Recording Techniques for Multichannel Stereo

by M. A. Gerzon, M.A.*

The following lecture was held at the ITA, 70 Brompton Road, London, S.W.3, on April 21, 1971. Chairman of the meeting was R. C. Driscoll, M.Sc.

IN THE last few years, many minds have been exercised by the problem of reproducing the all-round reverberation of the concert hall in the home. As became clear at an early stage, it proves to be necessary to place at least four loudspeakers around the listener for adequate results. The slump in classical record sales in the U.S.A. has encouraged a rather premature marketing of 4-channel equipment before satisfactory recording techniques have been developed. The result is that many commercial 4-channel tapes are of the 'gimmick' variety, with the listener placed in the middle of an orchestra whose instruments are concentrated at five discrete points. It is not intended here to discuss such 'gimmick' recordings further, as the microphone technique used is dependent purely on the whim of the producer.

Briefly there are two main aims in introducing multichannel stereo into the home: to reproduce sounds from all round the listener, and to reproduce that quality of 'richness' associated with reverberation arriving from all directions. In terms of these aims alone, current commercial 4-channel recording techniques can be quite successful, but it must not be forgotten that it is desirable to capture other qualities of sound as well, such as a sharp stereo definition, an accurate sense of distance, location in the vertical dimension, and a fair reproduction of the distinctive properties of the original live reverberation. In order to devise optimum recording techniques for four-speaker reproduction, it is necessary to have some understanding of how the ear derives this information from the sound reaching it.

It is well-known (see ref. 1) that the direction of a sound is determined either by the ratio of the amplitudes reaching the two ears, or by the time delay between the ears, or by a combination of these two effects. However, because of the approximate symmetry of the head about the axis joining the two ears, this mechanism is only capable of determining the angle off the axis from which the sound originates. The determination of the angle *round* the axis, i.e. the elevation of the sound and whether it comes from in front or behind, appears to rely on other mechanisms, most of which are not well understood. However, it has been shown that one mechanism for deriving this "axial" information is the interpretation of the way the sound at the two ears varies with small head movements. As the head is rotated, the sound gets louder in the ear moving towards the sound source, and quieter in the one moving away. By this means, head movements of a few degrees are sufficient for the ears to recover axial information, and such head movements occur unconsciously all the time.

Same Paper #2

In the concert hall, small head movements cause the volume of direct sounds to vary continuously in each ear, whereas because reverberation arrives uniformly from all round the listener, the relative intensity of reverberation reaching the two ears does not vary. It is therefore plausible that reverberation will take on the quality of 'richness' only if its relative intensity at the two ears does not vary with small head movements. This explanation of richness is supported by the rich sound obtained with stereo headphones, or with stereo reproduced through two loudspeakers placed one at each side of the listener; in both cases there is little variation of volume at the ears with small head movements.

good for Paper!

Suppose then that four loudspeakers are each reproducing four distinct reverberation channels of equal intensity, as would be the case for example with recordings made with four cardioid microphones placed at 90° to one another in a horizontal plane. With a given loudspeaker layout it is possible to compute theoretically whether small head movements will cause the volume of reverberation at each ear to remain unchanged. One can thereby determine the 'richness' given by various loudspeaker layouts, and determine optimal ones.

Full details of the method of calculation will be published elsewhere in due course, but the basic idea is to regard the loudspeakers as being 'delta-function' energy sources on a sphere around the listener, and to decompose this energy distribution into its zero, first, second, third and higher spherical harmonic components. The ear's energy polar response varies considerably with frequency, but this complication can be overcome by decomposing it into its spherical harmonic components which are approximately symmetric about the ears' axis. If one assumes that spherical harmonic components above the third are unimportant, then one can show that the variation in intensity of reverberation at the two ears caused by small head movements is minimal if the lower order (say up to third) spherical harmonic components of the reproduced energy distribution vary at most quadratically in the neighbourhood of the normal or average direction of the ears' axis.

Such an analysis gives results which accord with subjective impressions of richness, and reveals that the loudspeaker arrangements illustrated in fig. 1 give good richness, whereas those in fig. 2 give a less good richness, when the listener is facing approximately forward. Note

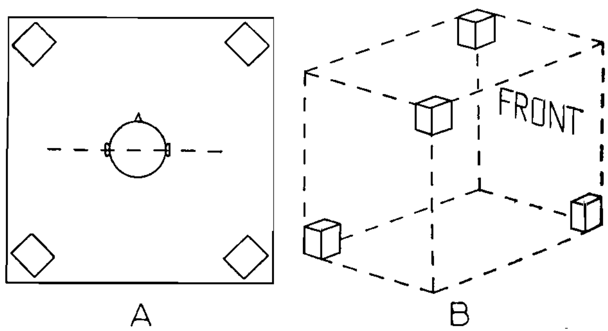

Fig. 1. 'Rich' loudspeaker arrangements.

that a system giving poor richness can often be made to give good richness, and vice-versa, simply by rotating the head through a suitable angle. It should be noted that the habit of elongating the conventional square arrangement of fig. 1a into that of fig. 2a causes an appreciable loss of richness, indicating the importance of careful speaker placement.

All this has a very important bearing on 4-channel recording techniques, as if some loudspeakers get less reverberation than others, there will be a loss of richness. Thus the recording engineer should endeavour to capture an equal amount of reverberation on all four channels, front and rear. Some commercial 4-channel recordings have inadequate front reverberation, although in this context note that it is becoming customary to record the rear reverberation channels 10 dB up relative to the front, and the listener must therefore always adjust the front-rear balance for best richness.

There are other important properties of live reverberation besides 'richness', although this fact is often overlooked. As the handling of reverberation is the *raison*

*Mathematical Institute, Oxford

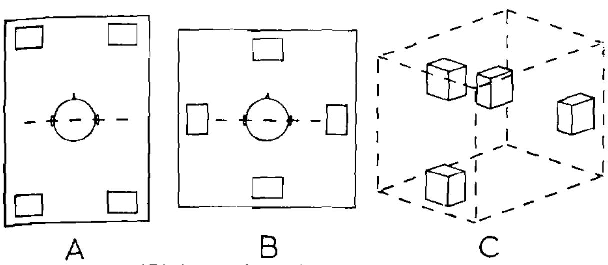

Fig. 2. Less 'Rich' loudspeaker arrangements.

d'être of multichannel stereo, it is necessary to look at why the ear is more accurate in its directional perception in live rooms with much reverberation than in the open air. This study also has lessons for the recording engineer, as on many commercial 4-channel recordings, the presence of the rear channels degrades the stereo sharpness.

Fortunately, a clue to the mechanism involved is provided by the observation that certain coincident microphone 2-channel recording techniques give a much more stable and precise stereo location than others, 90°-angled figure-of-eight microphones in particular revealing very fine details of placement. In order to explain this, I have calculated the distribution of reverberation energy coming from different positions across the stereo image for the various different coincident microphone techniques, and these distributions are illustrated in fig. 3. In this figure, the 'position' of a recorded sound is defined to be the quantity $\frac{4}{\pi}\arctan\left(\frac{L-R}{L+R}\right)$, where L is the amplitude of the sound on the left channel, and R on the right channel. For in-phase sounds, the 'position' is approximately proportional to the angle off centre of the reproduced sound image, and equals 1 for a sound on the left, 0 in the centre, -1 on the right, and ± 2 for a completely out-of-phase sound. It is found in practice that the precision of stereo location tends to be greatest when the distribution of reverberation energy across the stereo image is either flat or very slightly biased towards the centre.

Thus a basic rule for recording 4-channel stereo with good stereo location as well as richness in that the amount of reverberation coming from all directions around the listener must be evenly distributed. This is most easily ensured by using a fairly symmetric and not widely spaced microphone arrangement. Suitable microphone arrangements would consist of four cardioid or, preferably, moderately hypercardioid microphones pointing either in 4 directions at 90° to one another in a horizontal plane, or else pointing along the 4 axes of a regular tetrahedron, each microphone being at an angle of 109.5° from the other three.

A suitable alternative to the use of four horizontal cardioids is three coincident horizonal cardioids angled at 120° from one another, used with matrixing circuitry (see ref. 2). This arrangement enables any standard horizontal microphone directional characteristic to be obtained by matrixing, and permits a full basic 'four-channel' recording to be accommodated on only three tracks of the tape recorder. Experiments show that subjectively the best form of 'coincidence' is not to place one capsule above the other, but to place all three capsules in the same horizontal plane, as pictured in figure 4. With this arrangement, one capsule should point directly away from the centre of the orchestra. While this form of 'coincidence' prevents all horizontal sounds from being picked up without colouration, it does make the reverberation very much less coloured, which seems to be of subjective importance.

It is vital when using such matrixing techniques to use microphones whose polar response is accurately cardioid or hypercardioid, and which have a good polar phase characteristic (i.e. do not introduce phase shifts on sounds from some directions). This is also desirable for other coincident microphone techniques to ensure a uniform reverberation distribution around the listener.

continued on page 278

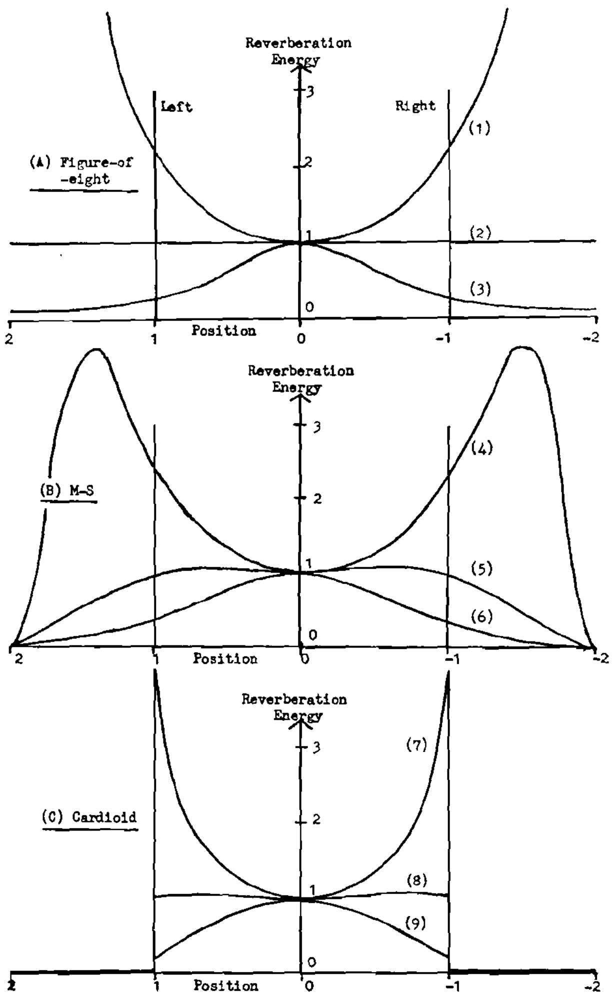

Fig. 3. Distributions of reverberation energy across the stereo image for various coincident microphone 2-channel techniques. (A) Crossed figure-of-eight microphones angled at (1) 120°, (2) 90° and (3) 60° apart. (B) M-S microphones with the S gain adjusted so that the angle between the null response on the left channel and the null response on the right channel is (4) 90°, (5) 141° and (6) 180°. (C) Crossed cardioid microphones angled (7) 180°, (8) 120° and (9) 90° apart.

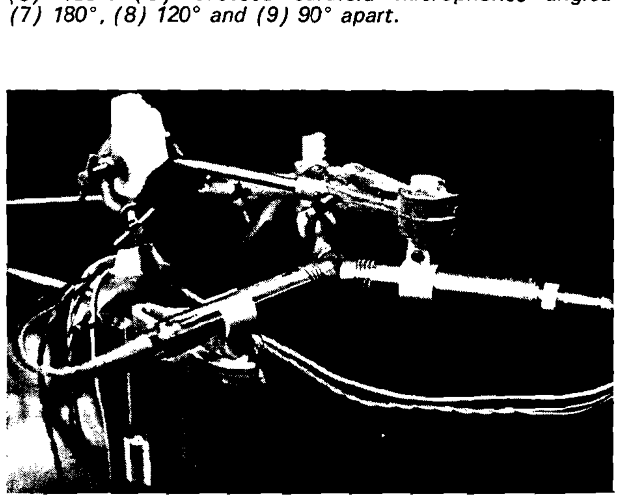

Fig. 4. 'Coincident' 120° angled cardioids.

Calrec cardioid capacitor microphones have been found to be particularly suitable in this respect, and all dynamic types are highly unsuitable.

Spaced microphone techniques present a number of problems with 4-channel stereo. If widely spaced microphones are placed at the corners of a square, this causes the reverberation to tend to appear to come only from the four loudspeakers on playback. Outward-pointing cardioids at the corners increase this 'isolation' of the four channels further, and no doubt this is why omnidirectional microphones are often used, especially for the rear channels. Unfortunately, the use of omni mikes causes a delayed direct orchestral sound to be reproduced from the rear speakers. If such a delay exceeds 15 to 30 millisecs, then a very distinct echo is heard. This time delay will be caused if the rear microphones are placed more than fifteen feet further back than the front ones, and is very audible on many commercial recordings. Such echoes appear to be much more audible when the sound surrounds the listener than in ordinary stereo, and often confuse the stereo image.

Perhaps a better method of preventing the reverberation from clustering at the loudspeakers with spaced microphone technique might be to use four additional centre-channel microphones, with one fed into each pair of adjacent channels. This has not been tried, but a suitable microphone placement might consist of eight outward-pointing cardioids at the vertices of an octagon. In any case, the recording engineer should pay adequate attention to the reverberation arriving from the sides of the listener, as many current recordings give the impression of listening in a tunnel, with everything coming from the front or back.

The following requirements can thus be stated for any system of 4-speaker surround sound: the recording must convey equal amounts of reverberation to all four speakers, the reverberation must be uniformly distributed all around the listener, and time delays of more than fifteen milliseconds (i.e. microphone spacings of more than fifteen feet) must be avoided. Also, it has been seen that essentially only two arrangements of four loudspeakers can give optimal richness, those of figure 1, although vertically squashed and mirror-image versions of the tetrahedral layout of fig. 1b will also do. These two loudspeaker layouts have the desirable feature that all speakers are at the same angle off the ears' axis, so that all speakers are masked to the same extent by the listener's head and hence contribute equally to the sound and stereo effect. Another important consideration is that the front pair of speakers are not angled too widely apart, so that no hole-in-the-middle occurs. In the proposed tetrahedral arrangement of figure 2c, the front speakers need to be 109.5° apart, which has been found to give a severe hole-in-the-middle. In the square layout of fig. 1a, the angle is 90°, which is tolerable but still on the high side. In the tetrahedral layout of fig. 1b, each speaker is at an angle of 54.7° off the ears' axis, so that as far as the ears' usual stereo location mechanisms are concerned, the effective angle between the front loudspeakers is only $180° - 2 \times 54.7° = 70.5°$, which is good.

It is clear that recordings made for the tetrahedral layout of fig. 1b are capable of reproducing sounds from all horizontal and vertical directions around the listener, and will also reproduce well on the square layout of fig. 1a without any height effect. It therefore seems sensible to make 4-channel recordings suitable for tetrahedral reproduction, as described in ref. 3, because of this compatability. Such tetrahedral recordings contain sufficient information for reduction down to any other system of 4-speaker reproduction that has been proposed.

Besides 'genuine' 4-channel systems, there have been a number of 2-channel 4-speaker surround sound systems proposed. Of course, exact reproduction of surround sound would require an infinite number of channels, and so such 2-channel systems are no less 'genuine' than four channels, as long as the system is linear, involves no 'synthetic' reverberation, and is capable of reproducing sounds from all directions.

One proposal due to Hafler and the present writer (see ref. 4) is to reproduce the left channel L of a 2-channel recording over the left side speaker of figure 2b, the right channel R over the right side speaker, the sum L+R over the front speaker, and the difference L−R over the rear speaker. While capable of reproducing sounds from all directions around the listener, this system has the two defects that the right side and rear speakers are out-of-phase, and the speaker arrangement is not capable of optimum richness.

David Hafler has also suggested that the front two speakers of the square layout of fig. 1a should reproduce ordinary stereo, and the rear speakers should reproduce L−R and R−L respectively. This system has a current vogue for use with both conventional stereo recordings and with special ones containing an out-of-phase ambience signal for the rear speakers. It has been claimed in several places that this system adds 'richness' to ordinary stereo because of 'random out-of-phase reverberation components between the microphones', but it probably works merely because the four speakers are each reproducing equal amounts of reverberation. When reproduced via this system, many commercial 2-channel recordings are found to have appreciable echo, indicating the use of widely spaced microphones. It is found that when 90°-angled crossed coincident figure-of-eight recordings are played, switching the rear speakers on causes a great increase in richness, but causes the sense of distance to disappear from the direct sound. This shows that richness and spacial perspective are quite distinct properties. The common use of the term 'depth' for what I have called richness is very misleading on this point. This system of reproduction gives very little richness with coincident crossed cardioid recordings, and is not much better with M−S, although here results will depend on the gain of the S microphone.

It may be shown that the sum of all the phase shifts between adjacent speakers when 2 channels are reproduced via matrixing circuitry via 4 speakers is 180°. This normally means that one pair of speakers must be out-of-phase one with respect to the other. This severe defec can be overcome by sharing out the phase shifts amonc the four speakers, either causing each speaker to be 45 out-of-phase with its neighbours, or more simply b making two pairs of speakers 90° out-of-phase. A suitabl circuit having a flat frequency response and giving tw outputs one approximately 90° out-of-phase with th other is given in fig. 5. It is a modified version of th 'SSB-Jr.' circuit known to radio hams, and has bee designed to be reasonably effective between 100 Hz an 20 kHz.

A version of the Hafler system using the desirab speaker layout of fig. 1a would reproduce the sign L−0.4R, L+0.4R, 0.4L+R, and −0.4L+R over t rear left, front left, front right and rear right speake respectively, with a +90° and −90° phase shift bei

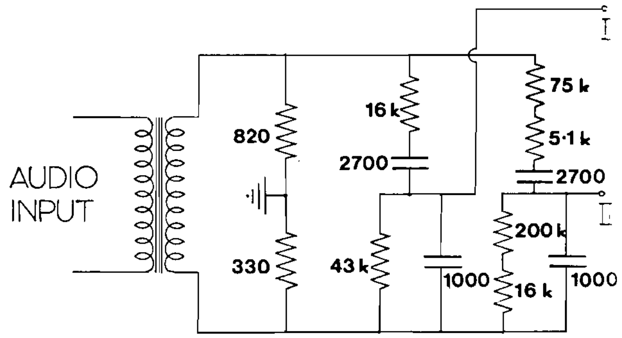

Fig. 5. Wideband audio 90° phase shift circuit. A components to be 2% tolerance or better. All resisto in ohms, all capacitors in picofarads. The two outpu (between I and earth, and II and earth) have an approx mate phase shift of 90° one with respect to the othe The load impedances should be as high as possible, an greater than 2 Megohm. In the extreme bass, the tw outputs are in phase with one another.

applied to one each of the rear speakers so as to render them in-phase and to increase the front-back separation. In order to avoid the overwide reproduction of conventional stereo that this gives, one might try compromises between this and the other fig. 1a — format Hafler system, and reproduce L−0.7R, L+0.2R, 0.2L+R, and −0.7 L+R with suitable rear-speaker 90° phase shifts. It is quite possible that experiments on these lines might reveal that 2-channel matrixing systems are perfectly adequate for all domestic surround-sound requirements in both the popular and classical fields. It is thus a matter of urgency that they are carefully tested before a 4-channel format is needlessly adopted. A general theoretical analysis indicates that Hafler-type systems have sufficient flexibility to cope with many possible future requirements, such as with-height tetrahedral reproduction.

It is unfortunate that it is impossible adequately to demonstrate many of the above points in a large auditorium, because of the large differences in the time of arrival of the sound from the various loudspeakers, and, to a lesser extent, because of the auditorium's own acoustics.

References

1. N. V. Franssen, Stereophony, Philips Technical Library, 1964.
2. M. A. Gerzon, Principles of quadraphonic recording Part I, Studio Sound, August, 1970.
3. M. A. Gerzon, Principles of quadraphonic recording Part II, Studio Sound, September, 1970.
4. M. A. Gerzon, Surround sound from 2-channel stereo, Hi-Fi News, August, 1970.

and D. Hafler, 2-channel quadraphony, Hi-Fi News, August, 1970.

Discussion

Granville Cooper: The difference between the two tetrahedral systems is that one is theoretical and the other is that of a practical man based on the fact that stereo, as we know it, has two speakers at the front so we just add two more speakers. Loudspeakers, being imperfect, draw attention to themselves because they 'ring'; if both in front are at the same height they are less obtrusive than if one is high and the other low. That is why my arrangement (your figure 2c) might be more successful in practice at least until loudspeakers are better!

Gerry Brook in America used a tetrahedral microphone system in which one microphone pointed vertically upward; have you analysed that system?

You said that American engineers indicated that the rear channels are 10 db less than those in front; I don't think this is so; what happens is that if, for instance, the subject is a choir, the peak power of the 'direct' speakers, being from the nearest microphone, is more than that of the reverberant sound heard from the rear speakers whose content is, of course, integrated. If all the microphones are somewhat distant, the peak powers will be similar.

Michael Gerzon: When I next demonstrate I will be using particularly uncoloured speakers for the reasons you have mentioned. There seems to be a conflict between the risk of colouration and the need to get the best quality of reproduction. Perhaps we will eventually use eight speakers one at each corner of a cubic pattern.

...y comment about increasing the area over which one gets a reasonable spatial effect?

Michael Gerzon: I only know that they should *not* be omnidirectional; if they were the sound images produced would be located at four precise points rather than spread around the listener as they should be. It seems unlikely that listening room reverberation will be a help in establishing the proper stereo image. On amplitude and time-difference grounds there is no optimum polar diagram which will give stereo stability but theory is not entirely adequate and there is need for further research. As regards listening area, there is still room for experiment. The tetrahedron I have proposed gives the biggest area in which the listener can remain actually within the speakers; this is a matter of solid geometry.

Peter West: I would like to draw attention to the law of diminishing returns; I very much doubt if, for many kinds of music, the increase in the quantity of apparatus both for recording and for replay is justified by the results that can be achieved.

Are you trying to recreate the acoustics of the place where the recorded performance took place or are you trying to produce pleasant listening?

Michael Gerzon: That is really a question for the recording engineer to answer. The most difficult thing is to recreate the original acoustic. One must explore the capability of the system before producing other environments by synthetic means. If one settles first for synthetic effect then the system may be capable of nothing else and will be unable to keep pace with changes in taste and musical or aesthetic requirements. The objective today should be to recreate the original acoustic.

R. C. Driscoll: Does a genuine four-channel system have the advantage over 'two plus two derived' channels?

Michael Gerzon: Four channels double the number of parameters available to recreate the original soundfield but an infinite number are still missing. Special microphone techniques, which I have described on paper, but not experimented with, should make it possible to recreate the soundfield both horizontally and vertically by means of derived matrices from a two-channel system. Is it worth it? If the results are fatiguing, as they have often been, the answer must be 'no'! But if, as happened when I was preparing this feature, a stereo record could appear compelling when fed to four channels then it is; the question is one of musical or aesthetic pleasure.

Peter West: More individual types of acoustic, like that of a church, will show up the effect most. Was tonight's demonstration influenced by the fact that the speakers were some distance from the walls of this large room whereas in a home they would be at the boundaries which would alter their radiation pattern?

Michael Gerzon: This would also add colouration and cause fatigue.

Sid O'Connell: You showed two diagrams which differed only in that in one the speakers were in the corners of the room and in the other they were at the centre of the sides. If one is trying to recreate a soundfield surely the two are equivalent provided the listener is in the middle of the field?

Michael Gerzon: The two arrangements are only equivalent if the number of speakers is much more than two or four.

Granville Cooper: Do you find, as I do, that the most effective material is that which is recorded in a studio with reverberation time long compared with that of the listening room?

Michael Gerzon: Possibly this is why this room, which has a reverberation time longer than a domestic situation, has given less impressive results than I have achieved elsewhere.

Mr. Powell: What experience have you had of matrixing multimicrophone stereo into your speakers; how well

very spectacular, especially with popular material; but sounds tend to splash around in a way disadvantageous to classical music.

R. C. Driscoll: We must thank Mr. Gerzon for his absorbing lecture and hope that he will tell us more about his investigations in due time.

Our thanks are due also to Peter Self for the tape equipment and to Messrs. Bell and Howell for the loudspeaker systems used tonight.