Modern Sound Control Technology in Large Auditories and in Small (Reference) Listening Rooms 3380 (F-3)

Gerhard Steinke
Audio Consultant
Berlin, Germany

Presented at the 93rd Convention 1992 October 1–4 San Francisco

AES

This preprint has been reproduced from the author's advance manuscript, without editing, corrections or consideration by the Review Board. The AES takes no responsibility for the contents.

Additional preprints may be obtained by sending request and remittance to the Audio Engineering Society, 60 East 42nd St., New York, New York 10165-2520, USA.

AN AUDIO ENGINEERING SOCIETY PREPRINT

MODERN SOUND CONTROL TECHNOLOGY IN LARGE AUDITORIES AND IN SMALL (REFERENCE) LISTENING ROOMS

GERHARD STEINKE

AUDIO CONSULTANT, BERLIN, GERMANY

ABSTRACT

Theses and experiences with the innovative "Delta Stereophony System" for larger and smaller rooms will be discussed, in connection with the problems of multi-channel sound reproduction for HDTV on the basis of the new CCIR standard.

1. AESTHETICAL TARGETS AND PRECONDITIONS FOR HIGH QUALITY SOUND SYSTEMS

For sound systems, i.e. the electroacoustical transmission and reproduction into large rooms as well as into small rooms (domestic rooms, reference listening rooms), the sophisticated aesthetical target should be

. in most of the cases, with or WITHOUT accompanying picture reproduction, as well as with live events,

. according to the specific genre and the spatial conditions

. with concentrated listening

to produce for the recipient a perfect listening event by which it can be created

. a CONVINCING IMAGINATION OF THE RELEVANT SOUND EVENT, of the acoustical environment as well as of the spatial coordination of the individual sound sources, also in the corresponding distance of the listener (in the original room),

. the impression OF BEING INCLUDED IN THE ARTISTICAL AND ACOUSTICAL ATMOSPHERE OF SOUND EVENTS (also in the case, if the events occur at another location)

in such a technical transmission and reproduction method that INTENSIVE FEELING and the CREATION OF SPECIFIC MOODS are possible.

With fulfilling all the conditions in the case of the electro acoustical transmission we speak of TELEPRESENCE, a term which is defined for HDTV transmissions specifically (in a reference distance from the screen), where the viewer/listener should have the impression of being included in the optical and acoustical perspective.

Main target of a transmission is either to create a LISTENING PERSPECTIVE, which has high similarity to the natural ACOUSTICAL PERSPECTIVE of an observer at the original location (transmission of a concert) or to compile new listening perspectives on the basis of components of the natural sound events, artificial sounds and processing methods (radio drama effects).

To realize this target the content of a performance is to be effected intensively by suitable transmission of directional, distancial and spatial (ambient) informations, so that a desired optimum aesthetically ratio of direct (D) and reflected (R) sound can be reached.

The electroacoustically created ratio D/R at a certain point of the room versus the time can be assessed as a measure of the effect of including the listener into the room at the corresponding distance with respect to the sound source. It has to be matched to the relevant situation and must be variable; therefore, D- and R-informations have to be recorded, processed, transmitted and reproduced in a discrete manner.

Methods like EIDOPHONY; ORTHOPHONY [1] etc., which intend to fullfil the conditions of the true fidelity transmission and reproduction of original sound fields do not allow the practicability which is necessary for the sound designer when dealing with formations which cannot have uniform loudness level distribution; the normal practice in large auditories, open-air theatres and in most of the TV-, Cinema and Broadcasting situations!

In the case of common picture and sound transmissions an additional condition should be observed, that is, the deviation of localization between the acoustical source and the visual picture should be not more than +/-6° [2] for smaller rooms; in large rooms with sound reinforcement (halls, cinema) this deviation can be permitted up to +/-15° (depending on the genre of the event; ideally: screen with 5 speakers).

The requirements to a high-quality sound system can be summarized in the following criteria which have to be ensured for an acceptable sound system comfort [3] [4]:

- highest sound quality, clearness (Deutlichkeit) and intelligibility by avoiding masking and falsification of the characteristics of the source

- performance of subjective limits for objective quality parameters (acc.to CCIR Recomm. 468, 644, 645; digital technique with format of more than 16 bits/s linear (because of the weighted signal/noise ratio of only 74 dB with 16 bits/s, which is not high enough for sound systems in large halls in most of the cases as stated in Rec.646); or in form of the ISO/MPEG bit-rate reduction systems

- uniform sound level distribution, well balanced sound colour (timbre) and balanced loudness level over the whole reception area

- coincidence of acoustical and optical perspectives in case of stationary and moving sound sources and effects by true directional and distancial localization for a suffcient large listening area

- good audibility within the area of action

- inclusion of the listener into the acoustical ambience and/or provision of simulation of complex sound field structures (as diffuse sound fields, "sound clouds") and special sound effects (moving panorama or surround sound etc.)

2. OPTIMUM SOUND REINFORCEMENT SYSTEM TECHNOLOGY FOR LARGE AUDITORIES AND OPEN-AIR THEATRES

The principal task for any designer of a system and the operating personell results from our consideration that in all cases of transmission and/or reproduction

> the aesthetical and acoustical structure of the original sound event (sound source or source simulation) should in no way be influenced by the electroacoustical means and measures used, but only supported as well as amplified and distributed at all listeners' places for excellent intelligibility!

This target can only be reached if the listener at any place of the room directly gets the first wavefront from any sound event; only then he also will get the first stimulus for distancial and directional localization and an impression of the individual natural sound source timbre and dynamics without markable masking or falsification of the properties of the source!

This well-known fact is neglected in most of the halls in the world having sound systems. The two-channel stereophony developed for recording/radio reproduction has been used wrongly in the sound reinforcement technique in large halls, but this solution with its strong dependence on the listeners's place in a very restricted area can only work to a certain degree in small (domestic) rooms (see section 3), as Fig.1 clearly shows.

In former publications this fact was often critisized (see also in [4)), but most of the owner/user of halls will not believe it, because of the simpler and non-expensive solution, which also non-technicans can handle with!

2.1 EXPERIENCES WITH THE DELTA STEREOPHONY SYSTEM

Up to now only with the help of the DELTA STEREOPHONY SYSTEM the most important presupposition can be ensured in large halls, but as it will be shown later, also for electronic cinemas and viewing sites.

On the principle of the Delta Stereophony System = DSS , we formerly informed at several AES Conventions and in [3,4] [5,6], together with different applications.
Fig.2 shows this priority principle required in section 1, where the delay times of the signals to the distributed loudspeakers staggered in such a way that always the original sound from the source Q arrives first at all listeners' locations and only after that the sound coming from the individual lodspeakers beginning with those having the smallest distance (and/or angle) to the connecting line between listener and source Q (To and T1,2,n = acoustic propagation time, T= electrical delay time).

The loudspeakers for the reinforcement system carefully distributed in computer simulation processing calculation bring the "power", the necessary amplification within 35-50ms after the original sound has been perceived, and in practice can have an increase in sound level of about 6 .. 8 dB.

In conventional sound systems a directional deviation between the visual and acoustical perception cannot be avoided. Moreover, centralized sound systems technique can only represent an information sound; its task is solely the amplification of the direct signal. With this technique irregular sound level distibution has to be accepted. This results in the often so-called "loudspeaker sound", because the original sound cannot reach the listener prior to the amplified (and mostly falsificated) sound; therefore the real original cannot be detected. More and more, the listener accustomed to CD's always expects this original sound impression, and does not accept such bad sound reinforcement systems, also the above mentioned and wrongly used two-channel systems.

The DSS was developed in 1975 for the application of new sound reinforcement systems in multi-purpose halls, under the heading of the author. It is world wide patented and was introduced firstly in Berlin 1976. The DSS was subsequently installed at Prague, Munich, Stade, Stuttgart and also in the Tokyo Metropolitan Art Center and is now in preparation for the Moscow Kremlin Palace.
Likewise, the system was applied with great success in open-air theatres such as Lake Festival Bregenz (Austria), Trachselwald (Switzerland), and Waldbühne Berlin.

In the latter case it was the very sensitive famous sound of the Berlin Philharmonic Orchestra which had been brought to each of the 22.000 listeners in a manner that they all had the impression that no electroacoustical medium was between the artists' play and the ears of the auditory. Only when we switched off the microphones (at the rehearsal), the original sound was weak and dry, the intelligibility unsufficient.

The principle of the DSS is relatively simple - but the realization under each special condition can only be successful with

- highly qualified and trained personell (engineers, tonmeisters)

- careful preparation of each performance, early coordination with scenery, art director, actors etc., sufficient time for rehearsals etc.

- high-quality (studio) equipment (see also 2.2).

Installations within halls need coordinated planning from the earliest beginning with user, architect, acoustican, sound designer, tonmeister etc., to get a harmonization of room and electroacoustics. In most of the cases (e.g. conference centre of Budapest, Karlsruhe etc.) the architecture is not conceived adequately for the later use of different genres in such a way that the sound design can be integrated successfully (fixing points of the loudspeakers, location of the mixing console etc.).

2.2 SYSTEM CONFIGURATION OF THE DSS

The basic arrangement of such a direction- and distance-oriented system as the DSS consists of the following components (Fig. 3):

- decentralized loudspeaker arrangement dimensioned for a uni form area covering sound distribution in the whole auditory

- support of original sound sources, if necessary, by so-called" source simulation speakers" for generating a "reference sound field" in the action area, for the determination of the direction and for the provision of a good hearing and contact for the actors themselves, supplemented by special delay time aligned monitors

- delay units for the generation of delay time-staggered signals for all the loudspeakers in the action area and the reception area

- summing-signal distributors (matrix) for the system-orientated routing of the differently delayed signals for the individual loudspeakers, whereby with a DSS-Processor, i.e. DSP 610/AKG [7] both the delay units and the matrix are combined in a compact digital device (Fig.4)

- multichannel sound console for the source area related mixing and allocation of all the sound signals (the often used two-channel stereo consoles are useless in any casel)

The use of several computer systems enables an efficient operation by means of pre-programmed set-ups (for the scene-by-scene setting of the individual parameters of the channels, running and distribution of the delayed signals, for the determination of the optimized delay times, for the tracking of single sound source movements, simulation of the sound field etc.)

Of course, as with any sound system with distributed loudspeakers, the reverberation time of the (multi-purpose) halls and especially, their stage houses should be as low with flat response as possible, if the sound system shall be successful and do not increase too much the reverberant sound. Good diffuseness is important, significant reflections should not be permitted.

3. OPTIMUM SOUND TECHNOLOGY FOR SMALL ROOMS

The philosophy and the the preconditions dealt with in section 1 are valid here in principle, too.
Of greatest importance are the listening conditions and the use of an universal, worldwide standardized multi-channel system.

3.1 LISTENING CONDITIONS

Minimum acoustical conditions and room-geometrical requirements for reference listening rooms, from which conditions for control rooms and refurnished living rooms can be derived were proposed by the author at the 84th AES Convention, Paris 1988 [8].
Within this considerations, the author had also proposed that it would be useful to represent the interaction of the listening room acoustics and the studio monitor loudspeaker behaviour in a single objective characteristic such as a reference sound field characteristic. At first considered to be a hypothetical ideal, it should be approached in an optimum way by improving the technical parameters and tolerances of the relevant realizations.
On the way to this solution this listening standard was indirectly defined by reference listening rooms and reference studio monitors, as the OIRT has recommended in OIRT-Recomm. 86/3 and 55/2. In recent practice it has proven as a good way - the reference listening room in the Broadcasting House Berlin [9] is now in use several years with great success.The experiences with this room were involved in the next step, the construction of the new reference listening rooms for multi-channel technique with DBP Telekom, Research Institute, Berlin.

Moreover, the used monitor RL 900/MEG(Musikelektronik Geithain, Germany), the top model of a monitor family (RL 900, 900A, 904 etc.)became the winner of an extensive IRT test for ARD/ZDF [10] because of its outstading objectice and subjectively assessed quality.

A summary of all this data and experiences was prepared by the author in a document for CCIR Tak Group 10/2 [11] to give the basis for listening test rooms to assess bit-rate reduction codecs.

What's with the reference sound field? The former CCIR Interim Working Party 10/12 established also by the author's proposal and now transferred into the new CCIR Task Group 10/3 "Subjective Assessments" after several meetings came to the conclusion that such a reference sound field could be described as a first approximation by a specification of an acceptable envelope for the reflection pattern as sketched in Fig.5. The frequency response is measured in time windows with about 2, 5 ,20 50 ms etc.

With such a specification not only the characteristics for listening test rooms for multi-channel sound techniques, but also guidelines for the arrangements in domestic rooms could be derived. The historical philosophy to look for an "medium domestic room characteristic" as the broadcasting reference, like several authors and also the IEC in specifications accepted, can never be a basis for the reference situation. If we look for optimum quality conditions we have to determine the necessary requirements for natural reproduction and to recommend also to the listeners who will use multichannel sound systems at home the required minimum to reproduce the intention of the artists and the sound designers.

Considering the listening conditions for the multi-channel sound system, described in section 3.2, we can only say up to now that an equal distribution of absorbing material on all the surfaces of the room is useful.
After concluding extensive experiments and tests it may be that it is sufficient to have the high damped area behind and beside the three frontal monitor speakers to get smallest size of phantom sources and to reduce the influence of early reflections on the direct sound, and to have a more diffuse reflective behaviour in the back part of the listening room where the speakers for the surround sound are installed (see as example in Fig.6). For ambient sound information it can be useful but for special effects it can be disturbing.

3.2 THE NEW CCIR RECOMMENDATION: 5-Channel (3/2) STEREOPHONY AND ASPECTS WITH RECORDING AND REPRODUCTION

After living for 30 years with a restricted two-channel stereophony system, a door was opened to the third dimension of the world of audio by means of the new CCIR-Recommondation:

"Universal Multi-channel stereophonic sound system with or without accompanying picture".

The conventional two-channel stereophony is criticized because its imperfections, which cannot ensure to reach high aesthetical and technical targets:

- great dependence on the position of the listener, i.e. only a small range of audibility ("tube") around the middle axis for perceiving correct directional impressions exists, especially for centre impressions

- the reproduction of listening events is restricted within a frontal angle of about 60° and the speaker basis plane

- spatial (ambient) listening events cannot be generated, therefore the listener cannot be included in the artistical and acoustical event

- conformity between picture and sound in the combination with a TV picture can be reached only to a small amount.

Therefore, the physical restrictions prevent any further improvement in the aesthetical sound domain!

During the sixties, proposals for including a centre loudspeaker and rear loudspeakers were made and a special stereo-ambiophonic transmission and reproduction system was developed with Deutsche Post in Germany [12], but the technical (transmission channels) and economical situation was unsufficient at that time. Also the later quadrophony was a flop because of wrong applications and targets.

Only the well-known ideas of R. Dolby (Dolby Stereo Surround) in matrixing techniques could be realized via the cinema industry, and this with great success, although only one surround channel was available. But now the new Dolby-SR-D-system solution is also in good coordance with the SMPTE and CCIR recommendations and supports the international trends.

Considering the given limits in domestic rooms, economical production and reception relations, a compatible system hierarchy also for multi-media applications, minimum time-delay between sound and vision - the great international standardization bodies EBU, SMPTE and others agreed with the compromise of the CCIR proposal to standardize a reference loudspeaker arrangement according Fig.7, with an option acc.to Fig.8.

Having in mind that with modern bit-rate reduction techniques also the transmission problems will be overcome, the system should be applied in the recording industry, cinema, and in the nearest future with the HDTV, so that we have to look for optimum compatibility because of the fact that most of the programme material of TV will come from the cinema.

Fig.9 shows the different reference arrangements for HDTV, Cinema and the conventional two-channel stereophony. The author prefers the "2 H - reference distance" for HDTV application, because only then high compatibility between cinema and TV can be ensured.

For the recording (pick-up) technique of the direct information for 3 frontal speakers a lot of new problems arise; new techniques have to be developed and the old 3-microphone pick-up technique with directional mixing, from the beginning of the stereophony times in the sixties, have to be studied again!

With the Stereo-Ambiophony System mentioned above, a special technology was developed, to get diffuse room informations in high quality, as it is shown in Fig.10, and applied in co-phase and anti-phase manner to the frontal as well as the rear/side loudspeakers, to create the feeling of involvement.

This technique is now 30 years old, and was patented formerly. But there was no publication in english language so that W.Woszczyk must believe that his ES microphone idea [13] is a competely new one - but only the application for the Dolby Surround Matrix is new. But nethertheless it is the hope of the author that this ideas should be distributed broadly.

The essential features of the Stereo-Ambiophony System are the delay of the direct signal against the room information and the pick-up method (with artificial head, usual room microphones A/B, M/S, X/Y, or now also with the new sphere surface microphones) at a location which has a large distance from the sound source, because only in the back of a hall etc. exists the optimum acoustical quality for room microphones, as Fig.11 shows, i.e., dense and high-energy reflections. The delay time can be varied to simulate different listening conditions at different listener positions (according to different envelope slope and density of room reflections within the first 80 ms (see also [14]).

The room signals have to be processed and mixed in M/S and X/Y form (see >Fig. 12) before adding to the frontal and surround speakers to get highest feeling of ambience!

The system was used in different manners - the first public demonstration was at the Tonmeistertagung 1966, Cologne.

It can be expected that the recording engineers will further develop this system and create new sophisticated technologies which can fulfil all the expectations and targets for the HDTV sound.That means also that binaural signal combinations, derived from artificial head signals or processors as Roland Space should be superimposed to the signal methods discussed above to bridge the distance between the listener and the simulation plane of the frontal speakers. This should also observed for experiments with 3-D video techniques.

3.3 MULTI-CHANNEL REPRODUCTION WITH LARGER LISTENING AREA

As investigations have shown it is possible to get a broad listening area and sound image stabilization with 4 frontal loudspeakers, but from technical and economical reasons ("household compatibility") such requirements could not have chances under domestic conditions. Therefore the compromise of 3 speakers was accepted.
Also under this condition the listening area is improved compared with the two-channel stereophony, but not large enough for more than 3...5 persons as demonstrations of HDTV presentations with the 3/2 system have shown.
Therefore in the Telekom Research Institute, Berlin, the application of the Delta Sterophony System was investigated and in the experimental studio a configuration as shown in Fig. 13 was set up.
Two different sound fields can be compared:
- the reference cofiguration accordg. to the CCIR-Recomm.,
- a larger listening area for about 25...30 listeners/viewers by adding 12 compact studio loudspeakers to the main frontal speakers.

The complete system is under study at the time being; but the results up to now show that the benefits of the DSS for larger rooms can be applied with great advantage also to the 5 (3/2)-channel system. Especially, the higher number of speakers increases the effect of including the listener in the acoustical atmosphere. The listeners sitting nearer to the surround speakers cannot localize them. in most of the cases.

The tests have shown that the DSS can solve the problems in an electronic cinema with multi-channel sound and give better sound level distribution compared with the normal cinema with only 3 strong frontal speakers and one-channel surround sound.

(Details will be published later)

The application of the DSS with HDTV multi-cannel sound effects an optimum sound reinforcement (perception) comfort with

- high sound quality and balanced sound level distribution

- high direction stability of the frontal imaging

- natural reproduction of ambient (effect/surrounding) information for a large involvement into artistical events

for all listening positions within smaller or greater reception areas, that means an optimum imagination of **TELEPRESENCE.**

REFERENCES

[1] Hensel,J.,Krause,M., Schaller,W., Orthophonie -ein neues Aufnahme- und Wiedergabeverfahren zur Abbildung räumlicher Schallfelder (Orthophony - a new recording and reproduction system for imaging of spatial sound fields).FERNSEH-UND KINO-TECHNIK, Berlin, 3, pp.165-170, Vol.46 (1992)

[2] Hoeg, W., Wagner, K., Aspects of stereophonic sound transmission in TV. Tech. Mitt. RFZ, Berlin, 4, pp.104-106 Vol. 16 (1972)

[3] Steinke, G.,Delta Stereophony - A Sound System with true direction and distance perception for large multipurpose halls,Journ. AES 7/8, p.500, Vol.31 (1983)

[4] Steinke, G., Fels,P., Hoeg,W.,Ahnert, W., True directional sound system - new applications of the DSS. Paper presented at the 82nd AES Convention, 1987, London

[5] Hoeg, W.,Steffen,F., Steinke,G., et al. Sound system with true directional and distance perception for large auditoria. 6th Conference of Acoustics, Budapest, 1976; Tech. Mitt. RFZ, Berlin, 2, pp.25-27, Vol. 20 (1976)

[6] Steinke, G.,Fels, P.,Hoeg, W.,The Delta Stereophony System (DSS) in the City Hall of Stade and in the open-air theatre Trachselwald. Paper presented at the 88th AES Convention, 1990, Montreux

[7] Nadler, W., The Delta Stereo Compact Processor DSP 610. Paper presented at the 82nd AES Convention, 1987, London

[8] Steinke, G.,Minimum requirements for reference listening rooms. Paper presented at the 81st AES Convention, 1988, Paris

[9] Lau, W., Huhn, K.,The new reference listening room in the Berlin Broadcasting House Tech. Mitt. RFZ, Berlin, 2, pp.38-44, Vol.31 (1987)

[10] Spikofsky, G., Selection of the new reference studio monitor (ARD/ZDF). IRT-Munich, Tech. Rep. No. 3115/90

[11] Steinke, G., Steffen, E., Listening Conditions for tests with bit-rate codecs. Doc. CCIR TG 10/2-6, November 1991

[12] Steinke, G., On the development of the Quadrophony Tech. Mitt. RFZ, Berlin, 2, pp. 59-64,ol.16 (1972)

[13] Woszczyk, W., "ES" Direct microphone encoder for surround sound recording. Paper presented at the 91st AES Convention, 1991, New York (Seminar Paper: Dolby Surround and..)

[14] Steinke, G. From Ambiophony to HDTV-Domestic Sound Systems FERNSEH-UND KINO-TECHNIK, 6, pp.297-306, Vol. 54 (1991)

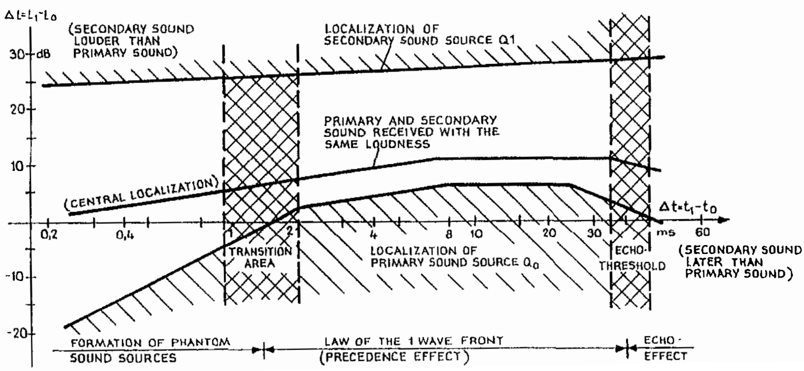


FIG. 1 VALIDITY DOMAIN OF THE LAW OF THE FIRST WAVEFRONT AND THE PHANTOM SOURCE INFORMATION [5]

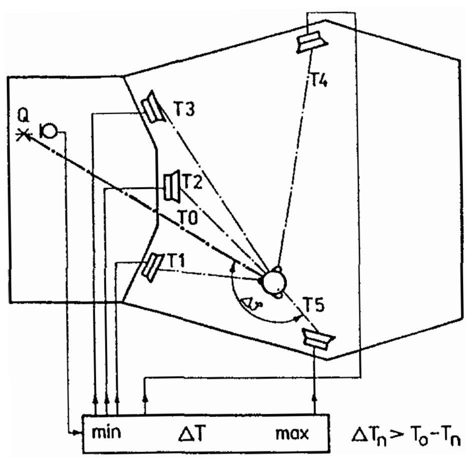


FIG. 2 ACOUSTICAL PRINCIPLE OF THE DSS

$\Delta\varphi$ = deviation angle between optical and acoustical impression without DSS

$T_{1,2,n}$ = propagation time of loudspeaker sound

T_0 = propagation time of original sound

ΔT = electrical delay time

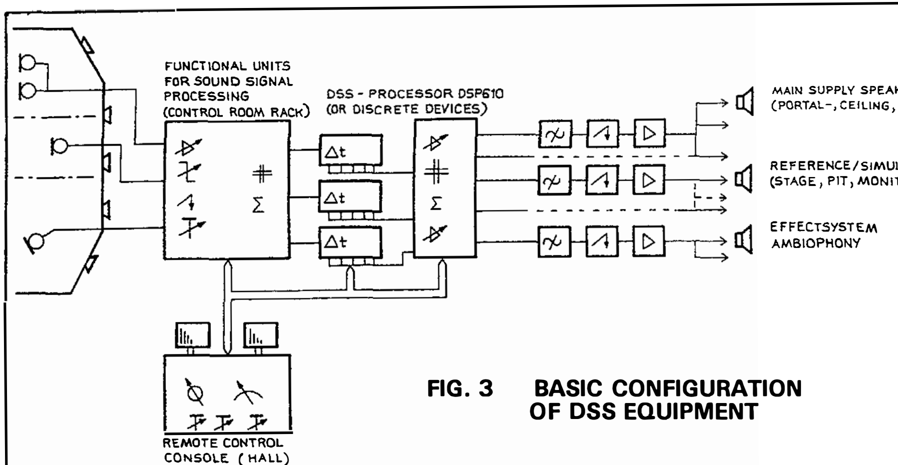


FIG. 3 BASIC CONFIGURATION OF DSS EQUIPMENT

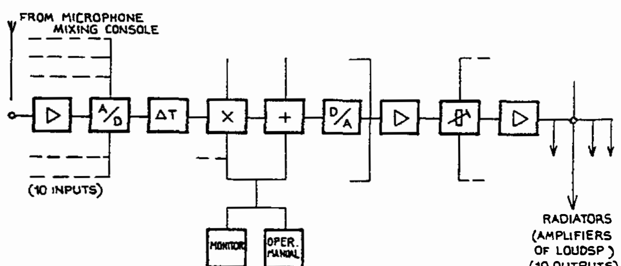
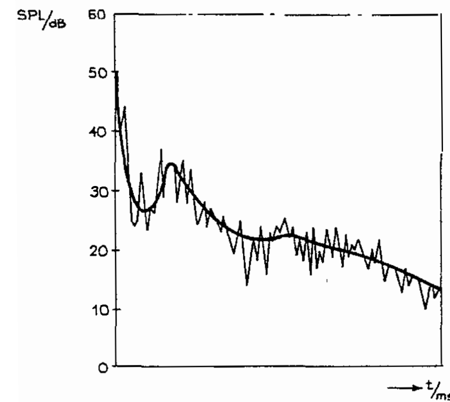


FIG. 4 PRINCIPLE BLOCK DIAGRAM OF A DSS PROCESSOR (SIMILAR TO DSP 610/ AKG)

FIG. 5

REFERENCE REFLECTION PATTERN FOR A GOOD LISTENING ROOM (Draft CCIR IWP 10/12)



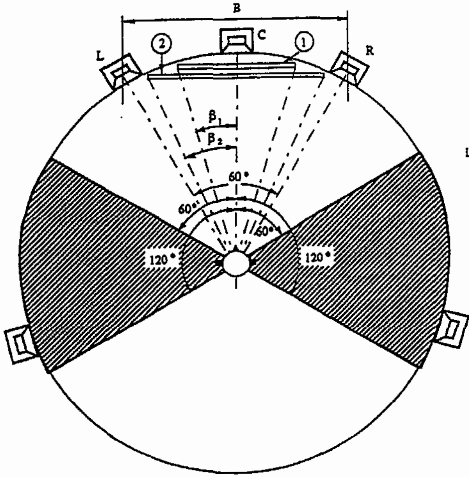


FIG. 7

Reference loudspeaker arrangement with loudspeakers L/C/R and LS/RS

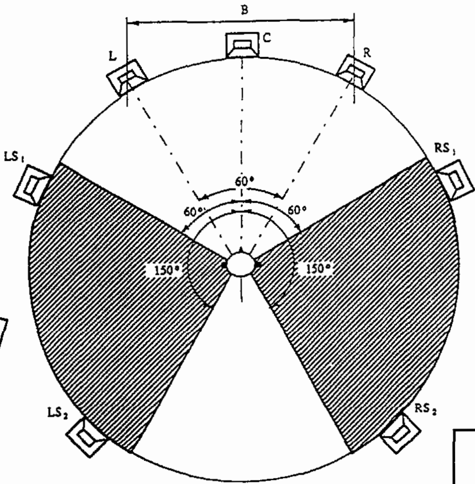


FIG. 8

Alternative loudspeaker arrangement for 3/4 (3 front and 4 surround) system with signals/channels: L/C/R and $LS_1/RS_1/LS_2/RS_2$

Screen 1 HDTV - Reference distance = 3H ($2\beta_1$ = 33°)
Screen 2 = 2H ($2\beta_2$ = 48°)

H: Height of screen
B: Loudspeaker baseline

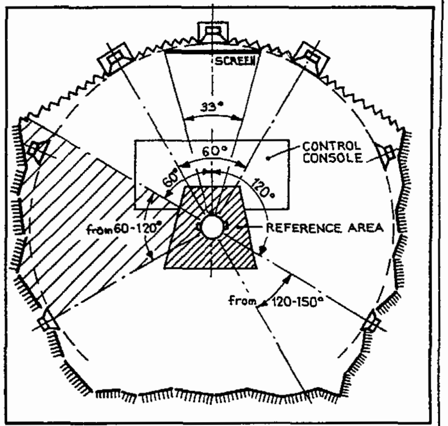


FIG. 6 MULTI-CHANNEL CONTROL ROOM FOR 3/4 - SYSTEM

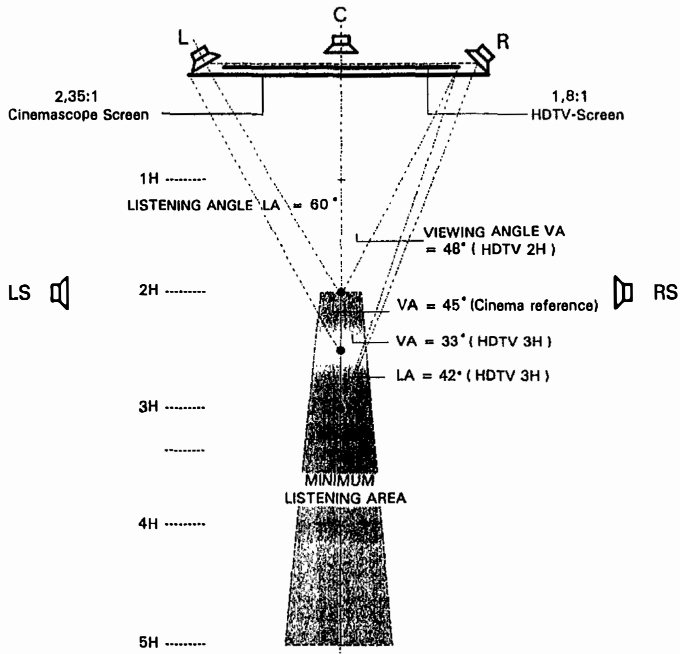


FIG. 9 REFERENCE ARRANGEMENTS: HDTV, CINEMASCOPE 2/0 SYSTEM, 3/2 SYSTEM

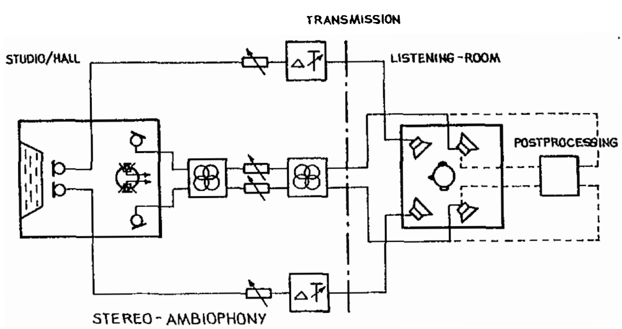


FIG. 10 PRINCIPAL CONFIGURATION OF THE STEREO AMBIOPHONY SYSTEM

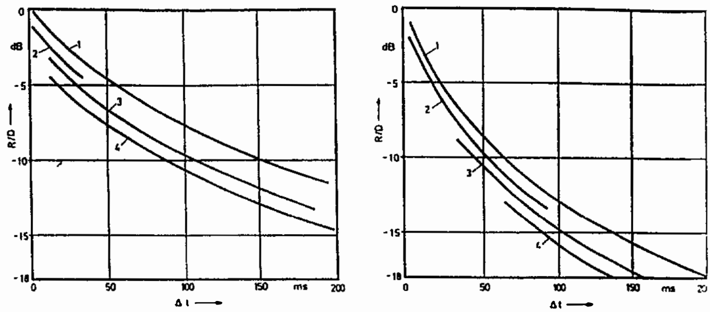


FIG. 11 R/D - RATIO AS A FUNCTION OF DELAY TIME Δt AGAINST DIRECT SOUND WITHIN A 10.000 M^3 RECTANGULAR ROOM
AT A DISTANCE OF SOUND SOURCE – LISTENER
a) 24 m b) 10 m

CURVE 1 – REFLEXIONS WITH T = ∞ (ECHOROOM)
2 – SINGLE
3 – DOUBLE } REFLEXIONS WITH T = 2,5 s
4 – TRIPLE

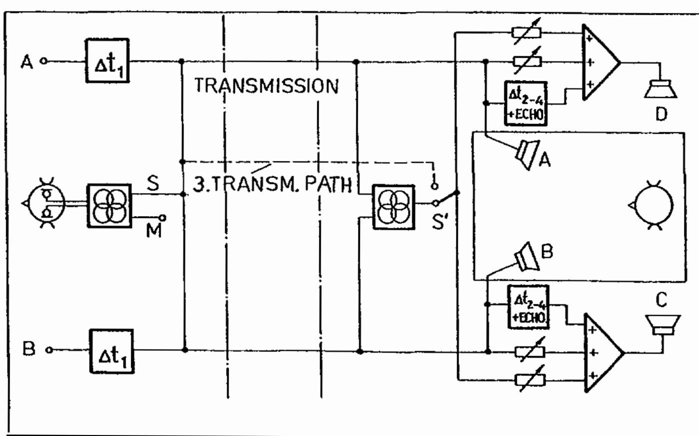


FIG. 12 STEREO AMBIOPHONY
WITH ADDITIONAL SIMULATION CIRCUIT FOR AMBIENT SOUND AND SIGNAL PROCESSING IN 4-2-4 AND/OR 4-3-4 MODE

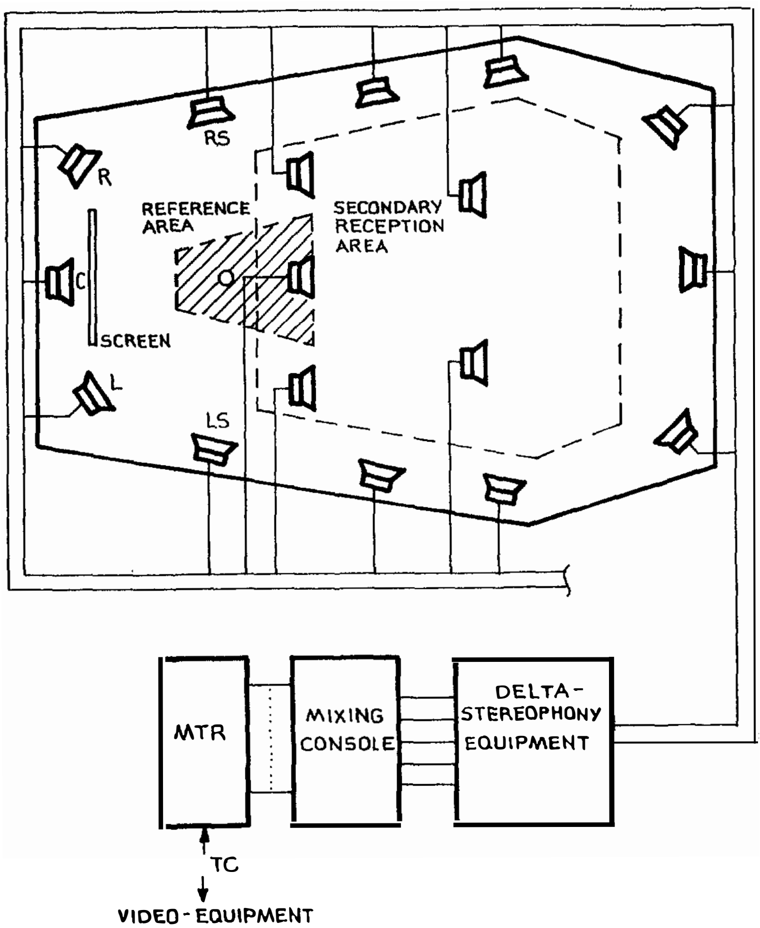


FIG. 13 HDTV SOUND REPRODUCTION WITH DSS