Stereophonic Image Sharpness

Results of experiments to discover how the width of the sound image is related to the image position; also how image position is related to interchannel level difference

*by H. D. Harwood**, B.Sc.

The main advantage of stereophonic over monophonic reproduction is the ability to hear sound images spread out in space as in real life, and the greater the number of different sources we can distinguish the better the system will satisfy us in this regard. This number is dependent on how sharp the images are and how evenly they are distributed across the acoustic space or stage, corresponding in television terms to focus of the beam and linearity of the scanning. In stereophony, using the usual coincident microphone sound pick-up, the spacing of the images is determined by the relative levels applied to the two loudspeakers; on the other hand the sharpness of the image depends on a number of factors, and in the end can only be determined by listening tests as the image is entirely a subjective phenomenon. If then we try to improve the stereophonic effect by changes in circuits, loudspeakers or acoustics of the listening room we must measure both the displacement and the sharpness of the image if we are to determine fully the results of the changes we have made. For example, it would be of little value to introduce a circuit which improved image sharpness over part of the stage width only to find that the linearity of image position had been distorted.

The possibility of improving the technical quality of broadcast programmes has been examined and the first requirement was to make a full and accurate assessment of the performance of the standard arrangement. To keep matters simple, images of a point source were used; ideally of course these should have no width at all, but in practice do have a finite size. An account is given here of experiments using male speech and designed to measure the relation between image position and interchannel level difference, to determine the variation of image width for differing image positions across the stage and to investigate the efficacy of a circuit, (described in one of the references) in reducing image width. The tests were with one exception carried out both in a listening room, thus simulating the average domestic acoustic environment, and in a free-field room, which has the advantage of providing not only one extreme acoustic condition but an environment that could be duplicated accurately elsewhere. The experiments on the image-width reducing circuit were carried out only in the free-field room.

**B.B.C. Research Department*

There are several references [1, 3, 4, 5, 6, 7] in the literature to the relationship between image position and interchannel level difference but, as will be shown, the conclusions vary considerably; very little quantitative information has been published on image width, and none on the way in which it varies for differing positions across the stage or with the acoustics of the listening room.

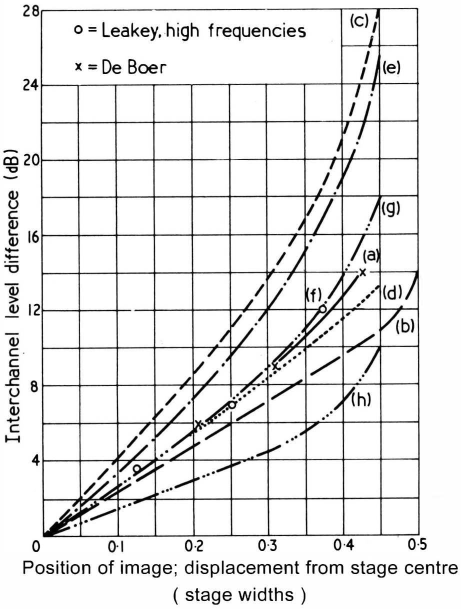
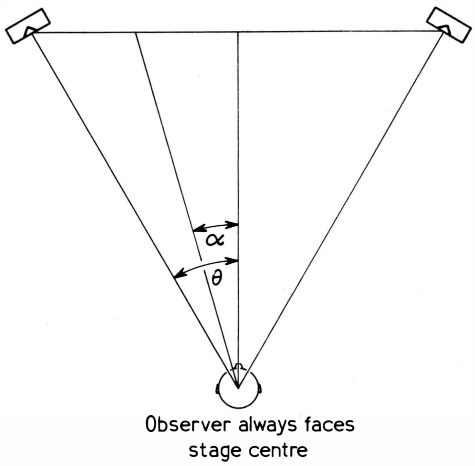


Fig. 1. Relation between image position and interchannel amplitude difference, by various authors: (a) De Boer; (b) Brittain and Leakey; (c) Clark, low frequencies; (d) Clark, high frequencies; (g) Wendt, $\frac{1}{3}$ octave band, f_c 316Hz; (h) Wendt, $\frac{1}{3}$ octave band 3160Hz.

Fig. 2. Relative positions of loudspeakers and observers.



The first reference in the literature to a measurement of the relation between image position and interchannel level difference is an article by de Boer[1]. No details are given of the acoustics of the listening room, the number of observers, nor of the programme employed, other than that it was from a gramophone record. His results are reproduced in curve (a) of Fig. 1. To account for the relationship shown by the curve he considered that diffraction around the head, resulting in difference in sound level at the two ears, was the main factor in providing directional information and that the difference in the time of arrival of the sound at the two ears played only a minor part. The interaural level difference due to diffraction varies considerably with frequency[2] and he weighted it according to the corresponding variation in energy density for speech, to arrive at the relationship between level difference and angle of incidence. It is, however, difficult to see how a single relation connecting image position with interchannel level difference can hold for all types of programme material if the hearing mechanism has to make a similar weighting for differing types of sound.

The next publication was by Brittain and Leakey[3] who employed excerpts, a few seconds long, of recorded speech with a frequency range up to 5kHz; the measurements were made in the open air on a site free of reflecting obstacles. The number of observers is not stated; their results are given in Fig. 1 curve (b).

In an article describing the "Stereosonic" system Clark *et al*[4] develop a theoretical expression determining the image position for low frequency sounds as a function of the interchannel level difference. For the experimental layout shown in Fig. 2, in which the subject faces stage centre, the theoretical relation is given as $\sin a = \{(L-R)/(L+R)\}\sin\theta$ and from data obtained with 4 observers this is claimed to be accurate up to about two thirds of the stage width. They state that the results of experiments carried out with variety of programme materials are in agreement with this relation for frequencies up to 700Hz, but that above this frequency the experimental displacements are greater than the predicted value and it is necessary for the value of sin a obtained from the above expression to be increased in the ratio of 1.4 to 1; to correct this they employ what they call a shuffle circuit. The paper by de Boer[1] is quoted in support of this requirement, but this support is not at all clear from an examination of the article in question. The theoretical relationships for both low and high frequencies are plotted as curves (c) and (d) in Fig. 1; it should be noted that the shuffle circuit is claimed to reduce the image width, presumably by an amount corresponding approximately to the difference between these two curves. In a subsequent article[5] one of the authors gives data to show that the acoustic environment has only a small effect on image position for a given interchannel level difference.

Leakey[6a, 6c] describes the limitations, even at low frequencies, of the sine law given above and derives another relation $\tan a = \{(L-R)/(L+R)\} \tan \theta$ which takes account of the variation of interaural time differences with small involuntary head movements; measurements he made in the open air with a number of observers varying from 4 to 12 and using a band of noise covering the frequency range 250 to 500Hz are in reasonable agreement with this law but he states that "to obtain a somewhat closer agreement with the practical results it is necessary to allow for the effect of signal attenuation around the head". Although he does not claim it, his formula implies that image position as a fraction of stage width is independent of the distance of the observer from the loudspeakers, and it can be seen from his results that this holds fairly well except at close range; on the other hand this cannot hold, except at small angles, for the formula given by Clark *et al* although they claim that it does. At high frequencies Leakey makes allowance for the shadowing effect of the head and assumes that it is the envelope function of the waveform reaching the ears which contains the directional information. The final expression is somewhat complex but agrees well with his own experimental results for a band of noise extending from 2 to 4kHz. Data calculated from these theoretical expressions for low and high frequency sounds are presented in Fig. 1 curves (e) and (f). To the extent that these curves differ from each other, Leakey also implies the need for some form of shuffle circuit although not to the degree indicated by Clark *et al.* In contrast to this, however, he also gives a curve[6b] showing the image displacement for "wide range speech", apparently 250Hz to 4000Hz bandwidth. This curve follows closely his curve for displacement of low frequency sounds only and does not show the deviation which might be expected if the high frequency components followed a different law.

Wendt[7] also carried out experiments in a room, whose characteristics were not stated, by a team of ten observers. He used various 1/3 octave bands of noise separated by half a decade, and found that for a given interchannel level difference the image displacement varied with the frequency of the test band; the results for the bands centred at 316 and 3160Hz respectively, i.e. similar to those of Leakey, are reproduced in Fig. 1 curves (g) and (h). From the slopes of these curves and the corresponding ones for other frequency bands, he obtained a weighting factor for middle and high frequencies somewhat resembling that obtained by Clark *et al.*

It will thus be seen that there is a considerable divergence not only in the results predicted by the theories mentioned but also in the experimental results.

The only data on image width appears to be that given by Clark *et al*[3] who state that image widths within about $\pm 2°$ were obtained for recorded music investigated in two separate bands: from the bass up to 700Hz and 600Hz upwards. This comment appears to cover differing image positions but the acoustic conditions are not stated.

Experimental Details

In the series of tests carried out by the author the observer faced stage centre in accordance with normal practice and with the requirements of the remarks given above. The arrangement of loudspeakers, scale divisions and position of subject is given in Fig. 3, which shows a plan of the listening room; the volume is approximately $85m^3$ and the average reverberation time is 0.3 second. The same arrangement was used in the free-field room, but in this case the reflection coefficient of the surfaces was less then 10% for frequencies down to about 80Hz. Twelve observers took part; they were all experienced in making subjective judgments.

To take account of any asymmetrical effects in the listening conditions, the loudspeaker system or the observer's reactions, tests were carried out with the image displaced both to the left and to the right of the central position. There is some evidence[8a] that the position of components at either of the extreme ends of the audio-frequency range is masked by that of components further within the band. To avoid this effect in the present series of measurements the test material employed was male speech, which is known to contain very little energy at either very low or very high frequencies; other advantages were the relatively constant level and continuous spectrum compared with other programmes. In the tests, recorded speech was played continously to the observer, until he had made his decision, at a maximum sound level reading of 74dB on an unweighted sound level meter; this level has been shown[8b] to be that preferred both by the team and by the general public. The observer was provided with a double attenuator with which he could change the relative levels of signal in the right- and left-hand channels, thus displacing the image position without changing the loudness.

The first set of experiments was designed to measure three things: the law connecting image position with interchannel level difference, the width of the image at various positions across the stage, and the minimum perceptible image shift. The observer was therefore asked to adjust the attenuator for the following conditions:

(a) Centre of image coinciding with each of the tapes 0 to 5 (Fig. 3)

(b) Left-hand edge of image coinciding with each of the tapes 0 to 5

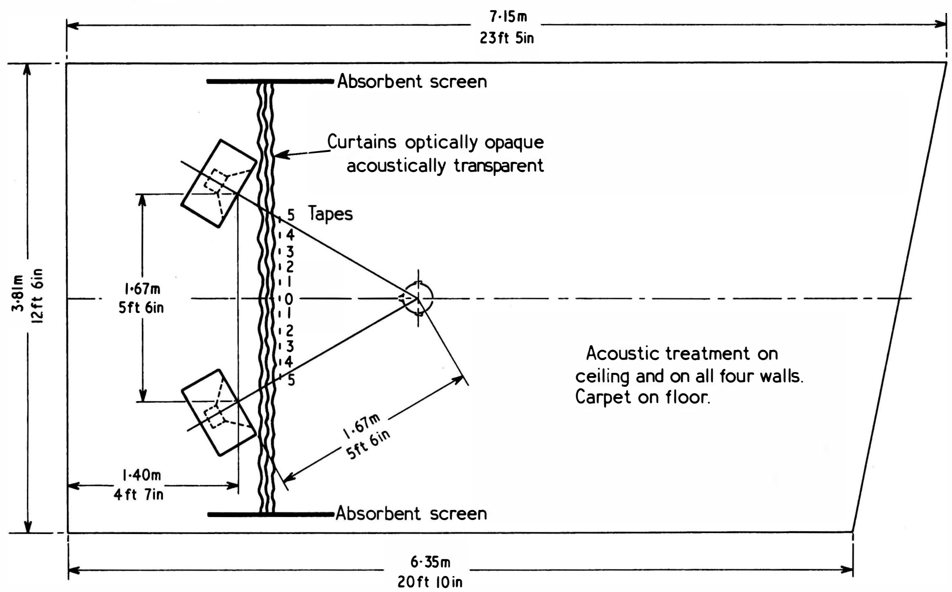
(c) Right-hand edge of image coinciding with each of the tapes 0 to 5

(d) Minimum perceptible shifts of image from centre to left-hand side and to right-hand side

The order in which tests were made was random.

The second set of measurements was designed to check the claim made by Clark, Dutton and Vanderlyn[4,5] mentioned earlier that the sharpness of the image can be improved by electrical means. According to those authors the part of the spectrum above 700Hz requires a smaller interchannel level difference for a given image displacement than the portion below this frequency. They designed the circuit shown in Fig. 4(a) having the amplitude frequency characteristic shown in Fig. 5, to be inserted in the difference channel of a sum-and-difference network, as shown in Fig. 6. The circuits

Fig. 3. Plan of listening room used in experiments.



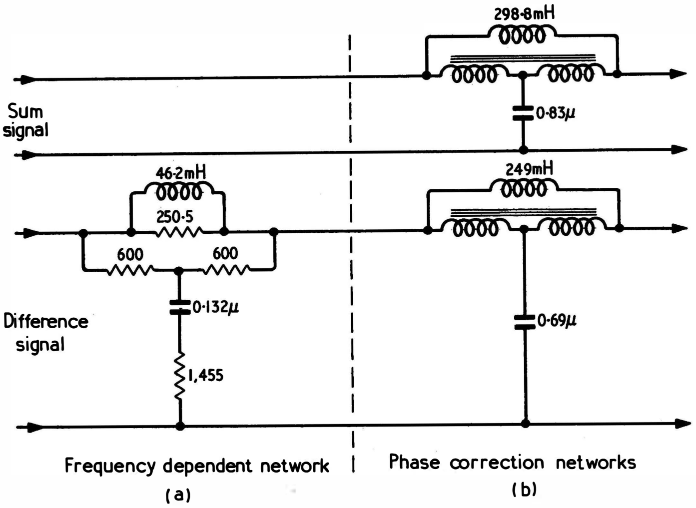


Fig. 4. Shuffle circuit intended for improving sharpness of image.

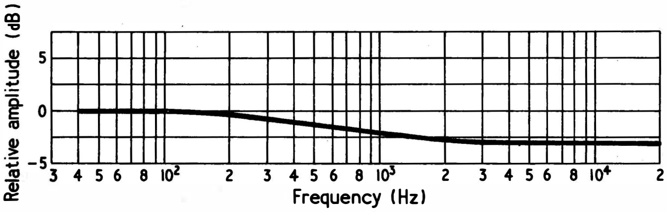


Fig. 5. Amplitude/frequency characteristic of shuffle circuit.

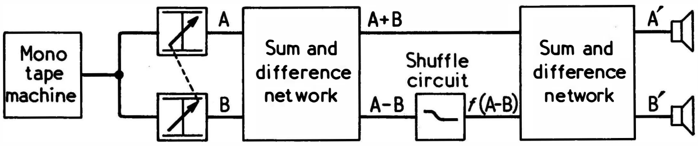


Fig. 6. Showing how shuffle circuit is inserted in sum-and-difference network of test apparatus.

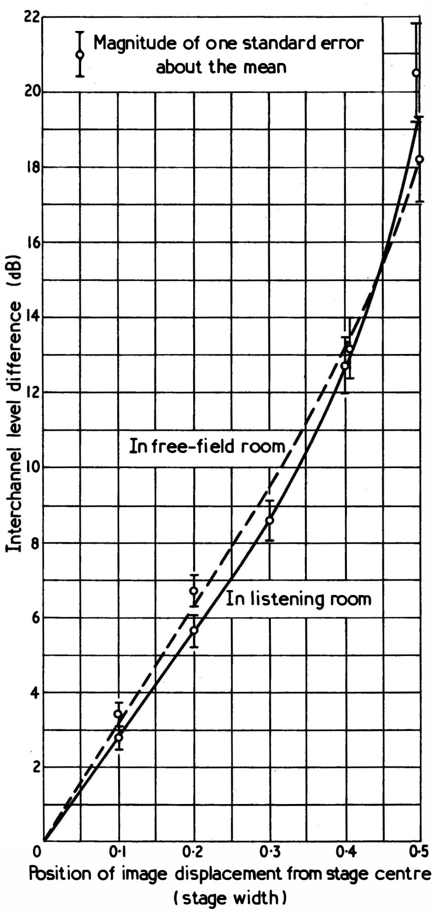
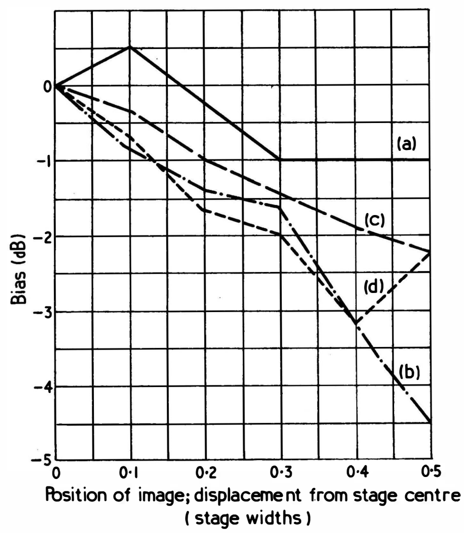


Fig. 7. Relation between image position and interchannel level difference, in free-field room and in listening room.

Fig. 8. Variation of bias with image position: (a) in free-field room, centre of image, average of team; (b) in free-field room, centre of image, average for individuals irrespective of sign; (c) in listening room, centre of image, average of team; (d) in listening room, centre of image, average for individuals irrespective of sign.



marked (b) in Fig. 4 are to compensate for the unwanted phase delay introduced by circuit (a); losses in the two circuits are equalized by means of attenuators not shown.

The efficacy of this device in reducing image width was tested in the free-field room by repeating observations (b) and (c) with the new device in circuit. For a central image there is no difference signal and therefore the added circuit has no effect; when the image is at one extreme of the stage the sound is radiated effectively by one loudspeaker only and the added circuit does not appreciably change the image width or position although it does change the sound quality. Observations were therefore confined to the case of an image at tape 3, where calculations indicated that the effect on image width should be quite noticeable.

The results of the first set of experiments are given in Fig. 7. The curves show the image position as a function of the interchannel level difference for the tests in the listening room and free-field room respectively. The points plotted are the median values for the team and the standard error for each value is also shown. It will be seen that the results in the two rooms are very similar and that the relationship is substantially a linear one except near tape number 5. This position will be seen from Fig. 3 to represent the extreme edge of the stage and in order that the image should be displaced to this extent, substantially all the sound must come from one loudspeaker. A linear relation therefore cannot be expected; in fact, owing to the finite image width smaller displacements must also be affected. Another factor involved is that most observers are found to have a bias in their observations so that the number of decibels required to shift the image to a tape on the left is different from that required for the corresponding tape on the right. Fig. 8 shows this bias averaged for the team (curves (a) and (c)) and the average of the absolute values (irrespective of sign) for the individual observers as a function of the image position both in the listening room and in the free-field room. Some persons found it impossible to displace the image centre to tape 5 on one side of the stage although on the other side they could displace it even beyond this position. The values shown in Fig. 7 for tape 5 are thus affected by the fact that subjects on one side of the mean have not been able to give a reading for the test, this applied particularly in the listening room, and the size of the standard error is also increased. On the other hand, the extreme value shown is in very good agreement with our figure of 19 dB given elsewhere[8a] for the interchannel difference required to give a minimum perceptible displacement of the image from tape No. 5.

Image Width. Figs. 9(a) and 9(b) show the relationship between width of image and image position for tests in the listening room and free-field room respectively. For tapes 0 to 4 these results were obtained from the team averages for the positions of the inner and outer edges of the image. For tape 5 it was not generally possible to displace the inner edge of the image to the tape, as this involved displacing the centre to a position

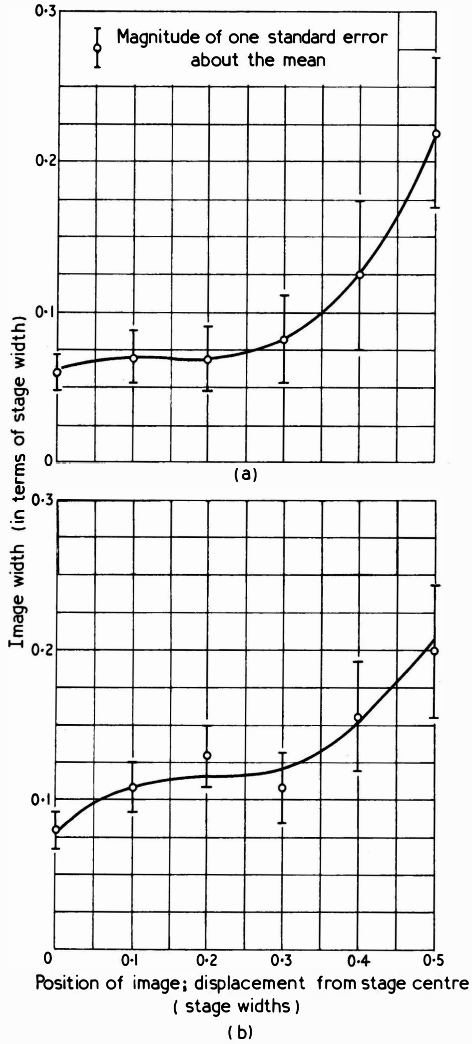
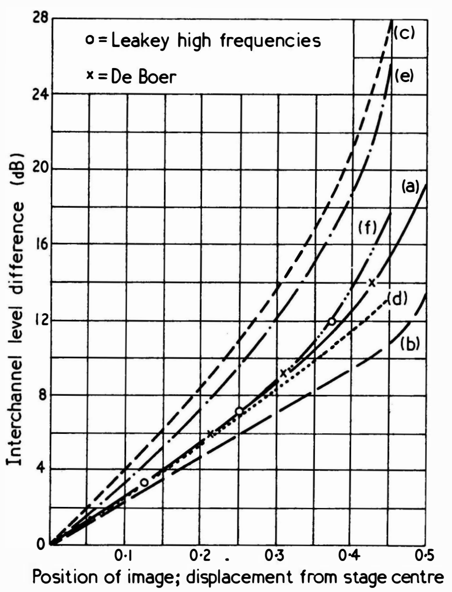


Fig. 9. Variation of image width with image position: (a) in listening room; (b) in free-field room.

Fig. 10. Relation between image position and interchannel amplitude difference: comparison of present results, shown by curve (a), with previous work. (b) Brittain and Leakey; (c) Clark, low frequencies; (d) Clark, hiagh frequencies; (e) Leakey, low frequencies; (f) Wendt, $\frac{1}{3}$ *octave band* f_c *316Hz.*



beyond the corresponding loudspeaker axis; the image width could not therefore be obtained directly. However, an examination of the results for the other positions showed that the position assigned by the team to the centre of the image was in fact very closely halfway between the two edges and it was assumed that this would probably hold at tape 5 also; to obtain the point plotted for tape 5 the difference between centre and outer edge of the image was therefore doubled.

It will be seen that up to 0.4 of a stage width the image is always narrower in the listening room than in the free-field room with a fairly constant ratio of 1 to 1.4; moreover, the standard errors are similar in the two rooms for the same image positions so this difference cannot be accounted for by the test being more difficult in one room than in the other.

At tape 5 the values of image width, having been estimated, are somewhat less certain, but the value is the same in both rooms and, as a matter of interest, amounts to about two-thirds of the loudspeaker width. For this image position the sound is radiated almost entirely from one loudspeaker and so it might be expected that as faults due to imperfect matching of the two loudspeakers are at a minimum the image would be sharpest. What is surprising therefore is that the image width at stage centre is only about one-third of that at tape 5; it is possible that if observers had always faced the image rather than the stage centre, this result would have been different for the off-centre positions.

The median value of the minimum perceptible shift for a central image in the listening room is 0.03 of a stage width, with a standard error of 0.007; the corresponding figures for the free-field room are 0.02 and 0.003 of a stage width. It will be seen therefore that the limit determined by the powers of discrimination of the ear has not been reached and it would seem that given suitable conditions the image width even for a central image could still be appreciably reduced.

The results were further analysed to see if there was any correlation between the width of image determined by a particular subject and the interchannel level difference necessary to displace the image position by a given amount. It is clear that for image positions near the stage centre such a connection must exist; for example, it is almost certain that a person who hears an image extending almost to tape 1 will require less interchannel level difference to displace the image to that tape than will someone for whom the image is narrower. The analysis showed this expected correlation for tape 2; for tapes 3 and 4, however, the correlation coefficients were too low to be significant. These low values are in part due to the smallness of the team, 12 persons, but it seems probable that even if a larger sample were taken the correlation would not be very marked.

Effect of Shuffle Circuit on Image Width. In the second set of experiments the median value for the image width at tape 3 in the free-field room when the shuffle circuit was used was 0.16 stage width, with a standard error of 0.025; the value obtained without this circuit and shown in Fig. 9(b) was 0.11 stage width, the standard error again being 0.025. It is seen that the effect of the circuit, far from reducing the image width, has been to cause an increase in width which is just significant. At this particular tape position $L/R = 9$dB and the 3dB step (Fig. 5) in the ratio of sum signal to difference signal $((L + R)/(L - R))$ introduced by thecircuit at high frequencies changes the ratio of the signals in the left and right channels by 3 dB also; from Fig. 7 it can be seen that an inter-channel difference of this amount will displace the image by 0.09 stage width. As, however, the image width has been increased by only 0.05 (from 0.11 to 0.16) stage width it appears that the frequency components causing the increased width could previously have been well inside the image area.

Comparison with Previous Work

The results given in the full-line curve of Fig. 7 for the measurements in the listening room of the relation between image position and interchannel level difference have been replotted as curve (a) in Fig. 10, together with some of those from Fig. 1. it will be seen that they agree extremely well with those of de Boer (shown by crosses) and with those obtained at high frequencies by Leakey (shown by circles). On the other hand, the values obtained by both Clark *et al* and Leakey at low frequencies are well removed from our results.

The difference between their high frequency and low frequency curves implies that in the absence of a shuffle circuit the width of image should vary in a corresponding manner and it is of interest to calculate this variation. The difference between the curves is plotted for both authors in Fig. 11(a) and (c) and the corresponding displacements of the edge of the image obtained with the aid of Fig. 10 is shown in Figs 11(b) and (d). If it is assumed that the image area originally contained all frequencies uniformly distributed, then the width should be increased by an amount corresponding to these displacements. Fig. 12 shows the result obtained when the width for a central image, which the shuffle circuit cannot affect, is superimposed on Figs. 11(b) and 11(d). It will be seen that the resulting curves do not bear a very close resemblance to that of the actual image width obtained and shown in Fig. 9(a) and in Fig. 12 and thus throw futher doubt on the validity of a shuffle circuit with these constants. The theories of Clark *et al* and Leakey both assume that the observer always faces stage centre, but it is not explicitly stated whether this condition did apply in their supporting experiments. They show that a change of head angle does affect the image position to some extent but no information is given as to the effect on image width.

It was shown earlier that the image width was less in the listening room than in the free-field room for a spectrum containing little energy below 100Hz. This is in contrast to the results of other tests [10] in which most of the energy was below 100Hz. In the latter

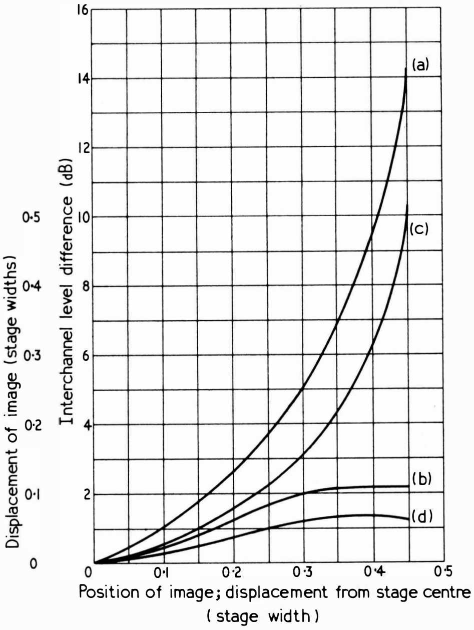
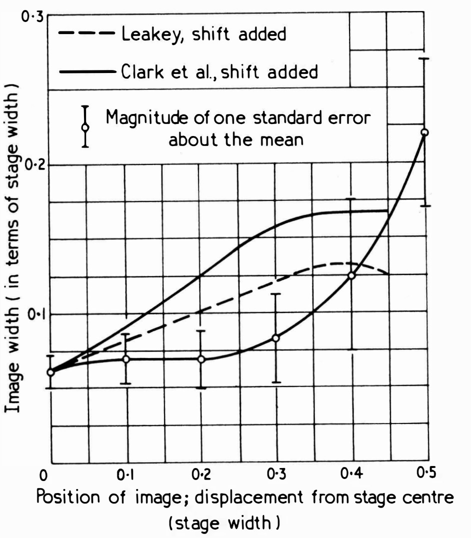


Fig. 11. Variation
image position: (a) interchannel level difference produced by shuffle circuit (Clark); (b) shift of image which would be produced at low frequencies by (a); (c) interchannel level difference required by Leakey; (d) shift of image which would be produced at low frequencies by (c).

Fig. 12. Measured and calculated variation image width with image position.



case the programme consisted of plucked double bass, drum and organ and the image at stage centre was wider in the listening room, although least so for the organ which in turn had more sound energy above 100Hz than did the other two programmes. It appears therefore as though at the bass end of the spectrum the effect of the acoustics of the listening conditions on the image width at stage centre varies somewhat with frequency.

Future Work

It has been shown that there is some agreement in the literature on the need for a form of shuffle circuit but the transition frequency and amount of the step appear to be based on rather scanty data. It should therefore be profitable to repeat Leakey's work with octave bands of noise over a wider frequency range in order to determine the optimum frequency characteristic and height of the step, and then to check whether in fact this relationship does appear to hold for a wide-band signal. It should be noted, in this connection, that neither of the theories proposed by Leakey and Clark and given earlier in this article takes account of the precedence effect whereby one sound reduces the apparent level of other sounds immediately following it. Leakey also admits that signals covering only a narrow frequency band give results differing from those having a wider band, and it is not clear whether signals extending over an octave will give the same results as programme.

In these tests it has always been assumed that the observer should be in a central position facing straight ahead. An examination of the data on the variation of diffraction around the head[2] and of the interaural time difference[1] with the angle of the incident sound wave indicates that the directional data supplied to the ears does not vary appreciably for angles of incidence up to $\pm 40°$ in the horizontal plane. The natural reaction, however, when attention is focused on a particular sound is to face the direction from which it appears to come, and the possibility should be examined that if such a movement is prevented this may have an effect on the image width, and thus afford some explanation of the results given in Fig. 9. Such an effect would in some ways be similar to the corresponding optical case in which visual acuity is higher for an image in the centre of the retina than for one at the periphery. For the tests described, image width is already smaller in the listening room than in the free-field room and this fact suggests that no improvement over the range covering speech frequencies would be obtained by increasing the ratio of direct to indirect sound. The use of directional loud-speakers might help, however, to reduce image width for listeners in off-centre listening positions, and possibly at the lowest frequencies for those in a central position. In this connection the effect of the width of the loudspeaker cabinet on image widths should also be examined.

Conclusions

The relation between stereophonic image displacement and the interchannel level difference has been obtained both in a typical listening room and in a large free-field room for observers facing stage centre. The results in the two rooms are very similar and show a substantially linear relationship over most of the stage width; the bias for the team and the average bias for individuals in these experiments have also been determined for the differing image positions chosen.

The image width shows an unexpected variation with position across the stage and is much smaller at the stage centre than at the edges. The variations between individuals were similar in the listening and free-field rooms, but the absolute widths were greater in the latter; measurements of the minimum perceptible shift at the centre position show that even for this position where the image is narrowest a reduction in width would still be observable.

The circuit claimed by Clark *et al* to reduce image width has been tested subjectively for one image position which it was thought would show such an effect most clearly. The measurement gave the unexpected result that the effect of the circuit has been to increase the image width by an amount which is statistically just significant.

Acknowledgements. The author wishes to express his thanks to the Director of Engineering of the British Broadcasting Corporation for permission to publish this article.

References

1. "Stereophonic Sound Reproduction" by K. De Boer. *Philips Tech. Rev.*, 1940, *5*, 4, p. 107.
2. "On minimal audible sound fields" by L. J. Sivian, and S. White. *J. Acoust. Soc. A.*, 1933, *4*, 4, pp. 288-321.
3. "Two channel stereophonic sound systems" by F. H. Brittain and D. M. Leakey. *Wireless World*, 1956, *62*, 5, pp. 206-210.
4. "The stereosonic recording and reproducing system" by H. A. M. Clark, G. F. Dutton and P. B. Vanderlyn. *Proc. Instn. Elect. Engrs.*, 1957, 104B, *17*, pp. 417-432.
5. "The assessment of two channel stereophonic reproduction" by G. F. Dutton. *J. Audio Engng. Soc.*, 1962, *10*, 2, pp. 98-105.
6a. "Some measurements on the effects of interchannel intensity and time difference in two channel sound systems" by D. M. Leakey. *J. Acoust. Soc. Am.*, 1959, *31*, 7, pp. 977-987.
6b. cf. Figs. 2 and 8(b) loc. cit.
6c. "Further thoughts on stereophonic sound systems" by D. M. Leakey. *Wireless World*, 1960, *66*, 4, pp. 154-160, 5, pp. 238-240.
7. "Versuche zür Ortung von Intensitats — Stereophonie" by K. Wendt. *Frequencz*, 1960, *14*, 1, pp. 11-14.
8. "Stereophony: the effects of cross-talk between left and right channels" by H. D. Harwood and D. E. L. Shorter. B.B.C. Engineering Monograph No. 52, 1964. *www.keith-snook.info*
9. "The effect on stereophonic reproduction of frequency-dependent crosstalk between left and right channels" by H. D. Harwood. Fourth International Congress on Acoustics, Copenhagen, 1963, Paper N35.
10. "Stereophony: the effect of interchannel differences in the phase/frequency and amplitude/frequency characteristics" by H. D. Harwood, W. I. Manson and D. E. L. Shorter. B.B.C. Engineering Monograph No. 56, 1964.

Circuit Ideas

As announced in last month's issue, a regular feature on original ideas in circuitry which have been found practical is to be started in *Wireless World*. Presented in the form of short notes, the items will be essentially functional "bricks" which somebody has found useful at some time. Performance, originality of realization and economy of components will be the most important criteria. Readers are invited to contribute to this series: the more ideas we get the better will be our selection.