] and at various levels of cost and complexity for all the major existing or proposed surround sound systems other than SQ. It is not possible to design an SQ decoder to satisfy the psycho-acoustic criteria established by the author in ref. 3.

Decoders including layout controls and frequency dependence are rather similar to that of figure 6 and have been designed for two channel C-format decoding, 2½ channel (as used in UD-4) and three channel decoding. As an example, **fig. 7** shows a basic decoder for the BMX system, with phase-matched filters centred on 350 Hz with gains as in Table 2. This is not the most refined version under development; improvements to reduce ‘phasiness’ will be announced shortly.

Other decoders have been designed adjustable for cuboid (box-shaped)

FIG. 7 DECODER FOR BMX ENCODED SIGNALS



with-height speaker layouts for periphonic reproduction. At the present time, this would be confined to experimental and in-studio use, where producers might find it worthwhile to explore the artistic possibilities of full-sphere directional effects well before they are pushed into premature commercial exploitation. Three and four channel C-format encoding has been designed for periphony compatible with existing or proposed horizontal C-format encoding, and it would even be possible to release periphonic material on disc (without announcement) to avoid inventory troubles at a future time.

Ambisonic technology also offers new opportunities for existing mono and stereo recording. The sound field information from a sound field microphone may be recorded in B-format on four channel tape, and be mixed down later to any coincident stereo microphone technique that may be required. Any image width and microphone polar diagrams may be selected off tape, along with any vertical angle of tilt. A control unit to perform these functions in an intuitive and easy to grasp fashion has been designed, and also gives adjustable 'quadraphonic' outputs for four speaker use. Due to their very good polar diagrams in the mid-treble frequency region, the sound field microphone technology also gives stereo with a particularly 'clean' and uncoloured quality of sound.

We have here only been able to skim over a few aspects of ambisonics studio technology. Other devices include an apparatus to convert certain existing types of artificial stereo reverb units to full surround reverb, and devices for compensating for and improving non-ideal microphone techniques that may have been used on any existing surround sound recordings.

Details of these and other devices will be released at a future time.

Ambisonic technology can, of course, be used as part of existing 'quadraphonic' systems (excluding SQ) by treating the A-format horizontal-only signals as if they were 'discrete' signals. Similarly ambisonic decoders may be used with existing quadraphonic systems (again excluding SQ). However, it will be appreciated that the inherent faults of the quadraphonic systems will be apparent in such cases, as the whole system, from microphones or pan pots to decoders and loudspeakers has to be designed correctly for correct results. When parts of the system are incorrect, ambisonic technology will not make the system any worse, and so a compatibility exists between ambisonics and quadraphonics – ie ambisonic material can be used for quadraphonic results, although the converse is not true – just as a poor colour film cannot be made good by good projector optics.
In conclusion, the ambisonic technology developed for the NRDC, and in particular by Professor Peter Fellgett, the present author and John Wright of IMF, is compatible with several existing and proposed encoding systems. It gives enhanced creative possibilities to the producer both by ensuring that what he hears will be substantially passed on to the consumer despite differing loudspeaker layouts and seating positions, and by giving him convincing side localisation and smooth 'circling' effects. In addition to existing 'interior' or 'in the head' effects, new 'waltz' (rotation) and 'width' effects are available as well as the first practical control over with-height periphonic effects, if required. When required, ambisonic equipment can cope with existing

recordings and equipment, but of course cannot remove the faults in existing material. Besides these 'creative' possibilities, the sound field microphone allows uniquely accurate recording, storage and playback of natural sound fields, with all their attendant advantages. The 'ambient labelling[4]' given by the sound field microphone in particular permits remarkably good sound localisation as well as the ability to separate by ear musical lines that would be completely masked by other lines in a pan pot recording.

References

1. Michael Gerzon 'The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound' Preprint of the 50th Audio Engineering Society Convention, London, 4th March 1975.
2. Michael Gerzon 'What's Wrong with Quadraphonics?' Studio Sound May 1974.
3. Michael Gerzon 'Surround Sound Psychoacoustics' Wireless World, December 1974, pp 483-6
4. P. B. Fellgett, 'Ambisonic Reproduction of Directionality in Surround-Sound Systems'. Nature, vol. 252, pp 534-8. December 13, 1974.
5. Michael Gerzon 'Principles of Quadraphonic Recording Part Two: The Vertical Element'. Studio Sound, September 1970.