BBC RD 1974/26

1969 THE QUEEN'S AWARD TO INDUSTRY
1974 THE QUEEN'S AWARD TO INDUSTRY

RESEARCH DEPARTMENT REPORT

QUADRAPHONY: techniques involved in four channel recording and reproduction

D.J. Meares, B.Sc.(Hons.)

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION July 1974

BBC RD 1974/26
UDC 534.76

QUADRAPHONY: TECHNIQUES INVOLVED IN FOUR CHANNEL RECORDING AND REPRODUCTION

D.J. Meares, B.Sc.(Hons.)

Summary

A series of tests has been carried out, and recordings made, to identify some of the problems that will be encountered in recording and reproducing quadraphonic programme material. Various microphone and loudspeaker arrangements have been evaluated with the main aim of re-creating, in the listening room, the acoustic qualities - ambience, spaciousness, etc. - of the original environment. In the case of orchestral music this aim is the re-creation of all the acoustic qualities of the best seat in a concert hall.

Recommendations are made with regard to future work on special microphones for quadraphonic use.

Issued under the authority of

Head of Research Department

Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION

July 1974

QUADRAPHONY: TECHNIQUES INVOLVED IN FOUR-CHANNEL RECORDING AND REPRODUCTION

QUADRAPHONY: TECHNIQUES INVOLVED IN FOUR CHANNEL RECORDING AND REPRODUCTION

D.J. Meares, B.Sc.(Hons.)

1. Introduction

Quadraphony, four-channel sound, is now considered commercially viable by recording companies and over the last eighteen months Research Department has been involved in a series of experiments to quantify the many different variables in a four-channel system. Some of these experiments are still unfinished but, as interest in quadraphony increases, it seems appropriate to discuss the problems that have already been encountered. This report covers only those problems which arose during the experimental work on recording and reproducing quadraphonic sound. Later reports will discuss properties of hearing relevant to quadraphony[1] and some of the problems of broadcasting four-channel material to the listening public.

Commercial quadraphonic recordings have exploited the two-dimensional (sometimes even three-dimensional) effects that the system is capable of producing, such as surrounding the listener with instruments or electronic noises, or even rotating instruments right around the listener's head. These effects are generally achieved by making a multimicrophone, multitrack recording and 'panning'* each instrument into the sound location that is required on reproduction. This produces a 'pan-pot' recording.

The orchestral recordings made by Research Department have been based on a different approach. The aim of these has been to re-create, in the home environment, the subjective balance of audio signals that exists in the best seats of a concert hall, with the orchestra in front of the listener and only reverberant information arriving from other directions. This has been achieved using a relatively simple microphone arrangement, i.e. a group of four, coincident, cardioid/hypercardioid microphones.

The same microphone arrangement has been used to record drama and the surround-sound effects of 'The Last Night of the Proms.' Such recordings, because of the microphone techniques used, are referred to in this report as coincident-microphone recordings.

Only one loudspeaker arrangement has been examined in detail. This is the square array, in which the loudspeakers are placed at the corners of a square, facing towards the centre. The listener sits at the centre of the square facing the mid-point of one side, which is conventionally regarded as the front of the sound stage. Other arrangements have been suggested, e.g. diamond and tetrahedral arrays,[2,3] but preliminary tests showed that they gave a quadraphonic performance inferior to that provided by the square layout. It is relevant to note that all commercial quadraphonic recordings are intended to be replayed using a square layout of loudspeakers.

This report discusses first the problems of accurate reproduction of quadraphonic material. It then covers the requirements of the recording equipment, particularly the microphones, with reference to the experience gained during the experimental recordings. Microphone techniques are one of the most important factors in the performance of quadraphonic systems, and the final section discusses a possible change to microphone characteristics which would produce an improvement in coincident-microphone recordings.

2. Listening conditions and equipment

2.1 Symmetry

A system of quadraphony can suffer from more intrinsic crosstalk between the four channels than is usual in stereo but because of the 'balance' of loudspeakers around the listener this cross talk need not be troublesome. If this balance is upset, however, the crosstalk can cause the images to shift from one point to another or even render the images so diffuse as to be unlocateable.

Multi-microphone recordings pan-potted down to four channels and played directly through four loudspeakers are not very critical with regard to the symmetry of the listening conditions. With a coincident-microphone recording, however, using microphones with cardioid responses, the unwanted crosstalk from the loudspeakers adjacent to that reproducing the wanted signal is only 6 dB down and, under these conditions, symmetry, for good placement of images, is more important. If a 4–2–4* matrix is involved in the system, crosstalk may be fed to

*A sound is 'panned' into a predetermined position in the reproduced sound field by mixing appropriate portions of the signal into the four master channels.

*'$l-m-n$' is a short hand notation indicating the number of signal channels at different points in a quadraphonic system. 'l' is the number of source channels, 'm' is the number of record/transmission channels and 'n' is the number of output channels.

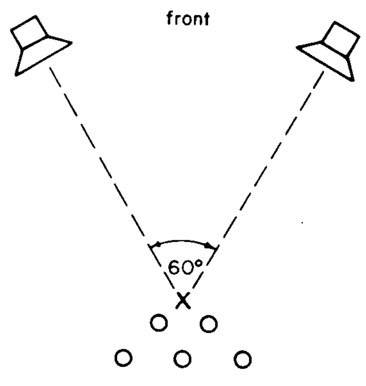
adjacent loudspeakers only 3 dB down on the wanted signal, independent of the origin of the material (i.e. pan-pot or coincident-microphone), and symmetry becomes essential. This symmetry applies to the loudspeakers, the listening position relative to the loudspeakers and the overall listening conditions.

2.2 Choice and placement of the loudspeakers

Experience with stereo has shown that sharp, well positioned images are obtained only if two, well matched, high quality loudspeakers are used; these are normally placed to subtend 60° at the listener (see Fig. 1).

In quadraphony each pair of loudspeakers subtends 90° at the listener, and the ears* are now called upon to create 'stereo images' at the sides and back of the head as well as at the front. Work already carried out on certain properties of hearing[1] has indicated that the 90° spacing has an adverse effect on image-sharpness and location. This does not cause gross errors in the front quadrant, except that a centre-front image tends to rise, from the horizontal, by some 20° to 30°, but the same work has also shown that the ear is not very effective in combining sounds from two loudspeakers at the side of the head (one towards the front, the other towards the rear) to form an image at the mid-point of the side.

*To avoid awkward expressions, 'ears' or 'ear' will be used to imply the combination of ear and brain.



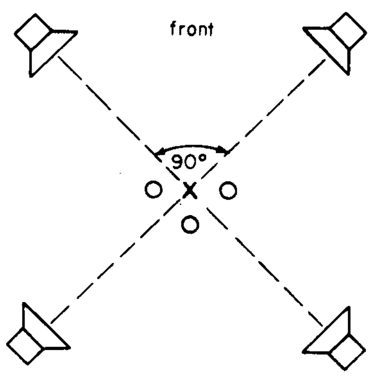


Fig. 1 - Optimum layouts for listening to stereo and quadraphony
(a) stereo (b) quad.
X = Best listening position
O = Other listening positions which do not upset the sound at X
NB: The number of these points will depend upon the size of the listening area, i.e. distance between loudspeakers.

Experience indicates that four identical* high quality loudspeakers should be used for the monitoring of all quadraphonic signals. This is especially important if a 4—2—4 matrix system is being used, because one of the unfortunate properties of such matrix systems is that they can generate rather large phase differences between the four decoded signals. At times during investigations involving 4—2—4 matrices the phase differences have been clearly audible on high-grade loudspeakers as unpleasant 'phasey' sensations in the reproduced quadraphonic sound. The author has attended commercial demonstrations, however, where the same material played through the same matrix equipment but reproduced on unknown loudspeakers has sounded quite acceptable; the loudspeakers (and perhaps the room in which they were placed) were apparently masking the defects of the matrix. This should not be taken as an excuse for lowering the specifications for studio monitoring conditions; on the contrary, it is a good reason for tightening standards if quadraphonic programmes are to match the high quality of present stereo productions.

Loudspeaker placement should also be carefully considered if meaningful results are to be obtained. As an extreme example of the problem, the sensations produced by quadraphony in a free-field room were found to be severely affected when one of the front loudspeakers was 5 cms (2 inches) too close to the listening position relative to a nominal 2·4 m (8 ft) radius for the other loudspeakers. This misplacement caused centre-front images to be shifted markedly towards the 'close' loudspeaker. When attempts were made to correct this by adjusting the relative signal levels it was found that an imbalance of 3 dB was needed to centralise the sound which produced a very phasey, diffuse image. Such a positional error would give rise to less than 0·2 dB difference in sound-pressure level at the listener and only 150 μs difference in time of arrival of the two sounds. It is, however, equivalent to a 90° phase error at 1·6 kHz, which would cause the unpleasant sound sensation.

Normal listening environments are not nearly as critical in their requirements as a free-field room, but consistent results can only be obtained when the loudspeakers are correctly placed.

2.2.1 Balancing quadraphonic loudspeakers

The best way to balance the four loudspeakers is an aural method (see Fig. 2). Loudspeaker A is first fed with a monophonic signal and adjusted to give a suitable listening level. The same signal is then also fed to B and

* Same make and model with well matched characteristics.

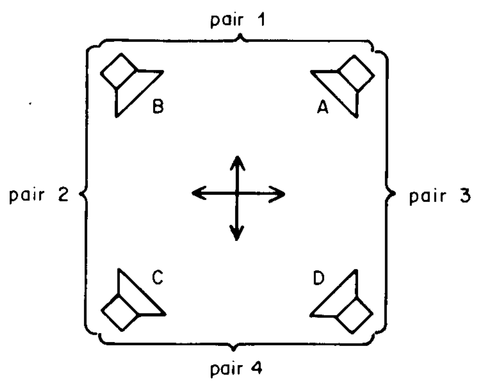


Fig. 2 - Balancing four loudspeakers

its level adjusted to generate a central image when listening on the centre line that divides the two loudspeakers. A is then disconnected and C energised; a balance is achieved as before, facing loudspeakers B and C. The same procedure is repeated for Pair 3 and if correct adjustments have been made Pair 4 should also give a similar result.

A method has been suggested in which the listener faces the front whilst balancing all four loudspeakers; however, for a centre-side image created by two loudspeakers, the ear appears to be more conscious of the front loudspeaker than the one at the back.* This method of balancing, therefore, is not recommended.

2.3 Listening position

Quadraphony in an absolute sense provides only one ideal listening position; with a square layout of loudspeakers this position is at the centre of the square. The quadraphonic sensation generated at this point is far superior to anything achievable elsewhere in the listening room, and it is to be recommended that any sound balancing engineer should be seated in such a position.

At other points in the quadraphonic square it is still possible to achieve a very pleasant quadraphonic effect but the directions of images are altered. In fact it is possible to stand on one side of the quadraphonic square and still receive the impression that the sound has retained its two-dimensional quality. Fig. 1 shows three secondary listening positions where a satisfactory effect is obtainable, but note these have been chosen specifically such that there is no-one directly between the best listening position and any of the loudspeakers; a head is extremely good at blocking out sound and ruining a quadraphonic balance.

2.4 Listening room requirements

As already mentioned, symmetry is the key to good reproduction and even the room can have a noticeable effect on the placement of quadraphonic images. In the early stages of the investigation into properties of hearing relevant to quadraphony, subjective tests were carried out both in a listening room and in a free field room. The listening room was designed to be acoustically similar to normal domestic lounge (reverberation time approximately 0·3 secs), but with the acoustic treatment more evenly distributed. The room is approximately 5·5 m (18 ft) long by 4·2 m (14 ft) wide by 3 m (10 ft) high and the bulk of the acoustic treatment is placed on three of the walls and the ceiling; there is also carpet on the floor. The fourth wall, at the back of the quadraphonic square, is the main source of asymmetry. There is a central window (1·2 sq.m., 13·7 sq.ft.) for visual communication with a control room, a door (1·8 sq.m., 20 sq.ft.) to the right of the window and a small amount of acoustic treatment (1·3 sq.m., 14 sq.ft.) to the left of the window. From the above description it can be seen that the listening room used for these quadraphonic assessments is more symmetrical than most domestic environments and certainly better than the normal control room associated with a sound studio. Even so, the effect of the window in the listening room has been detected and the quantitative results obtained have reflected the slight asymmetries.

A recommendation of perfect symmetry for control rooms dealing with quadraphonic signals would obviously be impractical, if not impossible, but gross asymmetries should be avoided. The loudspeakers, control desk and room should be arranged to be mutually symmetrical and if any imbalance has to be tolerated it is preferable to arrange this to be in the front-to-back direction rather than from left to right. Thus, for example, if a window is required between a control room and its studio, that window should be arranged to be in the centre of the front edge of the listening square, rather than on the left- or right-hand side of the square. Any large reflecting areas, e.g. equipment bays, ventilation trunking, should be avoided if possible, or separated from the listening area by acoustic screens. The latter arrangement will both prevent misleading reflections and afford a small degree of acoustic separation where the equipment bay is a source of interfering noise, e.g. transformer buzz.

3. Recording equipment

3.1 Tape recorders and associated equipment

Normally, at some stage of the production process, a four-track master of the quadraphonic material is produced; this may be derived directly from coincident microphones or from a particular balance of many microphones (in the latter case, an intermediate multitrack recording may be produced which is later balanced to four tracks).

There is, however, one main point to be borne in mind with regard to recording: this is that the signal-to-noise ratio should be maintained as high as possible, which implies, at the present time, that some form of noise-reduction technique be used to reduce the effect of tape hiss. During the early experimental recordings of quadraphony it was found that a disturbing level of tape hiss was reproduced over the rear loudspeakers in spite of

*See appendix

the fact that the rear-channel recordings had the same signal-to-noise ratio as the front channels. It was then noticed that, whilst facing the back of the quadraphonic square, the front loudspeakers, i.e. the ones that were now behind the listener, were those which subjectively produced the most tape hiss. This phenomenon was confirmed by several subjects and it is therefore concluded that the ears appear to accept a level of tape noise coming from in front which is higher than that which is acceptable from behind. To alleviate this problem later quadraphonic recordings were made using Type A Dolby noise-reduction units.

To find what effect the Dolby units might have on a quadraphonic recording eight Model 361 units were placed in series. Alternate ones were switched to the encode condition and the others were switched to the decode condition. Measurement of the phase and frequency responses gave results which were within the specification of 5° and ± 1 dB per encode-decode. On a listening test, comparing a monophonic signal before and after it had been passed through this chain of units, there was a just perceptible change in the tonal quality of one item of music, but on other items of music and speech there was no noticeable difference. It was therefore concluded that the units tested performed their task of noise reduction with no significant side effects. This conclusion has been confirmed by all the subsequent recordings of quadraphonic material, including fourth generation copies of some test material.

One other problem which may arise in quadraphonic recordings is that of phase errors due to tape weave. There is a need for this to be investigated further, particularly with respect to its effects where channel reduction techniques are used, especially 4—2—4 matrices with 'logic' decoding.*

3.2 Microphones for quadraphony

Several aspects of microphone technique were briefly investigated** before experimental orchestral recordings were made at the 1972 Promenade Concerts. Two of these were the comparison of coincident versus spaced-microphone configurations and the effects of changing the directional characteristics of the microphones.

For all the preliminary work, four AKG C12A capacitor microphones were used. These were tested in a free-field room and their characteristics were found to be matched to better than ±½ dB from 50 Hz to 3 kHz and better than ±1 dB from 40 Hz to 16 kHz.

Fig. 3 shows some of the microphone configurations tested. These configurations were set up, in turn, in a studio and the amplified microphone outputs were fed directly to four monitoring loudspeakers in a separate listening room, great care was taken to ensure that each channel had exactly the same gain. An engineer walked

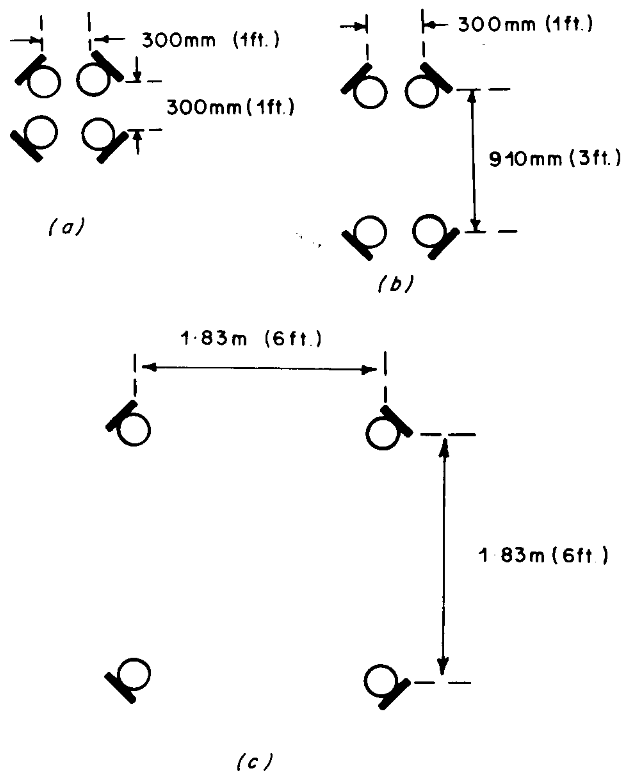


Fig. 3 - Microphone arrangements for the coincident v spaced microphone tests

round the microphones on a 3 m (10 ft) diameter circle, reading from a script, and a note was made of the locus and quality of his image in the listening room. The tests were repeated for the different microphone placements and for different microphone polar-responses; omnidirectional, cardioid and 'cottage-loaf' responses were used.

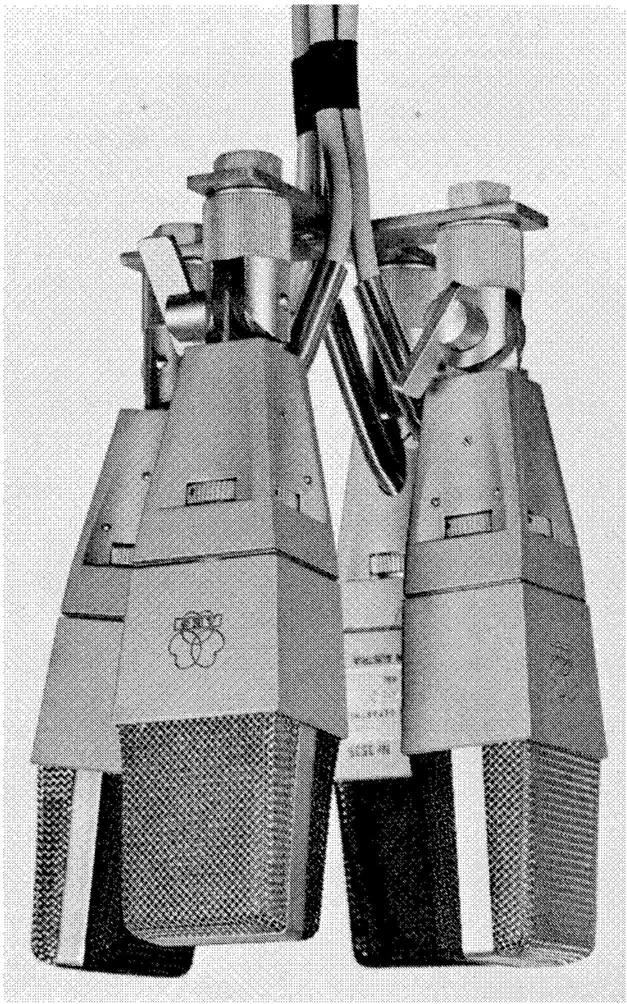
The most accurate results were obtained using layout 'a' (Fig. 3) with the microphones switched to either the cardioid or cottage-loaf response. Under these conditions a listener was able to locate the position of the engineer anywhere around the microphones to within $\pm 5^\circ$ of the true location.*

Both the spaced microphone layouts suffered from the same defect, namely that as the engineer walked around the circle embracing the spaced microphones, his sound image, which should have moved linearly between the corresponding loudspeakers, actually clung to the first loudspeaker until he was almost centrally placed. The image then moved rapidly to the second loudspeaker and stayed there until he was well beyond the axis of the second microphone; this is the well known stereo 'hole-in-the-middle' effect appearing in a quadraphonic form.

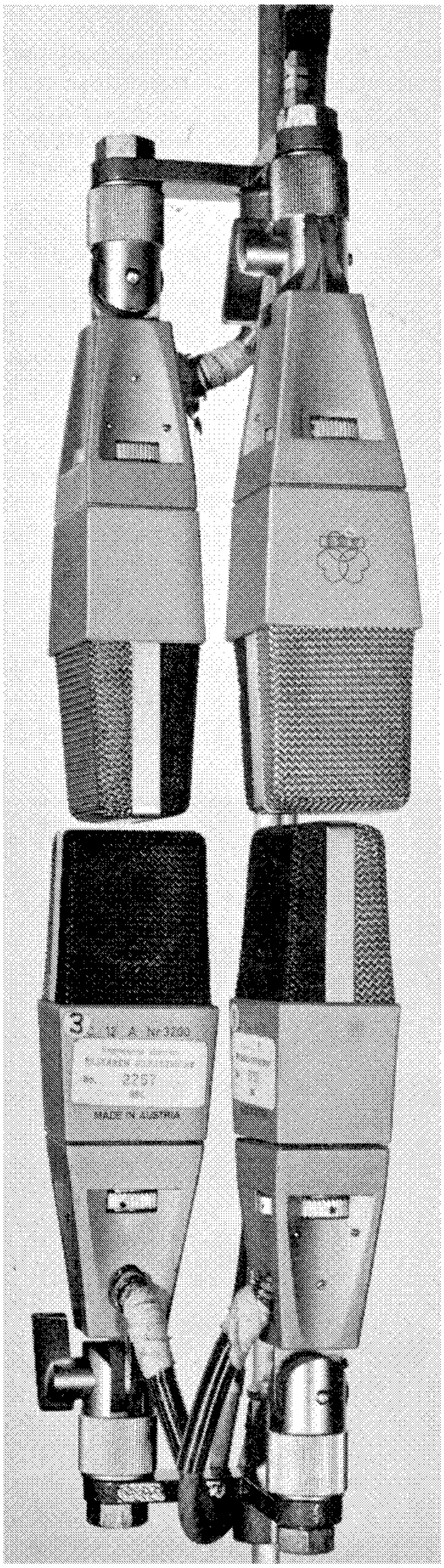
Based on these results it was concluded that coincident-microphone techniques produced better all-round results than spaced microphones. This was confirmed

*'Logic' implies non-linear circuits used to enhance separation under specified conditions.
**This work was carried out by H.D. Harwood.

*During these tests the listener was allowed to turn and face the sound image.



(a)



(b)

Fig. 4 - Microphone configurations

(a) Horizontal planar array (b) Two-plane array

by an orchestral recording made in the Town Hall at Barking where again coincident and spaced microphones were tested. Wherever the microphones were separated, the corresponding side of the quadraphonic square suffered a lack of positional discrimination during replay, whilst four coincident cardioid or cottage-loaf microphones gave a pleasant balance and distribution of sound images.

One property of the cottage-loaf response was however noted. The fairly large rear-lobe in the polar diagram (Fig. 6c) appeared to result in directional ambiguity across the diagonal of the square. The situation was a little confused by a full-width balcony in Barking Town Hall, relatively close to the microphone cluster. This gave rise to just audible reflections, even with cardioid microphones, but it was considered that the cottage-loaf recordings were slightly worse in this respect. Subsequent recordings, therefore, were made only with the cardioid microphone-response or with a response mid-way between cardioid and cottage-loaf,* and the diagonal ambiguity, although occasionally noticed by listeners with very acute hearing, has not proved to be a problem.

*The latter response will hereafter be referred to as the 'hyper-cardioid' response.

Having found that the coincident configuration gave the best quadraphonic results, two different clusters of microphones were tested during subsequent recordings, see Fig. 4. The horizontal square array was the first one to be devised. The cluster had a diagonal of 100 mm (4 inches), which is equivalent to a quarter wavelength at 825 Hz. The discrete (4–4–4) quadraphonic results obtained with this arrangement were very good but it was felt that systems involving 4–2–4 (and possibly 4–3–4) matrices might require even closer placement of the capsules. The two-plane array was therefore suggested, with maximum capsule spacing of 57 mm (2·3 inches), which corresponds to a quarter wavelength at 1·4 kHz. This was the microphone arrangement used throughout the Promenade Concert recordings.

3.3 Microphone placement and associated problems encountered during recording sessions

When using a coincident microphone array it is not always easy to find a position in the recording environment, the concert hall or studio, where a satisfactory balance of sound exists. For an orchestral recording, a satisfactory placement would be one in which a wide clear image of the orchestra was obtained together with a balance between direct and reverberant sound. In the case of the Prom-

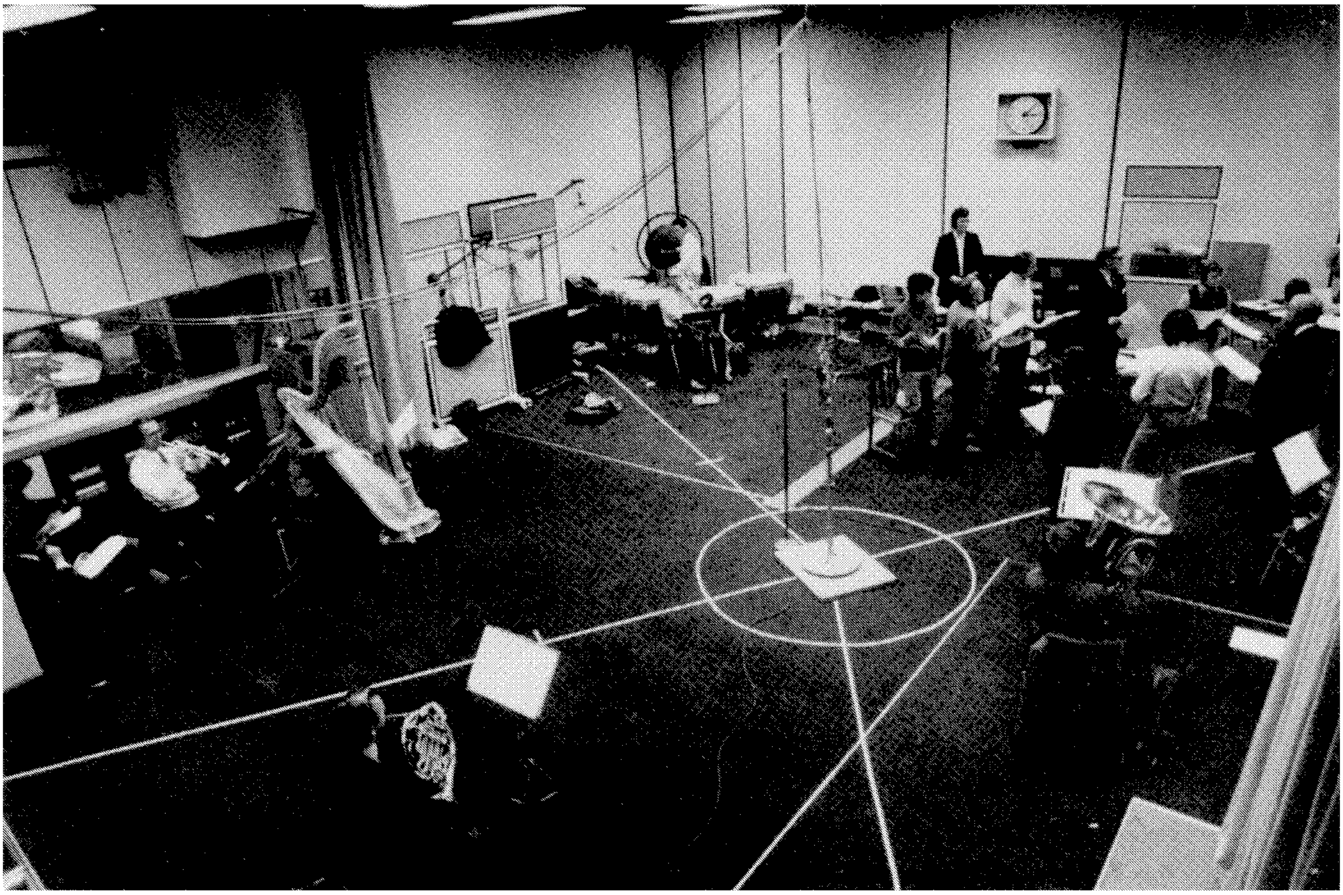


Fig. 5 - Recording of quadraphonic drama

enade Concerts three different microphone positions were tried before a satisfactory balance of direct and reverberant sound was obtained, and this, due to the distance of the microphones from the orchestra, slightly affected the orchestral image, which tended to be marginally narrower and more distant than in the equivalent stereo presentation. This will always be a problem, in certain environments, unless spot microphones are also used. It is considered, however, that this is an artistic rather than an engineering problem and so the use of additional spot microphones was not investigated.

The final position selected for the quadraphonic microphones at the Royal Albert Hall was approximately 7·2 m (24 ft) behind the conductor's rostrum and 6 to 9 m (20 to 30 ft) above the floor. This gave good recordings except for the previously mentioned point that the orchestra occasionally sounded rather distant. There is in the recordings thus obtained one factor which gives the listener a sense of participation in the concert, rarely achieved even by the best stereo production; that is audience reaction. The occasional cough from a member of the audience, coming from the correct direction during replay, gave the recording a sense of realism and, far from being distracting, seemed to add to the overall effect.

Unfortunately such all-round reception has the drawback that the microphones pick up unwanted sounds from all directions. This has not proved troublesome in the orchestral recordings, but a quadraphonic drama recording of 'Oedipus' (fig. 5) suffered from a high level of environmental noise; the collective effect of four microphones relative to one gives a 6 dB decrease in the ratio of signal to environmental noise. In the case of the Oedipus recording the situation was further worsened by the fact that the actors were further away from the microphones than is usual for a stereo drama production. The result was a recording with a relatively high level of low-frequency noise, picked up from the ventilation plant and other sources. The noise was easily removed from the recording by filtering out all signals below 70 Hz, but this indicates that close microphone techniques are required for drama unless much more stringent acoustic criteria are adopted from drama studios. Close microphone techniques imply that there will be very little room for movement by the actors, or even for placing actors in

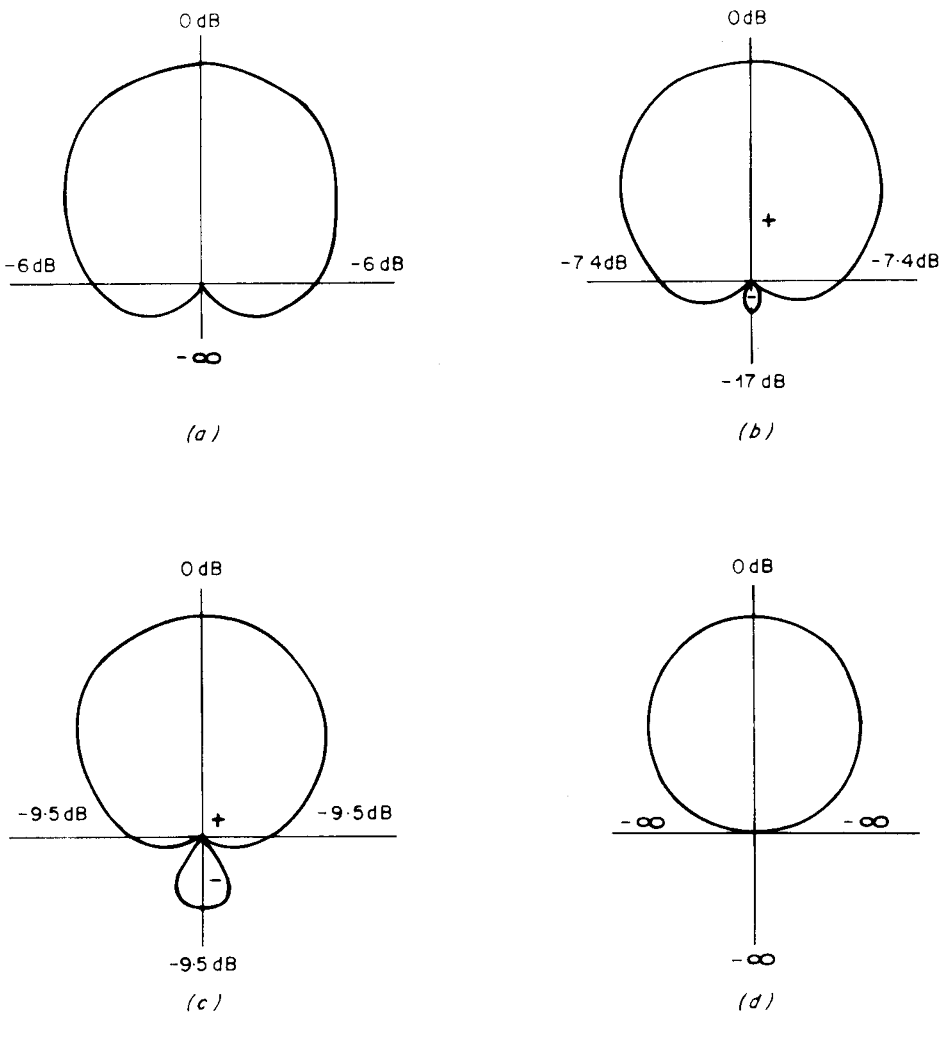


Fig. 6 - Analysis of microphone characteristics

(a) Cardioid: $V \propto 1 + \cos\theta$

(b) Hypercardioid: $V \propto 0{\cdot}75 + \cos\theta$

(c) Cottage loaf: $V \propto 0{\cdot}5 + \cos\theta$

(d) ½ figure-of-eight:

$V \propto \cos\theta$ *for* $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$,

and $V = 0$ *for* $\frac{\pi}{2} < \theta < \frac{3\pi}{2}$

the correct location round the microphones, and thus pan-pot methods of image placement may be more favoured for future drama productions.

4. Microphones: future development

During the experimental recording work cardioid and hypercardioid microphone responses were used and it was noted that the adjacent-channel crosstalk, due entirely to the polar responses of the microphones, affected the subjective impression provided by the recordings. The effect was not one of mislocated images, as in this respect the recordings were satisfactory; the crosstalk tended to 'pull the sound image in' towards the listener relative to the impression created by a pan-pot recording.

As can be seen from Fig. 6(a) a cardioid response is only 6 dB down when the source is 90° off axis, i.e. a noise from the right-front of a sound stage will generate signals in the left-front and right-back microphones only 6 dB lower than that generated in the right-front microphone. A switch to a hypercardioid response, as shown in Fig. 6(b), improves the adjacent separation slightly but at the expense of an antiphase rear lobe. A cottage-loaf response, Fig. 6(c), worsens the overall performance because of the greater size of its rear lobe. Fig. 6(d) shows a microphone response that would be ideal for quadraphony in that its output is

$$V \propto \cos\theta \text{ for} -\frac{\pi}{2} < \theta < \frac{\pi}{2}$$

and

$$V = 0 \text{ for } \frac{\pi}{2} < \theta < \frac{3\pi}{2}$$

Four of these could adequately cover the quadraphonic sound stage giving coincident-microphone signals with the same interchannel separation as pan-pot signals. In fact if a sound source moved around a truly-coincident, cluster of this form, the signals would vary in the same way as if they had been generated using the sine/cosine method for the pan-pot placement of sound images.

A useful approximation to this characteristic may be made by using a second-order pressure-gradient microphone. The response of this type of microphone is shown in Fig. 7. It can be seen that this has the required shape of front-lobe and that the response behind the microphone is everywhere more than 18 dB below the axial response. Four of these in a quadraphonic cluster would give omnidirectional sensitivity (in the horizontal plane), as in the case of four cardioid microphones; therefore, signals obtained using such microphones would give equally good directional accuracy of image, with improved sharpness and possibly a more 'open' impression.

It should be noted that these improvements will only be valid if a four-channel transmission system is available. Coincident cardioid microphones already make full use of the channel separation available in a three-channel system. There may, however, be practical advantages to be gained with the new type of microphone, such as an increase in the ease with which the cluster can be used in an acoustically difficult environment (see section 3.3).

There is, however, no commercial microphone available at present with the desired characteristic and the development of one, covering the audio band with high signal-to-noise ratio, would require a considerable amount of work.

5. Conclusions

The experimental quadraphonic recordings have supplied a wide range of material for qualitative tests to be carried out on possible broadcast systems. The experience gained has pointed out some of the possible problems of quadraphony. It appears to be more critical than stereo with regard to listening position and to be less tolerant of listening environment, if images are to be accurately located; such performance should be the objective, although for normal domestic reproduction these factors are far less important. Pleasing results, even though they may to some extent be inaccurate, can be obtained in a wide range of circumstances.

Only the square loudspeaker layout is recommended if accurate image-location is required, and symmetry in the listening environment is also recommended.

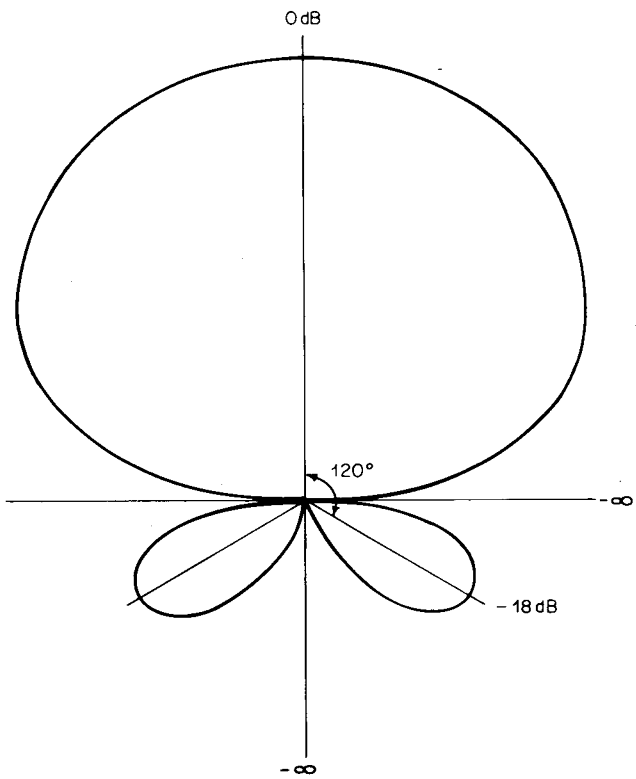


Fig. 7 - Polar response of second order gradient microphone $V = v_o\,(1 + \cos\theta)\cos\theta$

With regard to signal origination, two techniques have been outlined, the pan-pot method, using several (or many) separate microphones, and the coincident-microphone method, although only the latter has been discussed in detail. If the coincident-microphone method is to be used in the long term, the development of a microphone with a more suitable polar characteristic would be a significant step forward.

6. References

1. RATLIFF, P. Properties of hearing related to quadraphonic reproduction. BBC Research Department Report in course of preparation.

2. GERZON, M. 1971. Experimental tetrahedral recording; Part One: Studio sound, 1971, **13**, 8, pp. 369/8. Part Two: Studio sound, 1971, **13**, 9, pp. 471/5.

3. COOPER, G. 1970. Tetrahedral ambiophony, Part One. Studio sound, 1970, **12**, 6, pp. 233–4.

7. Acknowledgement

The author is grateful to many people in Radio O.B.s and Television O.B.s who co-operated in the making of recordings, at the 1972 Promenade Concerts and elsewhere.

Appendix

Work has been carried out to examine the properties of human hearing in order to determine what the ear/brain combination is capable of deducing from quadraphonic sound presentation. This work is to be reported in full in a later document.[1] One factor which is extremely relevant to this report, however, is the quality and position of the sound images generated at the side of the head by two equally energised loudspeakers, one towards the front and the other towards the back. Centre-side images have been found to be extremely sensitive to head movement and to shift rapidly with changes in the level-difference between the two loudspeakers. During tests in a free field room it was noticed that the side images were subjectively judged to be much further forward than expected. In fact on average the rear loudspeaker needed to be 10 dB louder than the front one to generate a centre-side image and a variation of ±½ dB moved the image through ±15°. This effect, however, is only applicable when the two loudspeakers are simultaneously energised by the same signal. Tests have also shown, that given an 'either/or' comparison, a listener can balance a loudspeaker at the front with one at the back to within 1 dB. Thus the ear/brain combination appears to exhibit a conditional directionality that only exists when identical sounds are presented in front of, and behind, the listener at the same time.

CLA/AM

BBC RD 1974/38

1969 THE QUEEN'S AWARD TO INDUSTRY
1974 THE QUEEN'S AWARD TO INDUSTRY

RESEARCH DEPARTMENT REPORT

Properties of hearing related to quadraphonic reproduction

P.A. Ratliff, B.Sc., Ph.D.

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION November 1974

BBC RD 1974/38
UDC 534.76

PROPERTIES OF HEARING RELATED TO QUADRAPHONIC REPRODUCTION

P.A. Ratliff, B.Sc., Ph.D.

Summary

An investigation of some of the properties of hearing relevant to the reproduction of an omnidirectional sound-stage has been undertaken. The work was designed to provide basic knowledge of certain properties of hearing, and to examine under critical conditions some of the phenomena which occur with quadraphonic (four-loudspeaker) reproduction of a sound field.

Three fundamental, sound-locating properties of the human auditory system have been determined, and a law has been established relating interchannel level and image location around the listener, using quadraphonic sound-source arrangement. Effects of unwanted signals in the reproduced field are examined, and also the effects of phase-shifts inserted between these signals, such as those which typically occur in the matrix quadraphonic systems currently under consideration by many workers. Results expose some psycho-acoustic myths, and account for some of the observed peculiar phenomena in practical matrix systems.

Issued under the authority of

Head of Research Department

Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION
November 1974

(PH-129)

PROPERTIES OF HEARING RELATED TO QUADRAPHONIC REPRODUCTION

PROPERTIES OF HEARING RELATED TO QUADRAPHONIC REPRODUCTION

P.A. Ratliff, B.Sc., Ph.D.

Terminology

For brevity in this report abbreviations for directions with respect to the listener are used extensively, along with a few other commonly used terms which are listed below.

A linear sixteen point direction scale for the horizontal plane is introduced in which direction '0' is always directly in front of the listener. Even numbered directions are also referred to by an alphabetic code indicating the directions in words, as shown in Diagram A.

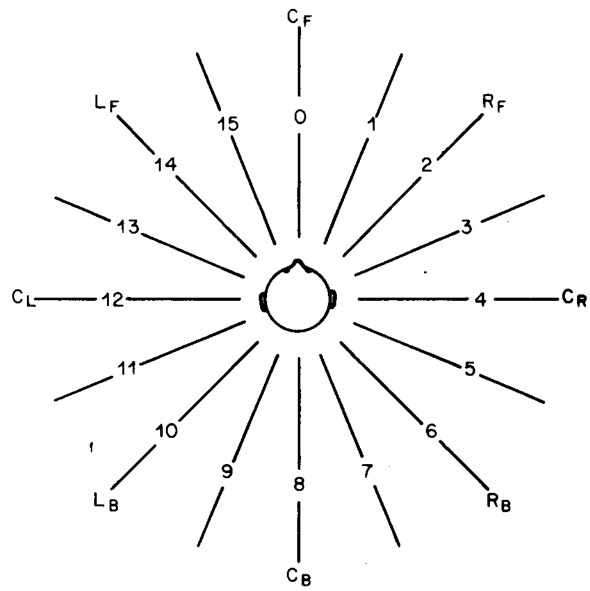


Diagram A

C_F centre-front, R_F right-front, C_R centre-right, R_B right-back
C_B centre-back, L_B left-back, C_L centre-left, L_F left-front

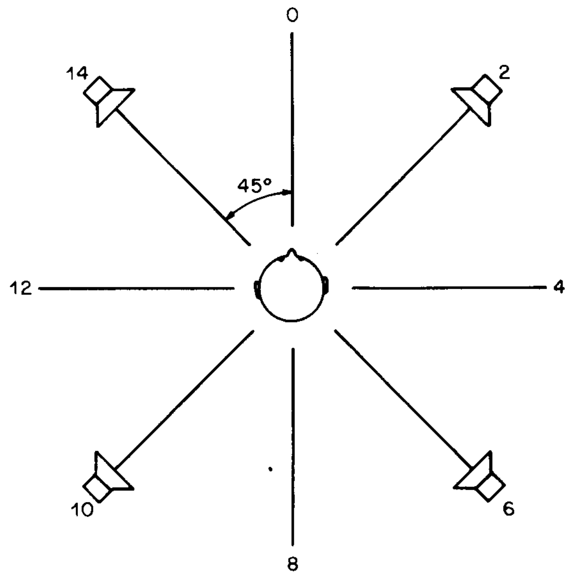
There are two well known quadraphonic arrangements of loudspeakers, the 'square' and the 'diamond' arrays as shown in Diagram B.

Much of this report deals with the 'square' array in which the loudspeaker positions are sometimes referred to as 'corner locations', and the span between any adjacent pair of loudspeakers is referred to as a 'quadrant', specifically defined as 'front' (14–2), 'right' (2–6), 'back' (6–10) or 'left' (10–14). Directions 0, 4, 8 and 12 are then referred to as centre-quadrant directions.

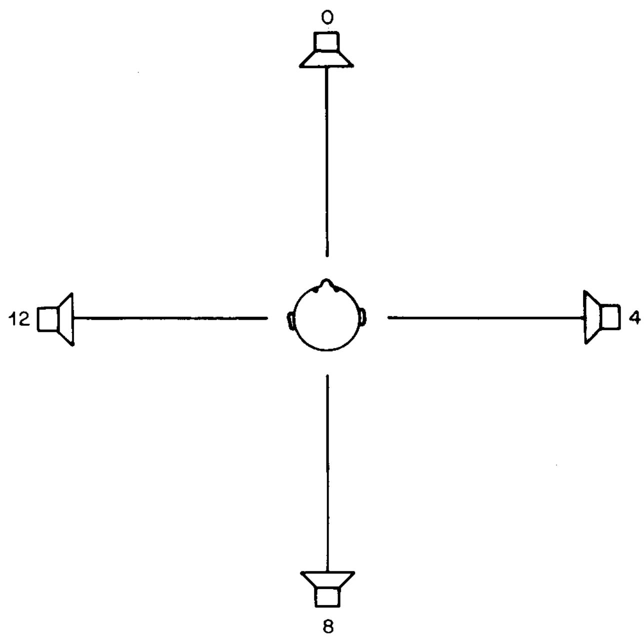
Other abbreviations used are listed below:

l. s.	loudspeaker
m.p.*l.*	minimum perceptible level
s.d.	standard deviation
s.p.l.	sound pressure level

(PH-129)



Square Array



Diamond Array

Diagram B

The following sound quality abbreviations are used:

B bass heavy
C close, near
D diffuse
F far, distant
G good, no adverse comment
H high)
L low) vertical position
NH normal)
J jumpy, horizontal image location jumps between loudspeakers
N nasal, bass lacking
P phasey, in the head sensation
s slightly (used in conjunction with above, i.e. sD, slightly diffuse)
v very (used in conjunction with above, i.e. vD, very diffuse).

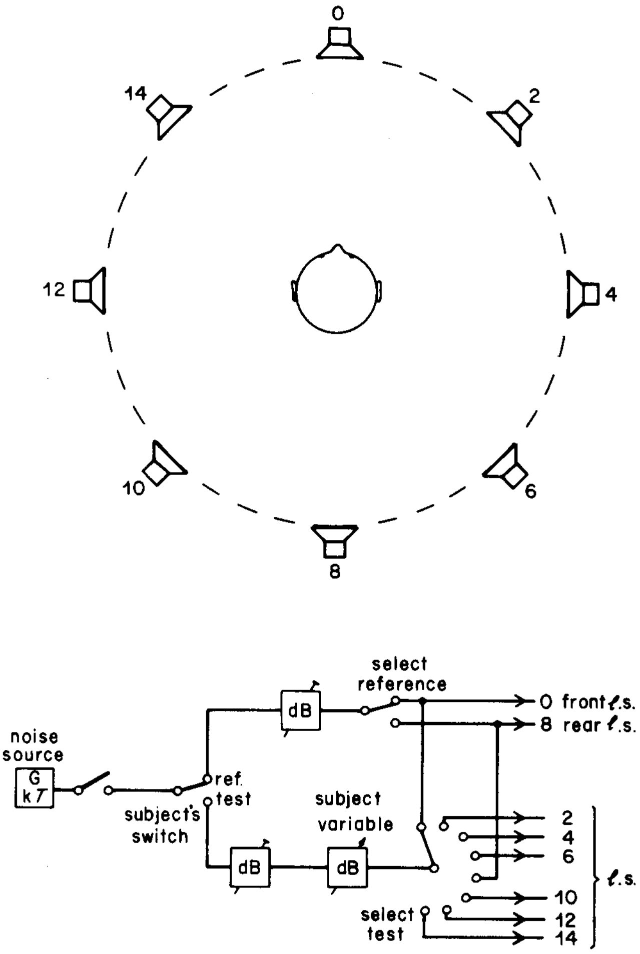


Fig. 1 - Arrangement for 'equal-loudness levels' experiment

1. Introduction

Increasing demands for standards on quadraphonic 'surround-sound' reproduction have led to the realisation that a greater fundamental knowledge of the properties of the human auditory system is required, if an optimum technical solution is to be obtained. This report contains experimental results on the angular (azimuthal) localisation and subjective quality of real and imaginary sound sources, and considers the effects of 'unwanted signals' in four-loudspeaker reproductions of uniquely localised images, typical of those produced by some recording techniques, and by matrix quadraphonic systems.

An understanding of the processes of human hearing is at present far from complete, and although there is a considerable quantity of literature on the subject (e.g. see references of Ref. 1), theories presently proposed[2,3] do not account satisfactorily for all observed phenomena. However, for present requirements, it is necessary to determine the fidelity with which a quadraphonic system can reproduce a sound field which *subjectively* satisfies the listener.

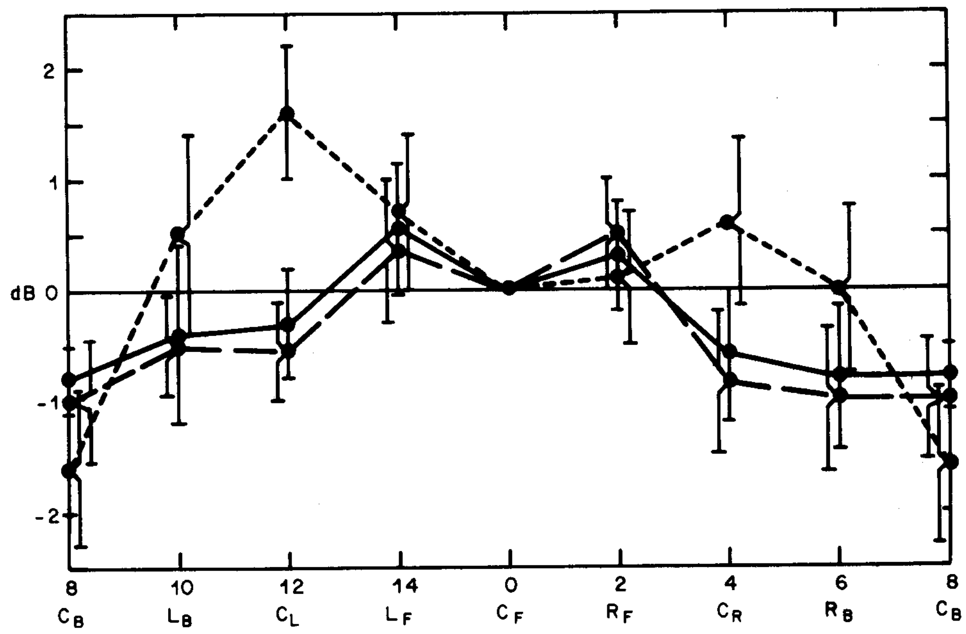
2. Equal loudness levels

It is well known that the sound levels at each ear differ, particularly at high frequencies, depending on the azimuthal direction of the source,[2] and this has been put forward to support an inter-aural intensity hypothesis of localisation.[4] It is thought unlikely, however, that this frequency-dependent difference is the sole determining factor. There is greater support for the inter-aural time-difference hypothesis,[5] although this is probably an over simplification of the localising process.[3]

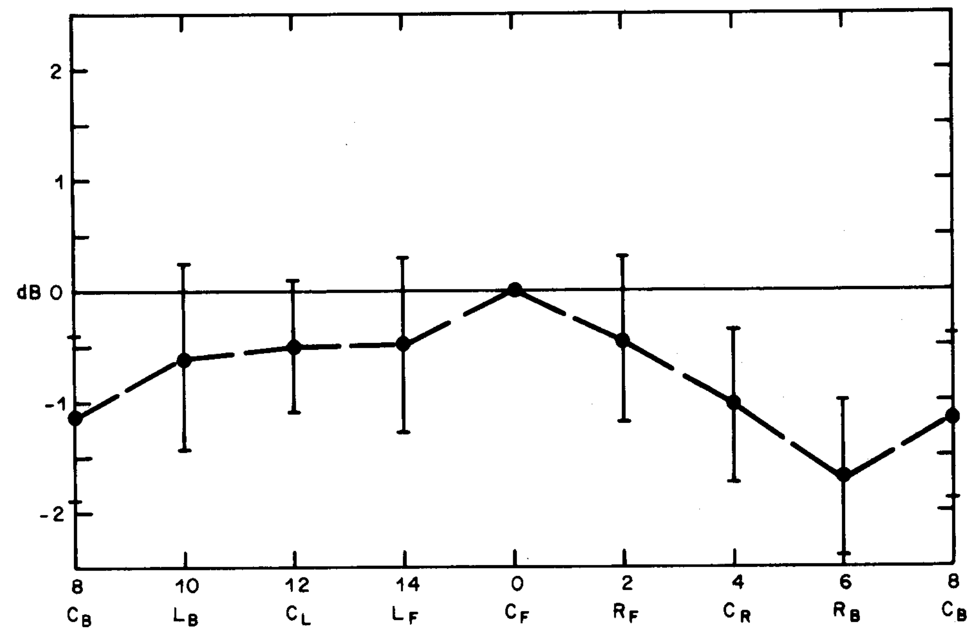
An experiment to determine the subjectively-assessed equal-loudness levels around an observer was conducted,* in which the observer was presented with sound from one of eight closely-matched loudspeakers arranged symmetrically around him. He was provided with a switch and a calibrated attenuator, and asked to adjust the attenuator until the sound-source was of the same loudness as a fixed reference, which could be selected by operating the switch. The reference loudspeaker was always immediately in front of, or behind the subject, who was seated in a chair with the nape of his neck against a thin wooden head-rest. The subject was asked to keep his head still, facing the front throughout the test; the arrangement is shown in Fig. 1.

The experiment was conducted in a specially designed 'average' listening room (about 70 m^3 capacity and average reverberation time of 0·35 sec.) using octave bands of pink-noise, centred on 230 Hz, 2 kHz and 7 kHz, reproduced at normal listening levels (about 70 dBA) from eight high-quality loudspeakers equally spaced on a circle 3·35 m in diameter. The results of tests using eight observers are shown in Fig. 2(*a*), and show little variation with azimuth. The experiment was repeated in a free-field room (surface reflection less than 10% at all frequencies above 40 Hz), but realistic results could only be obtained in the 2 kHz band, because the degree of loudspeaker matching was insufficient for observers not to be disturbed by subjective quality differences between them at other frequencies; the degree of subjective quality matching required under non-reverberant conditions such than an observer can make a loudness assessment with conviction is very high indeed,

* Work undertaken by T.W.J. Crompton.



(a)



(b)

Fig. 2 - Equal-loudness levels of octave bands of pink noise

(*a*) Listening room (*b*) Free-field room

——— 230 Hz) band — experimental result showing standard deviation

— — 2 kHz)

------- 7 kHz)

but small mis-matches are effectively masked by the reverberant fields present in typical listening rooms. The result of the free-field room experiment is plotted in Fig. 2(*b*), and again shows little variation with azimuth. It is concluded that the average auditory response is almost equally sensitive around the full azimuth circle, although there is a consistent trend for the back to be less sensitive than the front by about 1 dB, increasing slightly at high frequencies.

3. Absolute perception of direction

An important feature in 'surround-sound' reproduction systems is their ability to reproduce the directional properties of the programme to a subjectively acceptable standard,

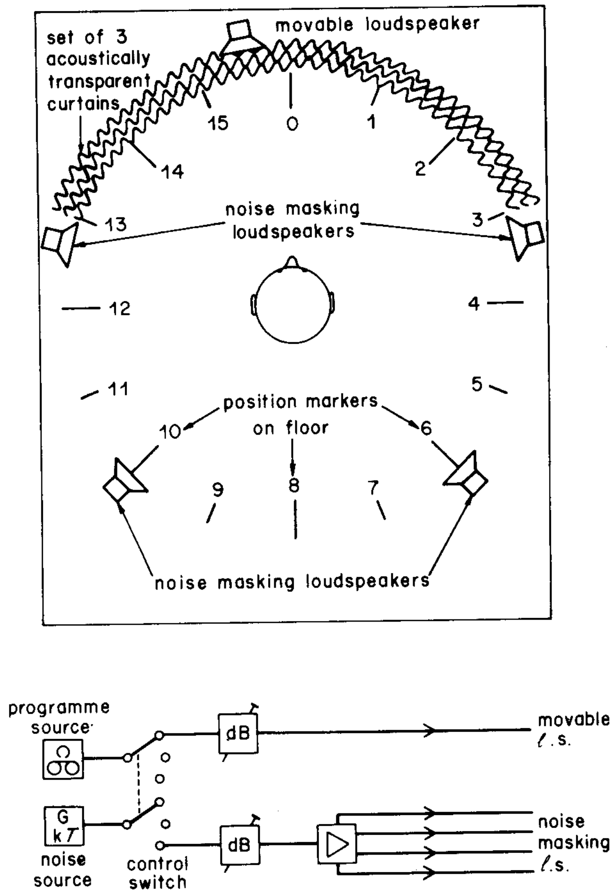


Fig. 3 - Listening room arrangement for localisation experiment

and to this end the absolute azimuthal accuracy of the auditory system was determined.* Initial tests were performed in the listening room with the arrangement shown in Fig. 3. The subject faced acoustically transparent curtains or sat at some multiple of 90° to this direction to examine each quadrant separately. This was necessary as there was insufficient room to completely surround the subject by curtains and examine the full compass in one experiment. The subject was asked to locate the sound-source (a moveable loudspeaker) in each of several tests, which were interposed with masking noise (from the four loudspeakers outside the curtains) to conceal any loudspeaker movement noises. The subject was given a chart (see Fig. 4) dividing the compass into sixteen 22½° segments (units), and was initially asked to place the sound-source in relation to this scale; markers were also placed on the floor to assist in angular awareness. The programme material for these tests consisted of a repeated 30 second excerpt of percussive music. This was found, in preliminary work, to be the most critical material for image localisation assessments.

* 'Absolute azimuthal perception' is defined as that related to the localisation of a single sound-source, and the term 'relative azimuthal perception' is used to denote that describing the relative localisation of one sound-source with reference to another that is closely spaced to it.

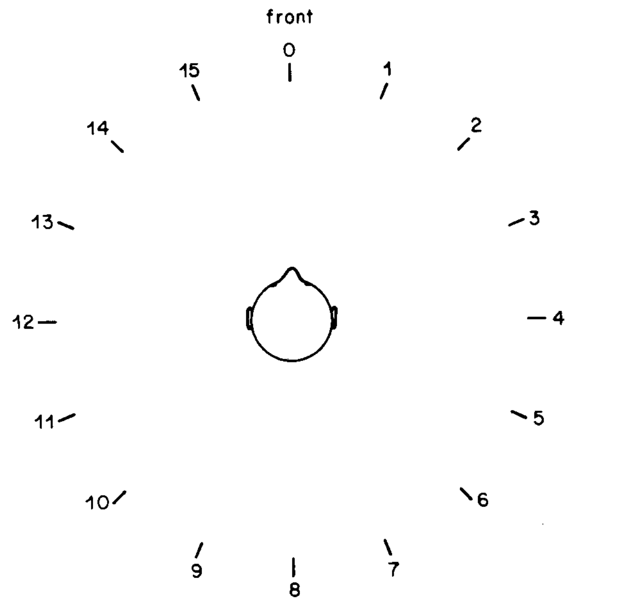


Fig. 4 - Location chart used in localisation experiments

A number of psychological problems were encountered in finding a suitable locating method. Quantisation of the position assessment occurred whether the sound-source was placed on a marker or between them, and since an increase in the number of marker positions would merely serve to confuse the subject, a second method was used where the experimenter moved the loudspeaker until the subject was satisfied that the source was where he had been instructed to locate it. During all such movement the loudspeaker was switched off and masking noise was switched on to avoid giving localising information by movement, and to dissociate each test. This second method proved more acceptable, and results of seven observers are shown in Fig. 5.

The front-stage quadrant is found to be fairly accurately defined (standard deviation, s.d. $\simeq \pm 2{\cdot}5°$) with marginal image expansion at the extremes of the quadrant. Centre-back (C_B) is similarly well defined, but away from this location greater uncertainty of source position arises. Greatest uncertainty occurs at left-back (L_B) and right-back (R_B), and considerable rear-image expansion occurs; about 11° at L_B and R_B. The amount of image-shift and the standard deviation are greater in rear-quadrant examination than in side-quadrant examination, for the same nominal source positions.

This is thought to be a psychological phenomenon, probably due to visual cues which can modify the ability of the brain to make unbiased decisions. For instance, the finite width of the curtains and the knowledge that the loudspeaker was always constrained to be located behind the curtains could have given rise to the discrepancies observed during the tests.

These results, and comments made by the subjects, indicate a considerably greater ease in the localisation of sound-sources which appear in the front 'visual' quadrant. The ability to 'see' the sound source (although it was, of course, concealed behind the curtains) appears to improve

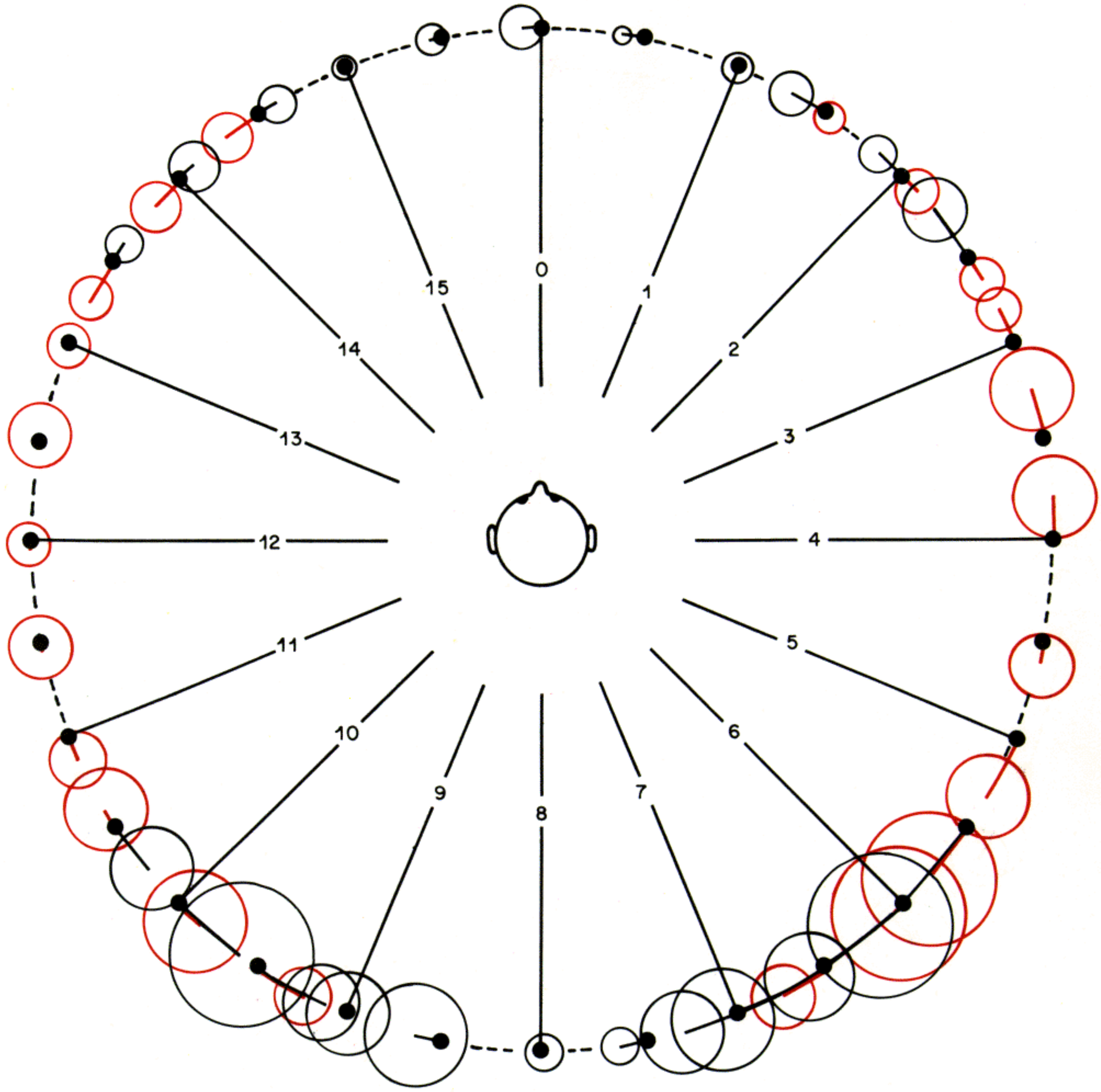


Fig. 5 - Absolute sound localisation in the listening room

image location

Results for front and rear quadrants

corresponding mean source positions showing standard deviation circles*

Results for side quadrants

* The circles denote the measured standard deviation (s.d.) of the source position. Its radius subtends the angular s.d. as perceived by the subject.

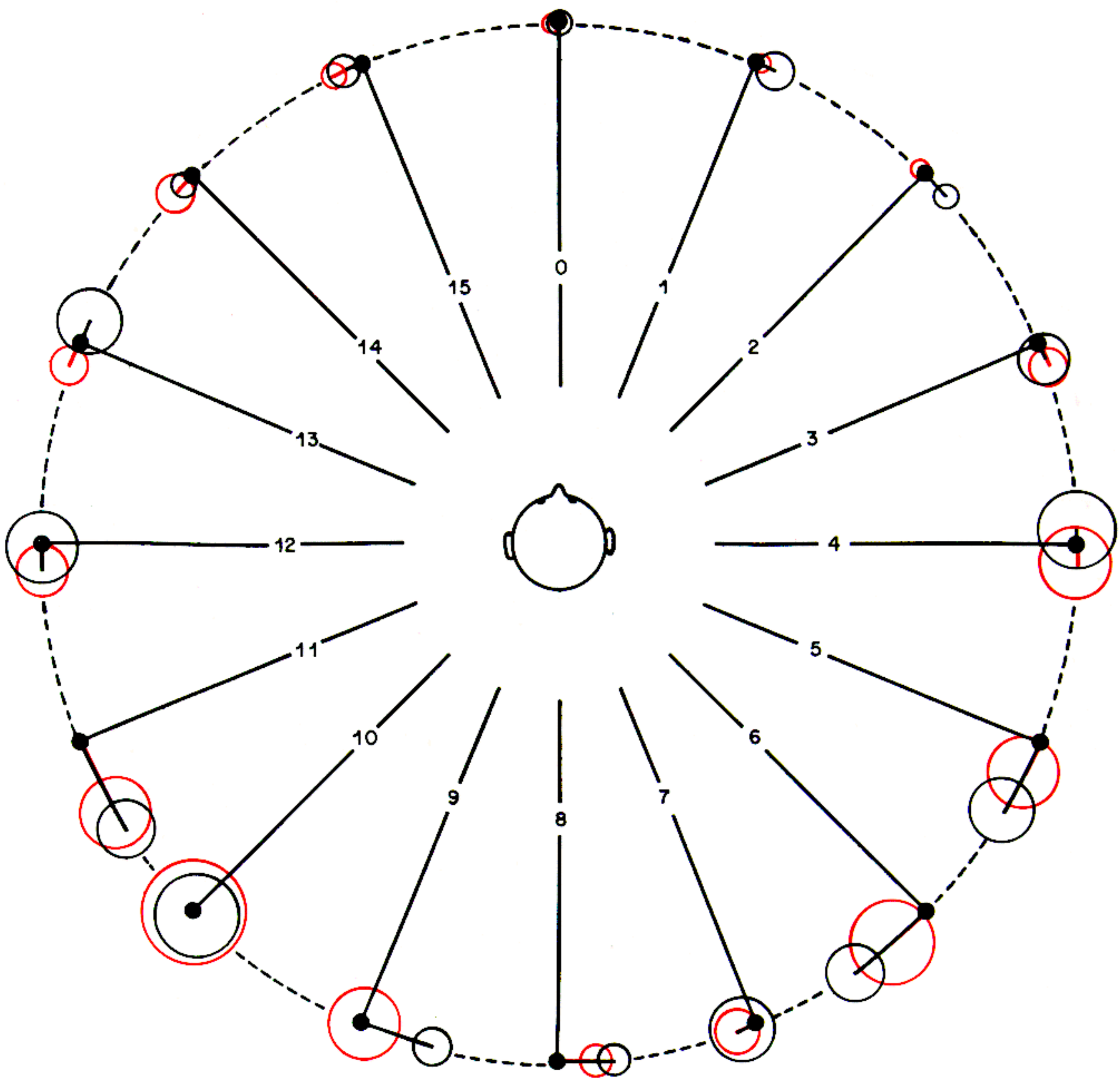


Fig. 6 - Absolute sound localisation in the free-field room

image location

Percussive music results

corresponding mean source positions showing standard deviation circles

Male voice results

the subjects' ease of sound-source localisation. This is possibly because the brain readily correlates information from multiple sensory inputs. In order to remove the effect of visual information, five subjects repeated the experiment in the front quadrant, blindfold, and their results showed remarkable similarities to those obtained for the rear quadrant previously. The uncertainty of position became similar to that in the rear quadrant, and image expansion again occurred, although only at the extremes of the front quadrant.

The anomalies observed made further use of the listening room undesirable, and so further work was conducted in the free-field room. There was sufficient room to construct a complete circle of curtains, 2 m in diameter, at the centre of which the subject was seated, with his head against the head-rest, facing centre-front (C_F). Similar position markers were affixed around the curtains, and the sound-source (a compact high-quality loudspeaker unit) was moved around a 1·5 m radius circle, beyond the curtains, at a height just below the subject's ears. Four loudspeakers in the corners of the free field room provided movement masking noise, as in the listening room. Illumination was provided only within the curtains to ensure that they formed a visually opaque screen to the subject. Results for seven subjects using percussive music are shown in Fig. 6 (in black), and are substantially more consistent than those obtained in the listening room.

Azimuthal acuity in the front semi-circle is good with little error, s.d.s only reaching about ±3·5° at the extremes (C_L and C_R). Again acuity at C_B is equal to that at C_F (s.d. ≏ ±1°), but a small left-hand image offset is observed. This is thought to be due to the entrance into the curtains being right of C_B, which may have pre-conditioned the subject's local impression of C_B. Away from C_B rear image expansion[1] is again evident, peaking at about 9° at L_B and R_B, with s.d.s of ±4·5°.

The percussive music programme excerpt used had considerable high-frequency spectral content, the first 15 seconds being mainly above 2 kHz and the latter 15 seconds mainly above 700 Hz. A programme excerpt was then

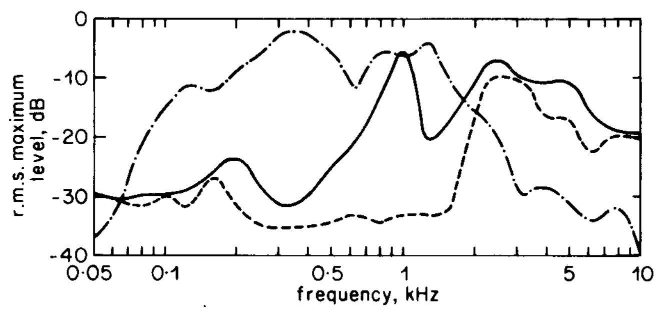


Fig. 7 - Spectral content of programme material used for subjective tests (as presented in the free-field room experiments)

——————— 30-second percussive music excerpt
– – – – – – – – first 15 seconds of percussive music excerpt
—— · —— · —— 30-second male voice excerpt

selected having a greater low frequency content and the experiments were repeated. This consisted of a 30 second news excerpt read by a trained male announcer, and had a spectrum essentially confined below 2 kHz (see Fig. 7). Results for this material (Fig. 6 (in red)) are similar to those to those using percussive music.

It is concluded that 'surround-sound' reproduction systems should be capable of accurately reproducing the absolute azimuths of sounds in the front semi-circle and at C_B, but some degree of latitude (±10°) is permissible in the rear semi-circle near L_B and R_B.

4. Relative perception of direction

Undoubtedly a more stringent requirement of 'surround-sound' systems is their ability to differentiate between two closely spaced sound sources, since under these conditions the human auditory response is involved in making a comparison. The experiments described in the previous section were repeated in the free-field room using a second loudspeaker as a reference sound-source, and the subject was asked to move the test sound-source so as to be directly in line with the reference. In this experiment the sources were a matched pair of compact loudspeaker units, the test-source traversing just above the reference, with their high frequency units placed adjacent to one another so as to minimise the subjective height difference (see Fig. 8). The subject selected the reference or test loudspeaker unit by means of a switch.

Results for both types of programme material are shown in Fig. 9, and it is seen that the azimuthal acuity is much more accurate than in the case of a single source, and now greatest uncertainty occurs at the sides of the subject (s.d.s ±3·5°) where audition becomes largely monaural. Localisation errors are not significant at any azimuth and there are no significant differences between the results obtained with the two programme excerpts.

Relative azimuthal acuity is thus very accurate, and this is clearly an important factor in 'surround-sound' reproduction. The listener may not be aware of true positional errors of various sound sources, but is likely to be much more critical of their relative positions.

It is also of interest to note that during the experiment a number of subjects experienced front/back ambiguity on several occasions. When both loudspeakers were approximately at C_B the subject sometimes perceived one or both to be in mirror-image locations near C_F, and even when informed of their error sometimes had great difficulty in perceiving the true locations. It would therefore appear that there is auditory ambiguity on the front/back centre-line through the head, which would normally be resolved by head movement[6] or visual cues.

5. Azimuthal image perception

Quadraphonic systems rely on an extension of the well-known stereophonic image principle,[7] which provides

Fig. 8 - Free-field arrangement for relative sound localisation experiment

a sound image located (normally) between two loudspeakers placed in front of the listener, depending on the inter-channel level-difference.[7,8,9] By use of four loudspeakers placed symmetrically around the listener an extension of this principle might be expected to provide complete azimuthal coverage, although this requires that the angle subtended at the listener between adjacent loudspeakers be 90°, rather than the 60° more usually preferred in stereophony.

Accordingly an experiment was devised to determine the 'interchannel level-difference law' of image localisation for adjacent loudspeakers placed in a quadraphonic array. The generally preferred 'square array' (see Terminology) was used although qualitative comments on the 'diamond array' are noted below. A similar arrangement to that shown in Fig. 8 was employed, with four matched high-quality loudspeakers placed on a 2·7 m radius circle at positions L_F, R_F, R_B and L_B. A small locating loudspeaker was used to provide the moveable reference sound-source and was arranged so that it did not significantly disturb the sound field generated by the loudspeakers in the quadraphonic array. Absolute loudspeaker levels were adjusted, initially, to be equal at the observer's head location, using a sound-level meter, and then the relative levels of an adjacent pair was adjusted, maintaining the total power delivered to the two loudspeakers constant (cf. Appendix), until the subject judged that the image created was azimuthally coincident with the reference source. Noise masking was again used between tests whilst the reference loudspeaker was moved, and the subject had a single C_F reference marker to look at.

Initial tests rapidly showed that the free field room was an unsuitable environment in which to form subjective sound images from such pairs of loudspeakers, particularly at the sides of the head, because of the lack of a reverberant sound field, and so a wooden floor was constructed to provide a single, uniform, reflecting surface. This improved the localisation of images considerably, although loudspeaker location with respect to the listener was found to be very critical. A 2% misplacement in radial distance of one

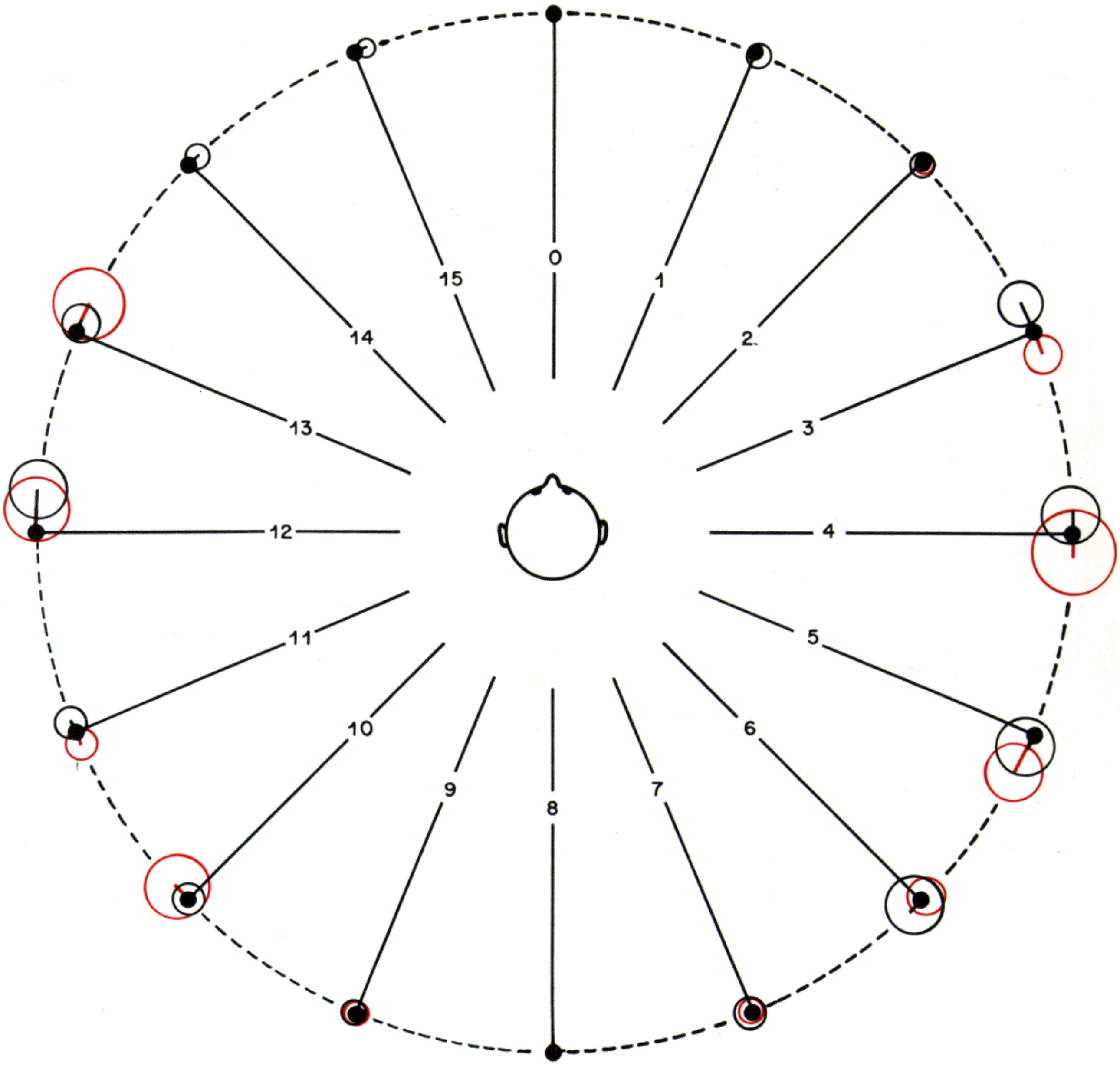
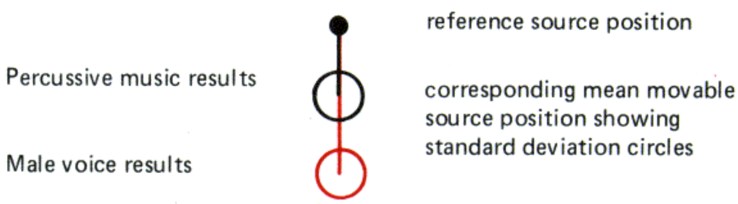
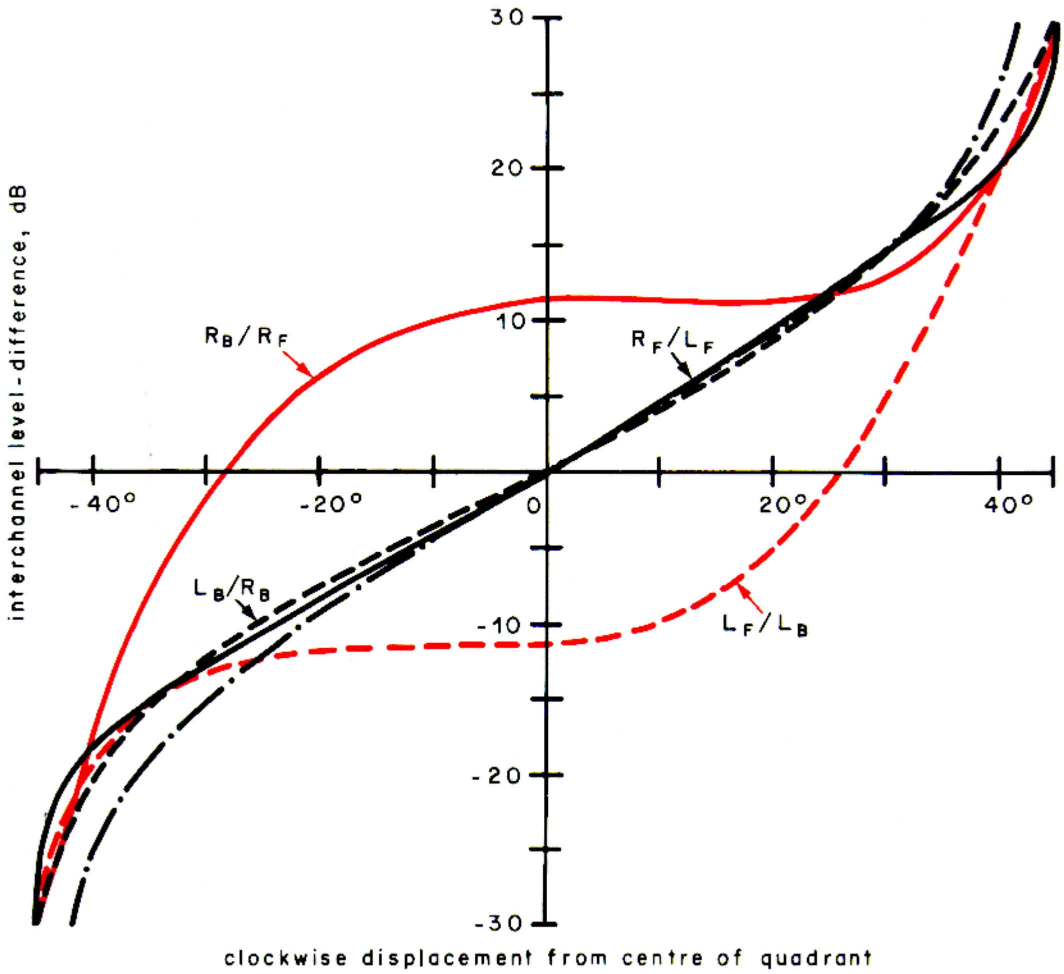


Fig. 9 - Relative sound localisation in the free-field room

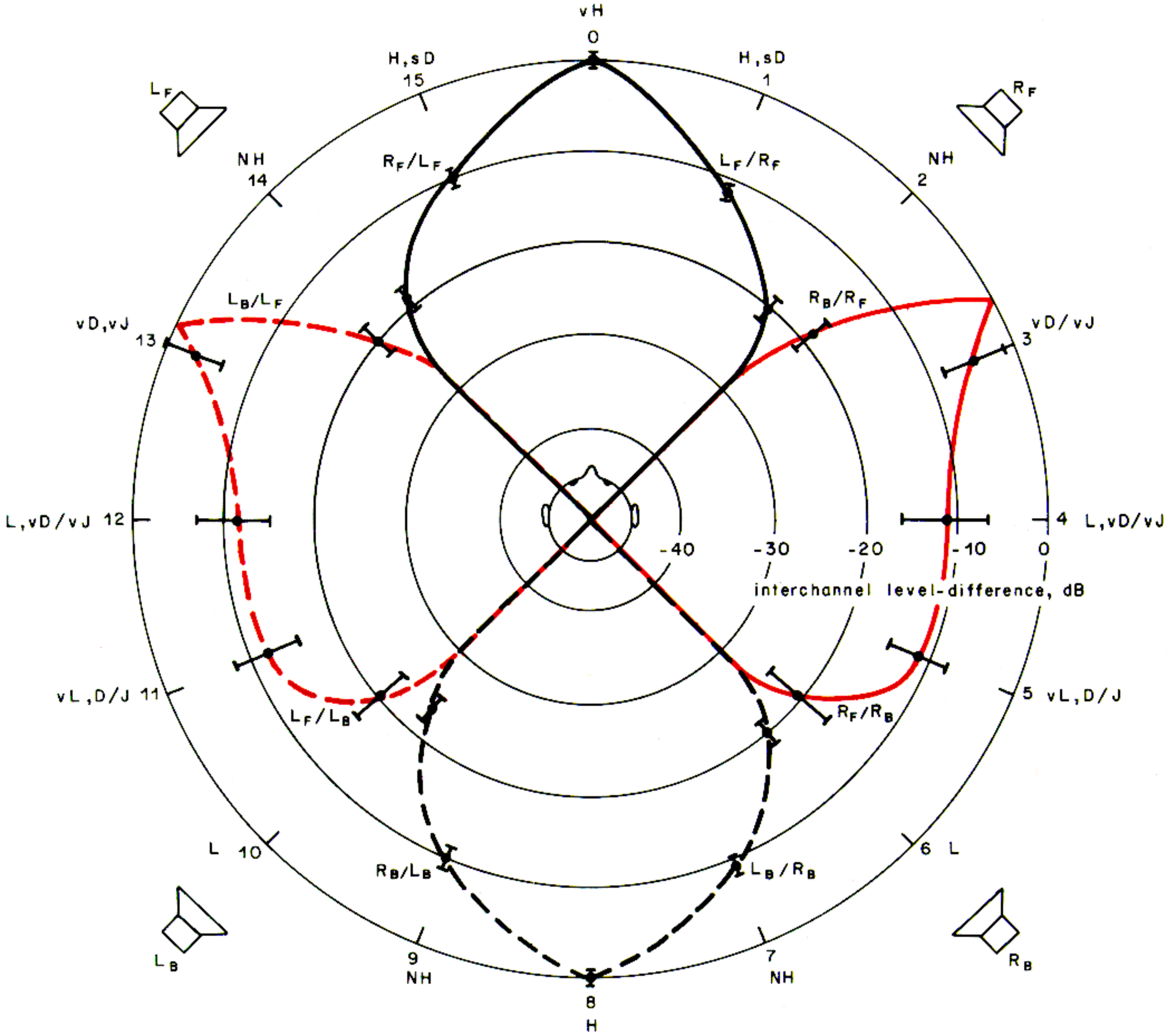




(*a*) Cartesian plot

Fig. 10 - Interchannel level-difference versus image location for adjacent pairs of loudspeakers in a quadraphone ('square') array, in the free-field room with a reflecting floor

— · —	Stereophonic law of sines[8]	D	diffuse
———	Front pair	H	high
– – – –	Back pair	J	jumpy
——— (red)	Right-hand pair	L	low
– – – – (red)	Left-hand pair	NH	normal height
I	Experimental result showing standard deviation	s	slightly
		v	very



(*b*) Polar plot

Fig. 10 - Interchannel level-difference versus image location for adjacent pairs of loudspeakers in a quadraphonic ('square') array, in the free-field room with a reflecting floor

Stereophonic law of sines[8]
Front pair
Back pair
Right-hand pair
Left-hand pair
Experimental result showing standard deviation

D diffuse
H high
J jumpy
L low
NH normal height
s slightly
v very

of the front loudspeakers (L_F and R_F) displaced a normal C_F image such that a 3 dB interchannel level-difference was required to correct it, whereupon the image exhibited a marked 'phasey' quality.

Results, the means obtained using seven subjects, are shown in Fig. 10 in both polar and rectilinear form. However, caution should be exercised in their interpretation, since large s.d.s were obtained for some locations, along, with adverse subjective comments (see polar plot). The front and rear quadrants are well defined and behave in the expected manner (cf. stereophony results[7,8,9]). However, the side quadrants exhibit a great degree of uncertainty, and subjects complained of either very diffuse or jumping images, with and without small head movements. It would appear that the 'stereophonic-image' phenomenon breaks down at the side of the head when predominantly one ear is excited, and the subject tends to hear each loudspeaker independently. Also there is preferential reception of the front loudspeaker, and about 10 dB more signal is required in the rear loudspeaker to give any impression of a centre-side image. There is a distinct threshold interchannel level-difference in this region, about which small variations cause the image to jump towards the front or back. However, the actual relative level at which this occurs varies greatly from subject to subject and from one occasion to another, as indicated by the large s.d.s for side image locations.

However, the effect is not so noticeable in the listening room, presumably because reflections cause directional information to be presented to both ears. It appears that the law determining image position is then largely dependent upon the phsyical properties of the room, and thus can be infinitely variable. For this reason, and lack of a totally symmetrical listening room, all the experiments on two-loudspeaker image-localisation were conducted in the free-field room with a single reflecting surface (i.e. a floor). Although such an environment does not give the same subjective impression as that of a listening room it does provide a far more critical and repeatable environment in which to determine the characteristics of various quadraphonic simulations; further, it gives good indications of possible short-comings, which may be subjectively disturbing under typical listening conditions.

Another feature of the image created by an adjacent pair of loudspeakers is the variation in image height around the compass. At C_F it is elevated by about 40° ('very high' in Fig. 10) and drops to eye-level height ('normal') at the front loudspeakers. Around the sides the image drops further becoming depressed by about 30° ('very low') at positions 5 and 11, such that the image appears to be at, or slightly below, floor level. Further towards the rear the image rises slightly, still being depressed by about 15° ('low') at the rear loudspeaker locations, and only rises slightly above eye-level (about 10° elevation) at C_B. The elevated front images are a seriously noticeable defect in quadraphony, although not serious in stereophony, and appear to be due to the increased angle subtended at the listener by the front loudspeakers. A theoretical hypothesis has been put forward to explain this effect,[8] but it does not agree well in magnitude, nor is there any reason to suggest an image rise as opposed to a fall.

Since side image localisation is poor with two-loudspeaker image synthesis it was considered possible that the 'diamond array' might prove more satisfactory. This configuration was briefly tested, and although image localisation was generally more similar in each quadrant, severe front/back ambiguity occurred as mirror-imaging about the C_L/C_R line. Accordingly this array was not considered further.

6. Effect of unwanted signals

So far only two-loudspeaker excitation has been considered in forming an image from a quadraphonic array, but in many quadraphonic systems three or even all four loudspeakers may be excited for a single point-source. In recording, for example, the use of four coincident cardioid microphones introduces unwanted* components (crosstalk) into a discrete quadraphonic reproduction, such that not only are the two adjacent channels to the source position energised, but also the two opposite channels carry components some 10—15 dB down.

Further experiments were conducted to give indications of the effects produced by exciting more than two loudspeakers in the array, and subjects were asked to determine the minimum perceptible crosstalk level for a number of selected situations, and to comment on localisation and quality changes brought about by excess crosstalk. Test image positions were chosen at either a loudspeaker (corner) location or mid-way between these (a centre-quadrant position), such that the wanted signal was applied either solely to one loudspeaker or equally to an adjacent pair. However, left/right symmetry was assumed and not all permutations were examined. Crosstalk signals were introduced into the diagonally opposite or adjacent pair of loudspeakers and the seven arrangements tested are shown in Fig. 11. The subject performed under the same test conditions as in the previous experiment, but was provided with a switch to add in the crosstalk, and an attenuator with which to vary its level. Having determined the minimum perceptible crosstalk level the subject was then asked to increase its level until it became equal to that of the wanted signals, describing the locus of the sound image as the increase was made. These results, the averages obtained using seven subjects, are shown in Fig. 11, and it is notable that in general about −20 dB of crosstalk is detectable, although it is considerably less in the C_F/C_B directions. As the crosstalk level is increased the images move in closer towards the subject, becoming bass heavy, and ending up rather unpleasantly within, or just above the subject's head.

* The signals from an adjacent pair of loudspeakers forming an image localised according to the 'interchannel intensity-difference' law (determined in the previous section) are termed the 'wanted' components and any radiated by the other loudspeakers are termed 'unwanted' or 'crosstalk' components. However, more generally, images formed by excitation of more than two loudspeakers are not necessarily undesirable, for instance, in the production of ranging effects.

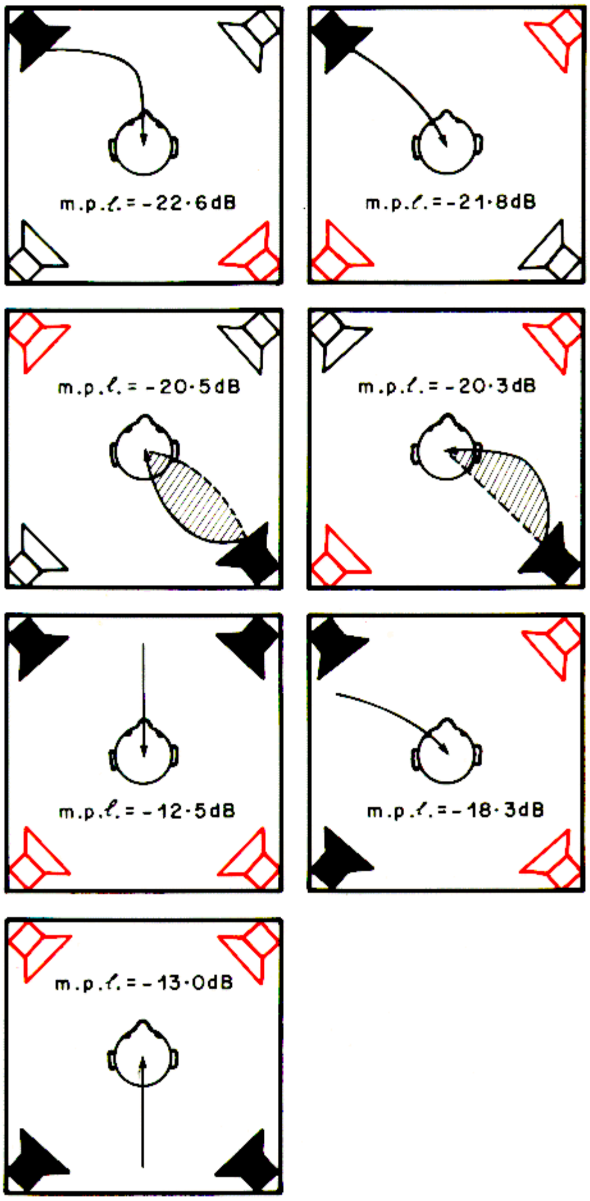
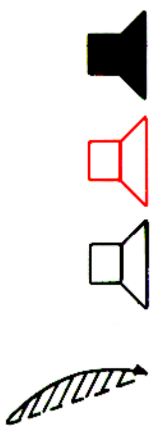


Fig. 11 - Minimum perceptible crosstalk level and image locus with increasing crosstalk for selected quadraphonic arrangements

m.p.*l.* is the minimum perceptible level of unwanted signal



wanted signal at reference level (0 dB)

unwanted (crosstalk) signal

no signal

locus of image position as crosstalk signal is increased from m.p.*l.* to 0 dB also showing area of uncertainty

7. Effect of phase differences

7.1. Occurrence

A number of quadraphonic systems employ only two transmission channels and matrix the four primary signals in differing amplitudes and phases into two composite signals. Decoding on reception produces crosstalk components with various amplitude and phase relationships relative to the wanted components, which themselves may exhibit phase differences. This section deals with a number of experiments devised to give some indication of the subjective effects experienced when such phase-shifted signals are presented.

7.2. Phase-shift between the wanted signals

In this experiment the minimum perceptible phase difference between the signals feeding an adjacent pair of loudspeakers was determined for selected image positions. The subject was tested under the same conditions as in the earlier image-locating tests, and was given a switch to compare the image with and without the inserted phase difference. The latter was reduced in 22½° steps (plus a final step to 11¼°) until the subject judged it only just perceptible. Seven image positions were tested, nominally at the centre of each quadrant and ±5° from each loudspeaker location; however, not all possibilities were tested since left/right symmetry was assumed. Fig. 12 shows results averaged from six subjects indicating the minimum perceptible phase difference for the selected image positions, and the corresponding azimuthal image-shifts which occurred. Centre-quadrant locations are most sensitive to phase differences, which is to be expected, but the side quadrant is considerably less sensitive than either front or back quadrants. Image shift follows the Haas precedence effect[10] at C_F and C_B, but when one loudspeaker is dominant the image always shifts towards it as the phase difference increases. Also at C_R (nominal position, equal signals to R_F and R_B loudspeakers) the image always moves forwards towards the front loudspeaker, presumably because the decorrelating effect of introducing phase-shift further enhances the forward source. General comments on the effects of excess phase-shift are that the images tend to move across the stage in a large arc, apparently moving further away from the observer as the phase difference increases, and becoming 'nasal' or bass lacking in quality. However, for images in a centre-quadrant location further increase in phase shift (>90°) causes the image to become diffuse, and finally results in the familiar unpleasant 'in the head' or 'phasey' sensations usually associated with stereophonic systems in which one loudspeaker has been phase-reversed.

7.3. Phase shift in the unwanted signals

The number of possible arrangements which could have been investigated is almost infinite, and so a very restricted set of tests was performed, based on the kinds of crosstalk components commonly introduced by existing matrix quadraphonic systems, and the tests previously reported in Section 6. The wanted signals were maintained in-phase and the unwanted signals were varied in 90° steps

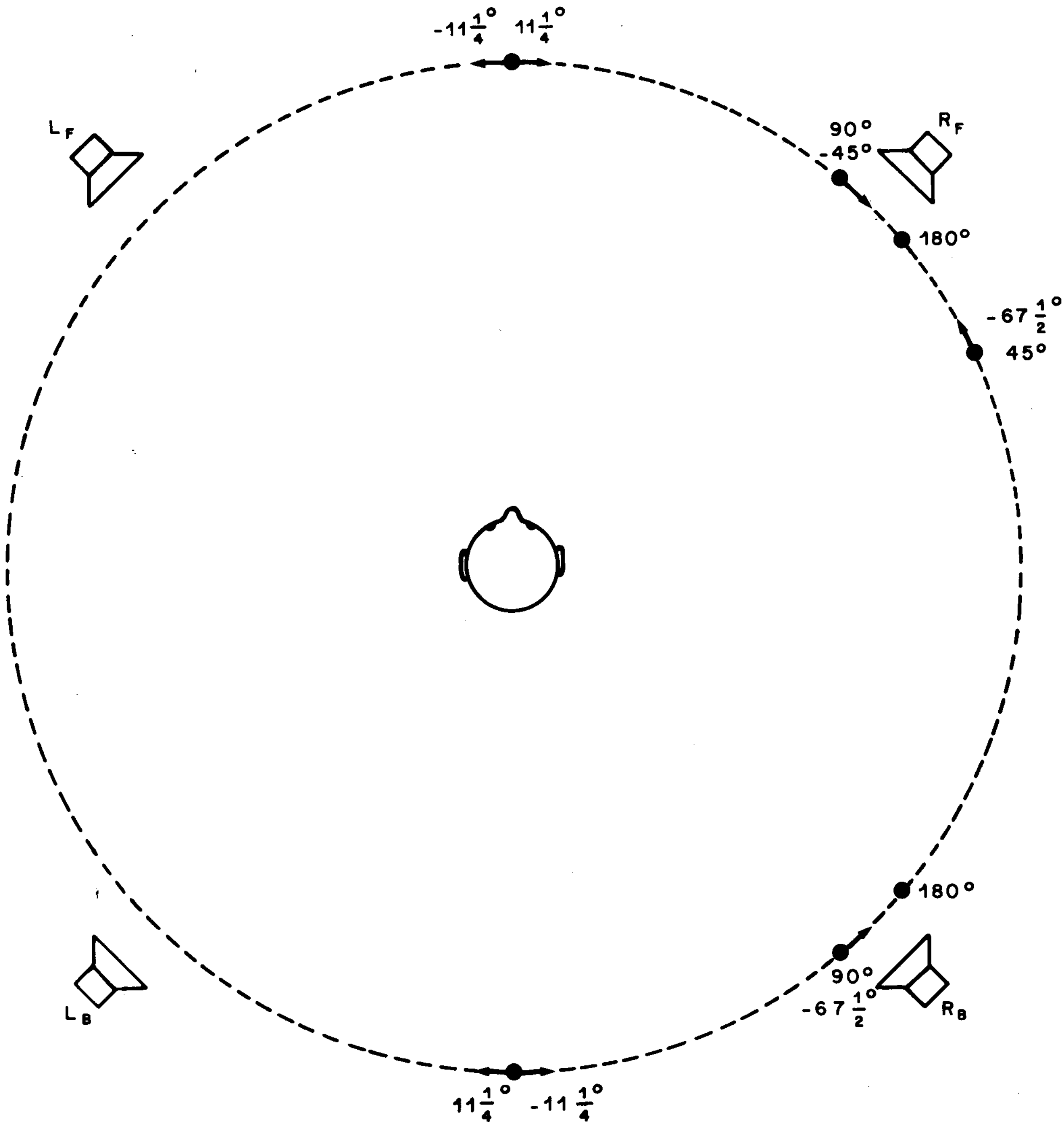


Fig. 12 - Minimum perceptible phase-difference between an adjacent pair of loudspeaker signals in a quadraphonic ('square') array showing image movement for selected positions

image location for in-phase signals

image location for minimum perceptible phase-shift

The sign of the phase-shift inserted is defined positive for the phase-leading loudspeaker clockwise of the nominal image position

(plus a minimum step of 45°), at a fixed level 10 dB below the wanted signals. The same seven loudspeaker configurations of Fig. 11 were used, and also an asymmetric crosstalk condition typical of the 'SQ' type of matrix was tested.

The subject was given a 3-way comparison of the wanted-signal image, and the composite-signal* image both with and without phase-shift applied to the crosstalk signals. This enabled him to identify readily the effects of the phase-shifts inserted. Results, the averages obtained with seven subjects, are presented in Fig. 13, and merit some explanation. Fig. 13(*a*) shows the effects of phase shift when the crosstalk signals are applied diagonally opposite the wanted signals, and Fig. 13(*b*) shows the results for corner images when two crosstalk signals occur in combinations similar to the 'QS' and 'SQ' types of matrix. Table I shows which loudspeakers were

* Wanted plus crosstalk signals.

Test Condition	Nominal Image Position	Crosstalk type	Relative loudspeaker levels (dB)			
			L_F	R_F	R_B	L_B
A	L_F	Symmetrical	0	–	–10	–
B	C_F	Symmetrical	0	0	–10	–10
C	C_R	Symmetrical	–10	0	0	–10
D	R_B	Symmetrical	–10	–	0	–
E	C_B	Symmetrical	–10	–10	0	0
F	L_F	'QS'	0	–10	–	–10
G	R_F	'SQ'	–	0	–10	–10
H	R_B	'QS'	–	–10	0	–10
I	L_B	'SQ'	–10	–10	–	0

Table I Test Conditions Referenced in Fig. 13

energised for each test condition (A to I) illustrated in Fig. 13(*a*) and (*b*). Crosstalk phase-shift was inserted both leading and lagging the wanted signals, and image position and quality comments are shown on concentric circles corresponding to a particular phase-shift. In cases where two crosstalk signals were present, phase-shift could be applied equally to both signals such that they were always in-phase, or alternatively with opposite sign such that one signal led, and one signal lagged, the wanted component. When both crosstalk signals were in-phase, or when only one crosstalk signal was present, the sign of the phase-shift was not subjectively detectable, and is therefore not indicated in the figure. However, when the crosstalk signals were phase-shifted in opposite directions from the wanted signals, a sign is appended to the phase-shift indicated in the figure; this is considered to be positive when the crosstalk signal located in the clockwise direction relative to the image position is phase-leading.

Generally, in-phase crosstalk produces images closer to the observer and bass heavy in quality (as was experienced in the earlier crosstalk tests), whereas anti-phase crosstalk produces phasey or nasal, bass lacking images. In the side quadrants in-phase crosstalk produces image shifts towards the front/back centre-line, and anti-phase crosstalk towards the left/right centre-line. With two crosstalks present simultaneously, in-phase crosstalk signals produce no image shift, but the image generally sounds phasey and diffuse. However, if one crosstalk signal leads and the other lags the wanted signals, less objectionable image qualities are observed, although some azimuth movement may be observed. Closer inspection of the results shows that, in general, crosstalk phase shifts of 45°, or +45° and –45° for two signals, produces the least disturbing subjective effect, except at centre-side where +90° and –90° is preferred.

Referring to Fig. 13(*c*) observation of the 'QS type' corner crosstalk (test conditions F and H) shows distinctly asymmetric results, although the exact locus of image movements is not well defined owing to the relatively small number of tests conducted. However, +45° and –45° again appear to give a satisfactory image quality with little image shift. With 'SQ type' crosstalk (test conditions G and I) the signal opposite the wanted signal was always in-phase or in anti-phase, and the adjacent signal phase is either + or –90° as indicated by the '0°/+90°' and '180°/–90°' nomenclature in the figure. Large image shifts or poor image quality is experienced at the front corners, the shift being dependent on the sign of the adjacent signal (Haas effect applied). At the rear corners the image shift is very small and quality not so impaired. However, in both cases anti-phase diagonal crosstalk is preferable to the in-phase form.

8. Comment on the results

These experiments have explored some of the fundamental properties of hearing which have particular bearing upon the engineering of 'surround sound' reproduction systems. Also, the effects of four loudspeaker or quadraphonic presentation of sounds have been investigated, and some insight into such limited point-source simulations of sound-fields has been gained.

Fundamental properties of hearing determined are:

(a) the human auditory system is approximately of equal sensitivity to isolated sounds from all azimuths (see Fig. 2);

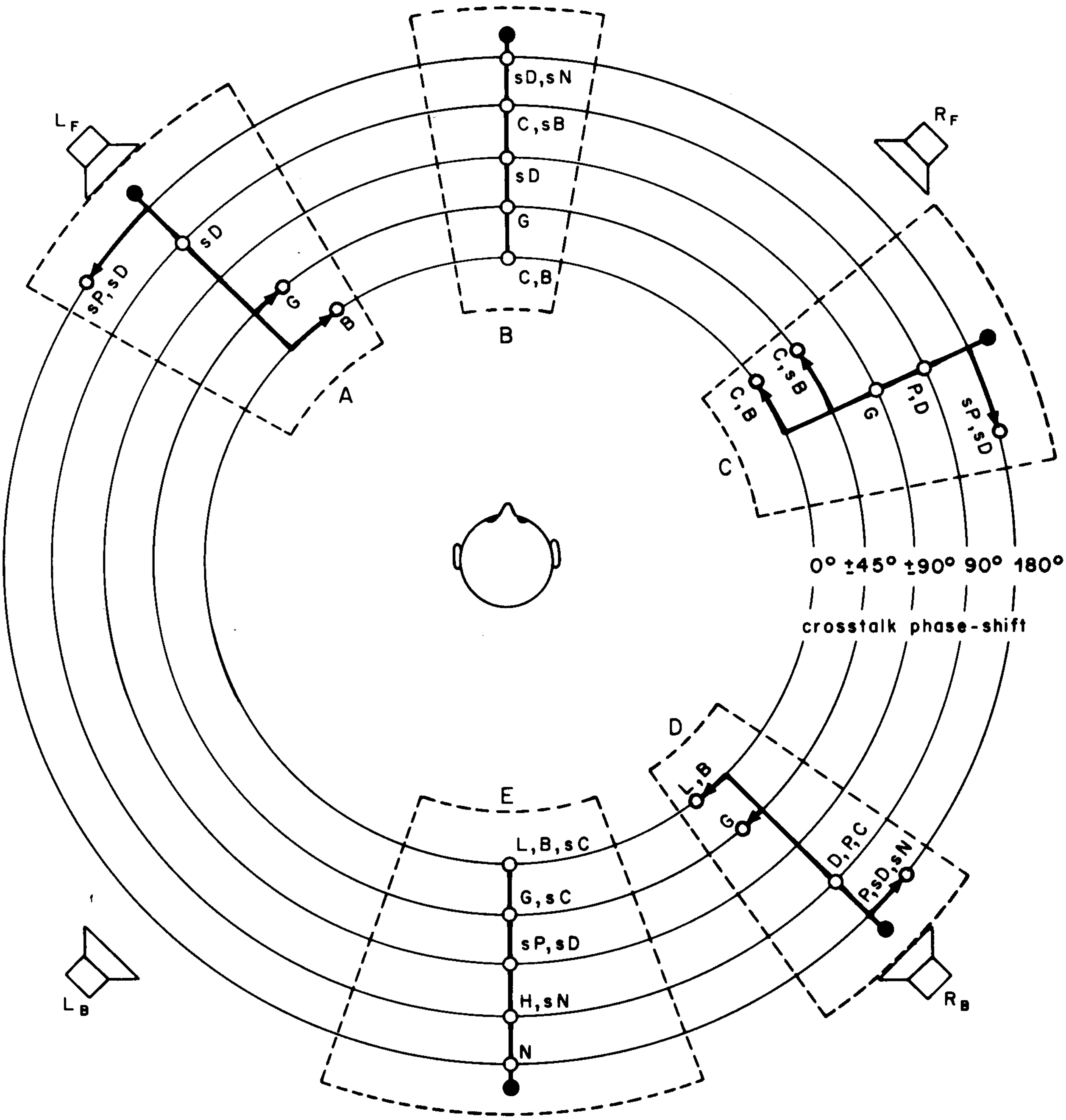


Fig. 13(a) - Effect of phase-shifted crosstalk signals on image location and image quality for selected quadraphonic arrangements:
Crosstalk image diagonally opposite wanted image position (test conditions A to E shown in Table I)

image location without crosstalk

image locations with crosstalk showing shift (the magnitudes of the crosstalk phase-shifts are indicated by the concentric circles)

B	bass heavy	H	high) vertical position
C	close, near	L	low) vertical position
D	diffuse	N	nasal, bass lacking
F	far, distant	P	phasey
G	good, no adverse comments	s	slightly

The phase-shift inserted into the crosstalk (unwanted) signals is defined positive for the phase-leading crosstalk clockwise of the nominal image position, and the equally phase-lagging signal anticlockwise. No sign indicates that all crosstalk signals are in-phase, but either lead or lag the wanted signals by the specified amount.

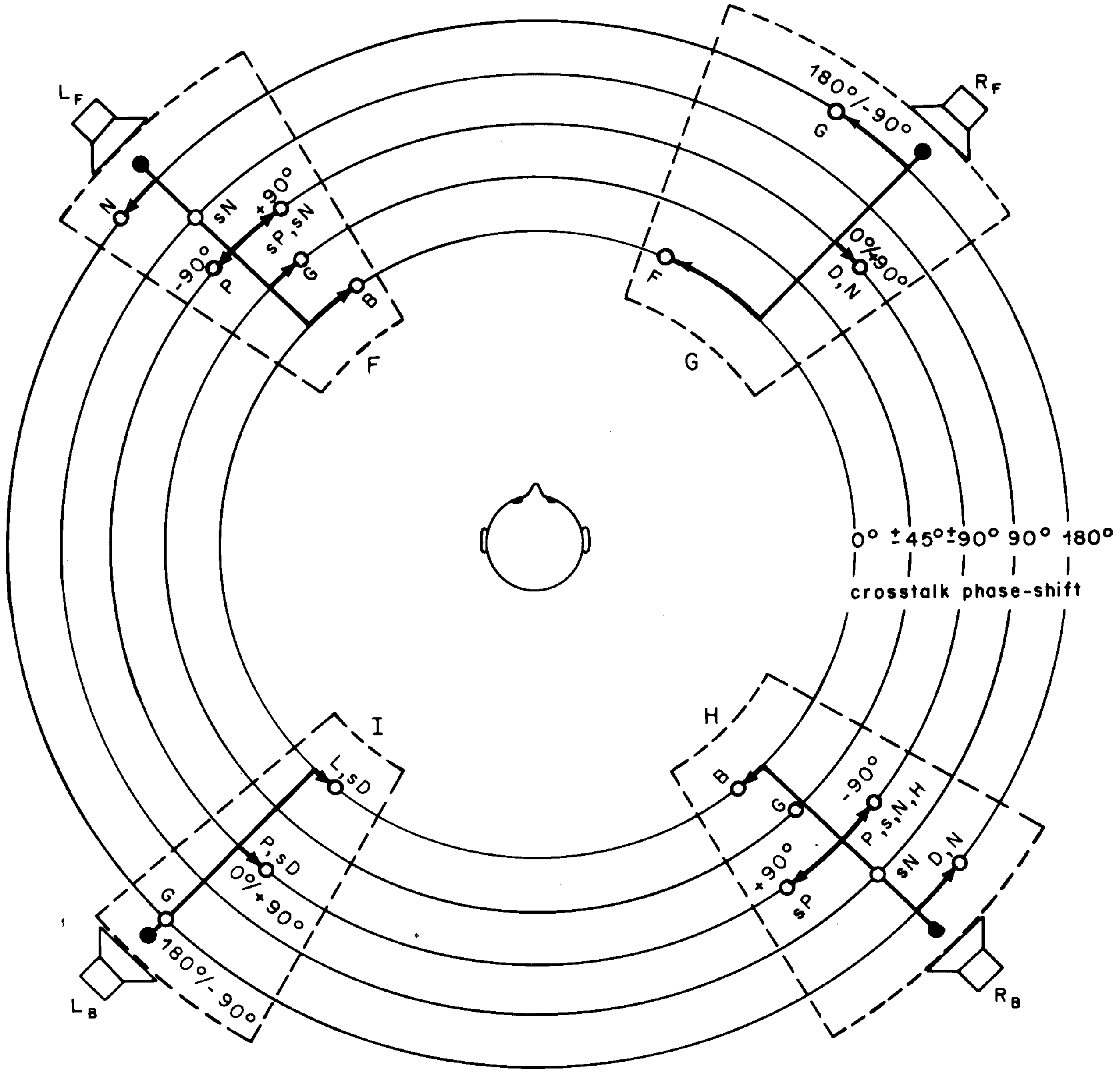


Fig. 13(b) - Effect of phase-shifted crosstalk signals on image location and image quality for selected quadraphonic arrangements:
Crosstalk images of the 'SQ' and 'QS' types at corner positions (test conditions F to I shown in Table I)

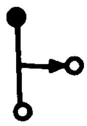


image location without crosstalk

image locations with crosstalk showing shift (the magnitudes of the crosstalk phase-shifts are indicated by the concentric circles)

B	bass heavy	H	high) vertical position
C	close, near	L	low)
D	diffuse	N	nasal, bass lacking
F	far, distant	P	phasey
G	good, no adverse comments	s	slightly

The phase-shift inserted into the crosstalk (unwanted) signals is defined positive for the phase-leading crosstalk clockwise of the nominal image position, and the equally phase-lagging signal anticlockwise. No sign indicates that all crosstalk signals are in-phase, but either lead or lag the wanted signal by the specified amount.

(b) isolated sound-source localisation in the horizontal plane is accurate to about 5° in the front semi-circle and at centre-back (C_B), but greater uncertainty ($\doteq$ 10°) exists, and considerable rear-image expansion occurs, in the rear semi-circle away from C_B (see Figs. 5 and 6);

(c) relative sound-source localisation in the horizontal plane is accurate to within 2° in front and rear quadrants, becoming more uncertain towards the centres of the side quadrants ($\doteq$ 5°) where audition is largely monaural (see Fig. 9).

The extent to which a quadraphonic 'square' array can reproduce a desired sound stage was investigated on the basis of an extension of stereophonic principles, and an interchannel level-difference law of image position for adjacent pairs of sources has been determined around the complete compass (see Fig. 10). However, the ability of the listener to realise an image so formed at the sides of the head was found to be questionable under non-reverberant conditions, and considerable front-source dominance* and independent reception of the front and rear sources was then evident. It was also found that room acoustics play an important part in specifying the interchannel level-difference law in these regions.

Fig. 10 shows the fidelity of back image localisation according to normal stereophonic principles,[8] which is not in agreement with the 'back image contraction' principle[11] claimed by workers elsewhere. However, the effect cited in favour of the latter principle may be explained by the apparent expansion of the real stage when the observer turns his back on it (see Figs. 5 and 6).

Perceptibility of unwanted (crosstalk) signals (Fig. 11) is high (at a relative level of about −20 dB) when one source is dominant (i.e. near corner locations), but is reduced (to a relative level of about −12 dB) for images in the centre-front and centre-back areas. Excess crosstalk produces images close to the subject and bass heavy in quality.

This conflicts with the 'front source dominance principle'[11] and indicates that the strongest source always predominantly defines the location of the image. Only with adjacent channel crosstalk will the image tend to lie to the forward side of the dominant loudspeaker. A diagonal crosstalk signal (coming from the loudspeaker opposite the dominant one) however, will always cause an image shift towards the front/back centre-line.

Investigation of the permissible phase-shift between adjacent pairs of loudspeakers (cf. results of Ref. 12) has shown greatest sensitivity at centre-front and centre-back ($\doteq$ 11°), with image shifts occurring according to the Haas precedence effect (see also 'quadrature image-shift principle'[11]). Elsewhere, however, the image moves towards the dominant source or, in cases where a side image is produced, front-source dominance already experienced with no phase-shift is further enhanced.

If crosstalk components are introduced, phase-shifted with respect to the wanted signals, the resulting image may exhibit a wide range of subjective effects, and the results of the restricted set of tests conducted show that certain amounts of phase shift, applied to the crosstalk signals in specified senses, can improve the image quality and cause minimal shift from the desired image position. A good compromise between the bass heaviness produced by in-phase crosstalk signals and the 'phaseyness' produced by those in anti-phase with the wanted signals can be obtained, generally by applying about 45° phase-shift to a single crosstalk signal, or about +45° and −45° phase-shift to a pair. However, the combination of +90° and −90° is preferred for centre-side images, whereupon the listener's ability to realise the image is greatly enhanced.

It is thought preferable to arrange the sense of paired phase-shifted crosstalk components such that the Haas effect aids image localisation towards the left/right centre-line. This will tend to counter natural tendencies to localise the image towards the front/back centre-line, although the effect is often not significant.

9. Conclusions

Although of limited nature, these results illuminate the fallacy of generalisation from a few observed phenomena,[11] and confirm that the mechanism of audition is indeed extremely complex. Accordingly, attempts to deceive the normal auditory processes (i.e. the re-creation of a total sound stage by multiple (4) point-source simulation) should be extensively studied and understood before the optimum engineering solution to the problem can be found.

Many workers have been searching for effective methods of reducing four or more studio signals into a smaller number of transmission channels (typically two) by simple linear processing techniques (matrixing), such that the received signals may be further processed to reproduce the original omnidirectional sound stage. Such signal processing typically provides quadraphonic signals of the type investigated in this report, and the results presented provide pointers for the design of more effective systems, and have been used to pin-point the causes of the undesirable features of some systems presently in existence.[13] In this way two-channel matrix systems have already been evolved which provide, in some ways, subjectively more satisfying results than current commercial matrix systems. Work continues in this field.

Also of importance is the subjective susceptibility to the geometry of the loudspeaker array. A 2% misplacement of one loudspeaker caused a considerable image shift from a nominally centre-front location, and it was also found that image localisation was very sensitive to head position during many of the tests. For 'surround-sound' systems to be of much value in a domestic listening environment, reasonable tolerances on the geometry of the loudspeaker

* Not to be confused with the 'front source dominance principle' of Reference 11, which states that the 'human hearing mechanism will judge the direction of sound arrival based upon the signals proceeding from the front loudspeakers' providing that one of these signals is considerably greater than any of the others.

array and the permissible listening area must be allowed, and it is therefore considered necessary to study these aspects further.

10. References

1. GARDNER, M.B. 1973. Some single and multiple source localisation effects. *J. Audio Engng Soc.*, 1973, **21**, 6, pp. 430 – 437.

2. FRANSSEN, N.V. 1964. Stereophony. Eindhoven, Philips Tech. Lib., 1964.

3. TOBIAS, J.V. 1971. Foundations of modern auditory theory. Vol. II, New York, Academic Press, 1972.

4. DE BOER, K. 1940. Stereophonic sound reproduction. *Philips Tech. Rev.*, 1940, **5**, 4, pp. 107 – 114.

5. SHAXBY, J.H. *and* GAGE, F.H. 1932. The localisation of sounds in the median plane. Med. Red. Council Spec. Rept. Series No. 166, London, HMSO, 1932.

6. WALLACH, H. 1940. The role of head movements and vestibular and visual cues in sound localisation. *J. Exp. Psych.*, 1940, **27**, 4, pp. 339 – 368.

7. CLARK, H.A.M., DUTTON, G.F. *and* VANDERLYN, P.B. 1957. The 'Stereosonic' recording and reproducing system. *Proc. Instn elect. Engrs.*, 1957, **104B**, 17, pp. 417 – 432.

8. BAUER, B.B. 1961. Phasor analysis of some stereophonic phenomena. *J. Acoust. Soc. Am.*, 1961, **33**, 11, pp. 1536 – 1539.

9. HARWOOD, H.D. 1968. Stereophonic image sharpness. *Wireless Wld.*, 1968, **1393**, 74, pp. 207 – 211.

10. HAAS, H. 1951. The influence of a single echo on the audibility of speech. (trans. from Acustica, 1951, **1**, pp. 49 – 58), *J. Audio Engng Soc.*, 1972, **20**, 2, pp. 145 – 159.

11. BAUER, B.B., GRAVEREAUX, D.W. *and* GUST, A.J. 1971. A compatible stereo-quadraphonic (SQ) record system. *J. Audio Engng Soc.*, 1971, **19**, 8, pp. 638 – 646.

12. KOHSAKA, O., SATOH, E. *and* NAKAYAMA, T. 1972. Sound-image localisation in multi-channel matrix reproduction. *J. Audio Engng Soc.*, 1972, **20**, 7, pp. 542 – 547.

13. CROMPTON, T.W.J. The subjective performance of various quadraphonic matrix systems. BBC Research Department Report in course of preparation.

Appendix

Equal Loudness Levels of a Spaced Pair of Loudspeakers and a Single Loudspeaker

A preliminary investigation was conducted into the equal loudness levels for an image formed by a pair of spaced loudspeakers and a single loudspeaker placed at the image location. A pair of loudspeakers was set up in the listening room, 3 m apart, such that they subtended an angle of 90° at the listener. A similar single loudspeaker was placed on the bisector of this angle at the same radial distance (1·7 m) from the listener, and the latter faced this loudspeaker. The listener was asked, by making a switched comparison, to adjust the level of the signal fed to the loudspeaker-pair such that the perceived loudness of this image matched that of the single loudspeaker. White and pink noise test signals were used, and the signal level to the single loudspeaker was fixed to give a sound pressure level (s.p.l.) about 65 dBA at the listener's head position.

Results averaged from six subjects show an attenuation of 5·2 dB (s.d. = ±1·2 dB) on the signal feeding the loudspeaker-pair over that feeding the single loudspeaker when using white noise, and 4·5 dB (s.d. = ±1 dB) using pink noise. The measured s.p.l. difference at the listener's head was +5 dB for the loudspeaker-pair radiating at the same levels as the single loudspeaker using either white or pink noise.

For comparison the experiment was repeated with the loudspeaker-pair subtending a 60° angle at the listener, as in normal stereophony. The subjective level difference was reduced slightly to 4·1 dB (s.d. = 1·3 dB) using white noise, but was still 4·5 dB (s.d. = 1 dB) using pink noise. The measured s.p.l. difference was also still 5 dB using white or pink noise.

SMW/AMM

BBC RD 1977/2

1974
THE QUEEN'S AWARD TO INDUSTRY

RESEARCH DEPARTMENT REPORT

QUADRAPHONY: developments in Matrix H decoding

P.S. Gaskell, B.A.
P.A. Ratliff, B.Sc., Ph.D.

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION February 1977

BBC RD 1977/2
UDC 681.84

QUADRAPHONY: DEVELOPMENTS IN MATRIX H DECODING

P.S. Gaskell, B.A.
P.A. Ratliff, B.Sc., Ph.D.

Summary

The Matrix H 2-channel quadraphonic encoding system has been designed to transmit the maximum amount of directional information consistent with mono and stereo compatibility. This report discusses methods of decoding such transmission signals for 'surround-sound' reproduction.

The basic form of decoder is a complex-coefficient linear matrix, but this has an inherent lack of separation between the output signals. 'Logic enhancement' techniques are discussed, which seek to improve interchannel separations for the principal sound-source, at the expense of secondary sources.

As an interim measure, a modified commercial logic-enhanced ('Variomatrix') decoder was studied, which led to the development of a purpose-built logic-enhanced decoder for Matrix H. The latter combines the virtues of both linear Matrix H decoding, and the variable-matrix logic enhancement technique.

All the decoders described are capable of providing a good surround-sound reproduction. In particular a purpose-built Matrix H logic-enhanced decoder has been shown to exhibit a performance very close to a discrete 4-channel system; the failings normally associated with logic-enhanced decoders are almost inaudible.

Issued under the authority of

Head of Research Department

Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION

~~January 1977~~
February 1977

(PH-169)

QUADRAPHONY: DEVELOPMENTS IN MATRIX H DECODING

QUADRAPHONY: DEVELOPMENTS IN MATRIX H DECODING

P.S. Gaskell, B.A.
P.A. Ratliff, B.Sc., Ph.D.

Terminology

For brevity in this report, abbreviations for directions with respect to the listener are frequently used; these are defined in Diagram A. The loudspeaker positions are sometimes referred to as 'corner locations', and the span between any adjacent pair of loudspeakers is referred to as a 'quadrant'. The four quadrants are specifically defined as 'front', 'back', 'left' or 'right', as appropriate. Signals associated with a particular loudspeaker are similarly designated by the appropriate direction symbol; the origination signals are shown unprimed, whilst the decoded signals are shown primed. Matrix encoded signals corresponding to conventional stereo left and right signals are denoted by L_T and R_T respectively. A list of symbols is given below.

List of symbols

L_F, R_F etc	quad origination signals
L'_F, R'_F etc.	decoded quad signals
L_T, R_T	matrix encoded 2-channel (stereo) signals
f	front } logic control signals
b	back } logic control signals
l	left } logic control signals
r	right } logic control signals
α_{pq}	linear-decode matrix coefficients
$r\underline{/\theta}$	polar representation of vector, of modulus r and argument θ degrees.

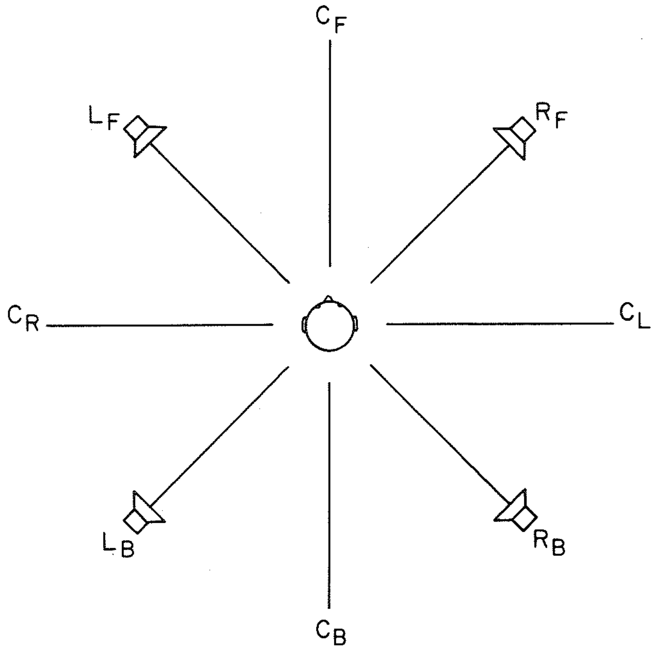


Diagram A - Quadraphonic loudspeaker array for reproduction of surround sound-stage, showing direction abbreviations

C_F centre-front, R_F right-front, C_R centre-right, R_B right-back, C_B centre-back, L_B left-back, C_L centre-left, L_F left-front

Three methods of Matrix H decoding are described in this report, a basic linear matrix decoder, a modified commercial 'logic-enhanced' decoder, and a purpose-built Matrix H 'logic-enhanced' decoder.

1. Introduction

The Matrix H 4-2-4 quadraphonic matrix system was designed primarily to overcome the mono and stereo compatability limitations of other proposed 4-2-4 matrix systems,[1] whilst retaining the ability to provide worthwhile quadraphony. In order that the quadraphonic reproduction should provide a significant subjective enhancement of the sound sensation together with a greater sense of realism and involvement for the listener, the decoding system employed must be effective in extracting the directional information contained in the Matrix H coded two-channel signals.

The decoding of Matrix H signals is not limited to one method and it is the purpose of this report to discuss some of the decoders that have been developed to date. It is highly probable, however, that developments will continue to be made in the field of decoding.

2. Matrix H linear decoding

2.1. Basic linear matrix

The fundamental method of decoding employs a basic linear matrix formed by taking the 'complex conjugate'* of the encode matrix thus:-

$$\begin{bmatrix} L'_F \\ R'_F \\ L'_B \\ R'_B \end{bmatrix} = \begin{bmatrix} 0{\cdot}940\ \underline{/-10^\circ}, & 0{\cdot}342\ \underline{/65^\circ} \\ 0{\cdot}342\ \underline{/-65^\circ}, & 0{\cdot}940\ \underline{/10^\circ} \\ 0{\cdot}940\ \underline{/25^\circ}, & 0{\cdot}342\underline{/-115^\circ} \\ 0{\cdot}342\ \underline{/115^\circ}, & 0{\cdot}940\ \underline{/-25^\circ} \end{bmatrix} \begin{bmatrix} L_T \\ R_T \end{bmatrix}$$

* so called because the column elements in the decode matrix are the complex conjugates of the row elements in the encode matrix.

This results in an overall transfer function for the Matrix H system, expressed in polar co-ordinates, of:-

$$\begin{bmatrix} L'_F \\ R'_F \\ L'_B \\ R'_B \end{bmatrix} = \begin{bmatrix} 1{\cdot}000\,\underline{/0^\circ}, & 0{\cdot}644\,\underline{/55^\circ}, & 0{\cdot}791\,\underline{/-40^\circ}, & 0{\cdot}193\,\underline{/163^\circ} \\ 0{\cdot}644\,\underline{/-55^\circ}, & 1{\cdot}000\,\underline{/0^\circ}, & 0{\cdot}193\,\underline{/-163^\circ}, & 0{\cdot}791\,\underline{/40^\circ} \\ 0{\cdot}791\,\underline{/40^\circ}, & 0{\cdot}193\,\underline{/163^\circ}, & 1{\cdot}000\,\underline{/0^\circ}, & 0{\cdot}644\,\underline{/-90^\circ} \\ 0{\cdot}193\,\underline{/-163^\circ}, & 0{\cdot}791\,\underline{/-40^\circ}, & 0{\cdot}644\,\underline{/90^\circ}, & 1{\cdot}000\,\underline{/0^\circ} \end{bmatrix} \begin{bmatrix} L_F \\ R_F \\ L_B \\ R_B \end{bmatrix}$$

The response of this system to the eight cardinal stage locations is illustrated in Fig. 1. Each small square describes the decoded output signal-relationships corresponding to the appropriate cardinal position; thus the top-centre square refers to the decoded outputs for a C_F (centre-front) encoded input signal. The numbers in the corners of each small square represent the relative phase-angles (in degrees) of the four output signals, arranged to correspond to the loudspeaker array around the listener (see centre-square). The numbers associated with the arrows indicate the 'separations' (in dB) obtained between the 'wanted' signal output and the 'unwanted' or crosstalk signals from the other outputs. Taking the top left-hand square as an example, this shows the decoded outputs obtained for an input signal encoded at the L_F position. The separation between L'_F and R'_F is 3·8 dB, between L'_F and L'_B is 2·1 dB, and between L'_F and R'_B is 14·3 dB. The relative phases of the crosstalk signals, compared with the 'wanted' L'_F output, are -55° for R'_F, $+40^\circ$ for L'_B, and -163° for R'_B.

It will be seen that low separation figures are obtained between adjacent outputs, and this characteristic is typical of two-channel linear matrix systems, since only two outputs can be completely isolated. However, in the case of Matrix H decoding, greater separation exists in the left/right direction than in the front/back direction. This is intentional since, for a forward-facing listener, it gives a greater accuracy of localisation around the sound-stage than a symmetrical distribution.

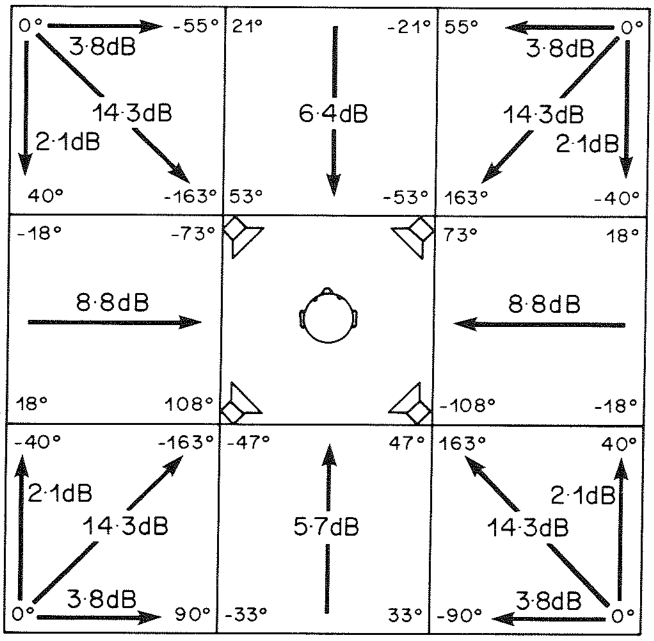


Fig. 1 - Performance chart for basic Matrix H linear decoder (see Section 2)

2.2. Phase-modified linear matrix

Earlier work on the properties of hearing relevant to quadraphonic reproduction[2] showed that the phase relationships between the output signals are also important. The output for a C_F encoded signal in Fig. 1 shows a 42° phase-difference between the 'wanted' signals, which, together with low separation to the phase-shifted rear-channel output signals, causes some diffusion of the C_F image. By phase-shifting the front-channel output signals, this effect can be substantially reduced without significantly degrading the output for other positions; centre-side positions are tolerant of increased phase-difference. In addition, image localisation at the front corners is improved.

The phase-modified linear matrix may be conveniently expressed in polar co-ordinate form:-

$$\begin{bmatrix} L'_F \\ R'_F \\ L'_B \\ R'_B \end{bmatrix} = \begin{bmatrix} 0{\cdot}940\,\underline{/-20^\circ}, & 0{\cdot}342\,\underline{/55^\circ} \\ 0{\cdot}342\,\underline{/-55^\circ}, & 0{\cdot}940\,\underline{/20^\circ} \\ 0{\cdot}940\,\underline{/25^\circ}, & 0{\cdot}342\,\underline{/-115^\circ} \\ 0{\cdot}342\,\underline{/115^\circ}, & 0{\cdot}940\,\underline{/-25^\circ} \end{bmatrix} \begin{bmatrix} L_T \\ R_T \end{bmatrix}$$

This has a performance shown in Fig. 2. It is seen that the separation figures remain unchanged, but since these are small, the comparatively small changes in the relative phase-angles of the crosstalk signals noticeably improve the subjective performance of the decoder.

The introduction of phase differences between the 'wanted' and the crosstalk signals has the effect of decorrelating the signals and of subjectively increasing their apparent separation. Thus the greatest phase-differences are arranged to occur for crosstalk signals opposite the 'wanted' source direction, although their magnitudes must be carefully controlled. There is an optimum balance between, on the one hand, a 'phasey' oppressive sensation and nasal sound-quality, when too great a phase-difference is employed, and on the other, a close and bass-heavy sound-quality, when too little phase-difference is employed.[2]

In the phase-modified decoding matrix, the various phase and amplitude relationships between the decoded

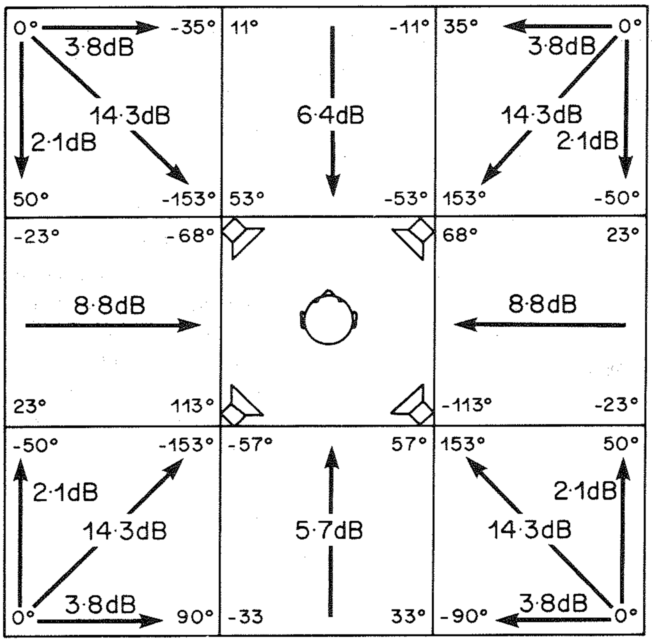


Fig. 2 - Performance chart for phase-modified Matrix H linear decoder (see Section 2)

signals have been optimised, and in analytical, single-source localisation tests this decoder gave accurate image-localisation for a centrally-positioned listener. However, when the listener moved away from the centre-position the directional information became diluted, although the sound sensation was still pleasant. When listening to programme material, listeners commented that, although an extremely pleasant sound sensation was produced, the sounds appeared to be rather close to the listener.

3. Logic enhancement

3.1. Introduction

It is apparent that for a larger usable listening area, signal separations are required which are greater than can be provided by linear matrix decoding. These can be obtained by applying so-called 'logic enhancement' techniques to the decoding process. The decoder is still based upon the linear matrix of the system, but 'logic' circuits are introduced which detect the principal (loudest) sound-source location and vary the decoding parameters to enhance its subjective localisation. In principle, a logic-enhanced 4-2-4 matrix system is capable of reproducing sources at any single location with the same fidelity as a 4-channel discrete system. However, it is a fundamental limitation of such matrix systems that sources at different locations cannot all be reproduced faithfully at the same time. Fortunately, there are ways of masking the deleterious effects of logic-enhancement by exploiting certain insensitivities of the human hearing system.

The overall performance of such a system is characterised by the basic linear matrix and the way in which the logic enhances the principal source. Logic decoders so far developed employ either 'gain-riding'[3] or 'variable-matrix'[4] enhancement; these two techniques will be described briefly.

3.2. Gain-riding logic enhancement

The gain-riding method appears to have been first proposed by Scheiber[3] and many variants have since been suggested. In its simplest form (see Fig. 3), logic circuits vary the gain of the linearly decoded signals to give greater separation between principal-source signals and the associated crosstalk signals; the logic signals are themselves derived from the linearly decoded signals. This process may be expressed as:-

$$\begin{bmatrix} L'_F \\ R'_F \\ L'_B \\ R'_B \end{bmatrix} = \begin{bmatrix} l_f & & & \\ & r_f & & \\ & & l_b & \\ & & & r_b \end{bmatrix} \begin{bmatrix} \alpha_{11}, \alpha_{12} \\ \alpha_{21}, \alpha_{22} \\ \alpha_{31}, \alpha_{32} \\ \alpha_{41}, \alpha_{42} \end{bmatrix} \begin{bmatrix} L_T \\ R_T \end{bmatrix}$$

where l_f, r_f, l_b, r_b are the logic signals and α_{pq} are the linear decode coefficients.

Tests[2] have shown that the minimum audible crosstalk levels for corner and centre-quadrant source locations are about 20 dB and 13 dB respectively. To realise this order of separation with a gain-riding logic system, the crosstalk channels must be attenuated by a similar amount (the basic matrix normally provides about 3 dB of separation). However, secondary sources (quieter than the principal source), at positions different from that of the principal source, will be attenuated similarly. This can lead to secondary-source 'gain-ducking' and 'image movement', although the residual crosstalk signals from the linear matrix sometimes serve to dilute the effect. Careful choice of the attack- and decay-times of the logic action can also help to mask the immediate perception of these effects, but only to a limited extent.

A practical decoder will normally be more complicated than is implied in Fig. 3 with features that attempt to overcome some of the undesirable effects of gain-riding enhancement. Some decoders include a logic-controlled

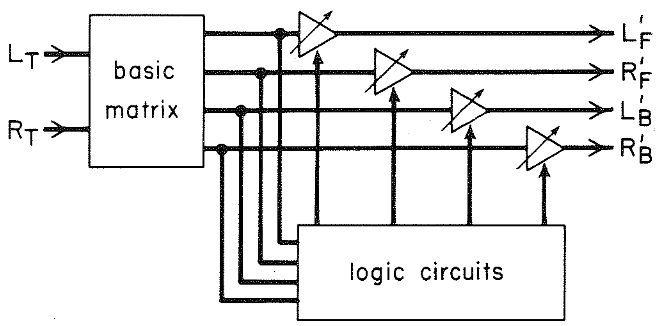


Fig. 3 - Logic-enhanced decoding — gain-riding

blend-circuit between two of the output channels, which operates in a way similar to the variable-matrix technique (described in the next section), albeit after the linear decoding process has taken place. Nevertheless, the overall performance is characterised by the major form of logic enhancement and by the basic matrix.

3.3. Variable-matrix logic enhancement

This technique was developed by Ito and Takahashi[5] and has been reported in several papers.[4,6] The principle involved is illustrated in block diagram form in Fig. 4.

It is appropriate at this point to discuss briefly a variable-matrix logic decoder first employed in the QS quadraphonic system. In this form of variable-matrix decoding, front/back and left/right 'directional detection' takes place in the logic circuits, and the decoding equations are given by:-

$$L'_F = [\ (1+f)\,(L_T-R_T) + (1+l)\sqrt{2}\,R_T\,]\ \angle 0^\circ$$

$$R'_F = [-(1+f)\,(L_T-R_T) + (1+r)\sqrt{2}\,L_T\,]\ \angle 0^\circ$$

$$L'_B = [\ (1+b)\,(L_T+R_T) - (1+l)\sqrt{2}\,R_T\,]\ \angle -90^\circ$$

$$R'_B = [\ (1+b)\,(L_T+R_T) - (1+r)\sqrt{2}\,L_T\,]\ \angle 90^\circ$$

where f, b, l, r are the logic signals.[6] When $f, b, l, r = \sqrt{2}-1$, this expression is identical to the linear decoding matrix for the QS system, and represents the operating point of the decoder. As enhancement takes place, the logic signals vary between the values 0 and $\sqrt{2}$ according to the directional information.

The merit of this technique is that high separation for a principal sound-source is achieved through cancellation of the appropriate pair of terms in the equations and, as a result, the logic signals ($1 + f$, etc.) need vary by only about 8 dB. With the gain-riding system the variation of the logic signals can be 20 dB or more for a similar order of separation. For a variable-matrix system, the variation in level of both the principal and secondary sources should, therefore, be less affected and, more importantly the degree of image movement should be less. Image movement is caused by the decode matrix being skewed to enhance the principal source, and secondary sources can, consequently, be mislocated. The displacement of the secondary image varies with the logic signals and, as the principal source changes, the secondary image is heard to wander.

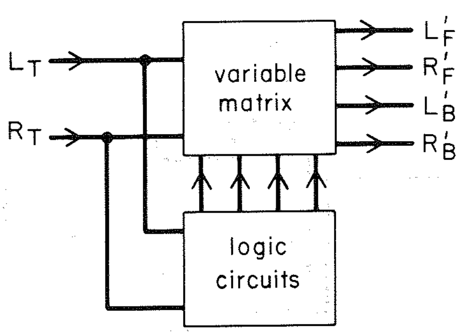


Fig. 4 - Logic enhanced decoding — variable-matrix

As with practical realisations of the gain-riding technique, variable-matrix decoders include features to mask or overcome the side-effects of logic-enhancement. A fast attack-time and slow decay-time are used in the logic circuits, as before, and band-splitting or linear blend-circuits may also be included. However, a brief analysis of the equations describing the decoding process indicates that fewer unpleasant effects may be expected with variable-matrix logic-enhancement than with gain-riding.

3.4. Subjective assessment

In any acoustic study, the performance of a system should be assessed subjectively. During the past few years, many practical decoders have been developed, based on both the two logic-enhancement techniques. Many of these have been assessed at Research Department in extensive subjective investigations. Single-source localisation tests (similar to those described in Ref. 7) were undertaken, and the more successful logic decoders were also assessed using a variety of quadraphonic programme material.

Both the analytical localisation tests and the assessments based on multi-source programme material showed a clear preference for the variable-matrix type of logic decoder.

It should be noted that the basic linear matrix, to which logic-enhancement is applied, also plays an important role in the overall performance of the decoder. The basic matrix of the best decoder of each type was assessed, therefore, in further subjective tests. Although neither was capable of reproducing an effective surround sound-stage, a slight preference was expressed for the basic matrix used in the variable-matrix decoder. This difference is not thought to be significant in accounting for the better performance of the variable-matrix.

The mathematical considerations of the two types of logic-enhancement appear to be fully endorsed by the subjective tests, and the variable-matrix technique should provide the more successful method of logic enhancement for Matrix H.

Two methods of logic-enhanced decoding for Matrix H have so far been investigated. The first was an adaption of a commercial QS–X2 'Variomatrix' decoder, by the addition of a 60° wideband phase-shift, and the second involved the application of variable-matrix logic enhancement to Matrix H linear decoding. These will be discussed in the following sections.

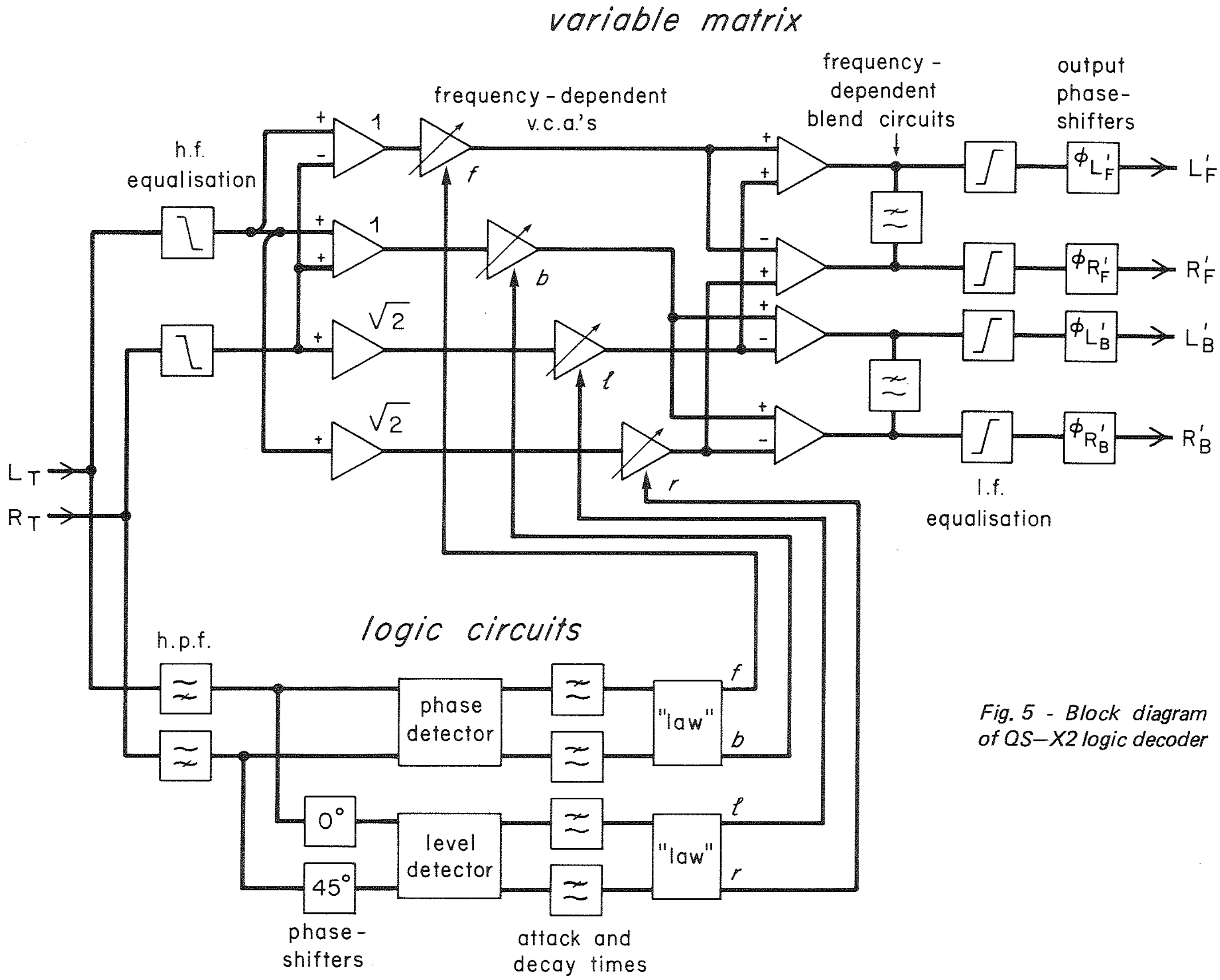


Fig. 5 - Block diagram of QS–X2 logic decoder

4. A commercial logic decoder modified for Matrix H

4.1. General description

In order to approximate the QS decoder (see Fig. 5) to the Matrix H encode, the phase of the R_T input signal was advanced by 60° relative to the L_T input signal, as shown in Fig. 6. In addition the logic action was adjusted to give slightly reduced separation figures. The block diagram shown in Fig. 5 may be considered in two parts, viz the variable-matrix and the logic circuits; these will now be discussed.

4.2. Analysis of variable-matrix logic enhancement

The decoding equations for the QS variable-matrix modified for Matrix H may be expressed as:-

$$L'_F = (1+f)\,(L_T - R_T\,\underline{/60^\circ}) \quad + \quad (1+l)\sqrt{2}\,R_T\,\underline{/60^\circ}$$

$$R'_F = -(1+f)\,(L_T - R_T\,\underline{/60^\circ}) \quad + \quad (1+r)\sqrt{2}\,L_T$$

$$L'_B = (1+b)\,(L_T + R_T\,\underline{/60^\circ}) \quad - \quad (1+l)\sqrt{2}\,R_T\,\underline{/60^\circ}$$

$$R'_B = (1+b)\,(L_T + R_T\,\underline{/60^\circ}) \quad - \quad (1+r)\sqrt{2}\,L_T$$

(an overall power-correction factor of 0·654, and also the output phase-shifts, have been omitted). The logic signals, f, b, l, r, may vary between the values 0 and $\sqrt{2}$ about a 'quiescent' or operating value of $\sqrt{2}-1$, at which point the equations approximate broadly to the basic linear decode Matrix H. When logic enhancement takes place, the control signal(s) corresponding to the direction of the principal sound-source increases in amplitude, whilst the control signal(s) corresponding to the opposite direction decreases.

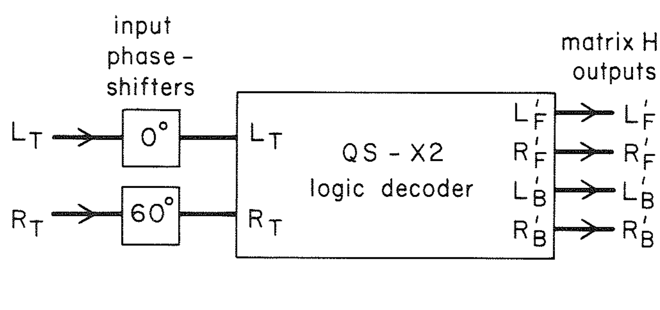


Fig. 6 - Block diagram of QS–X2 logic decoder modified for Matrix H

4.2.1. Single sources

The way in which single or principal sources are decoded will be treated first. Logic enhancement not only increases the desired interchannel separations, but also varies the level of the 'wanted' signals relative to their basic matrix values. Ideally, the overall input/output power law of a source should be constant with azimuth.

With the basic matrix, the crosstalk signals make a significant contribution to the total power (r.m.s. sum) and the power versus azimuth law is constant to within ±1 dB.

With logic enhancement, signal separations are increased and the crosstalk signals contribute only a small amount to the total power. Further, for corner locations, the level of the wanted signal is increased by the logic action by nearly 5 dB above that given by the basic matrix, whilst for centre-quadrant sources, the level remains the same. Nevertheless, the total power law is constant with azimuth to within ±2 dB (c.f. ±1 dB for the basic matrix), and the average power is 1 dB higher than that given by the basic matrix. These differences are small, and the variation in level is unlikely to be audible.

These points may be illustrated by two typical examples (C_F and L_F).

a) A C_F source is coded by Matrix H as

$$L_T = 0{\cdot}828 \underline{/0^\circ}, \quad R_T = 0{\cdot}828 \underline{/-48^\circ}$$

With logic enhancement, the logic signals will be approximately $f = \sqrt{2}$, $b = 0$, $l = r = \sqrt{2}-1$ ($\sqrt{2}-1$ is the quiescent value), and L_F is then decoded as

$$L'_F = (1+f)\ 0{\cdot}828\ (1-1\underline{/12^\circ}) + (1+l)\sqrt{2}\ .\ 0{\cdot}828\ \underline{/12^\circ}$$

giving $|L'_F| = 1{\cdot}67$ with logic-enhancement

and $1{\cdot}65$ with the basic matrix.

The same result holds for R'_F.

b) A L_F source is coded as

$$L_T = 0{\cdot}940 \underline{/0^\circ}, \quad R_T = 0{\cdot}342 \underline{/-75^\circ}$$

and the two logic signals f, l, associated with L_F decoding are both high at $\sqrt{2}$. The L_F decoded signal is therefore

$$L'_F = (1+f)\ (0{\cdot}940 - 0{\cdot}342\ \underline{/-15^\circ}) + (1+l)\sqrt{2}\ .\ 0{\cdot}342\ \underline{/-15^\circ}$$

giving $L'_F = 2{\cdot}60$ with logic enhancement.

This level is $((1+\sqrt{2})/\sqrt{2})$, i.e. 4·6 dB, higher than the basic matrix value. The same is true of other corner sources.

These two source positions may also be used as examples to show how high separation is achieved. For a C_F source, the crosstalk to L_B (equal to the crosstalk to R_B) is given by

$$L'_B = (1+b)\ 0{\cdot}828\ (1+1\ \underline{/12^\circ}) - (1+l)\sqrt{2}\ .\ 0{\cdot}828\ \underline{/12^\circ}$$

$$= (1+b)\ 1{\cdot}647\ \underline{/6^\circ} - 1{\cdot}656\ \underline{/12^\circ}$$

This has a minimum value of 0·173 when $b = 0{\cdot}000$ giving a maximum separation of 19·7 dB.

For a L_F primary source the crosstalk to R_F is given by

$$R'_F = -(1+f)\ (0{\cdot}940 - 0{\cdot}342\ \underline{/-15^\circ}) + (1+r)\sqrt{2}\ .\ 0{\cdot}940$$

$$= -(1+f)\ 0{\cdot}616\ \underline{/8{\cdot}3^\circ} + (1+r)\ 1{\cdot}329$$

If $f = \sqrt{2}$, the minimum value of R'_F is 0·214 given when $r = 0{\cdot}107$ and corresponds to a maximum separation of 21·7 dB.

These two examples show how the two terms on the right-hand side of the equations cancel to give high separation and how it is important for them to be approximately in anti-phase. Having defined the basic matrix it is possible to derive, in a similar way, the maximum separation figures for every principal-source position. In practice it is difficult to arrange the logic circuits so that the derived values of the logic signals, f, b, l, r are able to give these theoretical, maximum separation figures for every source location. However, this has not proved to be a disadvantage for the following reason.

As has been seen earlier, image wandering of secondary sources is caused by the decode matrix being skewed to enhance the localisation of principal sources. By reducing separation, and the range of the logic signals about the quiescent value, the amount by which the decode matrix is skewed, and hence the amount of image wandering, is reduced. In this way, the performance of the modified QS decoder was improved by reducing the separation for a corner source to about 14 dB in the front/back direction and 20 dB in the left/right direction.

4.2.2. Two sources

If two sources are present simultaneously at different locations in the quadraphonic stage, the situation is evidently more complex than for the single-source case. Even if one source is nominally louder than the other, programme signals have a wide dynamic range and the secondary source can often be louder than the principal source for short, but significant, periods of time. For this and other reasons (see Section 4.3.), the logic signals vary at a rate determined by the programme content but limited by the time-constants of the logic circuits.

An extreme case will be considered in which the secondary-source level is low relative to that of the principal source, so that the logic signals are independent of the secondary source. In this example, the principal source is located at L_F and the secondary source at

C_F. The logic signals corresponding to the L_F source are $f = \sqrt{2}$, $l = \sqrt{2}$, $b = 0$, $r = 0$ (nominal values); the secondary source is coded as

$L_T = k$, $R_T = k \angle{-48^\circ}$ (k being a factor less than unity).

These signals will therefore be decoded to give

$$|L'_F| = k \,|\, (1+f)\,(1-1\,\angle 12^\circ) + (1+l)\,\sqrt{2}\,\angle 12^\circ| = k\,3{\cdot}40$$

$$|R'_F| = k\,|-(1+f)\,(1-1\,\angle 12^\circ) + (1+r)\,\sqrt{2}\quad| = k\,1{\cdot}45$$

$$|L'_B| = k\,|\,(1+b)\,(1+1\,\angle 12^\circ) - (1+l)\,\sqrt{2}\,\angle 12^\circ| = k\,1{\cdot}45$$

$$|R'_B| = k\,|\,(1+b)\,(1+1\,\angle 12^\circ) - (1+r)\,\sqrt{2}\quad| = k\,0{\cdot}60$$

Clearly there is an imbalance; the secondary source image is 'pulled' towards the principal source at L_F, the crosstalk signals are increased, and the overall power of the secondary source is increased (3 dB higher than if it had been the principal source). In practice, the logic signals vary with programme content and one may hear the secondary source wandering between C_F and a position between C_F and L_F.

The pulling of secondary sources towards the principal source is a general feature of this type of logic enhancement and constitutes perhaps its most serious limitation. The worst example of image wandering conceivable is probably that which can occur when two sources of a similar level are located diametrically opposite one another; however such situations occur infrequently. Moreover, the ear appears to be relatively insensitive to secondary source movement, particularly if the attack- and decay-times of the logic action are judiciously chosen. Normally, image wandering is not found to be seriously objectionable.

Although emphasis has been placed on image wandering, the level of secondary sources also changes as a function of the logic action. It is found that, in most cases, the total power of the secondary source does not vary by more than about 3 dB.

4.3. Logic-enhancement circuits

The logic circuits are shown in block diagram form in Fig. 5. The encoded two-channel signals are high-pass filtered before being applied to interchannel level- and phase-detectors,* both of which incorporate high-gain, limiting-amplifiers and phase-discriminators. In the level-detector a phase-detector is preceded by a 45° phase-shifter together with sum and difference amplifiers. (The interchannel level (ratio) can then be estimated, independently of the absolute levels of the signals, by measuring the phase-difference of two derived signals; these are formed by taking the sum and difference of the encoded signals, after one has been phase-shifted by 45°). The combination of the limiting-amplifiers and the input filter prior to the detector gives an effective cut-off frequency of about 100 Hz. However, this is level-dependent since, for very small input signals, the limiting-amplifier has insufficient gain to drive the phase-detectors, and the output level of the detector falls to its quiescent value.

Both the level- and phase-detectors provide balanced d.c. outputs proportional to the phase-difference (90° representing the quiescent value). These constitute the logic signals and represent the principal-source location in the original quadraphonic stage. They drive voltage-controlled, variable-gain amplifiers, (v.c.a.'s) via filters that determine the attack- and decay-times of the logic action. The frequency content of the logic signals is related to the dynamic characteristic of the principal source, the relative levels of different sources, and the movement of a source around the stage. However, it is necessary to band-limit the frequency spectrum to mask the onset and decay of logic enhancement

The gain of the amplifiers in the matrix circuits varies with the d.c. level of the logic signals. Circuits prior to the v.c.a.'s determine the 'law' of the logic signals so that f, b, l, r excurse between the values 0 and $\sqrt{2}$ about the quiescent value of $\sqrt{2}-1$.

The v.c.a.'s themselves are frequency dependent. At low audio frequencies they have constant gain so that the decode matrix is independent of the logic action and reverts to its basic form. At high frequencies the logic action is partially bypassed and the separation is reduced, but still to a figure greater than that given by the basic matrix.

As has been seen there are three frequency-dependent stages in the logic circuits, i.e. the input-signal high-pass filtering, the logic-signal time constants, and the frequency dependence of the v.c.a.'s. All of these are mutually dependent and need to be carefully matched. The attack-time of the logic signals should be fast enough for a new principal source to be correctly located without transient mislocation or wandering. At the same time, the phase- and level-detectors require at least a few cycles of audio in order to derive the logic signals accurately, and at low audio frequencies, the necessary period is longer than the attack-time. The input signals are therefore high-pass filtered. Even so, the pass-band largely includes those frequencies where the energy of average programme material is at a maximum (i.e. about 100 Hz to 1 KHz),[8] and a fair estimate is made of the location of sources. Further equalisation of the input signals may improve this estimate.

Since the logic signals may vary up to a frequency determined by the attack-time, audio signals below this frequency must not be subject to the logic control, otherwise severe intermodulation distortion may occur. For this reason, the v.c.a.'s have constant gain at low frequencies. At higher audio frequencies, the modulating frequency (i.e. that of the logic signal) is only a small fraction of the audio frequency and the distortion under these conditions is almost inaudible,

* In this context interchannel level means the ratio of the levels of two signals, and interchannel phase means their phase-difference.

being transient in nature.

4.4. Output circuits

4.4.1. Blend and equalisation

Following the decoding stage, frequency-dependent blend-circuits are inserted between the corresponding left and right channels of the four audio output channels (see Fig. 5). At very low frequencies (less than 100 Hz) blend is almost total; thus low frequencies are localised on the front/back centre-line. With increasing frequency, left-to-right separation rises to a maximum at 1 KHz, but is again reduced at higher frequencies by the h.f. roll-off of the v.c.a.'s. Typical response curves are shown in Fig. 7. Because of the frequency-dependence of the v.c.a.'s, it is necessary to apply high and low frequency equalisation to maintain a uniform overall power response. An h.f. attenuation of 1·5 dB is effected at the input to the variable-matrix, and an l.f. attenuation of 3 dB is effected at the output.

4.4.2. Output phase-shifters

Finally, a phase-shifter is added to each output channel. These are simple, single-pole, all-pass networks that introduce constant phase-differences for only a relatively narrow band of frequencies around 700 Hz. They attempt to reduce the rather unpleasant sensation of 'phaseyness', but also appear to cause image blurring.

4.5. Subjective assessment

In subjective tests, the higher signal separations of this decoder afforded a more open sound than that given by the basic matrix, with sharper definition of the sound-stage. However, criticisms were made of image movement and mislocation of transient sounds, such as speech sibilants. Also ambience at the rear of the stage was too narrow, giving the effect of a 'tunnel of sound'. Nevertheless, the quadraphonic performance of this decoder, together with that of the QS system, was judged to be significantly better than that of other 4-2-4 matrix systems so far tested.

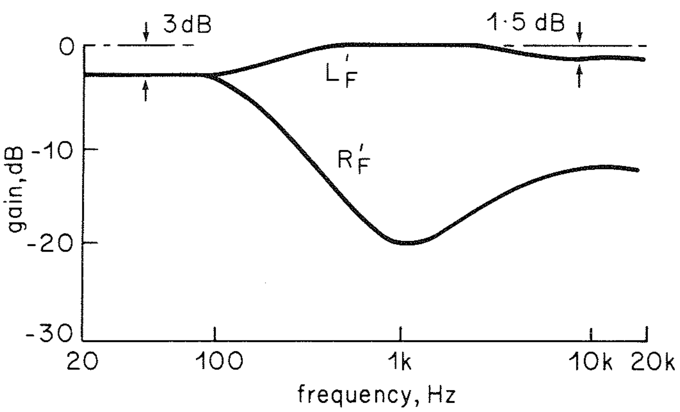


Fig. 7 - Typical separation curves for a L_F source given by the modified QS decoder

5. Matrix H logic decoder

5.1. Mathematical analysis

Although the quadraphonic performance of the modified commercial decoder was good, it still had limitations which it would be desirable to overcome. The design of a new decoder was therefore undertaken, in which logic enhancement was applied directly to the basic linear Matrix H.

The Matrix H decoding equations may be written as:-

$$L'_F = [\ 0{\cdot}940 f\ (L_T - R_T \underline{/75^\circ}) + l\ 1{\cdot}282\ R_T \underline{/75^\circ}]\ \underline{/-20^\circ}$$

$$R'_F = [-0{\cdot}940 f\ (L_T - R_T \underline{/75^\circ}) + r\ 1{\cdot}282\ L_T\ ]\ \underline{/-55^\circ}$$

$$L'_B = [\ 0{\cdot}940\, b\ (L_T + R_T \underline{/40^\circ}) - l\ 1{\cdot}282\ R_T \underline{/40^\circ}]\ \underline{/25^\circ}$$

$$R'_B = [\ 0{\cdot}940\, b\ (L_T + R_T \underline{/40^\circ}) - r\ 1{\cdot}282\ L_T\ ]\ \underline{/-65^\circ}$$

where, for linear decoding, the logic signals f, b, l, r are at their quiescent value of unity. By varying the logic signals in a similar way to that described for the modified commercial decoder, high separation for a principal source can be achieved.

These equations are realised in the decoder shown in block diagram form in Fig. 8. This decoder is similar in many ways to the modified commercial decoder and makes use of the same integrated circuits. The logic circuitry, however, incorporates a number of improvements, and more logic outputs are provided to drive extra variable-matrix circuits. The audio channels have been more extensively modified, not only to realise the exact basic Matrix H equations, but also to improve other aspects of the decoding.

Analysis of the variable-matrix can be performed in the way described for the modified commercial decoder, and maximum separation figures have been predicted. This has shown that adequate separation can be achieved for most source locations, but for a corner signal the maximum front-to-back separation is relatively low (13·6 dB) and this may displace the image slightly. Separation can be increased by slightly altering the front and back phase-angles of the R_T signal from their values of 75° and 40° respectively. With this modification, a better overall performance can be expected; this approach is under further investigation.

5.2. Input phase-shifters

Accurate wide-band all-pass networks provide appropriate interchannel phase-shifts at the input of the decoder. R_T is phase-shifted by 40° and 75° (relative to L_T) for the decoding matrix, and by 67° and 22° for the logic detection circuits.

5.3. Frequency-dependence of the variable-matrix

As already discussed, in the modified commercial decoder, the matrix is independent of logic action at low

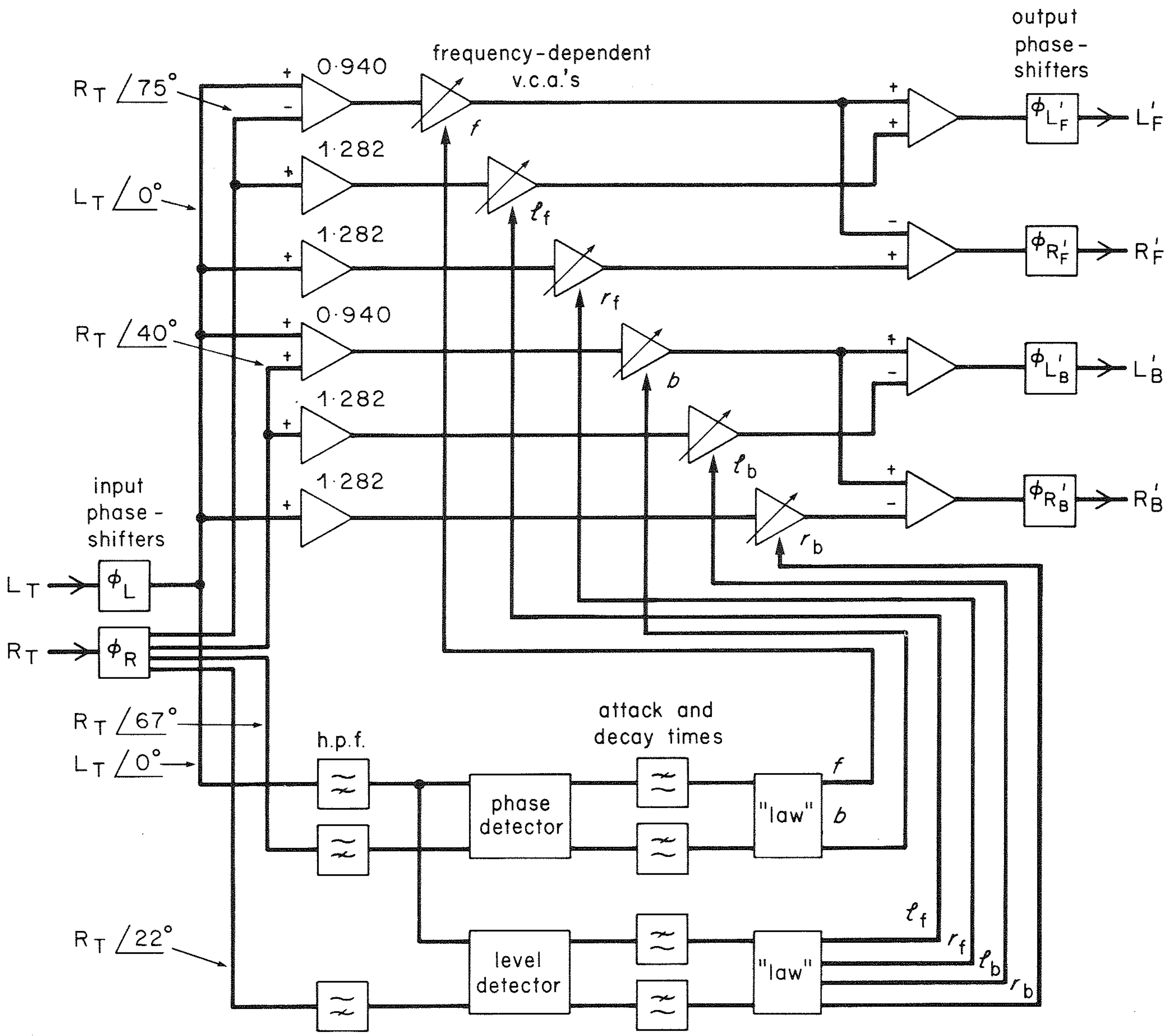


Fig. 8 - Block diagram of Matrix H logic decoder

frequencies and reverts to the basic matrix. However, at these low frequencies, the basic matrix is distorted by the left-to-right blend-circuits in the output channels, which localise low-frequency sounds on the front/back centre-line. No such blend-circuits are included in the Matrix H decoder, since basic Matrix H is capable of accurately localising sounds without logic enhancement.

This has the added advantage of maintaining high separation to a lower frequency (see Fig. 9) and as a result, the total energy of the crosstalk signals is less for the same maximum separation figure (at 1 KHz). This permits a reduction of the logic action so as to reduce image wandering, whilst still maintaining adequate separation.

The logic signals modulate the audio signals and the removal of the blend-circuits might be expected to result in audible intermodulation distortion (see Section 4.3.). However, in practice no evidence of this has been found, possibly because the distortion is of a transient nature.

At high frequencies (above 1 kHz), the high separation of mid-band frequencies is maintained by not restricting the high-frequency response of the v.c.a.'s in the variable-matrix, as in the commercial decoder (see Fig. 9).

5.4. Phase-correction of the output signals

Work on the effects of interchannel phase-differences on the localisation of quadraphonic images[2] has shown that even small phase-differences (of the order of 20°) can sometimes displace or blur an image. There is also evidence that adverse phase-differences can increase the audibility of image wandering in the following way. If a large phase-difference exists between two principal-source signals (for a C_F source, say), the image is displaced and

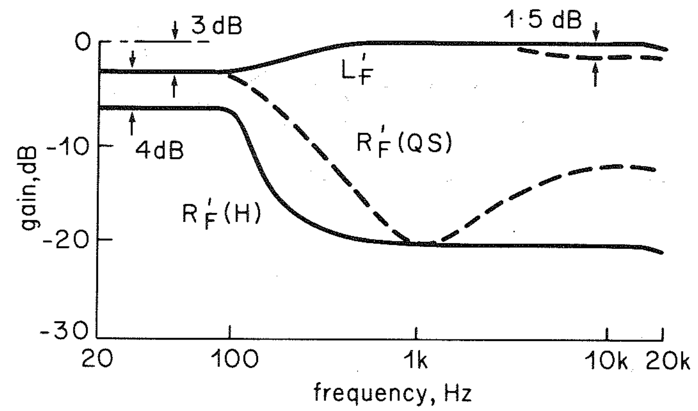


Fig. 9 - Typical separation curves for a L_F source with logic enhancement

H (see Section 5)

modified QS (see Section 4)

even small additional variations of phase can cause the image to wander. If, on the other hand, the phase-difference is small, the same variations will have a negligible effect.

Great care was taken, therefore, in the design of the output phase-shifters, to ensure that the proper phase relationships exist between two principal-source signals, and between a principal-source signal and a crosstalk signal. The phase-shifts used were slightly different from the basic matrix values (shown in the equations of Section 5.1.) in order to account for the logic action and the higher interchannel separations produced. They are accurate up to a frequency of about 4 kHz, it being unnecessary to maintain stringent tolerances at higher frequencies (unlike the input phase-shifters in the decoder). With the Matrix H logic decoder, images were found to be much sharper and better defined than those given by the modified commercial decoder; further, slightly less image wandering was also noted, although the sharper images might have been expected to emphasise this effect.

5.5. Discussion

On a brief subjective assessment of this decoder, significant improvements were found as compared with the modified commercial decoder. The principal improvements consisted of sharper images, a greater sense of 'openness' and better overall perspective, fewer sibilant mislocations, and a much greater tolerance to listener position.

As mentioned earlier, it is thought that this type of decoding, which uses different interchannel phase-angles for decoding the front and back channels, could be optimised by slightly altering the phase-angles from those used in the basic Matrix H decoding equations. Further developments can also be envisaged, for example, using more complex forms of logic-enhancement, or the incorporation of delay-lines in the audio channels to overcome low-frequency localisation and transient problems.

6. Simplified Matrix H logic decoder

The Matrix H decoder described above is a fairly complex device, as Fig. 8 shows. Methods of simplifying the circuits and of overcoming some of the residual impairments were investigated. This work continues, but a simplified decoder has already been developed with a performance comparable with, if not better than that of the Matrix H logic decoder described above.

A block diagram of this decoder is shown in Fig. 10. The logic circuits used are the same as those incorporated in the Matrix H logic decoder described in Section 5, but the basic matrix is modified to decode using a single phase-angle for both front and back channels. The decode equations become:-

$$L'_F = [\ 0{\cdot}940 f (L_T - R_T \angle 67^\circ) + l\, 1{\cdot}282\, R_T \angle 67^\circ]\ \angle -20^\circ$$

$$R'_F = [-0{\cdot}940 f (L_T - R_T \angle 67^\circ) + r\, 1{\cdot}282\, L_T\ ]\ \angle -50^\circ$$

$$L'_B = [\ 0{\cdot}940\, b (L_T + R_T \angle 67^\circ) - l\, 1{\cdot}282\, R_T \angle 67^\circ]\ \angle 25^\circ$$

$$R'_B = [\ 0{\cdot}940\, b (L_T + R_T \angle 67^\circ) - r\, 1{\cdot}282\, L_T\ ]\ \angle -95^\circ$$

where the basic matrix is given when $f, b, l, r = 1$.

Since R_T is phase-shifted by the same angle for the front and back channels, the variable-matrix circuits are reduced in complexity. At the same time, separations for corner sources are improved without significantly sacrificing other locations.

Although the basic matrix has been altered, it can be seen from the chart shown in Fig. 11 that its performance closely resembles that of the phase-modified linear Matrix H shown in Fig. 2. Moreover, the two linear matrices were compared subjectively using programme material and there was little apparent difference between them.

With the different basic matrix, the output phase-shifts are slightly changed from those used in the Matrix H logic decoder of Section 5. The values used give good phase-correction of the output signals, with both the basic matrix and logic enhancement. Fig. 12 shows a typical performance chart for the decoder, applicable to frequencies in the logic-enhanced band, and Fig. 11 represents the performance at low frequencies where logic

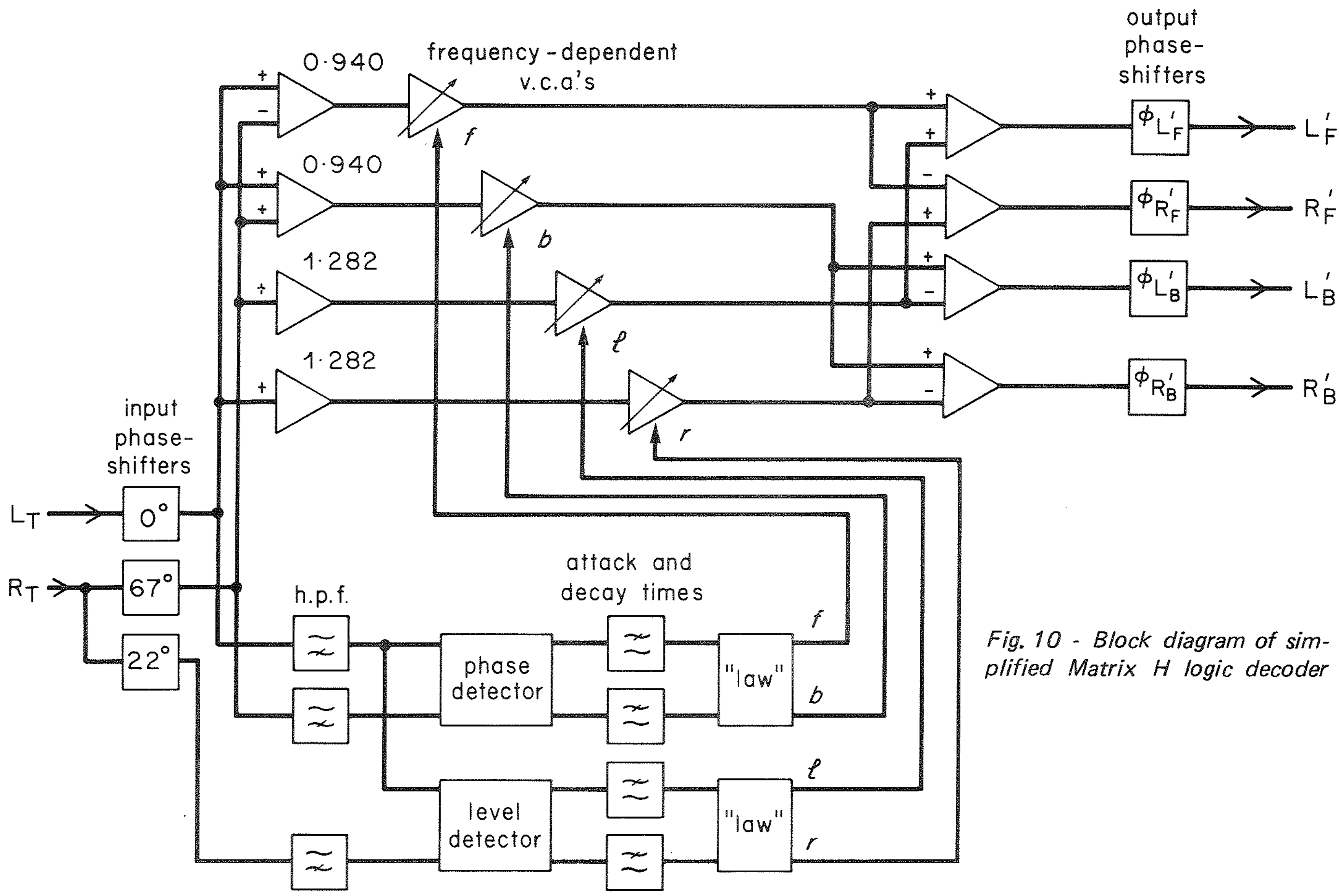


Fig. 10 - Block diagram of simplified Matrix H logic decoder

is not operative.

Subjectively the performance of the decoder was found similar in many ways to the Matrix H logic decoder of Section 5, with the addition that corner signals were better defined. Since the favourable qualities of the more complex decoder are also exhibited by the simplified version, more comprehensive subjective assessments were undertaken and these are discussed in the next section.

7. Subjective appraisal of Matrix H decoders

7.1. General

Throughout the development of the Matrix H quadraphonic system a considerable number of subjective tests have been conducted in order to assess the performance of this and other proposed matrix systems. These have included simple analytical tests, designed to examine the basic properties of the system, and programme listening tests to examine the qualitative performance and the ability of the system to reproduce complex source-signal arrangements.

7.2. Localisation tests

One highly informative and analytical set of tests involves single-source localisation, as described in Reference 7. The listener is asked to estimate the position and spread, or diffusion, of a sound-image produced by the system, with a source-signal encoded at any one of sixteen azimuth positions.

In such tests the basic Matrix H decoder was found to give good overall positional accuracy but the images were more diffuse than those of discrete 4-channel quadraphony. However, unlike most other systems, when decoded using a basic linear matrix the images were not unpleasant or 'phasey' in quality, and were reasonably stable with head movement. Some comments of 'closeness' of images were made, but otherwise the subjective results were found to be quite acceptable.

The commercial decoder modified for Matrix H gave better overall positional accuracy and considerably sharper images than the linear decoder, to the extent that the results were not significantly inferior to those of a discrete system. The 'closing-in' effect of the linear decoder was absent, but some comments were made that sibilants were localised at positions different to that of the main image. This is probably due to limitations in the transient performance of the logic enhancement.

The Matrix H logic decoder described in Section 5 has not yet been tested fully, but the simplified version (see Section 6) gave a further small improvement in overall positional accuracy compared with that given by the modified commercial decoder, and the overall performance closely matched that of a discrete system. In

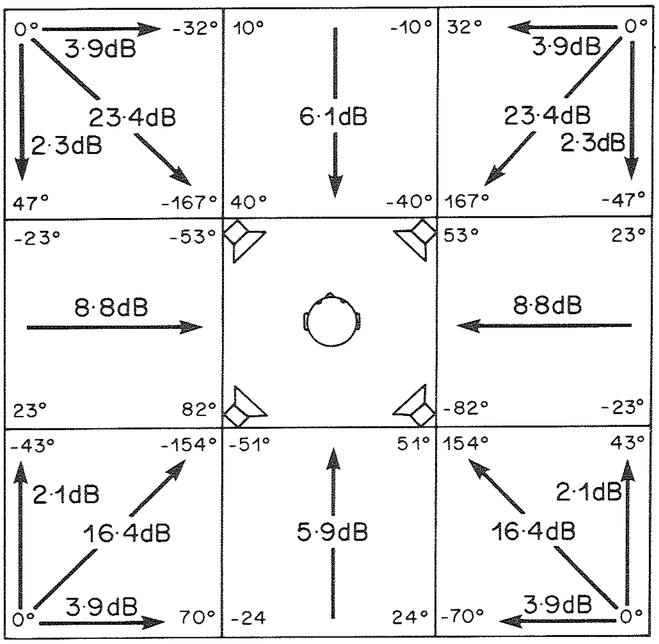


Fig. 11 - Performance chart for basic matrix in the simplified Matrix H logic decoder (see Section 6)

particular, although sibilant effects were not completely absent, they were much less noticeable than those for the modified commercial decoder.

7.3. Programme assessment

A wide selection of programme material, including serious music, light music, pop, comedy, and drama, has been used in assessing various 4-2-4 matrix systems. For assessments of this nature, a complete encode/decode combination is sometimes used when balancing the original programme material (obtained from, say, multi-channel recordings). By monitoring the encoded stereo signals and the decoded quadraphonic signals, the performance of the particular system can be optimised. In such circumstances, the four inputs to the encoder may not be the most suitable for a discrete 4-channel system. Further, and perhaps more importantly, a 4-2-4 matrix system balanced in this way may not provide optimum encoded signals (L_T, R_T) for an alternative form of decoder. Thus in tests assessing the subjective performances of the various Matrix H decoders, the programme material was balanced only for discrete quadraphony, and encoded by Matrix H. Given the same encoded signals, the decoders were then compared using discrete quad as a reference.

The basic linear Matrix H decoder gave good positional accuracy with multi-source material and good tonal quality. It gave an overall pleasing sound sensation that was, however, somewhat blurred and 'closed-in' when compared to discrete quad. Some instability of the sound-stage was apparent with considerable head movement about the listening position and, when the listener moved out of the central listening area, the sound-stage collapsed to the nearest loudspeaker more noticeably than with discrete quadraphony. Even so, a pleasing unoppressive sound was maintained, unlike most other 4-2-4 systems when decoded linearly.

With the modified commercial decoder a much more spacious sound was produced, generally with good tonal quality. Its performance was more similar to discrete quad than that of the linear decoder. Occasional sibilant mislocations were noted, mainly on speech, but these were not too objectionable. However, with serious music the ambience was often found to be too narrow at the rear of the sound-stage, and a narrowing of the front-stage also occurred when the main body of sound was located in the centre-front region of the stage. It was also found that, for complex programme material, sound-images seemed to be less clearly defined than with discrete quad, and there was an apparent excess of low-frequency energy in the centre of the stage. This was almost certainly due to the left/right blending in the commercial decoder at low frequencies. Some image movement was detectable, and in particular a dominant front sound-stage tended to pull forward secondary sound-images, located at the rear corners, to appear at the sides of the quad stage; however, this was seldom seriously objectionable. Some secondary image wandering could occasionally be detected by experienced listeners, but none of these deficiencies appeared to be severely detrimental to listener's appreciation of the quadraphonic sound. This decoder was more tolerant to off-centre listening positions than the linear decoder, but uncomfortable 'phasey' effects could be detected in some locations for some image positions, largely due to the limitations of the phase-correction circuits employed.

The simplified Matrix H logic decoder again produced a spacious sound of good tonal quality similar to discrete quad. The sound gave the impression of being significantly clearer, with a more 'open' perspective than that of the

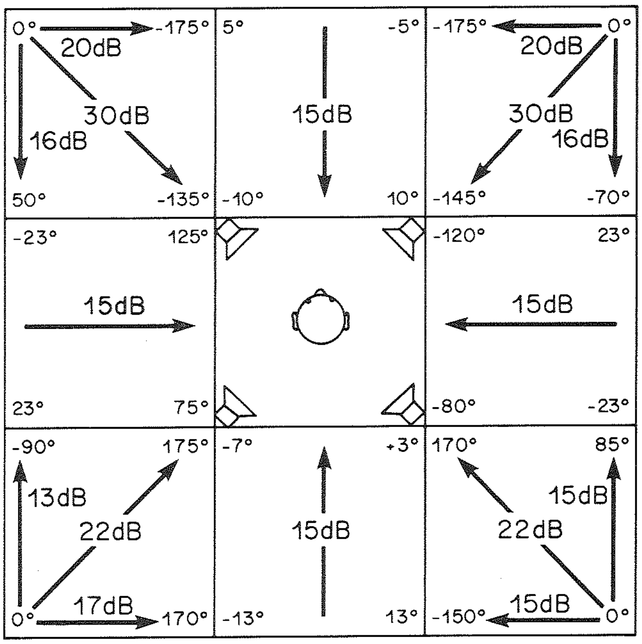


Fig. 12 - Performance chart for simplified Matrix H logic-enhanced decoder (see Section 6)

modified commercial decoder, and was judged to be very close to the discrete quadraphonic sound. Ambience-spread in the rear-stage was substantially improved, and had a more natural tonal quality. In addition, compression of the front-stage was much less obvious than with the modified commercial decoder. Sibilant effects were hardly noticeable, although occasional image movement could still be detected. The lack of low-frequency energy in the centre-stage region, using complex source material, was considerably preferred with this decoder, and this point was significant when listening for extended periods; a more 'comfortable' sound sensation was commented upon. Tolerance to off-centre listening appeared to be particularly good, very much like discrete quad, and the unpleasant 'phasey' sensations observed with the modified commercial decoder were absent.

A 3-way comparison test between the modified commercial decoder, the simplified Matrix H logic decoder, and discrete quad (used as a reference) was arranged after the initial assessment period. Nine studio managers from BBC Radio Broadcasting Groups were asked to assess and rate the two decoder performances on a continuous 0—100% quality scale, with discrete quad as a reference, necessarily defined as having a 100% rating. The listeners were unaware of the decoder options being used. They listened to a 30-minute tape containing a wide selection of programme items mixed for discrete quad. Overall, the simplified Matrix H decoder was rated at 77% as compared to discrete quad, and the modified commercial decoder was rated at 47%. However, this result pertained to tests where small differences in performance might be expected to be magnified; it should be noted that, in some earlier tests, where the original programme material was balanced for the Matrix H system using the modified commercial decoder, a much closer match was obtained to discrete quad. This match was considerably better than that for other matrix systems.

8. Conclusions

The Matrix H 4-2-4 quadraphonic system permits the use of a number of different decoding options, all of which are capable of producing worth-while quadraphony, but with different limitations.

The basic linear Matrix H decoder produces a pleasing surround-sound; it is simple and has the advantage of time invariance, but it cannot produce sharply-defined sounds, and gives a somewhat restricted listening area.

The application of logic-enhancement techniques to Matrix H decoding can provide a sharply-defined and more spacious sound-stage, similar to discrete quad. However, since the 'logic' circuits can only produce ideal decoding for one sound-source direction at a time, much of the success of this technique depends upon effective deception of the human hearing mechanism.

The variable-matrix technique of logic enhancement has been found to be most successful to date and, as an initial expedient, a commercial ('Variomatrix') decoder was modified for Matrix H decoding. This gave generally good results, although limitations due to the instrumentation of the decoder were apparent to the experienced listener.

The Matrix H logic-enhanced decoder combines the advantages of a good linear decode matrix with those of the variable-matrix enhancement technique, but is rather complex. However, a simplified version, which is comparable in complexity with the commercial decoder, and is based upon a slightly modified form of the basic matrix, has been shown to provide the best quadraphonic performance to date from a 2-channel matrix system. Although the limitations of logic enhancement still exist, their adverse subjective effects are, in general, well masked.

It is considered that the limit of the performance of Matrix H decoders has not yet been reached and, as discussed in this report, some aspects of decoding have yet to be optimised. In addition there are other techniques that may be applied to future decoders. Nevertheless, it has been shown that a high standard of quadraphonic reproduction may be achieved with the decoders discussed in this report.

9. References

1. MEARES, D.J. *and* RATLIFF, P.A. The development of a compatible 4-2-4 quadraphonic matrix system, BBC Matrix H. EBU Review, Technical, **159**, October 1976.

2. RATLIFF, P.A. Properties of hearing related to quadraphonic reproduction. BBC Research Department Report No. 1974/38.

3. SCHEIBER, P. Four channels and compatibility. JAES, **19**, (4), pp 267-279, April, 1971.

4. MOSELY, J., TAKAHASHI, S. *and* HIRANO, K. The principles and performance of the Sansui QS Variomatrix. Paper presented at 44th AES Convention, Rotterdam, March, 1973.

5. ITO, R. *and* TAKAHASHI, S. Variable matrix decoder for use in 4-2-4 matrix playback system. US Patent No. 3,825,684, July, 1974.

6. HELLER, D. Surround-sound decoders, Pts. 3 and 4. (QS Variomatrix). Wireless World, **82**, (1488-9), pp. 57-59 August, and pp. 53-56, September, 1976.

7. CROMPTON, T.W.J. The subjective performance of various quadraphonic matrix systems. BBC Research Department Report No. 1974/29.

8. MEARES, D.J. Statistics of typical programme sound pressure levels in sound studios and their control rooms, BBC Research Department Report No. 1973/37.

HR/JUC/AM