Principles and Problems of Stereophonic Transmission

by Gerhart Boré

GEORG NEUMANN

Laboratorium für Elektroakustik G.m.b.H.

Berlin SW 61, Segitzdamm 2 · Telefon 61 48 92

In the quest for more and more faithful reproduction by electro-acoustic means, after all the possibilities afforded by single-channel transmission have been exhausted, substantial improvements can only be achieved by multi-channel stereophonic transmission.

A few of the factors that play a part in this type of sound transmission have been known for about 30 years. Some of these, however, can only be understood and correctly reformulated in the light of recent discoveries in the physiology of hearing, and a number of questions still remain to be answered. In considering the practical means employed for stereophonic transmission, it seems that development hitherto has not led to a single optimum, but to a number of solutions of apparently equal value.

In this article, after a summary of the most important principles, a survey will be attempted of proven methods of transmission, and a few examples of their application will be given.

I. Principles

Spatial Hearing

Our ear enables us not only to appreciate sounds according to their intensity and duration, but to pin-point their origin in terms of direction and distance. To determine direction, the difference in times of arrival of the initial transient components of a sound at the two ears, is employed, and also, to some extent, the differences in intensity and tone-colour between right and left ears. Distance can also be determined with only one ear, and this relies on the ability of our hearing to estimate the curvature of the wavefront, which for point sources decreases with distance. As a means of measuring, the ear makes use of the phase difference between pressure and velocity, particularly in the low-frequency components and in low-frequency transients, which amounts to 90° at the point of origin and 0° for plane wave propagation. In the diffuse sound field of enclosed spaces, it is also possible to estimate the distance of the sound source by means of the intensity difference between the direct sound and its subsequent reflections from the boundary surfaces.

Possibilities of Single-Channel Transmission.

With the techniques used up to the present, employing only one transmission channel, the sense of distance can only be conveyed in this way, by the reverberant components. The art of the Sound engineer to-day lies largely in using this possibility to transmit a sense of space. By the skilful placing and mixing of so-called soloists' microphones, the direct sound (and also, if necessary, the loudness) of the various instrument groups, soloists, or speakers, can be increased and the performer thus brought "forward" to the listener within the total body of sound. The same effect can be produced by means of directional microphones, with the advantage that it is not necessary to approach so close to the sound source. Two special types of NEUMANN condenser microphones (M 49 and SM 2) even allow continuous remote adjustment of their directional characteristics, and therefore of the direct sound component, from the control room.

Transmission of directional impressions by means of a single channel is not possible. It is as if one heard the programme through a hole or tube. By using several loudspeakers or a spherical radiator it is indeed possible to widen the mouth of this tube, but not to increase its diameter. (Similarly, milk and coffee, once mixed, cannot be separated by using a coffee-pot with several spouts.) Only for spatially bounded sound-bodies, and for soloists, under definite circumstances the impression can be created that the sound source is present in the reproducing studio. Herein lies the advantage of the arrangements known as "3-D Technique".

"Head-referenced" Stereophony.

If one is prepared to use headphones for reproduction, a very good transmission of the directional component results when each earpiece is connected to a separate microphone and the microphones are mounted in place of ears in a "dummy head", perhaps of wood. One has only to keep one's head still otherwise the whole transmitting studio appears to rotate with it.

"Space-referenced" Stereophony and "Haas Effect".

If loudspeakers are used for the reproduction, each being connected by a separate channel to the corresponding microphone, it appears that a multiplicity of channels is not necessarily required, but that the increase from one to two channels already produces a large subjective improvement. The addition of a third or more channels is then only justified if moving objects are to be dealt with, and if conformity with a simultaneously presented optical image (sound-film) is required. In the following, two-channel transmission will be assumed throughout.

A general difficulty inherent in the use of loudspeakers is that the sound from each transmission channel cannot be confined to one ear. Each ear hears both loudspeakers. For a better understanding of this, a comparison with the preceding case in which a sound source is heard in a confined space is instructive. In spite of the great number of reflections, we generally succeed very well in localizing the source because, out of all the sounds received, we have learnt to use only the first components, arriving directly from the source, when estimating direction. The other components, reflected from different directions, cannot influence this estimate.

If, instead of a reflection, we hear the sound from a second loudspeaker, we cannot receive any directional impression from it if its sound is appreciably delayed with respect to that from the first loudspeaker (so-called "Haas effect"). The first loudspeaker only will be located and no stereophonic effect can be achieved.

Fortunately, this "precedence limit" for direction estimation is not insurmountable. It ceases to act after about 3 ms. Only up to this limit two components from different directions can be fused into a single impression of direction, which then lies between the two. However, the relationship between the difference in time of arrival and angle of localization is not linear; there is a crowding of the sound-events towards the loudspeakers, but this is not usually objectionable. In classical two-channel transmission using loudspeakers the caution must therefore be given:

> A relative time-delay of more than 3 ms or a path Difference of more than 1 yard between the two channels causes the sound to appear to come from one of the loudspeakers only.

In the transmitting studio this can best be taken into account by a suitable microphone arrangement. In the reproducing studio all the listeners should be seated equidistant from both loudspeakers, i. e. they should be placed on the axis of symmetry. Otherwise, the differences in time of arrival from the two loudspeakers has this disadvantage — that all sound sources appear to be drawn towards the nearer loudspeaker, and, for the side seats, will seem to come from this loudspeaker only. A smaller loudspeaker spacing increases the area in the listening room for which approximately correct localization is obtained, at the same time, the overall apparent width of the sound picture is reduced. A compromise must be made, which must take into account the shape of the reproducing room.

An effective method of increasing the useful area in the reproducing room consists in slowing up the leading edges of the direct sound: this somehow slightly reduces the sharpness of location. The exact connection is not yet understood. One must clearly seek to arrange that, prior to the full setting-in of the precedence effects mentioned above, a number of additional repetitions or reflections having only small delays arrive at the listener from each loudspeaker within 3 ms. Apart from this, the radiation characteristics should be constant at all frequencies.

The arrangement of hemi-spherically radiating combinations with carefully matched frequency and phase responses, each in one corner of the room, has proved satisfactory. The placing of the low-frequency radiators is less critical, since it is difficult to localize at frequencies below 300 c/s.

Influence of Intensity Differences.

It was not a great step to add intensity differences to the time-differences, or even to use intensity differences alone in the two channels, in order to convey directional impressions.

The result of such measures is not easy to predict, since the case of identical sounds arriving at an observer simultaneously does not occur in normal listening. On the contrary, in the case of speech, in estimating directions our hearing has become accustomed to ignore reflections which may be up to 5 dB louder than the original sound, and delayed up to 35 ms.

With the assistance of electro-acoustics it is possible, using two widely-spaced loudspeakers, to radiate identical in-phase signals, and then, solely by altering their relative intensities, to produce impressions of varying direction. It appears that an approximately linear relationship exists between the intensity ratio and the apparent angle of arrival, and that the localization is not quite so sharp and well-defined as in the classical method*. A certain amount of blurring is desirable. This must be tolerated in order that the useful area in the reproducing studio shall not be too small.

A comparison between the two stereophonic systems under the same experimental conditions: the previously described "classical" and the so-called "intensity-stereophony", for the purpose of evaluation, has not yet been possible. The only definite conclusion is that both methods can show a considerable enhancement of quality as compared with single-channel working. The listener has the option of concentrating on different directions within the total sound-picture, the sound acquires considerable transparency, and the residual noise and hum together with transmitted background noises appear less noticeable, and even a very high speaker level, intolerable on a single-channel transmission, is not considered annoying.

II. Practical Stereophonic Sound Transmission

Development has not yet gone far enough to allow one to judge which of the different systems will predominate. The examples that follow can only give a superficial survey of some of the methods that have already given good results.

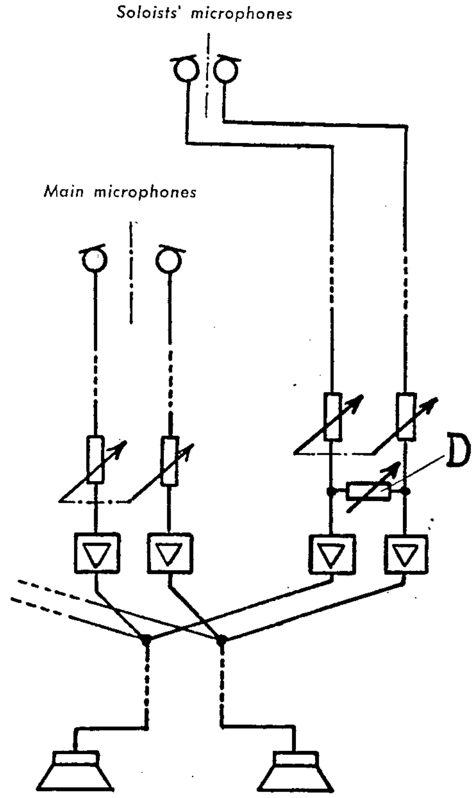

Fig. 1. Block Schematic of a Normal Two-Channel Transmission System.

Normal Classical Two-channel Transmission.

In the "classical" method, pick-up is carried out by two microphones accurately matched as to frequency response and polar characteristic. In practice, adequate matching is only attainable in the case of high-grade condenser microphones, free from subsidiary resonances within the transmission band. Cardioid microphones, which have proved particularly satisfactory in stereophonic sound film systems, are prefered. Thanks to their single-sided directional pattern, the direct sound, so important for localization, can be picked out of the general sound picture from a greater distance. However, microphones with other polar characteristics can be used, according to the nature of the room and method of reproduction. The microphones are used side by side at a spacing of 30—120 cm depending on the size of the working arc desired. If, as is not always necessary in stereophonic pick-up, additional soloists' microphones are employed, these should consist of p a i r s of microphones placed closer together. A variable attenuator D (Fig. 1), in the cross-connection between the soloists' microphones permits a diminution of the "base" along which the soloist appears to move to and fro, so that a small sideway movement of the soloist does not cause an apparent leap from one loudspeaker to the other. The use of a single microphone, the output of which, after amplification by separate amplifiers is arbitrarily mixed into the two channels by hand, is subjectively much less pleasing.

In order to enhance the left-right impression by additional intensity and tone-colour differences, a small baffle can be placed between the microphones, or they can be mounted on opposite sides of a sphere, a so-called "dummy head" of 10—30 cm diameter. For laterally displaced sound

* According to recent English experiments, an accurately linear relationship exists for frequencies up to 700 c/s; for higher frequencies, on the other hand, a displacement of the virtual source towards the speakers is found.

sources this causes a shadowing of the further microphone, leading to a more uniform representation of the sound-scene on the reproducing side; the further a sound source is displaced laterally, the less the incremental difference of path length to the microphones for further displacement. In this region the intensity differences carry the effect, and the above-mentioned apparent crowding of

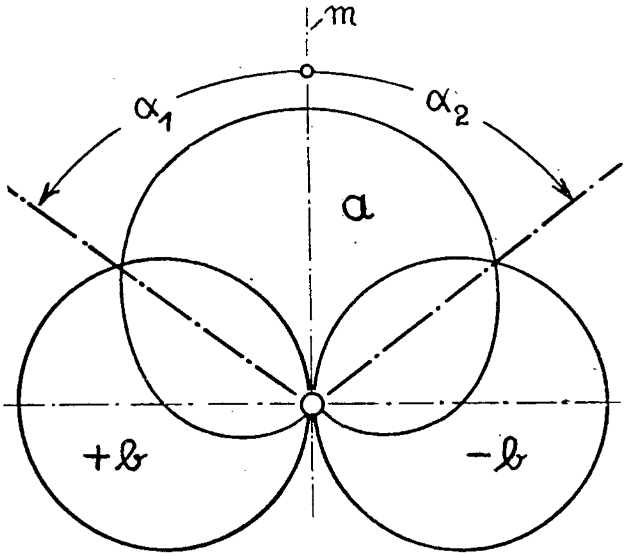

Fig. 2a. Polar Diagram of a Microphone Combination for MS-Stereophony.

sound sources towards the loudspeakers is avoided. In addition, the diffraction of the sound waves round the dummy head or baffle causes an attenuation of the higher frequencies on the far side.

"Intensity" Stereophony.

If the localization of sound sources is to result solely from intensity differences between the right and left loudspeakers, both microphones must necessarily be situated at the same point, in order to avoid differences of transmission time. They can thus be arranged very closely one above the other without disadvantage. Microphones with very well-defined polar characteristics must be used. They are rotated in different directions, so that each favours one half of the sound stage. This type of intensity stereophony has no practical value, since the discovery by Lauridsen of "MS-stereophony" achieves a substantially more elegant solution at about the same cost.

MS-stereophony means mid-side stereophony. One microphone with cardioid characteristic handles the whole sound picture, exactly like the principal microphone in a single-channel pick-up. A second microphone having a cosine characteristic is placed closely above or below the first and turned so that its null plane contains the principal axis of reception of the cardioid microphone (Fig. 2a). If the two microphone outputs a and b are interconnected so that the sum a + b and the difference a — b are formed, as shown diagrammatically in Fig. 2b, two channels result, in each of which one half of the pick-up area is preferentially received. The arrangement relies on the fact that the two principal axes of a pressure gradient microphone correspond to voltages of opposite polarity. The combining can, for example, be done in the manner shown, by using differential transformers.

If we assume that the instantaneous value of a sound from the left produces a positive voltage "b" in the cosine microphone, sound sources on the central axis "m" will give rise to the voltage "a" in the cardioid microphone only, thus producing a central impression. Sources making an angle α_1 with the central axis, give rise to a voltage a + b in the left loudspeaker and a — b in the right loudspeaker. With a = b only the left loudspeaker is energised, the source appears to the listener to lie in that direction. Similarly, sound sources at an angle α_2 appear to come from the right loudspeaker.

Smaller angles in the transmitting studio correspond to apparent directions between the loudspeakers in the reproducing room. The size of the angle $\alpha_1 + \alpha_2$ can be varied, within limits, by changes in relative gain of the microphone channels.

Sources which lie outside the included angle $\alpha_1 + \alpha_2$, will be localised more centrally. In this region the output from the cosine microphone predominates causing the loudspeakers to be driven in opposite phase and resulting in indefinite impressions of direction for a centrally situated listener. In this region no sound source should be set up. In this case too, as experience shows, if the quality is to be pleasant, soloists' microphones are eventually necessary, consisting of double microphones arranged as in Fig. 2c.

A significant advantage of MS-stereophony lies in the fact that one channel, namely the mid-channel, carries a satisfactory single-channel transmission. The decision as to whether a two-channel transmission shall be reproduced monaurally or stereophonically may thus be decided at the reproducing end. This advantage is not obtained with classical stereophony; with this, two channels are always necessary for the achievement of good reproduction. Mixing of the two channels is not successful, firstly be-

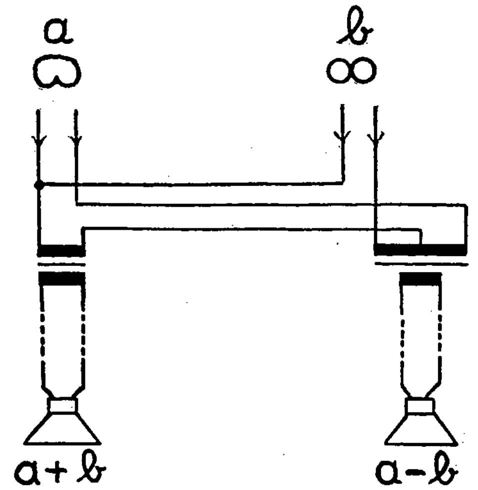

Fig. 2b. Electrical Sum and Difference Formation.

cause the combination of two microphones into a "doublet" gives an undesirable highly frequency-dependent polar characteristic, and secondly the microphones will not have been placed suitably for a single-channel pick-up.

If no value is placed on "compatibility" of intensity stereophony with single channel transmission, a pair of crossed gradient microphones with accurate cosine characteristics may be used for transmission with good results (Fig. 3). After double electrical addition and subtraction (according to Fig. 2b) similarly proportioned voltages result as for MS-stereophony. The microphone output voltages permit a good transmission of directional impressions only for sectors 90° wide to the front and back, and care is, there-

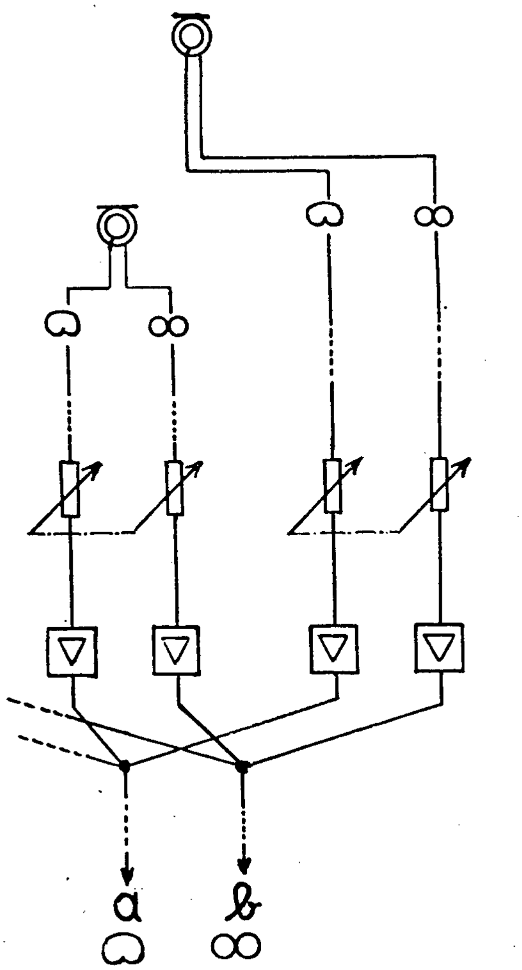

Fig. 2c. Method of Mixing-in Additional Microphones.

fore, necessary in placing. For sources lying outside this range of angles the directions will appear indefinite, because the loudspeakers will be driven by voltages wholly, or partially in antiphase.

III. Microphones for Stereophonic Transmission

Microphones for use in stereophonic pick-up must satisfy several additional requirements. Over and above the known requirements for single-channel working in respect of flat frequency characteristic and low harmonic distortion, together with a wide dynamic range, classical stereophony demands equality of frequency and phase characteristic in the transmission channels; otherwise, the transmitted directional impressions will be frequency dependent. It is found particularly disturbing if there are asymmetries in this respect for discrete frequency ranges. It is not easy to find matched samples from types of microphone in which the frequency response is achieved by a series of resonances spread throughout the transmission band, even if these microphones are quite satisfactory for single-channel transmission.

Condenser microphones have been found particularly suitable for the purpose. (Their accurate matching is, of course, only fully effective if, on the reproducing side, accurately matched loudspeakers are used.)

For stereophonic pick-up on classical principles, NEUMANN condenser microphones, specially selected for uniform frequency response, can be used. Transmission on

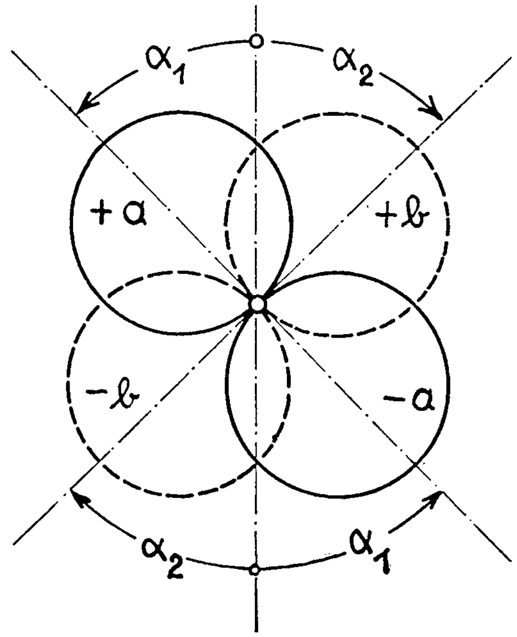

Fig. 3. Polar Diagram and Working Angles in the Case of Two Cosine Microphones.

the intensity stereophony principle requires microphones with strongly directional characteristics. In addition, they must have small physical dimensions so that they do not distort the sound field when mounted in close proximity.

These stringent requirements can only be met in practice by the use of condenser microphone capsules as pick-up elements. Since it is barely possible to put two single microphones sufficiently close together without some disadvantage, the NEUMANN stereo-microphone SM 2 was developed for these applications.

This double microphone contains two similar, closely adjacent, microphone capsules. Their principal axes can be turned away from each other. The capsules are pressure gradient operated with two diaphragms. Each system is capable of independent remote adjustment by variation of the polarising voltage, continuously from spherical through cardioid to cosine characteristic, so that with this microphone many different arrangements may be tried.

Berlin, December 1956.

Druck: Walter Grützmacher, Berlin SW 61

500. 6. 57.