TOWARD A MICROPHONE TECHNIQUE

FOR

DOLBY SURROUND ENCODING

Peter Cook

Faculty of Music
McGill University, Montréal
March 1991

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Music, Sound Recording.

© Peter Cook, 1991

TABLE OF CONTENTS

ABSTRACT

Dolby Surround technology offers consumers surround sound in their home via a 4:2:4 encode/decode matrix. Although originally intended for audio accompanying visual media, the system has potential as a music-only playback system.

The purpose of the thesis is to investigate this potential, particularly as it applies to acoustic music recording. Dolby Surround encode and decode technology and its relevance to acoustic music reproduction is reviewed. The classic stereo microphone techniques are discussed with particular attention paid to each one's theoretical ability to "encode" information for the Dolby Surround decoder. Practical limitations and benefits of these well-known methods are considered.

Recently proposed microphone techniques are reviewed in theory and in practice and are found to provide many solutions. Methods for optimizing the decoder technology for music reproduction are suggested. The paper is relevant to any acoustic recording application for a number of surround systems as well as for conventional stereo and mono.

RESUME

La technologie Dolby Surround rend le son "surround" accessible au consommateur par le truchement d'une matrice d'encodage/décodage 4:2:4. Bien que destinée au départ au traitement des bandes sonores de films, la technique peut aussi potentiellement être utilisée pour des systèmes de reproduction musicale.

La présente thèse a pour but d'étudier ce potentiel et son utilisation dans le domaine de l'enregistrement de musique acoustique La technologie d'encodage et de décodage Dolby Surround et sa pertinence pour la reproduction de la musique acoustique sont passées en revue. Les techniques classiques de prise de son stéréophonique et la capacité de chacune d'elle d'"encoder" l'information destinée au décodeur Dolby Surround sont discutées. Les avantages et les limitations pratiques de ces mèthodes connues d'enregistrement sonore sont examinés.

Les aspects théoriques et pratiques de techniques de prise de son récemment proposées sont examinées et certaines d'entre elles offrent des solutions intéressantes. Des méthodes pour optimiser la technologie de décodage en vue de la reproduction musicale sont suggérées. Les résultats présentés ici sont pertinents tant à l'enregistrement acoustique destiné à un certain nombre de systèmes "surround", qu'à l'enregistrement stéréophonique et monophonique conventionnel.

ACKNOWLEDGEMENTS

The author would like to thank David Gray of Dolby Laboratories Inc. for his helpful discussion early on in the preparation of this thesis, and especially for facilitating the use of a Dolby Surround decoder for this and other research.

Thanks go to colleagues and students who helped with the recordings and aided in the setup of the surround monitoring system.

Thanks also to André White, D'Arcy Gray and other musicians for their help in the recordings.

The author would like to express his appreciation to Floyd Toole for providing reference materials and offering encouragement early on in this research.

The author would like to thank his thesis advisor, Dr. Wieslaw Woszczyk, for his helpful comments on the text and for providing a stimulating learning environment.

Most importantly thanks to Bobbie Watson for her love, patience and support.

INTRODUCTION

Surround Sound Accompanying Visual Media

Dolby Stereo was introduced by Dolby Laboratories in 1975 [1]. This surround sound system was designed to enhance the movie-going experience by making possible an accurate stereo presentation and increased sense of atmosphere over the wide listening area of the movie theatre. The now *de facto* industry standard employs a 4:2:4 matrix with three loudspeakers -- left (L), centre (C), and right (R) -- in front of the audience (behind the screen) and a U-shaped array of surround (S) loudspeakers behind and to the sides of the audience.

The strong progress of video technology in the 1980's meant that many people were enjoying movies at home rather than in the movie theatre. With stereo audio tracks on video tapes the surround encoded movie soundtrack was available to the consumer. Recognizing the new market Dolby Laboratories introduced Dolby Surround in 1981 enabling consumers to enjoy basic movie surround sound at home. This "passive" decoder has L and R outputs, a delayed, band-limited S output and it may have an optional C output.

Dolby Pro Logic -- consumer technology with the same decoding power as the professional units used in movie theatres -- became available to consumers in the fall of 1987

[2]. In 1988 Sanyo developed an IC which has led to the proliferation of Pro Logic decoders in television sets, integrated amplifiers, receivers, etc and which has greatly accelerated the use of the Dolby Surround format in the home At the end of 1989 there were 3 million surround decoders in homes around the world. In Japan nearly 45% of all television sets sold that year included some form of built in surround decoder [3].

Along with hardware developments the software base has expanded dramatically. Hundreds of feature films with Dolby Stereo or Ultra Stereo audio have been transferred to videocassette and videodisc and may be purchased, rented or seen in stereo television broadcasts. Regular network television programs, major sporting events, and music specials are frequently broadcast in surround using the Dolby Surround or Stereosurround systems.

A New Application

Consumers who have invested in the decoder, extra amplification and loudspeakers will want to play their music-only software (eg compact discs) through the decoder. Results vary widely since the music was not recorded with Dolby Surround playback in mind.

With such strong support for the Dolby Surround standard it seems logical to produce compatible music software for the growing market. This thesis will investigate the potential of the Dolby Surround system as a playback format for acoustic music, that is music which is produced in a natural acoustic environment. In particular some of the classic stereo microphone techniques will be examined as to their potential as Dolby Surround encoders. Some suggestions will be made as to how these basic techniques may be altered and adopted to achieve a more suitable surround encoding system. Recent research on newly developed microphone techniques will be reviewed. Improvements in the decode chain for music reproduction will be recommended.

A Note on Terminology and Applications

The term "Dolby Stereo" refers to Dolby Stereo-encoded audio for movie soundtracks presented in theatres using Dolby Stereo cinema equipment. "Dolby Surround" programme material is that which will be played back in a consumer environment. The encoding equipment is the same as that used in Dolby Stereo productions. Shure's Stereosurround system and the Ultra Stereo system are based on the same technology and may be considered largely compatible with Dolby Surround.

For the sake of brevity the term "Dolby Surround" will be used throughout the paper except where reference is made specifically to another system. Left, centre, right, and surround are abbreviated to L, C, R, and S respectively.

The paper focuses on acoustic music recording for playback in the Dolby Surround format. However, the discussion will be relevant to those involved in any type of acoustic recording for many different formats including surround (matrixed and discrete), conventional stereo and mono.

RATIONALE

There are a number of very practical reasons for using the Dolby Surround encode/decode system for acoustic music recording and reproduction.

Support for Standard

It is widely anticipated that home entertainment systems will continue to evolve toward a "home movie theatre" concept, that relatively sophisticated sound systems will be a part of this system, and that music-only software will be reproduced over this same system [4].

With the visual based media -- video, television, HDTV, etc -- defining the orientation of the home entertainment system it makes sense to adopt a music reproduction format which is compatible. Indeed it is the goal of groups developing a sound system standard for HDTV to define "a new loudspeaker reproduction standard which is able to satisfy the requirements of HDTV sound whilst also providing optimum stereophonic quality to satisfy the 'purist' music-lover" [4]. This is a very attractive goal from the point of view of hardware manufacturers, software producers, and consumers.

A number of groups working on an HDTV sound system have recommended that the system be compatible with movie sound reproduction systems [5]. If home sound systems develop as expected it would seem very practical to produce music-only recordings intended for reproduction over this sound system.

Laser discs and videocassettes of classical music performances are currently making strong market gains. This is perhaps the ideal application of acoustic music recordings for Dolby Surround reproduction.

From a music industry perspective the format is both practical -- being relatively inexpensive and simple to use -- and a very well-established standard. It is also compatible with two-channel storage and transmission media. This is particularly important since two-channel audio is by far the dominant consumer standard now and is likely to remain so for some time to come.

Improvements Over Conventional Stereo

Dolby Surround appears to offer significant improvements over conventional stereo for acoustic music reproduction. In two-loudspeaker stereophony phantom images -- those not located at the left or right speaker --only work properly for a carefully centred listener. As the listener moves to one side the images collapse in that direction. The C speaker allows for the democratization of the "stereo seat" with image stability over a much wider listening area [3], [6].

It is also well-established that a "phantom" centre image is unsatisfactory with respect to timbre, "clarity", and spatial sharpness relative to a centre image provided by a virtual loudspeaker [5], [7]. The virtual centre speaker is described as being "easier to listen to" [8]. For acoustic music recording the L, C, R loudspeakers provide the basis for a solid front soundstage.

The surround channel provides for a solution to the spatial distortion inherent in conventional stereo where direct and reverberant sound come from the same direction. Madsen [9] suggested placing a loudspeaker on each side of the listener and feeding these the same information as the front loudspeakers but delayed 2.5 -10 ms. He reported performance equal to 4-channel recordings of the time and an enlarged listening area.

Hafler [10] suggested that ambience could be extracted by sending only the difference signal of the stereo pair to a rear loudspeaker. This difference signal could be expected to contain more reverberant, incoherent information and was to be delayed acoustically by placing the loudspeaker far behind the listener. Hafler's complete system, also with a derived centre channel has much in common with what is now known as the Dolby Surround system.

It has been established that lateral reflections in concert halls are largely responsible for our impression of spaciousness [11]. The surround channel offers an opportunity to recreate these important reflections thus enhancing the realism of the listening experience. Recent research has confirmed the validity of this approach [12], [13].

Discussion

Dolby Surround is better equipped than conventional stereo to reproduce the original acoustics of the recording venue. Some consumer decoders employ DSP to artificially generate reflections and reverberation in the playback environment. (See for example [14]). This can be very helpful for many existing recordings. However, creating a successful illusion depends a great deal on the spacial content of the original recording and on the skill of the consumer in selecting appropriate program parameters.

If recordings are properly encoded for Dolby Surround playback, control of the spatial content of the recording may be left in the hands of the recording producer. The philosophy of reproduction as opposed to artificial generation is generally more in line with the goals of acoustic music recording.

THE MOTION PICTURE MATRIX

Background

A surround sound system for movie theatre installations has several requirements. A wide sonic image is needed to match the large screen visual image but centred information, particularly dialogue and on-screen action, must stay centre for viewers the entire breadth of the theatre. In addition a surround channel is desirable for sound effects and ambience. This draws viewers into the action and places them in the acoustic space of the scene.

However, industry standards and technology dictated that the audio would be delivered on the optical sound track of the film. Splitting this track into two discrete channels was the practical limit for the medium.

A 4:2:4 matrix was adopted from the then recent quadraphonic technology [15]. This phase/amplitude matrix (described below) allowed four channels to be encoded onto the two optical sound tracks and then decoded on playback in the theatre.

The loudspeakers were deployed as shown in figure 1. The L and R speakers allowed for stereo sound effects and music and the C speaker for well-defined centre dialogue and other on-screen sounds. This left a single channel for the S signal To improve diffusion and audience coverage many loudspeakers were arranged in a U-shape to the sides and behind the audience. To ensure that everyone in the audience localized front originating sounds as coming from the screen, the S channel was delayed relative to the front speakers. Additional processing on the surround channel is described below.

The Encoder

Figure 2 shows the basic Dolby Surround encode scheme. Signals intended for the L and the R playback positions are sent directly to the Lt (left total) and Rt (right total) channels respectively. The C signal is reduced by 3 dB and sent equally to Lt and Rt. The S signal is also reduced 3 dB and sent equally to the Lt and Rt but in anti-phase; that is the Rt surround component is recorded 180° out of phase with the Lt component. Note that the signal sent to Lt is -90° and to Rt +90° to achieve the total 180° phase shift. This is done to avoid stereo compatibility problems with certain panning "moves". [16]

Due to limitations of the optical sound track and duplicating media additional processing is applied to the S signal. The high end rolls off above 7 kHz to reduce noise in the S speakers and to reduce C track bleedthrough caused by azimuth errors. A modified B-type noise reduction is also applied. This action compresses the S signal approximately 5-6 dB to further guard against artifacts.

Finally the S channel is also rolled off below 100 Hz. In the theatre large subwoofers are used to handle low frequency information. The roll off also helps protect the generally smaller S speakers.

The " Φ " on the L and R signals are all-pass phase shift networks. These are required to facilitate panning through the "interior" of the pan locus, that is from the C position to the S position.

Discussion

The heavy processing in the encoder is required to cope with limitations of analogue film media and transfer stages. Acoustic music is generally delivered to the consumer on digital media where problems of noise and proper azimuth alignment are insignificant. Much of the encoder's processing will have a negative impact on ideal music reproduction. This will be discussed further below.

For acoustic music recording the surround encoder can be completely bypassed. The stereo microphone systems themselves are used as encoders since they contain the required amplitude and phase information.

The Decoder

Figure 3 shows a Dolby Surround decoder. The Lt signal goes directly to the L output and Rt to the R. Lt and Rt are summed to create the C output.

The S channel is derived by subtracting Rt from Lt and as with the encoding stage it undergoes more extensive processing. There is a low pass filter at 7 kHz as well as complementary modified Dolby B decoding. In addition the S channel is delayed relative to the front channels.

As shown in figure 4 there is no loss of separation between the L and R channels since these are discrete channels. Nor is there any loss between C and S; the S signal (L-R) completely cancels in the C (L+R) and *vice versa*. However, between adjacent channels there is only 3 dB of separation -- rather unsatisfactory performance.

This problem is dealt with in one of two ways depending on the sophistication of the decoder. Consumer models labelled simply "Dolby Surround" are "passive" decoders. In these more basic systems the centre channel is eliminated improving separation and width across the front soundstage. In addition the delay and the low pass filtering on the S channel helps to improve subjective separation between L and S, and R and S. (See figure 5.)

All professional decoders and those consumer decoders labelled "Dolby Pro Logic" use a type of steering logic to improve separation between channels. The basic block diagram of the Pro Logic decoder is shown in figure 6 and the adaptive matrix is seen in figure 7. Shure's Acra Vector Logic decoder employs a slightly different approach. However it has been the manufacturer's intention to create an active decoder for use with Dolby Surround encoded software [17] and it has been considered successful in this regard by reviewer's of hi-fi equipment [eg 18].

Rather than gain riding techniques which were used in the quadraphonic era, more subtle cancellation techniques are used in Pro Logic. For example to reduce C signal leakage into the L, the decoder takes the R signal, inverts its polarity and blends it in to the L. Since the C component is equal and in phase in both the L and R channels this reduces the level of C in L. (See figure 8).

Dominant sounds are focused to their main point of origin and non-dominant sounds are redistributed among the remaining channels. This constant power scheme ensures that the overall loudness of the programme remains the same. The system depends a good deal on psychoacoustics, particularly masking. The idea is that a dominant sound limits the listeners ability to detect a change in directionality of non-dominant sounds.

To determine the dominant signal it is necessary to find the **relative level difference** as opposed to the absolute levels of the signals. Thus the logarithm of each signal (L, C, R, S) is derived and subtracted one from the other to produce a logarithmic control voltage. Signal dominance can be illustrated by plotting the L/R pair and the C/S pair as an X-Y coordinate as shown in figure 9. The signal dominance is a vector quantity with the magnitude representing the relative dominance and the angle, the encoded direction of the signal.

Pro Logic steering has three different levels of operation depending on the degree of dominance -- as opposed to the absolute level -- of the signal. At higher levels of dominance the decoder uses very fast attack times allowing it to provide directional enhancement for sounds occurring in rapid succession. At lower levels of relative dominance the decoder shifts to a slower, yet fully operational mode. This helps maintain a solid soundstage and avoids a blurry image which the fast mode may create. Finally, if no signal dominance can be detected the decoder goes into a relaxed state and no directional enhancement is implemented. Figure 10 shows the separation map produced by the Pro Logic decoder in the ideal situation of a purely dominant signal.

STEREO MIC TECHNIQUES AS SURROUND ENCODERS: THEORY

Introduction

Stereo microphone techniques, while generally used with the intention of two-loudspeaker stereo playback, actually gather much of the information required for Dolby Surround playback. In this sense a stereo microphone system may be considered as the encoder in a Dolby Surround encode/decode process.

For acoustic music recording the ideal microphone system would encode the direct sound sources and early reflections from the stage area to the front soundstage: L, C, and R. Ambience from the hall would be encoded toward and into the S channel. The mono surround channel is not capable of phantom imaging to the sides or behind the listener and in any case the best use of the S channel is for helping to recreate the spatial ambience of the recording venue. Side wall reflections are particularly important here since they are primarily responsible for creating a sense of spaciousness.

We have seen that the decoder detects position L to R across the front of the soundstage by amplitude differences and position from front through the interior to the S by phase differences. (See figure 11.) It is important to remember that the S channel is mono. As a sound is panned from the right to the surround it does not continue in a circle behind the listener but pulls in toward him/her through the interior and to the sides or rear (depending on layout of S loudspeakers.)

Keeping in mind both the ideal reproduction goal and the characteristics of the Dolby Surround decoder, how successful are the classic stereo microphone techniques as encoders? In this section some of these techniques will be analyzed, strictly in theoretical terms, as to their potential as Dolby Surround encoders. The analysis will be based on a pair's response in an anechoic or free space. In following sections practical experiences with the pairs will be reviewed.

Coincident Mic Techniques

In coincident microphone techniques the diaphragms are placed as close together as possible, theoretically at the same point in space. The only clues to localization with such a system are the intensity differences between the left and right pickups.

XY

Figure 12 shows an XY pair of bidirectional microphones crossed at 90°, the classic Blumlein configuration. A sound arriving from 0° incidence will go to the Lt and Rt channels with equal amplitude and in phase and will therefor encode to the C position. Moving clockwise around the pair sound sources are encoded more to the R and less to the L until a maximum amplitude difference between the two channels is reached at 45°. Beyond 45° the level encoded to the right channel decreases while the level into the opposite polarity rear lobe of the left microphone increases. This results in the sound moving from the right channel only then toward the S until at 90° the sound enters the Lt, Rt pair at equal amplitude but opposite polarity -- the S-only position. Remember that when decoded this signal will move from the R speaker and pull back in toward listener and into the surround loudspeakers.

Continuing around the pair the sound moves from the S position toward the L. From 135° to 180° the sound moves from the left channel back toward the C. A sound di ectly behind the pair enters the opposite polarity lobes of the two mics with the same amplitude. This is seen by the Lt, Rt pair as a common polarity, equal amplitude signal and thus encodes to the C. The pattern continues in a symmetrical fashion around the remainder of the encode positions.

In any coincident system all sounds arrive at the Lt, Rt pair in phase. Microphones with rear lobes encode this information with common or opposite polarity. This means that with coincident systems sounds are encoded either with 0° or 180° phase angle; other phase angles are not possible. Coincident pairs using microphones with no rear lobe -- such as crossed, zero order cardioids -- are only capable of encoding amplitude differences. Since there are no encode points with opposite polarity we could expect to hear nothing in the S channel when such a pair is decoded.

Insight into other combinations of angles and polar patterns may be had by examining theoretically equivalent MS pairs.

MS

Two-channel stereophony is usually thought of in terms of left and right channel information, as in the case of the left and right microphones of an XY pair. However, the stereo signal can also be described in terms of the sum and difference signals, that is L+R and L-R. The analogous microphone technique here is the MS (Mid, Sides) pair where the forward facing M microphone gathers the L+R or mono sum information and the sideways facing bidirectional mic collects the L-R or stereo difference information.

These basic principles of intensity stereo were first described by Blumlein in the early 1930's [19]. He showed how MS pairs could be converted to XY pairs and vice versa using simple sum-and-difference matrices.

Streicher and Dooley [20] have published an exhaustive list of MS to XY conversions showing equivalent polar patterns. Hibbing [21] uses graphs to illustrate theoretical recording angles and stereo imaging performance of XY and MS techniques. He also makes the case that in practice MS techniques can provide better results.

Julstrom's analysis [22] is more easily digested and is of particular interest to this study. It should be noted that this analysis is mathematically computed and assumes ideal first-order microphone polar patterns and perfect coincidence of the microphone diaphragms. Practical limitations not considered here will be discussed below.

Relating XY to MS

Conversion between XY and MS pairs is possible if all first-order polar patterns are available for each microphone of the XY pair and for the M of the MS pair. Any first-order polar pattern can be described by its polar equation A + B = 1, where A is the mic's pressure component and B is its velocity component. Figure 13 (from [22]) is a conversion chart showing all practical first-order coincident pairs. The polar pattern of the M mic is plotted on the vertical axis. The relative level of the S mic is shown on the horizontal axis.

Solid lines show an XY pair of microphones with a particular polar pattern -- eg crossed cardioids -- as the angle between the two mics is varied from 0° toward 180°. The dashed lines show a particular angle between an XY pair, eg 90°, as the polar pattern of the mics is varied from omnidirectional to bidirectional. Polar patterns or angles not specifically shown can be interpolated from the nearest data.

XY pairs are shown by the intersection of two lines and the equivalent MS pair can be read from this point. For example the solid line labelled crossed cardioids and the dashed line labelled 120° intersect at a point showing a Mid mic with a pattern between subcardioid and cardioid and a relative S level of about -5 dB. Working in the other direction an MS pair with a supercardioid Mid and a relative Side level of 0 dB would be approximately equivalent to an XY pair of crossed hypercardioids at 120°.

Polar Diagrams

The stereo polar diagrams of encoding positions shown in figure 14 (from [22]) are all based on MS pickup patterns. Equivalent XY pickup patterns may be read from the chart in figure 13. The original paper goes into some detail on the derivation of the encode positions on the polar diagrams. It is sufficient to note here that the encoding positions shown will not exactly equal their decode positions in Dolby Surround. However, the data allows for a valuable comparative analysis.

The view of the polar diagrams is from above with 0° incidence at the top of the plot. The sensitivity of a given pair to any direction in the horizontal plane is shown relative to the unit circle (equal to the on-axis sensitivity.) The left-hand column shows the encoding's Mid pattern which defines the pair's monophonic pickup.

The dots on the unit circle show positions which, prior to encoding, were located at 22.5° intervals around the microphone pair. Thus the dots give an indication of encoded angular distortion. The large dots on the unit circle correspond to the principal encoding positions of L, C, R, and S (labelled "(B)" for "back" in these diagrams).

The sensitivity curves have areas of thicker and thinner lines. The thicker lines between the L and R encode points define the recording angle of the pair. All sounds arriving within the recording angle arrive in phase at the Lt, Rt pair and will be reproduced on the front soundstage in Dolby Surround.

Except for those employing a cardioid Mid pattern, all MS pairs have an in-phase segment to the rear of the pair. This rear-facing recording angle will encode ambience from the back of the hall to the front soundstage. For mid patterns with A > 0.5 -- eg omni and subcardioid -- the ambience is encoded to the side it originated from. For mid patterns with A < 0.5 the ambience is encoded in the opposite channel.

The front and rear pickup angles are equal for pairs with an omnidirectional or a bidirectional Mid pattern. In all other cases the rear pickup angle is smaller and information it encodes is expanded on playback across the front soundstage.

The thinner lines define "opposite polarity recording angles". Sounds encoded at these points will be reproduced between the L or R speaker and the surround position. Sounds arriving at exactly the angles of incidence labelled "(B)" will reproduce only in the S channel.

From the conversion chart we see that a pair of crossed bidirectionals at 90° is the same as a bidirectional Mid mic and a relative Side level of 0 dB. Figure 14 shows results for this combination identical to those discussed above. Note that the Blumlein pair is the only configuration with equal sensitivity at all angles of incidence in the horizontal plane. In this case the sensitivity curve is equal to the unit circle.

Determining the suitability of a coincident microphone technique as a Dolby Surround encoder is now a simpler task. According to the previously stated goals we should look for a recording angle which keeps the direct sound sources within a front pickup angle, there should be minimal or no rear in phase pickup angle, and information to the sides and/or rear of the pair should encode to the S position. Unfortunately no such pair exists though one or two come reasonably close to meeting the requirements. This will be discussed in more detail below.

Spaced Microphone Techniques

By setting aside some of the practical limitations and concentrating on the theory it is possible, with some clever thinking and some good computer software, to produce a very helpful analysis of coincident stereo microphone techniques. This is a much more difficult task for techniques using spaced microphones.

Where coincident techniques provide only for phases angles of 0° or 180°, these and all other phase angles are possible with spaced techniques. A surround encoding polar diagram is no longer a straight-forward matter since the phase angle depends on many factors.

Spaced Omnis

Figure 15 helps to illustrate the situation for two spaced omnidirectional microphones. Sound sources located at a point equidistant from the two microphones will encode to the C position. Consider now two sound sources located 45° off-axis this centre line, one close to the microphones and one distant. The more distant sound will have a much smaller time-of-arrival difference than the closer sound and will pull toward the centre.

In cases where the spacing between the microphones is large compared to the distance between the microphones and the sound source there will be an amplitude difference as well as a time difference at the Lt, Rt pair. This is the case with the closer source in the above example. The additional amplitude difference will help to place this source more definitely at the L encode position. There is no such amplitude difference for the further source. All this leads to considerable imaging distortion depending on the distance of the sound from the microphone pair. We can see that image distortion will be different for sound sources at different distances and angles from the microphone array.

The situation for the Dolby Surround decoder is further complicated by the fact that the phase difference arriving at the Lt, Rt pair varies with frequency. The phase difference will be 0° for frequencies with wavelengths equal to the pathlength difference to the left and right microphones. These frequencies will encode to the front soundstage. A frequency one octave below this will arrive at the Lt, Rt pair 180° out of phase and thus encode to the S position. Another octave below this the sound is 90° out of phase and encodes exactly half way between the front soundstage and the S channel.

Thus the fundamental of a given musical note and its overtones will be encoded at different positions in the front to surround axis. The encoded position will be different for a different note and the relationships will change depending on the instrument's angle of incidence and its distance from the microphones.

With all phase angles possible and equal sensitivity to sounds from all directions spaced techniques will encode direct and ambient sounds to all positions.

Near Coincident Techniques

Near coincident techniques generally use a pair of directional microphones positioned at some angle and spaced by a small amount, typically on the order of the spacing between a person's ears. ORTF -- cardioids at 110°, 17 cm apart -- and NOS -- cardioids at 90°, 30 cm apart -- are well known examples but many combinations of angles, spacings, and polar patterns are possible.[1]

Near coincident techniques combine many of the advantages and some of the disadvantages, of coincident and spaced microphone techniques. There will be fewer complete phase reversals within the audio spectrum since compared to typical spacings for AB techniques the first opposite polarity situation occurs at a much higher frequency. Direct and reverberant sound may be encoded to all positions but the overall room response will be reduced to a degree determined

---

[1] Williams [26] has produced a thorough theoretical review which covers all stereo microphone systems but which is particulary useful in demystifying near coincident pairs. His data is quite specifically dedicated to conventional stereo reproduction but the analysis provides useful insight.

by the directionality of the microphones. The intensity differences created by the microphones' polar patterns will be particularly helpful in directing information to its correct position left to right.

STEREO MIC TECHNIQUES: PRACTICAL APPLICATION

The theoretical review of stereo microphone techniques has shown some of the potential of these techniques to act as encoders for the Dolby Surround format. We will now consider the practical performance of the complete system -- encoder and decoder -- based on critical listening to compact discs and recordings made using classic stereo microphone techniques during the course of this research.

Decoder Limitations

It will be worthwhile at this point to consider some of the performance limitations of using the Dolby Surround decoder for acoustic music playback. Initial critical listening to commercially available compact discs revealed a number of deficiencies. The most immediately obvious problem was the reduction of soundstage width which occurs when switching from stereo to surround. Almost all recordings lost some L to R width in surround while a number became almost completely mono. Recordings which subjectively seemed more spacious and had wider imaging performed best in Dolby Surround in this regard.

In several recordings the perceived width in Dolby Surround varied with the loudness of the program. In quieter passages the soundstage was almost mono but in louder passages the width would increase somewhat. On other recordings certain sound sources held their width rather well while others collapsed. For example violins would pull in to the centre but the lower-voiced instruments and ambience would remain in place.

Recordings varied greatly in how much information decoded to the S. Presence of a S signal varied from unnoticeable to moderate. Recordings in the latter category seemed to send a mix of direct and reverberant sound to the S. A few recordings sent plenty of signal to the S but the front soundstage collapsed to C. This tunnel effect makes for particularly unpleasant listening. In fact these recordings sounded much more spacious and pleasant in conventional stereo.

Switching to Dolby Surround changed the frequency response as well. For all recordings there was a perceivable softening of the high end and for some the low end increased when switched to Dolby Surround

Considerable improvements over conventional stereo performance were also noted in a number of recordings. One predictable though very impressive result was the stabilizing of the front soundstage at off centre listening positions. In all cases the centre image stayed very firmly in place, even for listening positions far off centre. For recordings which maintained reasonable width in Dolby Surround the confident, unambiguous soundstage made for very comfortable listening.

The recordings which decoded most successfully sounded very good in stereo and gained more than they lost when switched to Dolby Surround. Rather than looking in on the room the listener feels enveloped by sound. The front soundstage holds very well wherever the listener moves in the room, and in some cases sound sources move out into the room rather than staying on a flat line between the loudspeakers.

These preliminary listening sessions showed that recordings intended for stereo playback give widely varying results in Dolby Surround. Aware of the performance limitations of the decoder we are better able to select a suitable "encoder".

Of the problems described above the question of sufficient width and the quantity and quality of the S channel may be addressed by the microphone technique.

Coincident Techniques

When first working with coincident techniques the major preoccupation was with maintaining a usable width across the front soundstage. To ensure sufficient width we should look on figure 14 for polar diagrams showing sensitivity levels for the L and R encode positions equal to or greater than that of C. For Mid patterns with A > B fairly wide pickup angles are possible meaning that the pair could be worked relatively close to the source (eg see subcardioid with -1.15 dB side level.) For Mid patterns with A < B only smaller pickup angles are possible (eg see Blumlein.)

Next it is important to choose a pickup angle which will place the widest sound sources at the extreme L and R encoding positions. Conversely one could place the musicians to take maximum advantage of the chosen pickup angle. However one must be certain that the musicians on the far L and R do not move beyond the front pickup angle or they will spill into the S channel, again reducing the width of the image. This is particularly critical for setups involving a bidirectional Mid since the S encode position is relatively close to the edges of the pickup angle.

If the pickup angle has been set then the distance of the pair from the musicians has already been determined. If the pair is moved back the pickup angle must be narrowed or the image will collapse to C. If it is moved forward the angle must be widened or images will spill into the S. Of course if the pickup angle is adjusted one must be careful not to allow a situation where the C encode sensitivity is greater than the L and R.

All this and we have not yet considered the S channel. As we raise the side level to increase the amplitude difference between the L and R encode positions, the level of signal moving toward S rises. Approaching useful in-phase pickup angles the opposite polarity region's sensitivity becomes quite high, typically greater than for front sound stage encode positions.

Practical experience indicates that better results are obtained when the sensitivity of the actual S encode position ("B" in figure 14) is on par with the sensitivity of positions on the rest of the "opposite polarity pickup angle". This is the case with, for example, Blumlein and hypercardioid Mid with +.86 dB Side level.

For cardioid and lower directivity Mid patterns, however, the actual S encode position is either very low in level relative to other encode postions or does not exist. With these pairs a reasonable result can sometimes be achieved for Dolby Surround reproduction. The decoder is able to spatially separate the direct and reverberant information leading to better intelligibility. However, when we switch to stereo the ambience folds back in to the front often producing a heavy glut of phasey room resonance which masks the direct sound.

Figures 16a through f [from 22] show the relative levels of reverberation for different encoding positions. The various Side levels correspond to those shown in figure 14. The supercardioid Mid is now included.

According to Gerzon [23] the best distribution of reverb for *stereo* reproduction would be shown on these graphs as a line which is "either flat or very slightly biased toward the centre." For Dolby Surround encoding we might err towards a response with a dip in level at C and higher levels at S (labelled "(B)" in these graphs). This would help to keep information out of the heavily favoured C channel and force some more into the surround. Supercardioid, hypercardioid and bidirectional Mids with higher Side levels look promising and this corroborates well with the author's experience.

In practice it is difficult to achieve a pleasing balance of direct and ambient sound without compromising some other parameter particularly since the decoder is unforgiving of anything but the most accurately positioned and widest pickup angle. Once the pickup angle is set the ambient signal at this point in space must be used. If this is unacceptable the angle must be adjusted but without compromising the L and R channel separation. An ideal result is almost impossible to achieve except in highly favourable acoustic conditions and with time to experiment with placement of microphones and sound sources.

A Note About MS

It is often recommended that MS techniques be used rather than their XY equivalents [20], [21]. Practical microphones exhibit frequency response anomalies which vary with angle of incidence. With MS techniques, centre-stage sound sources arrive on-axis where the frequency response of a microphone is more likely to be accurate. Directional microphones typically become increasingly omnidirectional at lower frequencies. This can lead to a glut of low frequency information in the mono centre channel decidedly reducing any sense of spaciousness. It is easier to design a bidirectional mic with a consistent pattern at all frequencies. This means that the L-R signal in an MS pair will be picked up more accurately.

There are some trade-offs, however and they should be considered. The presence of a matrix means that more electronics are in the recording chain. In the heat of a recording session it is more difficult to keep track of what parameters are being affected when the relative Mid and Side balance is changed. Adjusting XY pairs is a much more intuitive process. Julstrom's analysis should be very helpful in making MS techniques more "user-friendly".

The benefit of remote control possibilities with MS pairs is often overstated. A small change in the relative side level has an effect on every parameter: the pickup angle, the way reverberant information is encoded, the relative sensitivity at all encode points, etc. A change in the relative M to S balance must therefor be accompanied by a change in either microphone position and/or sound source arrangement. These same problems are present in conventional stereo encoding. However, as we have seen, finding the optimum setup is particularly critical for Dolby Surround encoding.

Spaced Techniques

Spaced omnidirectional microphones are used as a main pair by many classical music recording companies. There are a number of benefits which also apply to Dolby Surround encoding. Pressure transducers have superior low frequency response, are less prone to off-axis colouration, and have superior transient response.

Spaced omni techniques are said to give a more spacious result. It has been argued that this spaciousness is really "phasiness" [24] but clearly there are also other factors at work. The *level* of bass response has been found to affect our perception of spaciousness [25] Since pressure transducers do not roll off at low frequencies one will perceive more bass and thus more spaciousness with spaced omnis The reflections and reverberation from all over the recording venue are encoded with a more consistent (less coloured) frequency response and with the same spatial "accuracy" as that of the direct sound.[2] By contrast most coincident systems reject certain information from the hall and encode the ambience to monophonic "spots" on the front soundstage or with a largely distorted stereo spread. This fact alone makes it easier to achieve a musically acceptable sound with spaced omnis. Ease of use is a considerable benefit and one not often offered by those who swear by the technique.

Dolby Surround decoding helps to stabilize centre images in recordings made with two spaced omnidirectionals. A third centre mic is often used and provides the engineer with some more control over the level sent to the S channel. The engineer is cautioned not to under use the centre mic -- in an attempt to improve width in Dolby Surround -- to the extent that the stereo image is compromised.

Width is improved if some form of amplitude encoding can be implemented. This is possible by at least two methods. If the microphones are placed relatively close to the sound sources there will be amplitude differences between the microphones. This implies that the sound source does not have significant depth and that a close perspective is aesthetically acceptable.

Microphones with some directionality will obviously help matters and this is possible with pressure transducers. The Neumann M50, prized by many classical recording engineers, is

---

[2] See Williams [26] for more on the stereo encoding of ambience.

a pressure transducer with its diaphragm mounted in a sphere. This results in an increasingly directional response at high frequen s. Woszczyk's spherical diffractors [27] have the same effect and can be fitted to pressure microphones with significantly less "grill and body work" than the M50. Such a microphone maintains its extended low frequency response and superior transient response while acquiring a shelf boost for frequencies above 1 kHz for on-axis sounds and a low pass filter acoustic response for off-axis sounds.

Near coincident techniques with their increased directionality offer improved amplitude encoding.

In some cases it may be desirable to create a lush surround-sound "feel" as opposed to a truly distinct room sound. Spaced omnis do the job very nicely. With random phase in the Lt, Rt pair and little or no amplitude encoding the steering logic will be suspended.

Spaced techniques encode direct and ambient sound with equal sensitivity to all encode positions. Obviously this technique can not meet the stated goals of keeping sound sources on the front soundstage and ambient information to the rear. Nonetheless spaced omni recordings can give a very satisfactory result when decoded. The delay and the high frequency roll off in the S channel help to fool the ear into perceiving the S signal as ambience.

SOLUTIONS -- THE MICROPHONE AS ENCODER

Some of the classic stereo microphone techniques are capable of providing a pleasant result when played back by a Dolby Surround decoder. However, we have seen that these techniques impose many limitations on the engineer and often some aspect of the recording is compromised to achieve another effect. A small number of coincident techniques are able to encode directional information in a useful manner. However, these techniques are very limited in their flexibility and it is very difficult to achieve satisfying results. Spaced and near coincident techniques are easier to implement but are not able to encode direct and ambient information to specific positions

Clearly what is needed is improved flexibility, in particular individual control over the direct sound encoded to the front soundstage and the ambient sound encoded to the surround.

The "S-Only" Encoder

One solution is to use one microphone system to encode the front soundstage and another to encode the surround information. The "S(urround)-only" encoder is the S(ides) component of the MS technique. By removing the L + R component we are left with only L - R, a signal which appears with equal amplitude and opposite polarity at the Lt, Rt pair. The laterally-oriented bidirectional microphone will favour side wall reflections which have been shown to be particularly important to our perception of spaciousness. The microphone is placed farther back in the hall at a position which will pick up an appropriate ambient sound.

A microphone of any polar pattern could be split, polarity inverted and used as an S-only encoder. An omnidirectional microphone will collect reflections from all over the room and will have a better bass response (though this could become a liability if room rumble is a problem.) The choice of polar pattern will be largely determined by acoustic conditions.

The *s*ide- (or *s*urround)-only mic gives the engineer easy control over the type and amount of reverb in the hall. Now that a surround encoder for ambience is available, the engineer requires an amplitude encoder for the front soundstage. The most straight forward solution is to use a pair of coincident cardioid microphones since this system is purely an amplitude encoder. However, as we have seen it is difficult to achieve any reasonable width with this system. More options for a front stage encoder will be discussed below.

The recording engineer now has great control and flexibility over the sound. The front soundstage encoder can be independently controlled and positioned in the optimum position for picking up direct sound and early reflections from the stage area. Likewise the surround encoder can be placed for optimum sound quality and balanced independently. The original goal for the surround system has been met.

Compatibility

Typically a relatively small number of listeners will hear the recording in surround so it is important that the encode scheme deliver good results in stereo and mono. In stereo the front soundstage is played back as usual between the two loudspeakers. Being an opposite polarity signal the ambience is quite diffuse and unlocalizable in stereo. It is perceived as coming from some direction away from the plane of the loudspeakers. In effect one of the most bothersome limitations of conventional stereo is largely overcome. Of course opposite polarity ambience is found in many conventional stereo encodings but its implementation sometimes seems rather arbitrary, almost accidental.

A further advantage of using separate front and surround encoders over a single coincident system is that the ambient information is not coherent with the direct sound. This separates the direct signal from its reverberation temporally leading to improved intelligibility.

In mono reproduction the ambience in the S signal is completely cancelled. While this can improve clarity on a low-fi system the spatial *perspective* of the recording is altered, sometimes drastically. The problem is easily remedied by "leaking" some of the ambient surround information into the front soundstage. This technique is frequently used in the movie industry albeit with the surround encoder. With the S-only technique this is accomplished by adding an omnidirectional M to the distant mic at a low level. Surround reproduction is not significantly compromised. Some information from the front and rear of the room will be encoded to the front soundstage. However, we can interpolate from figure 14 that an omnidirectional Mid with a high Side level has very high sensitivity at opposite polarity encode positions and quite low at ones with common polarity.

Enhanced Sides

The Enhanced Sides technique [28] makes use of two coincident cardioid microphones at 180° with the polarity of one of the microphones reversed. Figure 14 includes a polar diagram for two cardioids with common polarity at 180° (omnidirectional Mid with a Sides level of 0dB). If we invert the polarity of one of the microphones there will be no common polarity signal at any angle of incidence. If we were to redraw the stereo polar diagram using Julstrom's model the entire sensitivity curve would be a thin line and the C encode positions would become S.

The ES technique negates many of the limitations imposed by the encode/decode matrix. We can see that there are two opposite polarity encode angles, one in front of the pair subtending an angle of 180° and another in the rear, also 180°. No signal will be encoded to the C position. At the same time sensitivity at the L and R encode positions is quite high. The small dots show that there is wide angular distortion which is very helpful in forcing energy toward the S position.

This system encodes sounds arriving from the sides of the mic very firmly at the L and R encode positions. This is because of the high difference in amplitude created between Lt and Rt channels. Sounds arriving from in front of the mic and ambience from the rear are encoded with a wide spread between the L and R encode points but toward the S position instead of across the front soundstage. Unlike the situation with the S-only mic, directional information about the ambience is encoded, reinforcing directional cues from the amplitude encoder.

The fact that front originating sounds are encoded at the S position would appear to be contradicting the desired directional encoding. In fact this "distortion" is very helpful in pulling instruments away from the plane between the front loudspeakers. The instruments seem to be in the listening room instead of suspended on an unseen wall in front of the listener.

The ES technique is very effective in creating a stable ambient environment stretching from the L and R speakers and around to the surrounds. There is now a comfortable opening at the C speaker where centre stage sound sources can be positioned. This may be accomplished in a number of ways as discussed above.

Compatibility

Again in stereo the ambience is "localized" out away from the loudspeakers, effectively widening the overall image and improving intelligibility.[3]

Instead of cancelling in mono reproduction the ES technique produces a laterally-oriented bidirectional pickup. The overall ambience level is lower improving clarity but the original perspective and some important room information are maintained in mono.

Some Comments on Application

The ES technique has proven very successful in providing a solid spatial environment in which to build up a soundstage. It can be used further back in the recording venue to encode sidewall reflections and reverberation. Direct sound can also

---

[3] See [28] for additional psychoacoustic advantages of the ES technique in stereo reproduction.

be encoded to great benefit by closer placement of the ES microphone. The image of the source will be very wide and will pull out into the listening room away from the plane of the loudspeakers.

Boundary placement of the microphone improves signal-to-noise ratio and increases the system's directivity. Diffuse sound level rises 3 dB (addition of incoherent signal) and direct sound by 6 dB (addition of coherent signal).

More than one ES microphone can be used in the same recording. A setup might use one boundary layer ES mic on the floor in front of the musicians, another against a wall behind them, and another further back in the hall for reverberant pickup. Or more than one ES microphone could be used at different locations in the hall to encode more complex information.

An experimental piano recording using a combination of techniques produced very good results. A pair of pressure microphones fitted with diffractive attachments and setup in a quasi-ORTF configuration was used to encode the centre soundstage. An ES mic further back encoded sidewall reflections to the hard L and R positions and general ambience toward S. Finally an S-only mic was placed even further back in the hall to encode the more distant reverberation only into the S channel. Using an S-only mic farther back in the hall ensures that the ambience collected here will not "compete" for localization space with the sidewall reflections encoded by the closer ES mic.

A number of compact discs which have been recorded using the ES, Baffle Boards and other techniques are about to be released by the McGill Records label.

Other Useful Techniques

Woszczyk [29] conducted tests with various stereo microphone systems set in the diffuse field of a concert hall with speech and music sound sources. Listeners evaluated the results for spatial involvement and imaging accuracy in conventional stereo and Dolby Surround playback. The ES technique was included in this study and its many benefits were confirmed.

"Baffle boards" -- two pressure transducers mounted on the sides of high density particle boards angled at 45° -- were also very highly rated. This system, which should be used close to the source, can spread direct sound sources wide across the front soundstage and give a very convincing spatial effect in the surround channel.

One limitation is reduced flexibility in controlling direct and ambient pickup. This can be helped by using the baffle boards in combination with other techniques such as ES. Timbres and reverberation were sometimes judged to be coloured in this system. Baffle techniques with lower diffraction effects such as the Sphere microphone [30] show promise in resolving this problem.

During the preliminary listening sessions some ambisonic UHJ-encoded CD's from Nimbus gave interesting results. Nimbus uses a "home made" Soundfield microphone for their recordings. [31] A forward-facing and a side-facing bidirectional mic, and a pressure mic are all set up in a coincident array. With the addition of the omnidirectional pickup any first order Mid pattern can be achieved by properly combining it with the forward facing bidirectional. These recordings are encoded with the intention that they be played back through an Ambisonic decoder for surround sound. [32] There are some side effects through the Dolby decoder but reproduction is definitely superior to stereo.

These recordings have very wide images in stereo and this width holds up quite well in Dolby Surround reproduction. As a bonus the S channel often contains only reverberant information. In stereo these recordings can sound phasey and have unstable images. These problems are largely overcome when played through the Dolby Surround decoder. Of course these recordings should ideally be decoded with a proper ambisonic system. The point made here is that coincident mic techniques can be made to give worthwhile results.

Spot microphones are essentially amplitude panned mono encoders. They can give helpful extra cues to the dominance detection circuits. Spot microphones would generally be panned wider than for typical stereo applications to help counter the decoder's inclination to pull sounds to the centre. Problems with phase are largely avoided by the large distance between the close and the main pickup microphone.

Obviously there is lots of room left for experimentation and combining various existing techniques can produce excellent results. John Eargle's recordings for Delos stood out for their Dolby Surround performance in the preliminary listening sessions. Eargle generally uses a near coincident main pair, flanking omnidirectionals, occasional close support pairs, and an ambient pair of cardioids looking at the side walls and spaced about 1.5 m apart. [33]

Additional Practical Comments

Ideally a decoder and surround monitoring system would be available for the recording session. This allows the engineer to check for pumping and other steering artifacts which do not appear in stereo and to ensure sufficient front stage width. Of course it is useful to hear what is happening in the surround channel. However, the recently developed techniques described above are quite reliable in their action. After some practice with these "encoders" and with the decoder a competent engineer can produce predictable results while monitoring in stereo. The engineer can rely on his/her ears and mono checks to work in stereo and be confident of a good result.

One of the great benefits of working in Dolby Surround is that the engineer becomes much more conscious of the important L-R component of his/her stereo recording. This is where the ambience of the hall should be encoded since a good spatial impression is much more effectively portrayed in surround sound and in conventional stereo when ambience is proportionally high in the L-R component.

SOLUTIONS -- THE DECODE UNIT

The original Dolby Stereo decoder was conceived specifically for use in cinemas. It is important that Dolby Surround decoders be compatible for proper reproduction of encoded movie soundtracks. However, much of the circuitry of the decoder is dedicated to solving problems inherent in analogue visual media. This can impose unnecessary performance limitations on music-only software particularly since this software is likely to be delivered in a digital medium.

An appropriate solution would be to have a "music" mode on the decoder. This would use the same speaker layout as Dolby Surround but would eliminate unnecessary processing and perhaps introduce some improvements for music reproduction.

Optional control of some parameters on more sophisticated decoders allow listeners so inclined to make adjustments.

Surround Output Processing

Delay

As mentioned above the surround delay can be helpful in making sound in the S channel seem more like ambience, particularly when spaced microphone techniques are used. Consumers should be able to adjust the delay time to a setting appropriate to the individual recording. For example when an ES mic is set back in the hall to collect ambience there is an inherent delay built into the S channel of the recording. The additional delay of the decoder often creates a undesirable echo effect.

LPF

Users should be able to adjust the high frequency roll-off in the surround channel as well. High frequency components which deviate which are partly or completely out of phase will be sent to the S channel and removed from the overall sound. This may have an adverse affect the perceived "openness" of the sound.

However some degree of filtering may be helpful in certain recordings to help disguise direct sound as ambient sound. The electrical roll-off simulates the acoustical roll-off caused by absorbtion in a typical concert hall. High frequency lateral components contribute to the perception of image broadening [34], an effect which may be undesirable in some recordings.

Modified Dolby-B Decoding

This processing merely exacerbates the problem of pumping and image shift, a particularly bothersome effect in acoustic music reproduction (see below). In fact this process is not even compatible with programme material which is conventionally encoded for surround sound. The compressed signal going into the encode side of the matrix does not equal the derived S signal expanded on the output. [35] This processing should be eliminated from the decoder.

Stereo Surround

Many people in the audio community take exception to the idea of a mono surround channel being used for music reproduction, eg [36] and [37]. On most programme material mono ambience simply contributes to an overall collapse toward centre. However if the ambience is properly encoded, as it is by some of the techniques described above, this should be less of a concern.

Some manufacturers use some mild processing on the S output to produce a quasi-stereo effect. In one study many different techniques were tried in an attempt to decorrelate the mono surround and thus improve spaciousness. Complementary comb filtering, time delays, and phase shift networks between left and right surround speakers were all found to be inadequate [38]. Holman reports that a slight pitch shift between left and right surround outputs produces a greatly enlarged spatial impression. This could easily be implemented on consumer decoders, particulary those which decode in the digital domain. However, this is a *simulation* of spaciousness rather than a reproduction and some would have trouble accepting the idea of pitch shift introduced in a music reproduction system.

Another solution might be to devise a special steering logic which would derive additional control voltages from the L, R, and S signals. This could be used to steer information between the L and S toward a left-surround output and that between R and S toward a right-surround output. At least one manufacturer is working on an as of yet undisclosed procedure (different from the one just described) for deriving a stereo surround. [39]

Pumping

On some programme material the steering logic of the decoder produces significant image shift. This is particulary obvious following a strong transient in a highly reverberant environment. If the amount of directional enhancement could be controlled the user could adjust it to suit the programme.

In a digitally implemented decoder the output could be delayed until the signal has been analyzed and the logic sorted out which information was to be steered where. When the audio is accompanying a picture synchronization between the audio and the visual image imposes an upper limitation on this processing time. This limitation does not exist in music-only reproduction and this fact should be exploited.

Derivation of steering logic control signals could include a more rigorous analysis of the audio signal and could take psychoacoustic factors into account. Time constants for the steering logic should also be optimized for typical musical programme material.

Loudspeaker Layout

Even the most elaborate home theatre environment is significantly different from that of the cinema. The loudspeaker setup can be optimized for music reproduction in the home. A number of references offer good discussions on this topic [12], [38], [5]. The general consensus is that if two surround loudspeakers are used that they be placed on either side of the listener rather than behind for maximum spatial impression. (See figure 17).

It is important that the centre loudspeaker be the same distance from the listener as the left and right speakers. In typical setups the speakers are placed in a straight line for practical reasons. Sounds from the centre loudspeaker will reach the listener first and the front soundstage will be even more biased toward the centre. This could easily be compensated for in digital decoders by an adjustable delay on the centre channel output.

It is imperative that the three front loudspeakers be well matched to ensure a smooth front soundstage and even timbres. Surround speakers should match the timbre of the front speakers and should be have a wide directivity characteristic to improve the impression of spaciousness.

Level Controls for C and S Output

It is important that the level controls for the four speaker outputs be properly balanced for accurate reproduction of properly encoded material. However, it may be helpful to supply some level controls to compensate for inadequate programme material. C output level could be reduced to improve width across the front soundstage and/or the S output level brought up increase the ambient level.

DISCUSSION

Dolby Laboratories have always stressed that their system was designed to improve audio accompanying visual media. In the past they have actively discouraged the use of the system for music-only productions [40]. Recently, however, a couple of compact discs bearing the Dolby Surround logo have appeared. The fact that the first disc consisted of remixed versions of movie overtures perhaps made the move easier to rationalize!

Shure, by contrast, seems eager to exploit this and other markets. Their Stereosurround encoder is essentially compatible with the Dolby Stereo system but has been "optimized" for the production of programme material intended for the home environment. The surround channel in this system is full bandwidth and does not include noise reduction [35].

Meanwhile the hi-fi music press has been practically screaming for properly encoded music-only software to play on the elaborate new decoding systems they review [37] [41], [42]. Engineers involved in recording acoustic music are rightfully suspicious of extra circuitry and processing. This circuitry is required on surround encoders to meet the demands of audio accompanying visual media. Here it has been demonstrated that the Dolby Surround Encoder can be completely bypassed.

If one set out to design the ideal *music* surround sound system the result would not be Dolby Surround. Nevertheless, the system offers significant improvements over conventional stereo. Practical reasons for adopting this particular surround sound system have been discussed above. Perhaps the most significant factor is that Dolby Surround has penetrated the home market to a degree far exceeding any other surround system. This will likely pave the way to general acceptance of surround technology and improved systems for music reproduction in the future.

Evidence of this trend can be found today. A number of manufacturers have included "music" modes and some user adjustability in their decoders. Many of the decoder improvements mentioned above are found in the Lexicon family of processors. Some of these units also include processing for binaural recordings, cross-talk cancellation, stereo shuffling, as well as concert hall simulation. (Features of the CP-1 are described in [12]).

Recently a Home THX system has been introduced in an effort to improve compatibility between the sound intended for the movie theatre audience and that reproduced in a home environment. [43] It remains to be seen whether this system will benefit music-only reproduction. Perhaps the systems main contribution will be the concept of standardizing the home listening environment.

CONCLUSION

Dolby Surround is a well-established surround sound format with significant and rapidly expanding acceptance in the consumer market place. The system offers considerable improvements over conventional stereo including a stable front soundstage across a wide listening area and a surround ambience channel. While these improvements are particularly beneficial to acoustic music reproduction very few recordings have exploited the medium.

It has been demonstrated that while some of the classic stereo microphone techniques are capable of encoding for the system, they are either very cumbersome to implement or unable to encode direct and ambient energy with spatial accuracy.

Recently developed microphone techniques overcome limitations of the matrix and are capable of ideal encoding of a stable spatial environment. Recordings made with these techniques offer improved stereo and monophonic performance as well as full surround encoding.

It has been established that no special processing should be used, nor is it required, when encoding acoustic music for the system. Considerable improvement in the decoder technology can be made, particularly in optimising it for musical programmes.

A number of compact discs recorded with these new techniques are currently being released and feedback from the market place should demonstrate the validity of adopting the Dolby Surround system for acoustic music recording.

REFERENCES

[1] Dolby Laboratories Incorporated. Dolby Surround: A Listener's Guide. San Francisco, Aug 1987. S87/6813/7881.

[2] Dressler, Roger. Dolby Pro Logic Surround Decoder: Principles of Operation. Dolby Laboratories Information, S8357/8540/89. San Francisco, 1988.

[3] Sawaguchi, Masaki *et al*. "Surround Broadcasting in Japan," The Proceedings of the AES 9th International conference: Television Sound Today and Tomorrow, Detroit, Feb 1991, pp 163-174.

[4] Theile, Gunther. "HDTV Sound Systems: How Many Channels?" The Proceedings of the AES 9th International conference: Television Sound Today and Tomorrow, Detroit, Feb 1991, pp 217-232.

[5] Theile, Gunther. "Further Developments of Loudspeaker Stereophony," Presented at the 89th AES Convention, Los Angeles, Sept 1990. Preprint #2947.

[6] Torick, Emil. "A Triphonic Sound System for Television Broadcasting." J SMPTE, 92, 8 (Aug 1983) 843-848.

[7] Ausperger, George L. "Loudspeakers in Control Rooms and Living Rooms," Presented at the AES 8th International Conference: The Sound of Audio, Washington, DC, May 1990.

[8] Holman, Tomlinson. "Surround Sound Systems used with Pictures in Cinemas and Homes (Revised)", Presented at the AES 8th International Conference: The Sound of Audio, Washington, DC, May 1990.

[9] Madsen, E. Roerbaek. "Extraction of Ambience Information from Ordinary Recordings." JAES, 18, 5 (Oct 1970): 490-496.

[10] Hafler, D. "A New Quadraphonic System," Audio, 54 (July 1970): 24.

[11] Barron, M. "The Subjective Effects of First Reflections in Concert Halls -- The Need for Lateral Reflections." Journal of Sound and Vibration, 15 (1971): 475-494.

[12] Griesinger, David. "Theory and Design of a Digital Audio Signal Processor for Home Use." JAES, 37, 1/2 (Jan/Feb 1989): 40-50.

[13] K. Ohgushi et al. "Subjective Evaluation of Multi-Channel Stereophony for HDTV," IEEE Transactions on Broadcasting, BC-33, 4 (Dec 1987): 197-202.

[14] Meyer, Brad E. "Magic Space," Stereo Review, 53, 8 (Aug 1988): 60-66.

[15] Scheiber, Peter. "Four Channels and Compatibility." JAES, 19, 4 (April 1971): 267-279.

[16] Sommerwerck, William. "The Dolby Labs Surround-Sound System for Motion Pictures." MCS Review, 4, 1 (Summer 1982): 5-13.

[17] Julstrom, Stephen. "A High Performance Surround Sound Process for Home Video." JAES, 35, 7/8 (July/Aug 1987): 536-549.

[18] Riggs, Michael et al. "Shure HTS-5200 Surround-Sound Processor," High Fidelity, 38, 2 (Feb 1988): 33-35.

[19] Blumlein, A.D., "Improvements in and relating to Sound-transmission, Sound-recording, and Sound-reproducing Systems," U.K. Patent 394,325, June 14, 1933. Reprinted in JAES, 6, 2 (April 1958): 91-98, 130.

[20] Dooley, W. and Streicher, R. "M-S Stereo: A Powerful Technique for Working in Stereo," JAES, 30, 10 (Oct 1982): 707.

[21] Hibbing, M. "XY and MS Microphone Techniques in Comparison," JAES, 37, 10 (Oct 1989): 823-831.

[22] Julstrom, Stephen. "An Intuitive View of Coincident Stereo Microphones." Presented at the 89th AES Convention in Los Angeles, Sept 1990. Preprint #2984.

[23] Gerzon, Michael. "Recording Techniques for Multichannel Stereo." British Kinematography Sound and Television, 53, 7 (July 1971): 274-279.

[24] Lipshitz, Stanley P. "Stereo Microphone Techniques: Are the Purists Wrong?," JAES, 34, 9 (Sept 1986): 716-744.

[25] Barron, M. and A.H. Marshall. "Spacial Impression due to Early Lateral Reflections in Concert Halls: The Derivation of a Physical Measure." J Sound & Vibration, 77, 2 (1981):211-232.

[26] Williams, Michael. The Stereophonic Zoom: A Variable Dual Microphone System for Stereophonic Sound Recording, published by the author, 1990.

[27] Woszczyk, Wieslaw. "Diffraction Effects in High Quality Studio Microphones," Presented at the AES 86th Convention, Hamburg, March 1987. Preprint #2792.

[28] Woszczyk, Wieslaw. "A Review of Microphone Techniques Optimized for Spatial Control of Sound in Television." The Proceedings of the AES 9th International conference: Television Sound Today and Tomorrow, Detroit, Feb 1991, pp 133-142.

[29] Woszczyk, Wieslaw. "A New Method for Spatial Enhancement in Stereo and Surround Recording." Presented at the 89th AES Convention, Los Angeles, Sept 1990. Preprint #2946.

[30] Theile, Gunther. "On the Naturalness of Two-Channel Stereo Sound." The Proceedings of the AES 9th International conference: Television Sound Today and Tomorrow, Detroit, Feb 1991, pp 143-149.

[31] Smith, Howard J. "The Soundfield Microphone." db, 12, 7 (July 1978): 34-36.

[32] Gerzon, Michael. "Ambisonics in Multichannel Broadcasting and Video." JAES, 33, 11 (Nov 1985): 859-871.

[33] Conversation with John Eargle at AES Conference in Washington, DC, May 1990.

[34] Blauert, Jens and Werner Lindemann. "Auditory Spaciousness: Some Further Psychoacoustic Analyses." JASA, 80, 2 (Aug 1986): 533-542.

[35] Gilbert, Mark and Robert B. Schulein. "Stereosurround -- A Compatible Multichannel Encoding/Decoding Process for Audio and Audio/Video Applications." Presented at the 87th AES Convention, New York, Oct 1989. Preprint #2855.

[36] Canby, Edward Tatnall, "Promising Progeny." Audio, 71, 8 (Aug 1987): 18+.

[37] Holt, J. Gordon. "Surrounded by Sound." Stereophile, 11, 5 (May 1988): 5-7.

[38] Holman, Tomlinson. "New Factors in Sound for Cinema and Television", Presented at the 89th AES Convention, Los Angeles, Sept 1990. Preprint #2945.

[39] Conversation with David Griesinger at AES Conference in Washington, DC, May 1990.

[40] Conversation with David Gray in Montreal, Sept 1988.

[41] Ranada, David. "Product of the Year Awards." High Fidelity, 38, 12 (Dec 1988): 46-47.

[42] Toole, Floyd E., and Ian G. Masters. "The Final Illusion: Space and Ambience." Sound & Vision 3,6 (Dec '87): 10-12.

[43] "THX Surround Sound Comes Home." AES Boston Section Newsletter, Feb 1991.

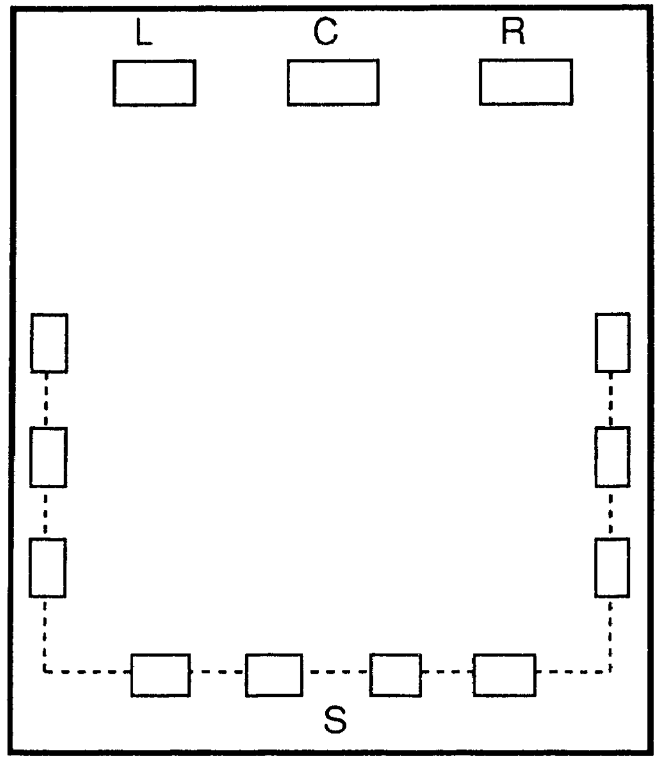


Fig 1. Loudspeaker layout
for Dolby Stereo.

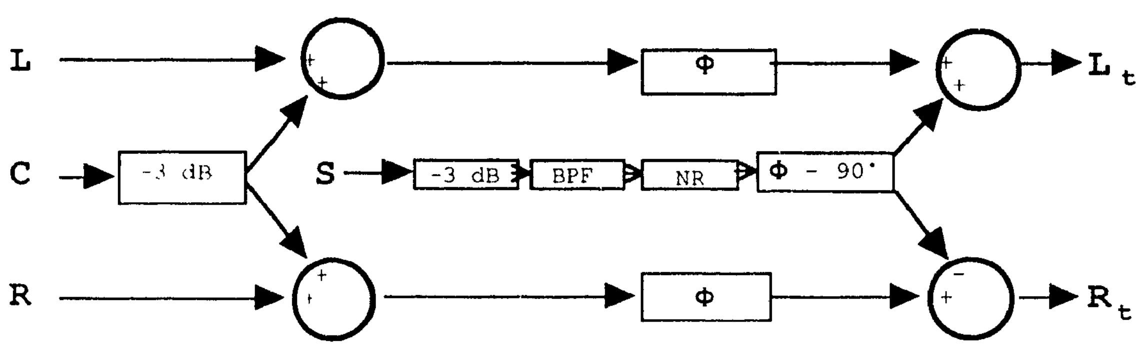


Figure 2. Dolby Surround
Encoder

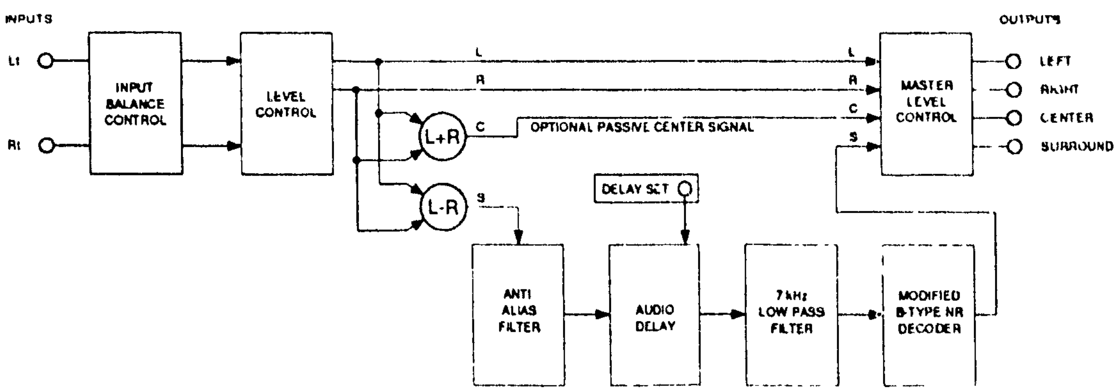


Fig. 3 Passive surround decoder block diagram

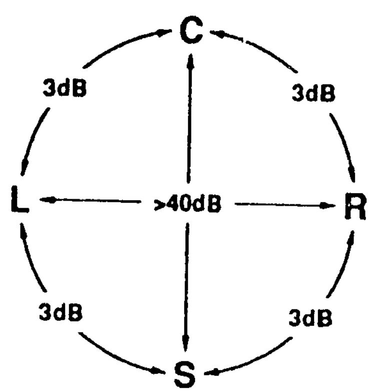


Fig. 4. Four output separation map.

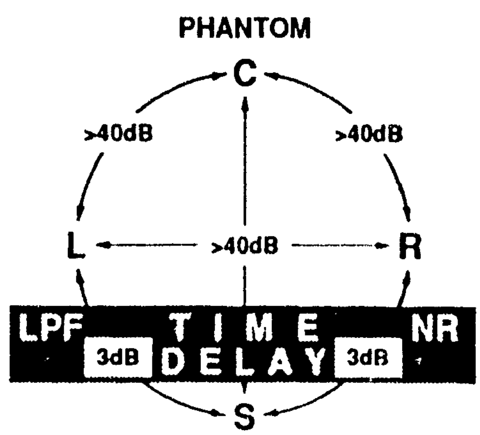


Fig. 5. Three output separation map.

Figures 3,4,5 reprinted by kind permission of Dolby Labs from "Dolby Pro Logic Surround Decoder: Principles of Operation." by Roger Dressler [2].

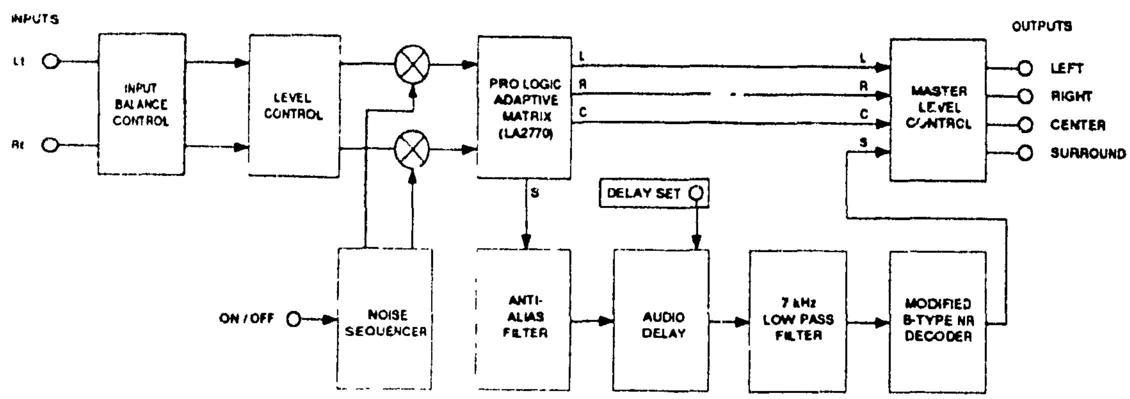


Fig.6 Pro Logic Decoder
Block Diagram

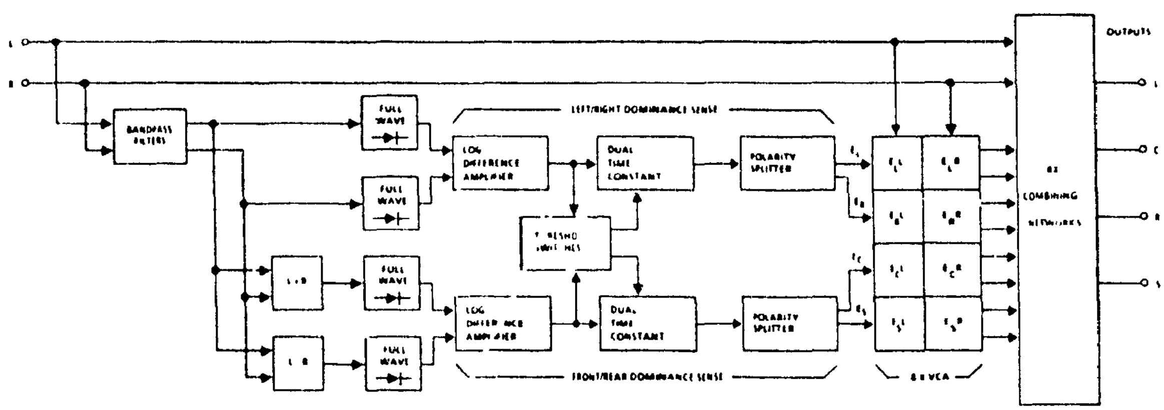


Fig 7 Pro Logic adaptive matrix

Figures 6 and 7 reprinted by kind permission of Dolby Labs from "Dolby Pro Logic Surround Decoder: Principles of Operation." by Roger Dressler [2].

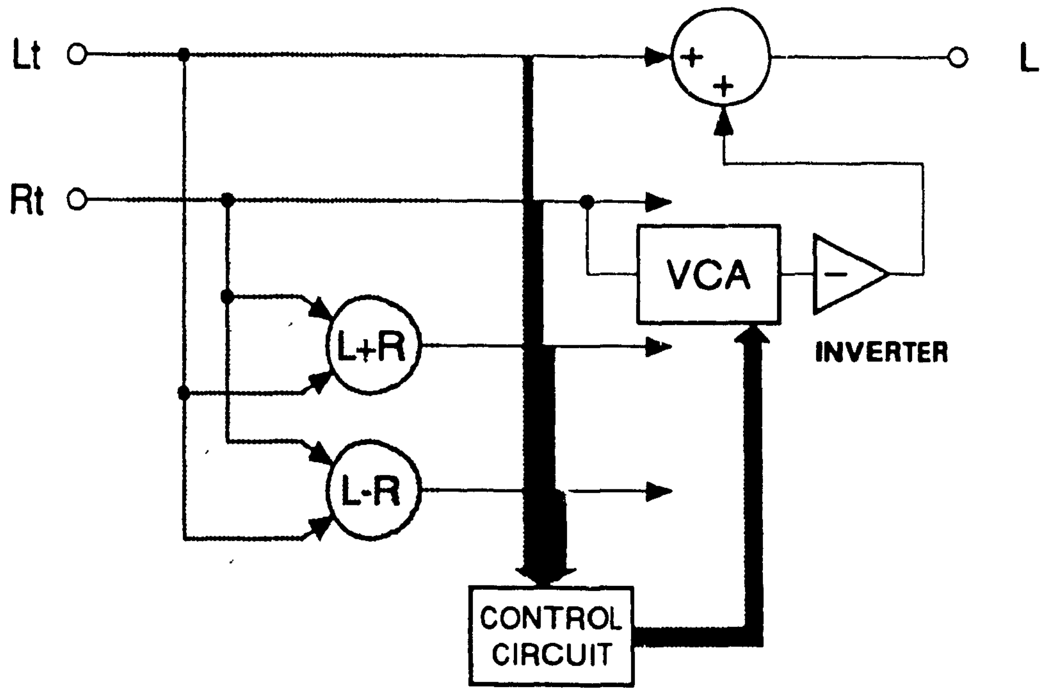


Fig.8 Signal cancellation concept

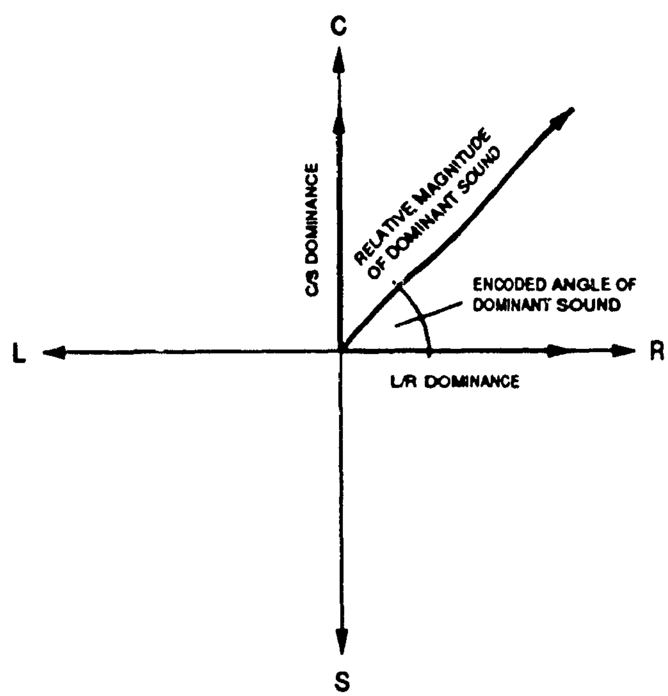


Fig.9 Resolving the dominance vector

Figures 8 and 9 reprinted by kind permission of Dolby Labs from [2]

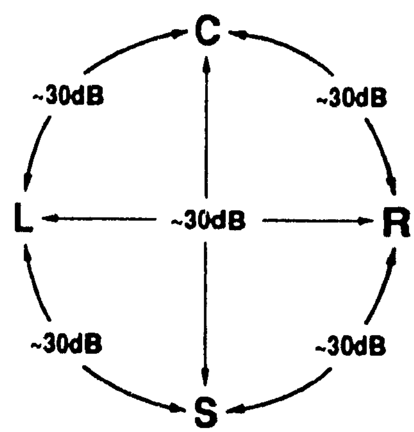


Fig.10 Pro Logic output separation map.
Reprinted by kind permission of Dolby
Labs from "Dolby Pro Logic Surround Decoder:
Principles of Operation." by Roger Dressler.

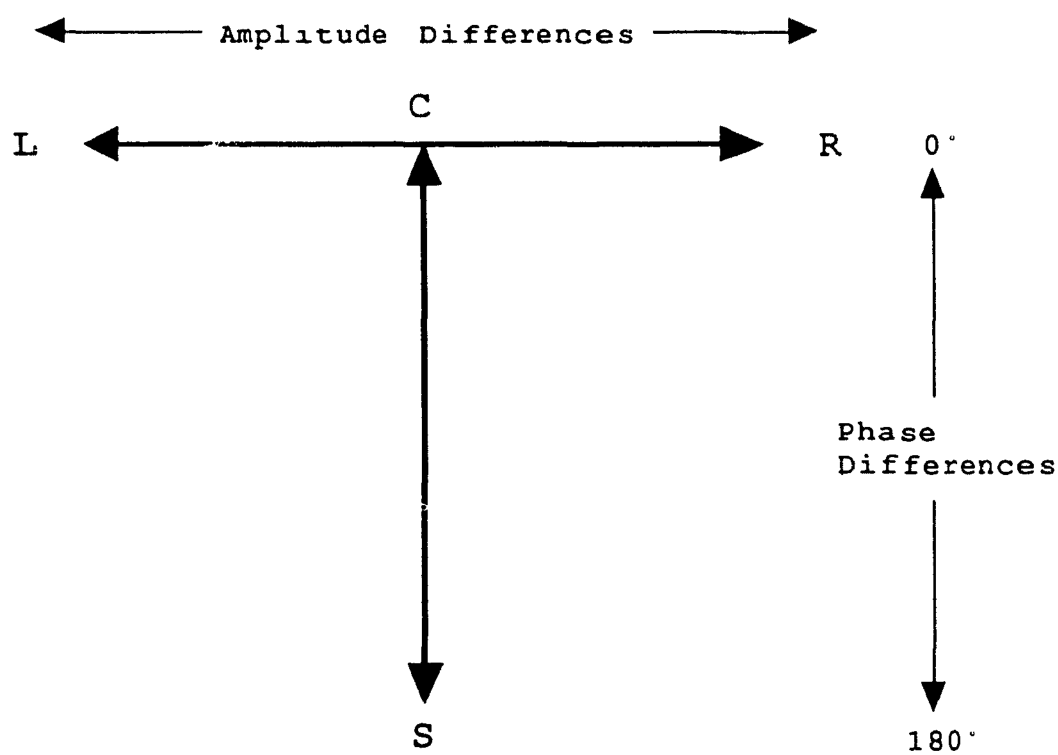


Fig.11 How the decoder senses
direction

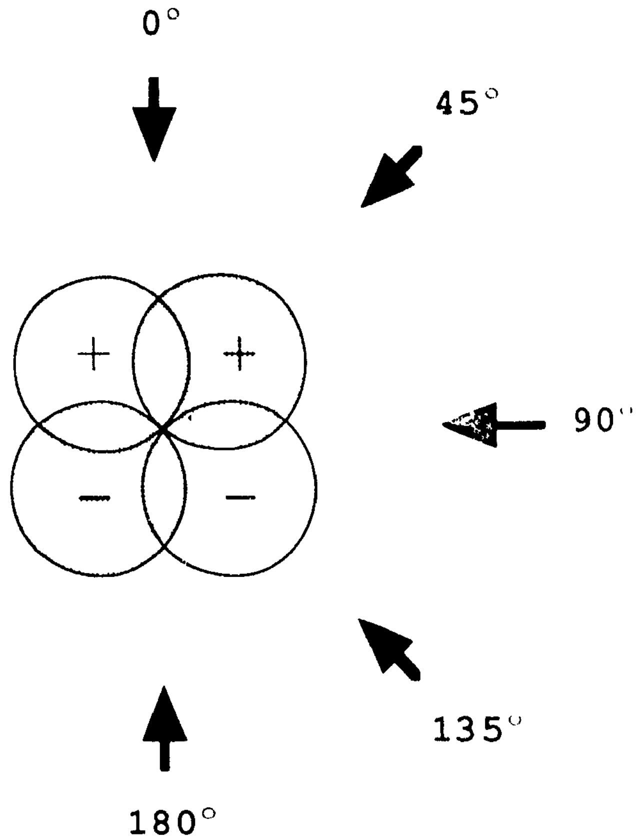


Fig.12 The Blumlein technique

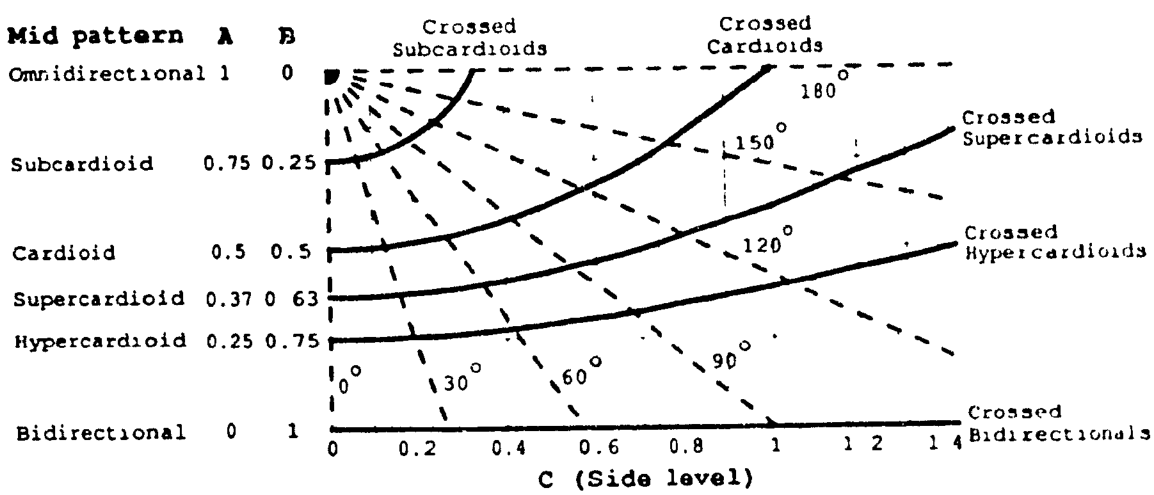


Fig. 13 Relating XY pairs to MS microphones

Figures 13 and 14 are reprinted from [22] Julstrom, Stephen. "An Intuitive View of Coincident Stereo Microphones." Presented at the 89th AES Conventionin Los Angeles, Sept 1990. Preprint #2984

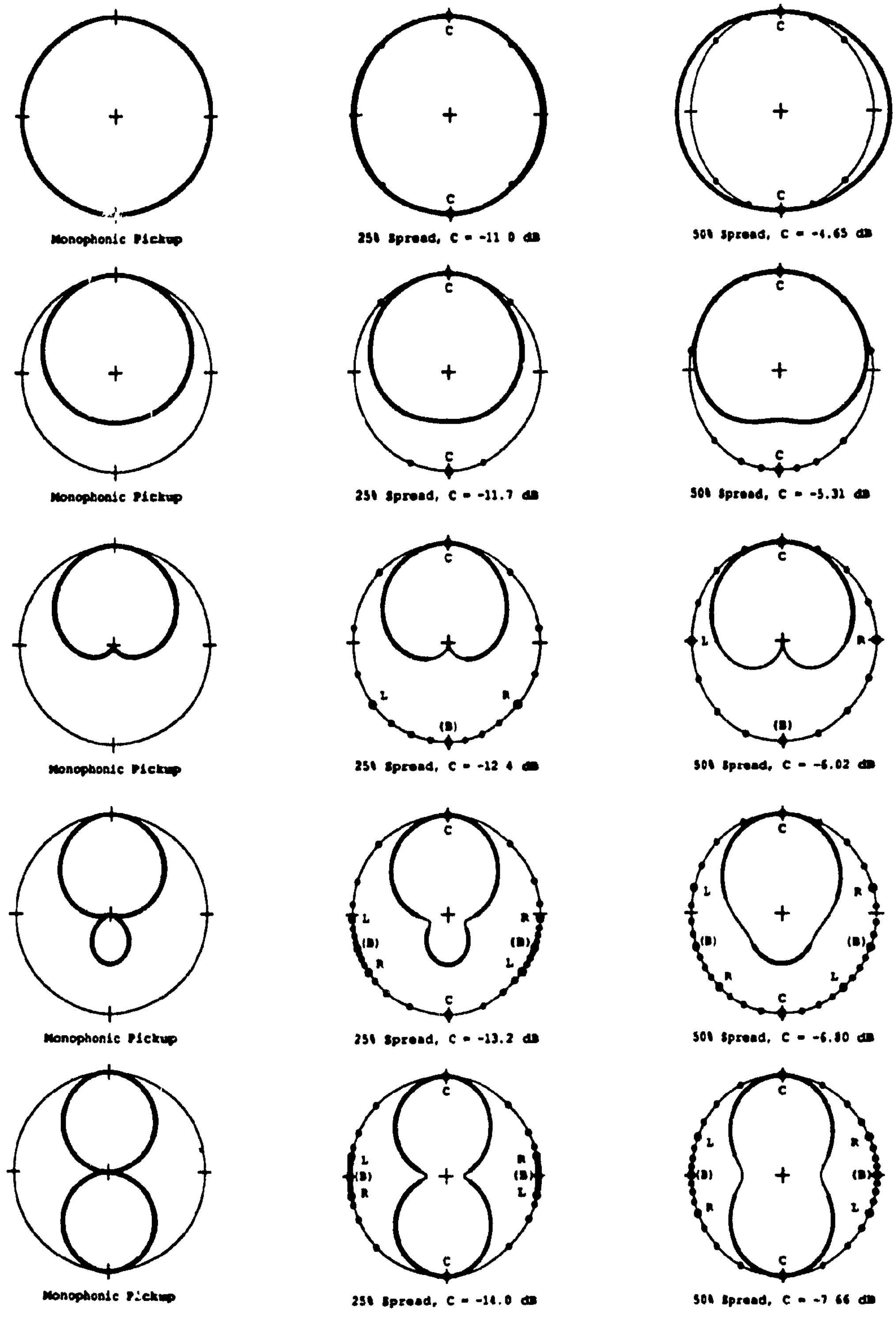
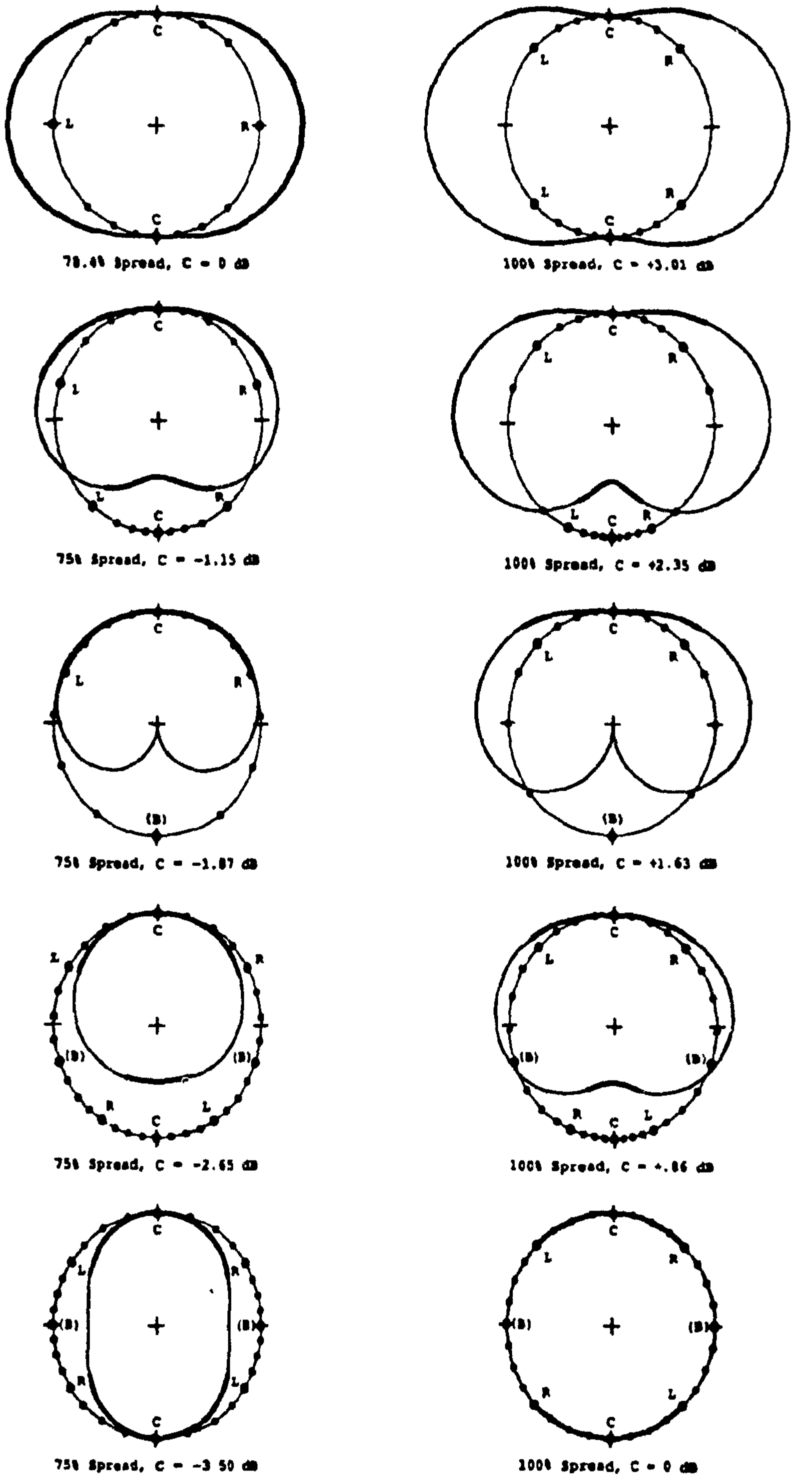


Fig 14. Two-dimensional stereo polar diagrams for various stereo encodings covering

Reprinted from Julstrom [22] by Kin



Omnidirectional

Subcardioid

Cardioid

Hypercardioid

Bidirectional

ng the range of Mid patterns and the range of front quadrant stereo spreads to 100%.

nd permission of JAES

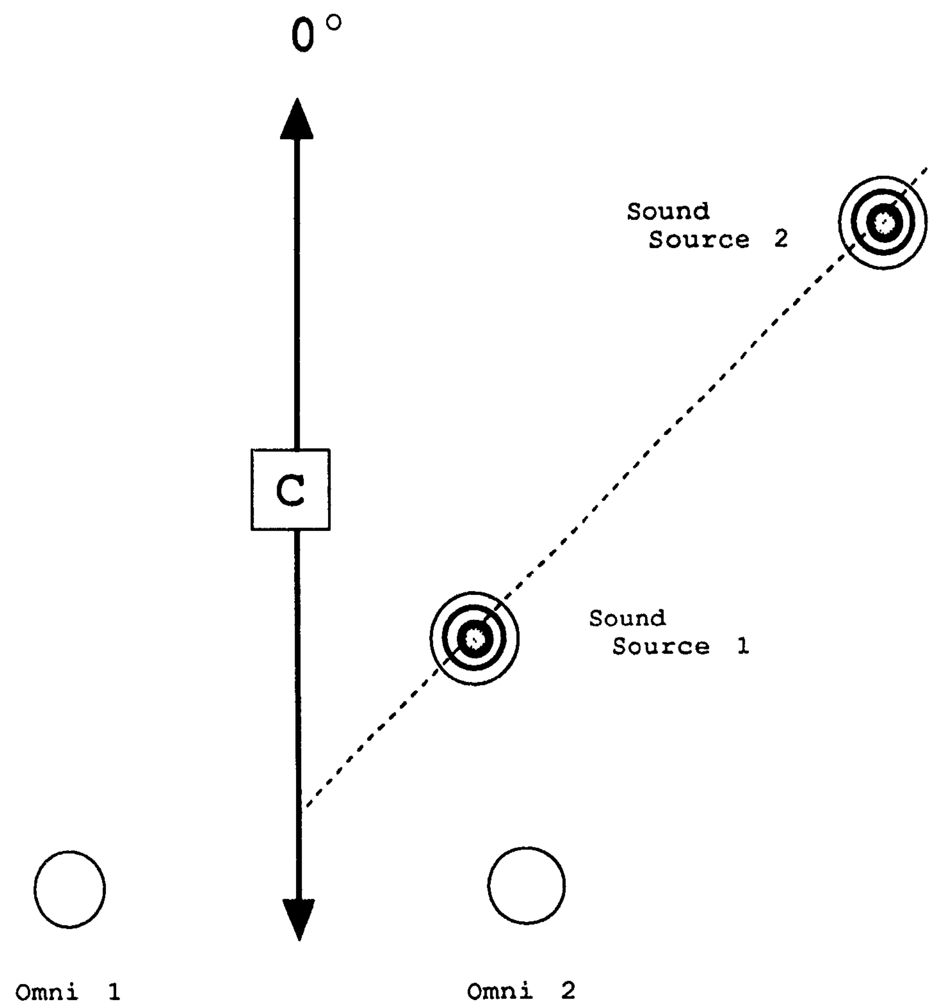


Fig.15 Directional encoding with
the spaced omni technique

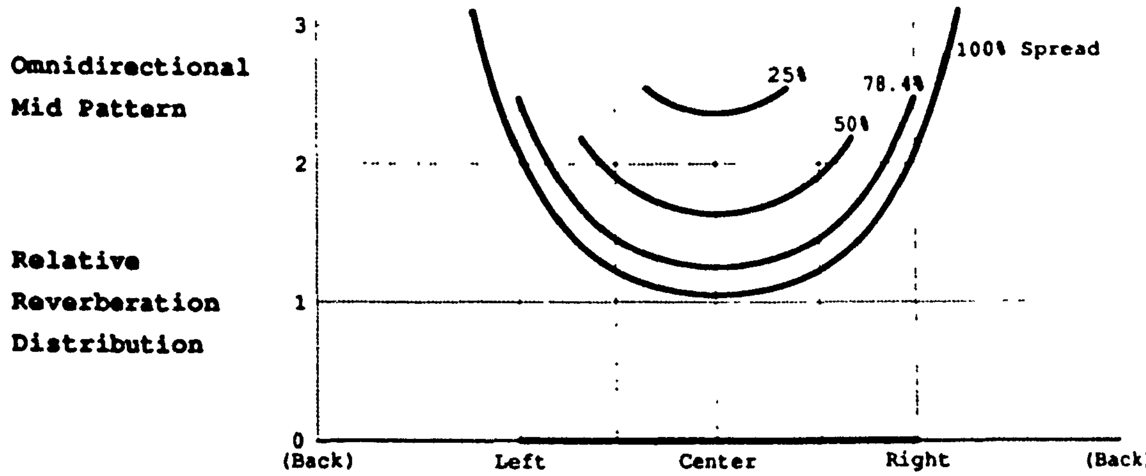


Fig.16a Directional encoding of reverberation for omnidirectional Mid–based stereo encodings with selected Side levels corresponding to the indicated front quadrant stereo spreads.

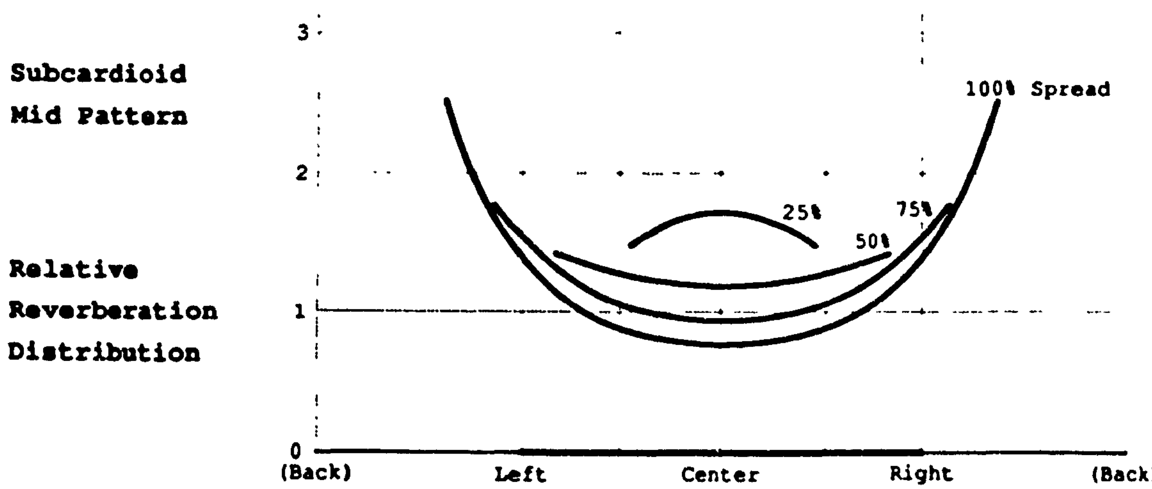


Fig.16b Directional encoding of reverberation for subcardioid Mid–based stereo encodings with selected Side levels corresponding to the indicated front quadrant stereo spreads.

Figures 16a through f are reprinted from Julstrom [22] by kind permission of the JAES.

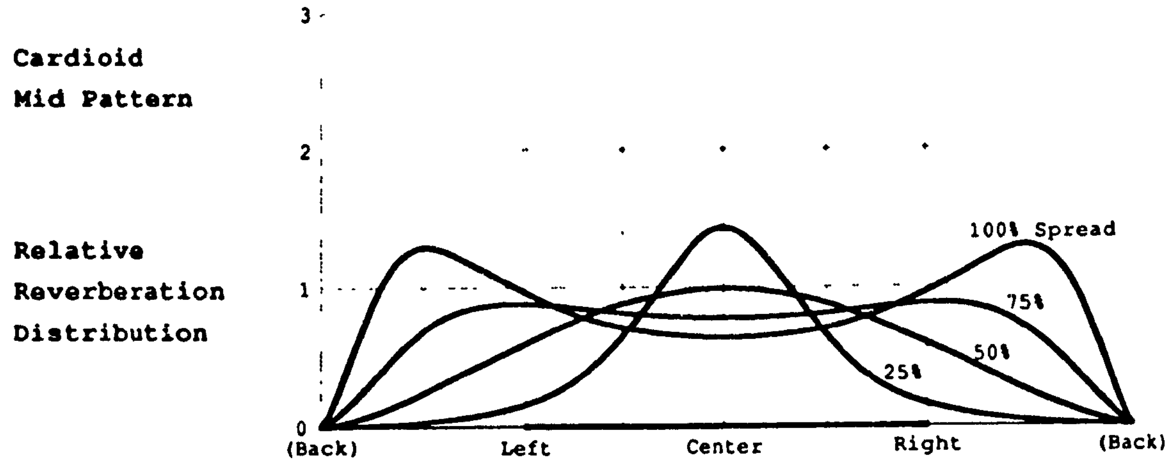


Fig.16c Directional encoding of reverberation for cardioid Mid– based stereo encodings with selected Side levels corresponding to the indicated front quadrant stereo spreads.

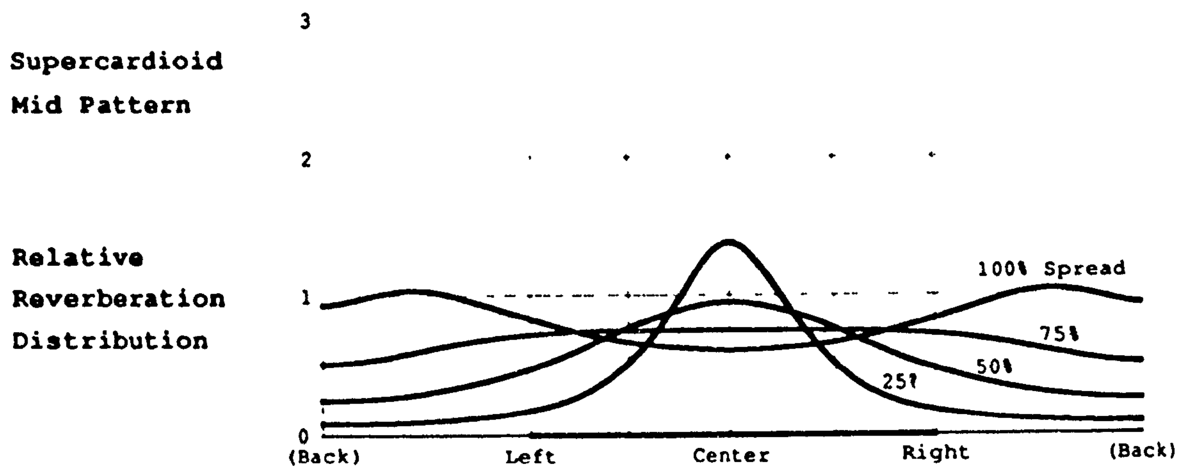


Fig.16d Directional encoding of reverberation for supercardioid Mid–based stereo encodings with selected Side levels corresponding to the indicated front quadrant stereo spreads.

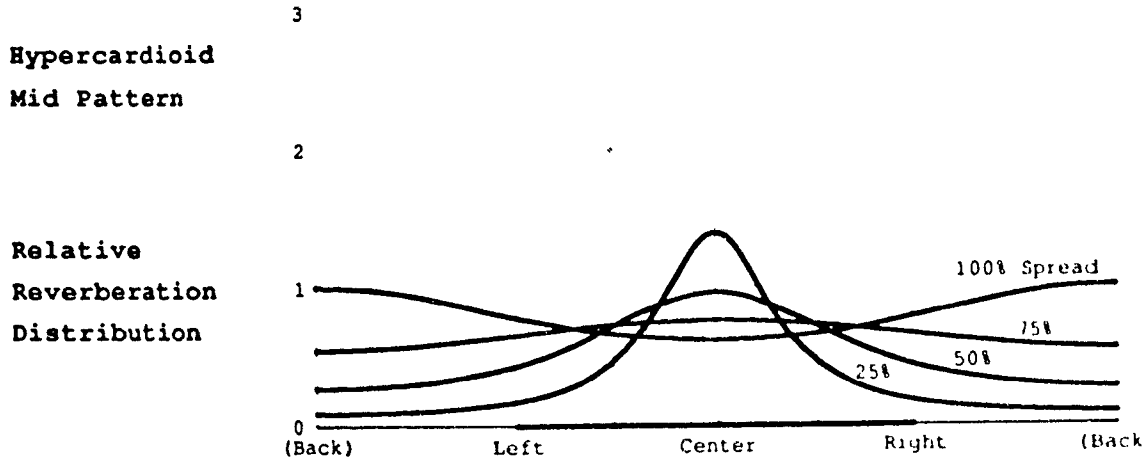


Fig.16e

Directional encoding of reverberation for hypercardioid Mid- based stereo encodings with selected Side levels corresponding to the indicated front quadrant stereo spreads

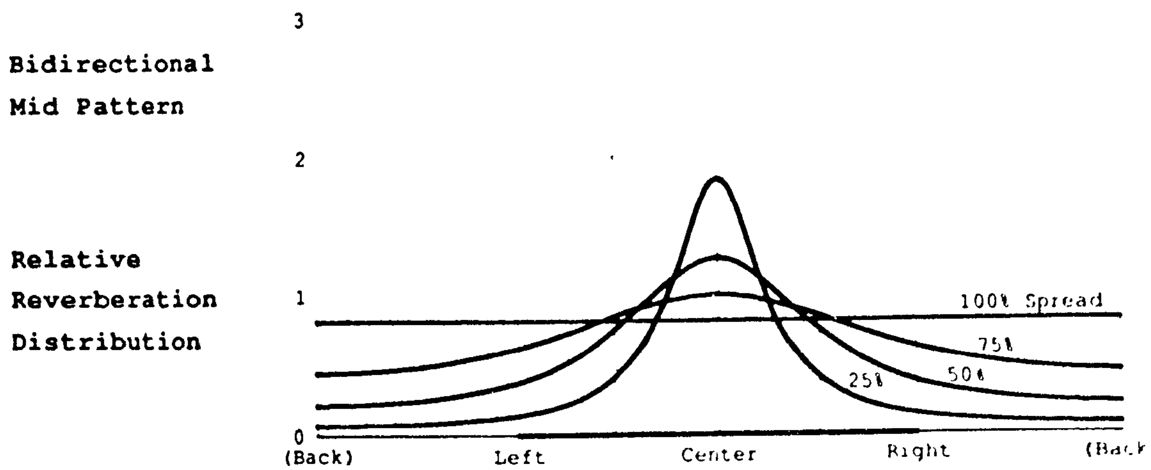


Fig.16f

Directional encoding of reverberation for bidirectional Mid-based stereo encodings with selected Side levels corresponding to the indicated front quadrant stereo spreads

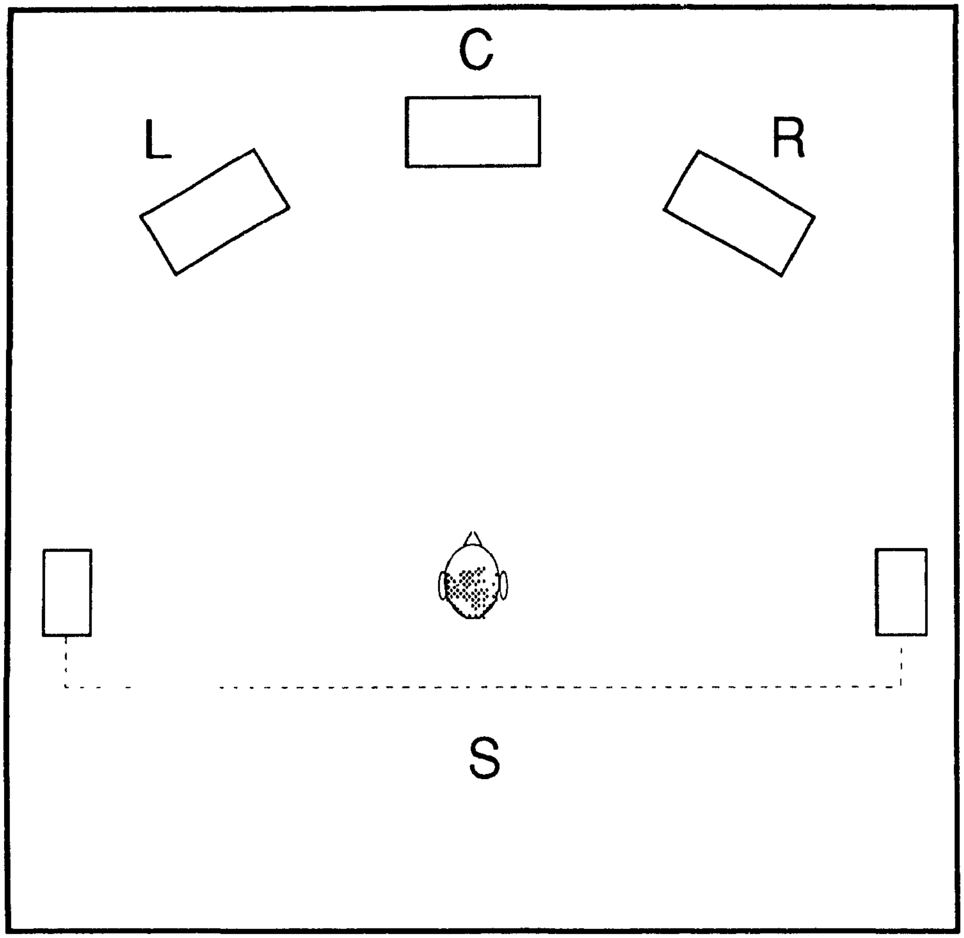


Fig.17 Improved Dolby Surround
Loudspeaker Layout